# AIR POLLUTION AND DIABETES MELLITUS

A Dissertation

by

RAED KHALID M ALOTAIBI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PUBLIC HEALTH

| | |
|---|---|
| Chair of Committee, | Jennifer A. Horney |
| Committee Members, | Xiaohui Xu |
| | Natalie Johnson |
| | Haneen Khreis |
| Head of Department, | Xiaohui Xu |

August 2020

Major Subject: Epidemiology and Environmental Health

# ABSTRACT

Background: Air pollution is a leading contributor to the burden of disease globally. Recent systematic reviews suggested that air pollution may cause diabetes mellitus among adults. However, the number of studies included in these reviews were small and the confidence intervals were wide. Leading health practitioners have called on studies to quantify the burden of disease due to air pollution and examine health disparities associated with such burden. In response, our study aims to update previous reviews with recent studies and to quantify the burden of diabetes due to air pollution in the United States (US) while examining health disparities.

Method: We conducted a systematic review and meta-analysis of studies examining exposure to air pollution in the form of Nitrogen dioxide ($NO_2$), Black Carbon (BC), and Ultra Fine Particles (UFP) and the risk of developing diabetes mellitus among adults. Using joined the concentration-response function of the pooled estimate with air pollution, census, and diabetes prevalence and incidence across the US to produce burden estimates. We explored health disparities across geographical and social stratum. Finally, we developed accessible interactive maps and tables to visualize and explore the burden of disease across counties.

Results: Our search yielded 21 studies included in our analysis. We found that exposure to $NO_2$ increased the risk of developing diabetes among adults OR = 1.05 [1.04-1.05, I2 = 95%] per 10 $\mu g/m^3$ unit increase. For BC and UFP, we could not reach a similar conclusion since the number of included studies was small. We estimated that around 5,978,048 prevalent and 213, 641 incident diabetes cases may be attributable to air pollution exposure representing 28.1% and 11.0% of all diabetes prevalent and incident cases, respectively. The fraction of attributable cases

were higher in urban areas compared to rural areas, and in census blocks with a predominantly Asian population and lower-income groups.

Conclusion: This study updates the current knowledge of exposure to air pollution and the risk of developing diabetes mellitus, quantifies the burden of disease to air pollution exposure, explores the health disparity associated with the burden of disease, and presents interactive tools that make our results accessible.

# DEDICATION

This dissertation is dedicated to my parents Fatima Althubaiti and Khalid Alotaibi. To my loving wife Amal Almutairi and children Khalid, Faisal, and Nouf, whose never-ending support made this possible. Also, to my siblings Matar, Kholoud, Razan, Bashair, Faris, Saif, Zayed, Lulu, and Aisha.

ACKNOWLEDGMENTS

I would like to thank the many people who supported me in this endeavor. I thank my parents for their inspiration, drive, and support. I thank my wife, Amal, for her never-ending love and push to get this over with.  Thanks to my children Khalid, Faisal, and Nouf who were always banging the buttons on my laptop to get my attention.

I thank my committee chair, Jennifer Horney, for her encouragement and inspiration. I also would like to thank my committee, Haneen Khreis, Natalie Johnson, and Xiaohui Xu for their knowledge, expertise, and assistance.

A special thanks goes to Abdulla Aljoudi for being a mentor, friend, and brother. Thanks to all my friends and colleagues who made this journey memorable and fun. I am grateful for your friendship.

## CONTRIBUTORS AND FUNDING SOURCES

### Contributors

This work was supervised by a dissertation committee consisting of Professors Jennifer Horney and Xiaohui Xu of the Department of Epidemiology and Biostatistics, Professor Natalie Johnson of the Department of Environmental and Occupational Health, and Dr. Haneen Khreis of the Texas A&M Transportation Institute. Each committee member invested valuable time and critique to this project.

The risk of bias form for Appendix – B was provided by Dr. Inyang Uwak with adjustments by the student. The search strategy for the systematic review was in collaboration with Dr. Margaret Foster of the Texas A&M University Libraries.

All other work conducted for the dissertation was completed by the student independently.

### Funding Sources

Graduate study was supported by a scholarship from the Saudi Arabian Cultural Mission and Imam Abdulrahman Bin Faisal University.

## NOMENCLATURE

| | |
|---|---|
| DM | Diabetes Mellitus |
| IDF | International Diabetes Federations |
| US | United States |
| CDC | Centers for Disease Control and Prevention |
| BMI | Body mass index |
| ROS | Reaction oxygen species |
| TRAP | Traffic-related air pollution |
| $NO_x$ | Nitrogen oxides |
| $NO_2$ | Nitrogen dioxide |
| $O_2$ | Oxygen |
| $CO_2$ | Carbon dioxide |
| CO | Carbon monoxide |
| HC | Hydrocarbons |
| PM | Particulate matter |
| BC | Black carbon |
| UFP | Ultra-fine particles |
| GBD | Global burden of disease |
| SES | Socioeconomic status |
| RR | Relative risk |
| OR | Odds ratio |
| HbA1C | Hemoglobin A1C |
| LUR | Land use regression |

| | |
|---|---|
| NHGIS | National Historical Geographic Information System |
| USDSS | United States Diabetes Surveillance System |
| BRFSS | Behavioral Risk Factor Surveillance System |
| EPA | Environmental Protection Agency |
| GIS | Geographical information systems |
| CRF | Concentration-response function |
| AF | Attributable fraction |
| AC | Attributable cases |
| TMREL | Theoretical minimum risk exposure level |

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# 1. INTRODUCTION

## 1.1 Background

### 1.1.1 Diabetes Mellitus

Diabetes mellites (DM) are a group of metabolic disorders characterized by elevated blood glucose levels (hyperglycemia) over prolonged periods that occurs due to a defect in insulin secretion, action or both (WHO, 2019). Diabetes presents with multiple signs and symptoms including excessive thirst, excessive urination, excessive hunger, weight loss, and blurred vision. In many cases, the diagnosis goes undetected for several years. Diabetes increases the risk of all-cause mortality and the development of adverse health outcomes spanning multiple organ systems. Adverse outcomes of the cardiovascular system include coronary artery disease, myocardial infarctions, stroke, atherosclerosis, and loss of limbs (ADA, 2003; Fox et al., 2004; Lotufo et al., 2001; Lundberg et al., 1997; Miettinen et al., 1998; Murabito et al., 1997). Adverse outcomes of the neurological system include peripheral neuropathy, depression, dementia, and Alzheimer's disease (R. J. Anderson et al., 2001; Biessels et al., 2006; Boulton et al., 2005). Diabetes is also a leading cause of renal damage, end-stage renal disease, retinopathy, and blindness (Brancati et al., 1997; Fong et al., 2004). Diabetes causes immunosuppression increasing the risk of contracting infectious diseases including tuberculosis. Around 15% of tuberculosis infections globally are linked to diabetes (Kim et al., 1995; Stevenson et al., 2007; WHO, 2019).

### 1.1.2 Epidemiology

Diabetes is increasing globally. The International Diabetes Federation (IDF) estimated 451 million people (18-99 years) were living with diabetes in 2017 (Cho et al., 2018). Half of the

451 million cases aren't aware they have diabetes. The prevalence of diabetes varies across regions. In Africa, it is around 4.4%, while in North America and the Caribbean it reaches 11%. With current trends, it is expected the number of cases will reach 693 million by 2045. Moreover, the IDF estimated that in 2017, 374 million people were living with prediabetes, also known as impaired glucose tolerance (Cho et al., 2018). In the United States (US), the Centers for Disease Control and Prevention (CDC) estimated that in 2017, 30.3 million people were living with diabetes and 84.1 with prediabetes. One in every four individuals with diabetes in the US is not aware that they are diabetics, while nine out of every 10 do not know they have prediabetes (CDC, 2017b).

### 1.1.3 Burden

The health impact and cost of diabetes is a growing burden. In 2017, around 5 million global deaths among adults (20-99 years) were attributable to diabetes, representing 9.9% of the global all-cause mortality among the same age range (Cho et al., 2018). In terms of health care expenditure, the IDF estimates that around $850 billion has been spent globally in 2017 due to diabetes among those aged 18-99, while expenditure is expected to increase to $958 billion by 2045 (Cho et al., 2018). In the US, the estimated total cost due to diabetes increased from $248 billion in 2012 to $327 billion in 2017. Of the total US cost due to diabetes, direct medical cost increased from $176 billion (2012) to $237 billion (2017) which includes hospital inpatient care, prescription medication, diabetes supplies, and healthcare office visits. The cost of lost work time and wages increased from $69 billion (2012) to $90 billion (2017)(ADA, 2018; CDC, 2017b).

### 1.1.4 Pathophysiology

The body maintains blood glucose levels by regulating several hormones, most notably insulin which lowers the blood glucose, and glucagon which increases blood glucose levels. When blood glucose levels increase, the β-cells in the islets of Langerhans in the pancreas are triggered to secrete insulin in the blood. Insulin lowers blood glucose levels by promoting the transfer of glucose from the blood into fat and muscle cells. Diabetes occurs when there is either a deficiency of insulin secretion, or the body becomes insensitive to secreted insulin. Multiple mechanisms can lead to diabetes including genetic predisposition, auto-immune diseases, inflammation, and environmental factors (WHO, 2019).

### 1.1.5 Diagnosis and Classification

Diabetes is diagnosed using either of the following methods: a) random blood glucose ≥ 200 mg/dl (≥ 11.1 mmol/L) with the presence of signs and symptoms of diabetes, b) fasting blood glucose ≥ 126 mg/dl (≥ 7.0 mmol/L), c) two-hour postprandial blood glucose ≥ 200 mg/dl (≥11.1 mmol/L), or d) Hemoglobin A1C ≥ 6.5%. If blood glucose levels are elevated without signs and symptoms of diabetes, repeated testing is warranted to confirm the diagnosis (ADA, 2019b; WHO, 2019).

Diabetes is a heterogeneous disease with hyperglycemia as a common feature. There are multiple subtypes of diabetes with different etiology, natural history, pathophysiology, disease consequences, and treatment (Leslie et al., 2016). However, due to resource limitations, a simple system for classifying diabetes takes into account the clinical and management setting only (ADA, 2019a; WHO, 2019). The main classifications of diabetes are a) type 1 diabetes, b) type 2 diabetes, c) other or special types of diabetes, and d) gestational diabetes. Type 1 diabetes is caused by an autoimmune reaction leading to the destruction of insulin-secreting β-cells in the pancreas. Insulin secretion eventually drops until glucose control becomes impaired(Atkinson et al., 2014). The degree of insulin secretion depletion can happen progressively over time or abruptly. In many cases, diabetes is first identified when patients present with ketoacidosis which is a life-threatening condition due to insufficient levels of insulin in the body (Wolfsdorf et al., 2009). Patients with type 1 diabetes are insulin-dependent and need daily insulin intake to maintain normal blood glucose levels for survivability(Daneman, 2006). The incidence and prevalence of type 1 diabetes are generally not known. However, studies have shown the prevalence is increasing worldwide due to the increase in survivability of patients with type 1 diabetes because of the wide availability and accessibility of treatment (Dabelea et al., 2014; You et al., 2016). Although type 1 can develop at any age, it occurs more frequently among children and adolescents (Dabelea et al., 2014). Type 2 diabetes occurs when the body becomes insensitive to insulin due to a dysfunction in β-cells. In many cases, this is followed by a drop-in insulin secretion over time. Although patients with type 2 diabetes initially do not require insulin treatment to survive, it is often the case that insulin secretion becomes increasingly deficient making exogenous insulin intake necessary (Weyer et al., 1999). Type 2 diabetes accounts for more than 90% of diabetes cases globally (WHO, 2019). Type 2 diabetes is more common in adults, however, it is becoming increasingly diagnosed among children (WHO, 2016b). Special and other types of diabetes are rare and include the following: monogenic diabetes syndrome, diabetes initiated due to illnesses affecting the function of the pancreas, chemical-induced diabetes, and other types (ADA, 2019b; WHO, 2019). Finally, gestational diabetes is defined as diabetes first diagnosed during the second or third trimester of pregnancy.

### 1.1.6 Etiology

Diabetes can be caused by both modifiable and non-modifiable risk factors (Baker et al., 2011; Forouzanfar et al., 2016; Howells et al., 2016; Vazquez et al., 2007). Non-modifiable risk factors include genetic predispositions, family history, age, and gender. Modifiable risk factors include lifestyle (eating habits, physical activity), increased body mass index (BMI), and smoking. Recently, there has been emerging evidence that indicates exposure to air pollution might increase the risk of developing diabetes (Eze et al., 2015; Landrigan et al., 2018; Wang et al., 2014). The mechanism in which air pollution impacts human health is by oxidative stress(Health Effects Institute, 2010). Oxidative stress occurs when an imbalance between pro-oxidants and anti-oxidants occurs, leading to the release of reaction oxygen species (ROS) also known as "free radicals". ROS alter the biological structure of the body and cells through the activation of signaling pathways that trigger inflammation leading to target cells and organ damage (Sies, 1997). In-vitro and In-vivo studies show that exposure to air pollutants promotes the release of ROS (Olefsky et al., 2010; Shoelson et al., 2006). Oxidative stress also damages systems linked to glycemic control, possibly leading to diabetes. Mice exposed to $PM_{2.5}$ develop several body changes including visceral inflammation (Sun et al., 2009), altered energy metabolism (C. Liu et al., 2014), and increased hippocampal inflammation which may lead to dysregulation of metabolic control and insulin resistance (Fonken et al., 2011) (Table 1-1)

### 1.1.7 Traffic-Related Air Pollution and Exposure Assessment

Ambient air pollution is sourced from industry, mining, agriculture, electric generation, and motor vehicle combustion (Forouzanfar et al., 2016; Prüss-Üstün et al., 2016). It is a major source of the burden of health. A more specific type of ambient air pollution is Traffic-related air pollution (TRAP). TRAP is primarily sourced from motor vehicles in areas with a high aggregation of motor vehicles and people. TRAP is a major source of ambient air pollution (Health Effects Institute, 2010). Motor vehicles emit large quantities of chemicals from combustion and non-combustion processes. Combustion processes result from the burning of a fuel source. Chemicals emitted include carbon dioxide ($CO_2$), carbon monoxide (CO), nitrogen oxides ($NO_x$), hydrocarbons (HC), particulate matter (PM), and other chemicals. non-combustion emissions are from wear and tear of the vehicle, tires, brakes, road, oil spills, and resuspension of air particles from the ground by moving vehicles. Emissions of non-combustion sources include heavy metals, organic materials, PM, among other chemicals (Health Effects Institute, 2010). Epidemiological studies use several methods to measure TRAP exposure including a) assigning exposure status through buffer zones using a distance to road and/or traffic volume metric, and b) measuring or modeling concentrations of chemicals emitted by traffic as surrogates or indicators of exposure. Several chemicals are frequently used as indicators of exposure including CO, $NO_2$, Ozone, elemental or black carbon (BC), PM, and ultra-fine particles (UFP). Chemicals have different characteristics with varying degrees of specificity to TRAP as a competing source of air pollution. Of these chemicals, $NO_2$ and black carbon are more specific compared to others like PM (Health Effects Institute, 2010).

## 1.2 Study Rational

### 1.2.1 Systematic Review and Meta-Analysis

Air pollution is a leading cause of mortality and morbidity around the globe (Landrigan et al., 2018). Multiple studies indicate that exposure to air pollution causes multiple noncommunicable diseases and increases all-cause mortality (Bateson et al., 2004; Beelen et al., 2009a; Beelen et al., 2007). Non-communicable diseases include cardiovascular diseases (Brook et al., 2004), respiratory diseases (H. Anderson et al., 2011, 2013; Brauer et al., 2002; Gowers et al., 2012; Khreis et al., 2017), renal (Bowe et al., 2018b), and other non-communicable diseases (Health Effects Institute, 2010). Studies examining exposure to air pollution and the development of diabetes mellitus have increased recently (Balti et al., 2014; Eze et al., 2015; Wang et al., 2014). Wang et al. (2014) conducted a review of 10 cohort studies examining the effect of long-term exposure to particulate matter <2.5 $PM_{2.5}$, $PM_{10}$, and $NO_2$ and the risk of developing type 2 diabetes before 2014. The studies included controlled for important risk factors including age, gender, BMI, smoking status, physical activity, and socioeconomic status. The review concluded that there was positive evidence of the adverse effect of long-term exposure to air pollution and the risk of developing type 2 diabetes. Balti et al. (2014) conducted a review examining exposure to $NO_2$, $PM_{2.5}$, and $PM_{10}$ and the risk of type 2 diabetes. The review included 10 studies concluding that there was a positive association between pollutants and diabetes. Eze et al. (2015) conducted a review examining only studies conducted in North America or Europe. The review included 7 studies and concluded that there is a positive association between exposure to $NO_2$ and $PM_{2.5}$ and the risk of developing diabetes. However, the number of studies included in all the previous reviews for each pollutant was limited in number and confidence intervals of the effect measures were relatively wide with a high level of heterogeneity. Since 2014, several studies have been published examining the association between air pollution exposure and the risk of developing diabetes. The rationale for our study is to incorporating recently published data by conducting a systematic review and meta-analysis thus updating the current state of knowledge regarding exposure to air pollution and the risk of developing diabetes.

### 1.2.2 The Burden of Diseases Assessment and Health Disparity

According to the global burden of disease (GBD) report, there were an estimated 6.5 million deaths in 2015 attributable to air pollution exposure. Of these deaths, 4.2 million were attributable to ambient air pollution (Forouzanfar et al., 2016). The trend of mortality due to ambient air pollution is projected to increase by more than 50% by 2050 without the necessary intervention (Lelieveld et al., 2015). The costs associated with adverse health outcomes due to air pollution are growing and burdening economic systems around the world. However, air pollution prevention can reverse the impact on the economic system. In the US it was estimated that for every dollar spent on preventing and mitigating air pollution a return of $30 was realized (Environmental Protection Agency, 2011). The burden of disease from air pollution affects individuals and countries of lower socio-economic indexes to a greater magnitude compared to individuals and countries of a higher socio-economic index. It is estimated that more than 89% of deaths due to ambient air pollution occurred in low-income and middle-income countries in 2015 (Forouzanfar et al., 2016).

The literature on the burden of air pollution on human health is limited. A recent commission by health experts "The Lancet Commission on pollution and health" has called on studies to quantify the burden of air pollution on health (Landrigan et al., 2018). Recent advances in air pollution monitoring techniques and the availability of easily accessible air pollution and health data at fine geographical levels make assessing the burden of disease possible on a large scale. Our rationale for conducting a burden of diabetes due to air pollution is to fill the knowledge gap on the magnitude of the burden of diseases due to air pollution and to assess whether health disparities exist. We also believe that the burden of disease data should be easily accessible and thus we aim to create interactive tools that are easy to use and are accessible for researchers and the general population to explore the burden of air pollution.

## 1.3 Aims and Objective

Aim 1: Assess whether exposure to air pollution increases the risk of developing diabetes mellitus among an adult population.

Objective 1.1: Conduct a systematic review and meta-analysis on studies examining exposure air pollution measured in the form of $NO_2$, BC, or UFP and the risk of developing diabetes mellitus among adults.

Aim 2: Quantify the burden of disease of diabetes mellitus due to air pollution exposure in the United States.

Objective 2.1: Estimate the number of prevalent and incident diabetes cases attributable to air pollution exposure in the United States.

Objective 2.2: Estimate the fraction of prevalent and incident diabetes cases attributable to air pollution exposure in the United States.

Objective 2.3: Compare the burden of diabetes attributable to air pollution exposure by state, county, and urban vs rural areas.

Aim 3: Evaluate the health disparities of the burden of diabetes due to air pollution in the United States.

Objective 3.1: Compare the burden of diabetes attributable to air pollution exposure by median household income and race.

Aim 4: Create accessible interactive tools to visualize and explore the health burden across the United States.

Objective 4.1: Create an interactive map showing the burden of disease by county.
Objective 4.2: Create an interactive lookup table summarizing the burden of disease by county.

Table 1-1: Toxicological mechanism

| **Pollutant Exposure** | **Effect** |
|---|---|
| Oxidative stress. | |
| Black carbon and diesel-exhaust PM | Increased expression of heme-oxygenase (oxidative-stress-response gene) (Koike et al., 2006). |
| $PM_{2.5}$ | Curbside $PM_{2.5}$ had a higher ROS generation than $PM_{2.5}$ from an urban background location (Baulig et al., 2004). |
| Gasoline exhaust | Increased expression of mRNA for proteins related to oxidative stress (Lund et al., 2006). |
| Immune | |
| Air pollution | Potentiated inflammation might lead to insulin resistance in mice (Sun et al., 2009). |
| Liver | |
| $PM_{2.5}$ | Mice exposed to $PM_{2.5}$ developed non-alcoholic steatohepatitis (NASH) - like phenotype, hepatic steatosis, inflammation, and fibrosis. Also, impaired glycogen storage, glucose intolerance, and insulin resistance (Zheng et al., 2013). |
| Adipose and metabolism | |
| $PM_{2.5}$ | Mice exposed to $PM_{2.5}$ had altered energy metabolism, $O_2$ consumption, $CO_2$ production, and thermogenesis. These changes were accompanied by insulin resistance, visceral adipose tissue, and inflammation. Also, the expression of inflammatory genes leading to decrease expression of brown adipose tissue (C. Liu et al., 2014). |
| CNS | |
| $PM_{2.5}$ | Increased hippocampal inflammation which is hypothesized to result in dysregulation of metabolic control (Fonken et al., 2011). |
| $PM_{2.5}$ | Increased sympathetic activation may involve hypothalamic inflammation (Ying et al., 2013). |
| Other | |
| Ozone | Increased plasma endothelin-1 resulting in endothelial dysfunction and vasoconstriction (Vincent et al., 2001). |
| Diesel | Impaired blood flow and systemic regulation of vascular dilatation responses (Mills et al., 2005). |
| CO | Competes with $O_2$ resulting in tissue hypoxemia and impaired cellular function (Allred et al., 1989). |

# 2. SYSTEMATIC REVIEW AND META-ANALYSIS

## 2.1 Introduction

Diabetes is a group of chronic metabolic diseases characterized by elevated blood glucose levels caused by defective insulin secretion from the pancreases, insensitivity to secreted insulin, or both (WHO, 2019). Diabetes has several modifiable non-modifiable risk factors including genetic predisposition, an increase in body mass index, smoking, and a sedentary lifestyle to name a few. Recent studies have linked diabetes with exposure to air pollution (Brook et al., 2008b). The mechanism believed to be responsible for such a link is oxidative stress, in which air pollution causes the release of free radicles which then trigger inflammation of organs that are involved in the glycemic control within the body including altering the energy metabolism, hippocampal damage, liver changes, and nervous system alterations (Table 1-1). Systematic reviews and meta-analysis have shown that air pollution exposure in the form of $NO_2$, $PM_{2.5}$, and $PM_{10}$ are possibly associated with an increased risk of developing diabetes among adults (Balti et al., 2014; Eze et al., 2015; Wang et al., 2014). However, the number of included studies were small, and the confidence intervals were wide. We aim to update previously published meta-analysis by including more recent literature on air pollution exposure and the risk of developing diabetes mellitus in the form of $NO_2$. We will also search for other pollutants not examined in previous reviews including UFP and BC.

## 2.2 Methods

### 2.2.1 Study Question and Eligibility Criteria

The objective of the study is to answer the following question: “Does exposure to air pollution in the form of $NO_2$, UFP, or BC increase the risk of type 2 diabetes mellitus among adults?”. We used the following eligibility criteria:

- Population**:** Studies that included a non-institutionalized adult population ≥18 years of age.
- Exposure: Studies investigating exposure to air pollution in the form of $NO_2$, BC, or UFP measured as a continuous exposure.
- Type of study: Observational studies including cohort, case-control, or cross-sectional studies.
- Outcome: Studies that report or assume type 2 diabetes mellitus as the outcome will be included. Diagnosis of diabetes can either be self-reported, lab-based, or using secondary data sources.

### 2.2.2 Search Method

We searched the electronic databases MEDLINE, EMBASE, Transportation Research Information Services (TRIS) Database and the OECD's Joint Transport Research Centre's International Transport Research Documentation (ITRD) for literature published up to Oct 30, 2019, using the following terms: “nitrous oxide”, “nitrogen dioxide”, “$NO_2$”, “black carbon”, “carbon black”, “soot”, ultrafine particles”, “ultrafine particulate”, “UFP”, “UFPS”, “diabetes”, and “diabetes mellitus” (see Appendix). Reference lists of included studies were hand-searched for additional eligible studies.

### 2.2.3 Study Selection

We conducted a title/abstract screening for duplication and eligibility using Rayyan – a software to organizing and working with systematic reviews (Ouzzani et al., 2016). Studies measuring human exposure to $NO_2$, BC, and UFP and risk of diabetes mellitus were eligible for a full-text review. We excluded studies for the following reasons:

- Not observational (i.e. reviews, reports, or letters).
- Only examined childhood or maternal exposure (i.e. pregnancy), institutionalized populations (i.e. nursing homes and prisons), industrial, agricultural, or indoor-only exposures.
- Conducted on non-human subjects (i.e. animals, or tissue).
- Exposure to $NO_2$, BC, UFP was not measured, or estimated, for the individual (i.e. using national mean levels of exposure).
- Diabetes status was not assessed.

### 2.2.4 Data Extraction

We extracted detailed information on study design, population characteristics, exposure assessment, outcome assessment, and risk of bias information from each included study (see below for a description of each criterion). Study authors were contacted in case data were missing for the analysis.

- Study design: years of study, location, type of study, study objective, funding source, source population.
- Population characteristics: age of the population, female (%), total participants, number of cases.
- Exposure assessment: pollutants, exposure definition, exposure assessment method, exposure summary.

- Outcome assessment: type of diabetes examined, source, outcome definition, outcome measure reported.

### 2.2.5 Risk of Bias Assessment

We adapted the navigation guide methodology (Johnson et al., 2014) for the following; a) assessing the risk of bias, b) rating the quality of evidence, and c) measuring the strength of evidence. The risk of bias of each included study was assessed individually across eight domains; recruitment, blinding, exposure assessment, confounding, incomplete outcome data, selective outcome reporting, conflict of interest, and other sources of bias. The rating of each domain was assigned as either “low risk”, “uncertain”, or “high risk” using a predetermined form (see Appendix). To assess the overall quality of the body of evidence an initial rating of “moderate” was assigned (Balshem et al., 2011; Viswanathan et al., 2012) then either upgraded or downgraded based on several considerations including:

- Downgrading: risk of bias across studies, indirectness, inconsistency, imprecision, publication bias.
- Upgrade: the large magnitude of the effect, dose-response, confounding minimizes effect, overall quality of evidence.

Finally, the overall strength of the body of evidence was evaluated based on the following four factors: a) quality of the body of evidence; b) direction of effect estimate; c) confidence in effect estimate; and d) other attributes of the data that may influence certainty (IARC, 2006; Sawaya et al., 2007).

### 2.2.6 Statistical Methods

We used a fixed-effects model as the main pooling method for the effect estimates. The null hypothesis was as follows:

$$\text{Ho: The common effect size} = 1.$$

We also reported the pooled estimates using a random-effects model as a sensitivity measure, in which the null hypothesis was as follows:

$$\text{Ho: The mean effect size} = 1$$

The fixed-effect model assumes only one source of variation, variation between the observed mean and a true mean shared by all studies. The random effect model has two sources of variation; a) variation between the observed mean and a true mean for each study, and b) variation between the true mean of a study and a grand mean (Borenstein et al., 2010). We reported heterogeneity with the $I^2$ metric and the between-study variance with the $Tau^2$ using the DerSimonian and Laird method (DerSimonian et al., 1986).

### 2.2.7 The Combined Effect, Weighting Scheme, and Uncertainty Measure

The combined effects were estimated by taking the weighted mean effect across all studies ($\overline{Y}$) using the following formula (Borenstein et al., 2010):

$$\overline{\mathrm{Y}} = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

Where

- $W_i$ = inverse variance weight for study i
- $Y_i$ = Observed effect of study i
- k = the number of studies

The inverse variance weight ($W_i$) has two sources of variation; the within-study variance ($V_i$), and the between-study variance ($T^2$) using the following formula:

$$W_i = \frac{1}{V_i + T^2}$$

The within-study variances ($V_i$) is estimated as follows:

$$V_i = \frac{\sigma^2}{n}$$

- $\sigma^2$ = variance of individual observations in the sample
- n = sample size

The between-study variance ($T^2$) is estimated as follows:

$$T^2 = \frac{Q - df}{C}$$

In which (Q) is the weighted squared difference between the observed effect ($Y_i$) and the weighted mean effect ($\overline{Y}$):

$$Q = \sum_{i=1}^{k} W_i (Y_i - \overline{Y})^2 = \sum_{i=1}^{k} \frac{(Y_i - \overline{Y})^2}{V_i}$$

The degrees of freedom (df) is as follows:

$$df = k - 1$$

While the denominator (C) is as follows:

$$C = \sum W_i - \frac{\sum W_i{}^2}{\sum W_i}$$

We estimate a 95% lower and upper confidence intervals using the standard error ($SE_{\overline{Y}}$) of the weighted mean effect ($\overline{Y}$) as follows:

$$SE_{\overline{Y}} = \sqrt{V_{\overline{Y}}}$$

where ($V_{\overline{Y}}$) is the meta-analysis error variance:

$$V_{\overline{Y}} = \frac{1}{\sum_{i=1}^{k} W_i}$$

and lower and upper 95% confidence intervals:

$$\text{Lower and Upper 95\% CI } = \overline{Y} \pm 1.95 * SE_{\overline{Y}}$$

### 2.2.8 Effect Measure Selection

We considered the reported odds ratios as equivalent to risk ratios. In our main analysis when studies reported more than one effect measure, we selected the "main model" or "fully adjusted model". If these were not reported than the choice of the model was as follows:

- If more than one exposure model was reported, we chose the land-use regression model followed by the dispersion model.
- If multiple exposure durations were reported then longer exposure durations were chosen over shorter exposure durations.
- The most restrictive model in terms of adjustments was selected.
- The most inclusive in terms of population (i.e. both genders vs one gender, all age groups vs limited age group).
- Single pollutant model over multipollutant models.

$NO_2$ concentrations reported in "ppb" were converted to "µg/m$^3$" using the following formula (WHO, 2016a):

$$\text{Concentration } (\mu/\text{m}^3) = 0.0409 * \text{concentration(ppb)} * \text{molecular weight}$$
$$\approx \text{concentration(ppb)} * 1.88$$

Further, reported effect measures were standardized by converting the reported exposure increments to standardized increments as follows:

- $NO_2$; Per 10 ug/m$^3$
- BC; Per $10^{-5}$/m
- UFP; Per $10^4$ count/cm$^3$

Using the following formula:

$$\text{OR}_{(\text{standardized})} = \text{reported OR}^{(\text{standardized increment/reported increment})}$$

### 2.2.9 Subgroup and Sensitivity Analysis

We conducted subgroup analysis for the following variables; minimum age of inclusion in the study, gender, location of the study, exposure model used, incidence vs prevalence, and diabetes ascertainment source. In case studies reported more than one effect measure that includes subgroups (i.e. reported effect measure by gender and an effect estimate by separate exposure models) we extracted the reported effect measure for each subgroup separately. Sensitivity analysis was conducted using a funnel plot and a linear regression test of funnel plot asymmetry also known as Egger's test (Egger et al., 1997). We also tested whether adding a new study would potentially shift the combined effect estimate to a) overlap the confidence interval, and b) cross the null value (Johnson et al., 2014). All analysis was conducted using R (R Core Team, 2019) and the "meta-package" (Schwarzer et al., 2012).

## 2.3 Results

### 2.3.1 Search Results and Study Characteristics

The database search yielded 243 articles of which 9 were duplicates (Figure 2-1). 193 articles were excluded after the title and abstract screening. Full-text review of 41 articles returned 22 studies that met our inclusion criteria and 21 included in the quantitative analysis. Of the 22 included studies, 10 reported a longitudinal design (Andersen et al., 2012b; Bai et al., 2018; C. Clark et al., 2017; Coogan et al., 2016; Eze et al., 2017; Hansen et al., 2016; Honda et al., 2017; Krämer et al., 2010; Lazarevic et al., 2015; Renzi et al., 2018), 11 cross-sectional (Dijkema et al., 2011; Eze et al., 2014a; Howell et al., 2019; Li et al., 2017; F. Liu et al., 2019; O'Donovan et al., 2017; Orioli et al., 2018; Riant et al., 2018; Shin et al., 2019; Yang et al., 2018), and 1 case-control (Brook et al., 2008b) (Table 2-1). Regarding the location of studies, 7 of the studies were conducted in North America (4 in Canada; 3 in the US), 11 studies in Europe

(1 in Germany, 2 in the Netherlands, 2 in Italy, 2 in Denmark, 2 in Switzerland, 2 in the United Kingdom, and 1 in France), 3 studies in Asia (2 in China, and 1 in Korea), and 1 study in Australia. The earliest study was available in 2008 (Brook et al., 2008b), and the latest in October-2019 (F. Liu et al., 2019). Study exposure measurement periods started from as early as 1990 (Krämer et al., 2010) to 2017 (F. Liu et al., 2019). The total sample size of all included studies was 6,357,054 ranging from 704 (Li et al., 2017) up to 2,496,458 (Howell et al., 2019). The total number of reported diabetes cases was 748,812 ranging from 73 reported cases (Li et al., 2017) up to 292,086 (Howell et al., 2019). Also, 4 studies only recruited female participants (Coogan et al., 2016; Hansen et al., 2016; Krämer et al., 2010; Lazarevic et al., 2015) while the remaining studies recruited both males and females (Table 2-1).

Diabetes was defined using several criteria including self-report, physician diagnosis, hospital admission/discharge, anti-diabetic medication prescription and/or intake, and lab testing (Table 2-2). The lab testing included the following criteria: Non-fasting blood glucose > 11.1 mmol/L (≥2 g/L); Fasting glucose ≥7.0 $mmol \cdot L^{-1}$ (≥1.26 g/L); Two-hour glucose ≥11.0 $mmol \cdot L^{-1}$ (≥2 g/L); HbA1c of ≥ 6.5% (48 mmol/mol). 7 studies indicated that type-2 diabetes was the main outcome of interest while the remaining studies either did not indicate the type or did not differentiate between type-1 & 2 (Table 2-2).

Regarding exposure of interest, 21 studies examined exposure to $NO_2$ (Andersen et al., 2012b; Bai et al., 2018; Brook et al., 2008b; C. Clark et al., 2017; Coogan et al., 2016; Dijkema et al., 2011; Eze et al., 2017; Eze et al., 2014a; Hansen et al., 2016; Honda et al., 2017; Howell et al., 2019; Krämer et al., 2010; Lazarevic et al., 2015; F. Liu et al., 2019; O'Donovan et al., 2017; Orioli et al., 2018; Renzi et al., 2018; Riant et al., 2018; Shin et al., 2019; Strak et al., 2012; Yang et al., 2018), 4 studies reported exposure to BC or Soot (Bai et al., 2018; C. Clark et al., 2017; Krämer et al., 2010; Strak et al., 2012), and 2 studies examined exposure to UFP (Bai et al., 2018; Li et al., 2017). Exposure was assessed using multiple modeling methods including LUR, monitoring station, dispersion models, distance to road, emission inventories, traffic flow at the nearest road, hybrid models, and others (Table 2-2).

### 2.3.2 Risk of Bias Assessment

The risk of bias was generally low across the 22 included studies (Figure 2-2). We found that incomplete outcome data was the most common type of bias followed by exposure assessment. Studies that relied only on self-reporting of diabetes or secondary data sources were assigned a high risk of bias compared to studies that actively ascertained diabetes diagnosis through testing. Studies that used exposure assignment through air quality monitors only were assigned a high risk of bias compared to studies that used a validated air pollution model.

### 2.3.3 Statistical Analysis

We included 49 effect measures in our pooled analysis from the 21 studies across all pollutants (Table 2-3). The summary effect in odds ratios (OR) of studies reporting exposure to $NO_2$ and risk of diabetes mellitus (n = 20 studies) using a fixed-effect model was 1.05 [1.04-1.05, $I^2$ = 95%] per 10 μg/$m^3$ increase (Figure 2-3). By study design, the pooled effect of incident diabetes as the outcome (n = 8) was 1.02[1.01-1.02, $I^2$=95%] per 10 μg/$m^3$ increase, and for prevalent diabetes (n = 13) was 1.05[1.04-1.05, $I^2$=95%] per 10 μg/$m^3$ (Figure 2-4). By gender, the OR for females (n = 14) was 1.02[1.01-1.03, $I^2$=90%] per 10 μg/$m^3$, and for males (n = 10) 1.03[1.02-1.04, $I^2$=90%] per 10 μg/$m^3$ (Figure 2-5). By minimal age of inclusion, studies that included a population of >=18 years (n = 11) had a pooled OR of 1.03[1.03-1.04, $I^2$=95%] per 10 μg/$m^3$, >=40 years (n = 6) reported an 1.06[1.05-1.06, $I^2$=96%] per 10 μg/$m^3$, and >=50 years (n = 3) reported an OR of 1.13[1.07-1.19, $I^2$=44%] per 10 μg/$m^3$ (Figure 2-6). By location, studies conducted in North America (n = 6) reported an OR of 1.06[1.06-1.07, $I^2$=97%] per 10 μg/$m^3$, in Europe (n = 10) an OR of 1.01[1.01-1.02, $I^2$=86%] per 10 μg/$m^3$, and in Asia (n = 3) an OR of 1.10[1.07-1.12, $I^2$=93%] per 10 μg/$m^3$ (Figure 2-7). By exposure model, studies using LUR models (n = 9) had a pooled effect of 1.05[1.04-1.05, $I^2$=97%] per 10 μg/$m^3$, dispersion models (n = 3) an OR of 1.02[1.01-1.04, $I^2$=67%] per 10 μg/$m^3$, air monitors (n = 4) an OR of 1.07[1.05-1.10, $I^2$= 71%] per 10 μg/$m^3$, and other models (n = 7) an OR of 1.14[1.10-1.18, $I^2$=81%] per 10 μg/$m^3$ (Figure 2-8). By outcome definition, studies using self-reported diabetes (n = 6) reported an OR of 1.05 [1.03-1.06, $I^2$=84%] per 10 μg/$m^3$, using lab results (n = 7) reported an OR of 1.20 [1.16-1.25, $I^2$=73%] per 10 μg/$m^3$, and studies using secondary data sources (n = 7) reported an OR of 1.05 [1.04-1.05, $I^2$=98%] per 10 μg/$m^3$ (Figure 2-9).

The summary effect (OR) of studies reporting exposure to BC and risk of diabetes mellitus (n = 4 studies) using a fixed-effect model was 1.02 [1.01-1.03, $I^2$ = 87%] Per 1 ($10^{-5}$/m) increase (Figure 2-10). Not enough studies were available to warrant subgroup analysis using BC as the outcome of interest.

The summary effect (OR) of studies reporting exposure to UFP and risk of diabetes mellitus (n = 2 studies) using a fixed-effect model was 1.06 [1.04-1.07, $I^2$ = 78%] Per 10,000 count/$cm^3$ (Figure 2-11). Not enough studies were available to warrant a subgroup analysis using UFP as the outcome of interest.

### 2.3.4 Quality of the Body of Evidence

We tested the effect size needed to shift the confidence interval to a) overlap the null value, and b) move below the null value. We assumed a study with a standard error of (0.0038) equal to the smallest standard error in our metanalysis (Howell et al., 2019). To overlap the null value, a study with an effect estimate of 0.87 and a 95% CI of [0.86-0.88] is needed. To mover below the null value, a study with an effect size of 0.84 and a 95% CI of [0.83-0.85] is needed. We did not find signs of publication bias (asymmetry) on the funnel plot (Figure 2-12). Furthermore, the Egger's test (Egger et al., 1997) was insignificant ($t = 0.629$, $df = 19$, p-value = 0.537), indicating no evidence of asymmetry. In conclusion, we assigned an overall rating of "moderate" for the quality of the body of evidence (Table 2-4). We did not upgrade or downgrade the rating based on any of the criteria.

### 2.3.5 Strength of the Body of Evidence

We assigned a "Sufficient" rating for the overall strength of the body of evidence-based on the following considerations (Table 2-4):

- Quality of body of evidence: moderate
- The direction of effect estimate: Increasing exposure to $NO_2$ resulted in an increased risk of diabetes.
- Confidence in the effect estimate: an introduction of a new study is unlikely to change the confidence interval of the pooled estimate towards a null value or beyond.
- Other compelling attributes of the data: none.

## 2.4 Discussion

### 2.4.1 Main Results

We utilized an adapted version of the navigation guide (Johnson et al., 2014) to determine whether exposure to air pollution in the form of $NO_2$, BC, and UFP increases the risk of diabetes among adults. Our search yielded 20 studies that examined exposure to $NO_2$, 4 studies examining exposure to BC, and only two examining exposure to UFP. We have concluded that there is sufficient evidence of an association between exposure to $NO_2$ and risk of diabetes among adults based on several considerations (Table 2-4); moderate quality of the body of evidence that included several well designed and conducted studies with a low risk of bias, an effect estimate with a positive direction where the risk of diabetes is increasing with increasing exposure to $NO_2$, a pooled effect with a narrow confidence interval with a direction of effect that is unlikely to reverse or reach the null value with an addition of a new study, and a consistent direction of effect estimates among smaller studies except for a few. We were not confident in making the same conclusion regarding exposure to BC nor UFP on the risk of diabetes due to the small number of studies included in the analysis.

When assessing causality an important factor to consider is consistency in effect under similar circumstances. Although most of the included studies had an effect estimate in the positive direction, the magnitude was variable with a high level of heterogeneity. In the next part of the discussion, we will discuss where the heterogeneity could be coming from including confounding, subgroup analysis, and bias in measurement.

### 2.4.2 Confounding

The most controlled variable across included studies was socioeconomic status (Figure 2-13). However, the definition and criteria for choosing socioeconomic factors varied across studies. Marshall et al. (2014) conducted a literature search of peer-reviewed articles examining exposure to air pollution and environmental injustice and found 307 articles of which (88%) of those conducted in the US showed a higher than average risk or exposure to air pollution among racial minorities and/or groups of lower socioeconomic status defined as poor, lower education, or a combination of both. Hajat et al. (2015) conducted a meta-analysis of studies addressing unequal exposure of environmental hazards on a certain population and concluded that most North American studies have shown that areas where populations of a lower SES dwelling experienced a higher concentration of criteria pollutants. However, the results were mixed for the European region, while other parts of the world showed a similar trend to north America. When examining the risk of diabetes across social strata Agardh et al. (2011) conducted a systematic review and meta-analysis of 23 studies with 41 effect estimates (16 cohort and 7 cross-sectional) which were conducted across the US (n=10), EU (n=7) and other regions (n = 5). The outcomes measured were across educational attainment, occupation, and income. Results of the pooled effect showed a positive association across all the outcomes of education, occupation and income and risk of diabetes with an RR of 1.41(1.28-1.55, $I^2$= 65.5%) for education, 1.31(1.09-1.57, $I^2$=52.8%) for occupation, and 1.40(1.04-1.88, $I^2$ = 71.9%) for Income. The effect was also positive and statistically significant for education across the different regions (US, Europe, Asia, Latin America, Africa).

When comparing the pooled effect estimate of the fully adjusted models with the unadjusted (crude) models and found a smaller effect size for the adjusted models with a narrower confidence interval, and less heterogeneous than the crude models (Figure 2-14). Although residual confounding likely remains due to the nature of observational studies, we concluded that we could rule out confounding as an explanation of the effect found between exposure to $NO_2$ and the risk of developing diabetes with reasonable confidence.

### 2.4.3 Comparison with Previous Studies and Subgroup Analysis

The positive association between exposure to air pollution and the risk of diabetes was consistent with previous studies. Wang et al. (2014) performed a meta-analysis of 6 studies exploring exposure to $NO_2$ and risk of diabetes and reported a RR of 1.12[1.02-1.23, I2 = 63.5%] per 10ug/$m^3$ increment. The pooled estimate was higher with a wider confidence interval. This can be attributed to a smaller number of studies included compared to our analysis (6 vs 21 studies). Eze et al. (2015) also conducted a meta-analysis that included 4 studies examining exposure to $NO_2$ and risk of diabetes and reported an elevated RR of 1.08[1.00-1.17] per 10 ug/m3 increment.

Our results showed a positive and significant association between exposure to $NO_2$ and the risk of diabetes for both males and females. For females, the results were consistent with Eze et al. (2015) who reported an effect estimate of 1.15[1.05-1.27] and Wang et al. (2014) who reported an effect estimate of 1.09[1.02-1.15]. However, for males, Eze et al. (2015) and Wang et al. (2014) did not find a significant association. We found a positive association for studies reporting prevalence and incidence using a fixed-effect model. Eze et al. (2015) and Balti et al. (2014) reported a positive association of 1.12[1.05-1.19] and 1.13[1.04-1.22], respectively, across longitudinal studies. However, the results were based on a small number of studies. Balti et al. (2014) also reported a positive association across two cross-sectional studies 1.16[1.00-1.35]. By location, The pooled effect was positive across all locations. Studies conducted in Europe were less heterogenous compared to other locations. By exposure assessment methods, the pooled effect was positive across all studies using different methods. Stratifying by exposure assessment did explain part of the heterogeneity across the studies. By outcome ascertainment, the pooled effects were all positive across all methods. However, the effects for studies using an active ascertainment method to detect cases showed a larger magnitude of effect compared to studies using secondary sourced for case ascertainment. Studies with active ascertainment on average had fewer samples and used a case definition that is likely to be more sensitive compared to studies using secondary ascertainment which were more likely to include a larger number of samples and use a more specific definition of diabetes (Table 2-1 & Table 2-2). Studies recruiting older age groups reported a larger magnitude of effect compared to studies that included younger age groups. One possible explanation is increasing risk with age due to cumulative exposure if individuals remained in higher exposure areas throughout their life compared to individuals who lived in less exposed areas throughout their life.

### 2.4.4 Limitations

The statistical methods used in a pooled analysis assumes that no measurement error occurs, and the only source of error is a random error only (i.e. sampling and randomization error) (Carroll et al., 2006). However, observational studies are known to contain non-random errors (systematic error or bias) that occur from various sources including selection bias, information bias, and residual confounding. Although pooling the effect reduces the random error it increases the magnitude of systematic error as a proportion of the total error, thus statistical significance, in this case, might not imply a cause and effect but indicates the need to investigate the sources of the systematic error (Rothman et al., 2008).

Across the pooled estimates there was high heterogeneity which persisted across the subgroup analysis despite some reduction. We expect that variation of the effect estimate among observational studies to occur due to the different methods used in the design of such studies. Sources of heterogeneity can be explained partially by differences in population sources, characteristics, the number of variables and methods of adjustment across models, exposure assessment methods, and finally how the outcome was defined and ascertained.

There were several limitations in exposure assessment. First, studies varied in their air pollution measurements including using different instruments, varying duration of measurements, and modeling techniques. For example, some studies used air monitor readings, while others used statistical models like land-use regression and dispersion which also vary in how they model air pollution concentrations. Studies also varied in their methods of assigning air pollution concentration levels with some using mean concentration levels while others used median levels and whether a lag time between exposure and outcome was considered or not (Table 2-2). Second, it is not clear whether assigning an exposure level based on residential location reflects well with personal exposure levels. Individuals might not spend most of their time at the residence location or might spend more time in areas of higher air pollution concentrations (e.g. occupational settings). Finally, not all studies considered a cumulative exposure effect of air pollution on the risk of developing diabetes. Bias from the stratifying of confounders may also occur with an unknown direction of effect. For example, collapsing income levels from a continuous variable into categories can introduce differential misclassification even if the measurement error was nondifferential with an unknown direction of effect (Flegal et al., 1991).

There were several limitations in outcome assessment. First, a few studies differentiated between type I and II DM. However, type I diabetes represents a small fraction of cases among adults and most studies assumed type II. Second, studies varied in their assessment of diabetes outcomes. For example, some studies used a fasting glucose measurement while others included an HbA1C. Third, studies assessing diabetes using self-report and secondary data sources might suffer from outcome misclassification. The degree and direction of misclassification would depend on how prevalent undiagnosed diabetes in a population is and whether undiagnosed diabetes is differential or not across the population and confounder strata. For example, according to the national diabetes statistical report in 2020, undiagnosed diabetes represented 21.5% of total diabetes cases in the US (CDC, 2020). The undiagnosed diabetes percentage varied in magnitude across race, age, and educational level. Finally, differences in background rates of diabetes between two populations can produce a variation on the effect measure even if exposure to air pollution added a constant amount of risk (Greenland, 1987).

## 2.5 Summary and Conclusion

In summary, we have conducted a systematic review and meta-analysis of studies examining exposure to $NO_2$, BC, or UFP and the risk of developing diabetes mellitus among adults. We have concluded that there is sufficient evidence of an association between exposure to $NO_2$ and risk of diabetes among adults based on a moderate quality of evidence, an effect estimate with a positive direction, a pooled effect with a narrow confidence interval with a direction of effect that is unlikely to reverse or overlap the null value with an additional study, and a consistent direction of effect estimates among smaller studies. Our pooled effect suffered from a high level of heterogeneity despite stratifying across multiple variables. However, the direction and significance of the pooled effect remained positive throughout the subgroup analysis. We were not able to reach a similar conclusion for the other pollutants BC and UFP because of the limited number of studies for each. Future studies of the effect of exposure to other pollutants are needed to assess the effect of air pollution exposure across different sources on the risk of developing diabetes mellitus.

## Table 2-1: Source and population

| Author | location | Years of study and follow-up | Study objective | Population (n), age (years), and gender (%) of participants |
|---|---|---|---|---|
| Kramer et al. (2010) | Ruhr district, Germany | 1990-2006 (16 years) | Longitudinal: Examine the association between traffic-related air pollution and incident type 2 diabetes. | • n = 1,775,<br>• 54-55 years,<br>• Female (100%) |
| Dijkema et al. (2011) | Westfriesland, Netherlands | 1998-2000 | Cross-sectional: Examine the relation between long-term exposure to traffic-related air pollution and type 2 diabetes prevalence among 50 to 75-year-old subjects living in Westfriesland, the Netherlands. | • n = 8,018,<br>• 50-75 years<br>• Female (51%) |
| Brook et al. (2008a) | Hamilton and Toronto, Canada | 1992-1999 | Case-control: Investigate the association between DM prevalence and exposure to traffic-related air pollution (nitrogen dioxide). | • Hamilton (n) = 5,228<br>• Toronto (n) = 1,260<br>• ≥40 years<br>• Female (54.8%) |
| Renzi et al. (2018) | Rome, Italy | 2008-2014 (6 years) | Longitudinal: Evaluate the association of long-term exposure to particulate matter (PM), nitrogen oxides (NOx) and ozone (O3), with baseline prevalence and incidence of type 2 diabetes in a large administrative cohort in Rome, Italy. | • n = 1,425,580<br>• ≥35 years<br>• Female (54.6%) |
| Andersen et al. (2012a) | Denmark | 1993-2006 (9.7 years) | Longitudinal: Study the association between long-term exposure to traffic-related air pollution and the incidence of diabetes. | • n = 51,818<br>• 50-65 years<br>• Female (52.6%) |
| Eze et al. (2014b) | Switzerland | | Cross-sectional: Explore the association between air pollution and prevalent diabetes, in a population-based Swiss cohort. | • n = 6,392<br>• 29-73 years<br>• Female (51.3%) |
| Lazarevic et al. (2015) | Australia | 2006-2011 (5 years) | Longitudinal: Assess the effect of long-term exposure to ambient air pollution on the prevalence of self-reported health outcomes in Australian women. | • n = 14,563<br>• 31-90 years<br>• Female (100%) |
| Coogan et al. (2016) | US | 1995-2011 | Longitudinal: Assess the association of the traffic-related nitrogen dioxide (NO2) with the incidence of diabetes in a longitudinal cohort study of African American women. | • n = 43,003,<br>• ≥30 years<br>• Female (100%) |
| Hansen et al. (2016) | Denmark | 1993-2013 | Longitudinal: Examine the association between long-term exposure to PM2.5 and diabetes incidence | • n = 24,174<br>• ≥44 years<br>• Female (100%) |
| C. Clark et al. (2017) | British Columbia, Canada | 1994-1998 | Longitudinal: Examine the influence of long-term residential transportation noise exposure and traffic-related air pollution on the incidence of diabetes using a population-based cohort in British Columbia, Canada. | • n = 380,738<br>• 45-85 years<br>• Female (54%) |
| Li et al. (2017) | Boston, US | 2009-2012 | Cross-sectional: We hypothesized that high UFP exposure near busy roadways may be associated with cardiovascular disease and its risk factors | • n = 704<br>• >40 years<br>• Female (58%) |
| Honda et al. (2017) | US | 2004 | Longitudinal: Investigate the associations between airborne fine particulate matter (PM2.5) and nitrogen dioxide (NO2) and HbA1c levels in both diabetic and non-diabetic older Americans. We also examined the impact of PM2.5 and NO2 on prevalent diabetes mellitus (DM) in this cohort. | • n = 4,121<br>• ≥57 years<br>• Female (53.7%) |
| Strak et al. (2017) | Netherlands | 2012 | Cross-sectional: Investigate associations between long-term exposure to multiple air pollutants and diabetes prevalence in a large national survey in the Netherlands. | • n = 289,703<br>• ≥19 years<br>• Female (52.6%) |
| Eze et al. (2017) | Switzerland | 2002-2011 | Longitudinal: Investigate the independent effects of noise (road, aircraft, and railway noise and specific noise characteristics like the number and temporal variation of noise events), and NO2 on diabetes incidence. | • ≥18 years<br>• Female (52.7%) |
| O'Donovan et al. (2017) | Leicestershire, UK | 2004-2011 | Cross-sectional: Investigate the association between air pollution and type 2 diabetes, while reducing bias due to exposure assessment, outcome assessment, and confounder assessment | • n = 10,443<br>• 25-75 years<br>• Female (47.1%) |

Table 2 1: Source and population (cont.)

| Author | location | Years of study and follow-up | Study objective | Population (n), age (years), and gender (%) of participants |
|---|---|---|---|---|
| Yang et al. (2018) | Liaoning province, China | 2009 | Cross-sectional: Explore the associations of long-term exposure to ambient particulate matter (PM) and gaseous pollutants with diabetes prevalence and glucose-homeostasis markers in China. | • n = 15,477<br>• 18-74 years<br>• Female (47.3%) |
| Orioli et al. (2018) | Italy | 1999-2013 | Cross-sectional: Evaluate the association between area-level ambient air pollution and self-reported DM in a large population sample in Italy. | • n = 376,157<br>• >45 years<br>• Female (53.7%) |
| Riant et al. (2018) | Lille and Dunkirk, France | 2011-2013 | Cross-sectional: Investigate the relationships between long term exposure to air pollution at the place of residence, diabetes biomarkers, and prevalent diabetes in two cities with a relatively low level of pollution. | • Lille (n) = 1,403<br>• Dunkirk (n) = 1,338<br>• 40-65 years<br>• Female (52.2%) |
| Bai et al. (2018) | Toronto, Canada | 1996-2012 | Longitudinal: Investigate the associations between exposures to ultrafine particles and nitrogen dioxide (NO2) and the incidence of diabetes and hypertension in a population-based cohort | • n = 1,056,012<br>• 30-100 years<br>• Female (53%) |
| Shin et al. (2019) | Korea | 2003-2012 | Cross-sectional: Examine the associations between PM10, NO2, CO, SO2, and O3 and CMD using data collected from the Korean Community Health Survey. | • n = 100,867<br>• ≥19 years<br>• Female (50.1%) |
| F. Liu et al. (2019) | Henan province, China | 2015-2017 | Cross-sectional: Evaluate the associations between long-term exposure to particulate matter with an aerodynamic diameter ≤1.0 μm and ≤2.5 μm (PM1 and PM2.5), nitrogen dioxide (NO2), and type 2 diabetes prevalence and fasting blood glucose levels in Chinese rural populations. | • n = 39,191<br>• 18-79 years<br>• Female (60.6%) |
| Howell et al. (2019) | Ontario, Canada | 2008 | Cross-sectional: Assess how walkability and traffic-related air pollution jointly affect the risk of hypertension and diabetes. | • n = 2,496,458<br>• 40-74 years<br>• Female (51.8%) |

## Table 2-2: Exposure and outcome

| Author | Diabetes type and source | Summary of outcome definition | Pollutants | Exposure definition | Exposure summary |
|---|---|---|---|---|---|
| Kramer et al. (2010) | Unspecified-Incidence; Questionnaire | • Self-report of physician-diagnosed diabetes after 1990. | NO2, Soot, PM10, PM2.5 | 5 year mean levels (1986-1990) using monitoring stations nearest to the residence with an 8-km grid.<br>• Annual mass of PM and $NO_2$ emission inventories (1994) with a 1-km grid.<br>• LUR modeling of $NO_2$ and soot concentration using 1-year measurement in 2002.<br>Distance from a home address at baseline to the next major road with >10,000 cars per day | Median ($25^{th}$-$75^{th}$ percentile)<br>Monitoring stations (μg/m3):<br>• PM10 46.9 (44.0–54.1)<br>• NO2 41.7 (23.3–48.2)<br>Traffic emission inventory (tons/year/km2):<br>• PM 0.54 (0.22–1.09)<br>• NO2 12.0 (5.4–24.4)<br>Land-use regression:<br>• Soot 1.89 (1.67–2.06) ($10^{-5}$ m)<br>• NO2 34.5 (23.8–38.8) (μg/m3)<br>Distance < 100 m from the busy road:<br>• No diabetes (15.6%)<br>Incident diabetes (17.7%) |
| Dijkema et al. (2011) | Type 2-Prevalence; Questionnaire, blood test | • Self-report of the previous physician-diagnosed diabetes; and<br>• If the risk of diabetes was high, further testing based on 1999 WHO guidelines for the diagnosis of type 2 diabetes. | NO2 | • LUR modeling of $NO_2$ concentrations in 2007<br>• Distance to the nearest road with ≥5,000 vehicles/24 hrs.<br>• Traffic flow at the nearest main road (vehicles per 24 hrs.)<br>Total traffic per 24 hrs. on all roads within 250 m buffer | Median ($25^{th}$-$75^{th}$ percentile)<br>• $NO_2$ ($\mu g*m^{-3}$): 15.2 (14.2-16.5)<br>• Distance nearest main road (m): 140 (74-220)<br>• Traffic flow nearest main road (vehicle/24hrs): 7,306 (5,871-9,670)<br>Traffic within 250 m buffer ($10^3$/24hrs): 680 (516-882) |
| Brook et al. (2008a) | Unspecified-Prevalence; Health databases | Diagnosis of diabetes made by:<br>• two or more claims by a general practitioner; or<br>• one claim by a specialist; or<br>• hospitalization | NO2 | LUR modeling of NO2 using field measurement between 2002 and 2004. | Median ($25^{th}$-$75^{th}$ percentile)<br>$NO_2$ (ppb):<br>• Hamilton: [Male] 15.2 (13.9-17.1); [Female] 15.3 (14.0-17.0)<br>Toronto: [Male] 23.0 (20.8-25.0); [Female] 22.9 (20.8-24.7) |
| Renzi et al. (2018) | T2DM-Incidence; Health databases | • Qualified for health care for diabetes<br>• Hospital admission with a diabetes diagnosis (ICD-9)<br>• Prescribed hypoglycemic medication at least twice in one year | NO2, PM2.5 absorbance, PM10, PM2.5-10, PM2.5, NOx, O3, Traffic noise | • LUR modeling of NO2, PM2.5 absorbance, PM10, PM2.5-10, PM2.5, and NOx using annual mean levels at baseline, 2008, and 2010.<br>Dispersion model of O3 using summer daily (8h) and seasonal (2005) levels in a 1-km grid [the Flexible Air Quality Regional Model (FARM)]. | Mean (SD)<br>Average annual air pollution level at baseline<br>• $PM_{10}$ (μg/m3): 36.6 (5.2)<br>• PM2.5–10 (μg/m3): 16.9 (3.4)<br>• PM2.5 (μg/m3): 19.6 (1.9)<br>• PM2.5 absorbance ($10^{-5}$/m): 2.7 (0.5)<br>• NO2 (μg/m3): 42.4 (10.4)<br>• NOx (μg/m3): 83.9 (24.4)<br>O3 (μg/m3): 97.4 (6.5) |

Table 2 2: Exposure and outcome (cont.)

| Author | Diabetes type and source | Summary of outcome definition | Pollutants | Exposure definition | Exposure summary |
|---|---|---|---|---|---|
| Andersen et al. (2012a) | Unspecified-Incidence;<br>Health databases | Through NDR inclusion with:<br>• Hospital discharge (ICD-10 or ICD-9; or<br>• Chiropody; or<br>• Five blood glucose readings within one year; or<br>• Two blood glucose readings per year for 5 consecutive years; or<br>• Purchase of diabetes medication within 6 months.<br>Also, including only confirmed cases by excluding those in the NDR solely for a blood glucose test. | NO2, NOx | • Danish AirGIS human exposure modeling system:<br>○ Mean NO2 and NOx since 1971.<br>○ Mean NO2 since 1991.<br>• 1-year mean NO2 at baseline.<br>• 1-year mean NO2 at follow-up.<br>• Traffic proximity to a major road (≥10,000 vehicles/day) within a 50-m radius.<br>Traffic load (total kilometers driven by vehicles) within a 100-m radius. | Median (IQR)<br>NO2 (mg/m3)<br>• 1971 to end of follow-up: 14.5 (4.9)<br>• 1991 to end of follow-up: 15.3 (5.6)<br>• Baseline (1 year): 15.4 (5.6)<br>• End of follow-up (1 year) [median (IQR)] 15.2 (5.7)<br>Traffic load within 100 m at baseline (103 vehicle km/day): 0.34 (1.3)<br>Major road within 50 m at baseline [n (%)]: 4,184 (8.1) |
| Eze et al. (2014b) | T2DM-Prevalence;<br>Health assessment, blood test | • intake of any anti-diabetic medication; or<br>• Self-report of physician-diagnosis; or<br>• Non-fasting blood glucose of >11.1 mmol/L; or<br>• HbA1c of >6.5% or 48 mmol/mol. | NO2, PM10 | • The dispersion model of PM10 and NO2 using mean ambient levels in 1990 and 2000.<br>Hybrid model (dispersion and LUR) of NO2 over 10 years preceding follow-up survey. | Mean (SD)<br>• 10-year mean PM10 [μg/m3]: 22.3 (7.4)<br>10-year mean NO2 [μg/m3]: 26.8 (11.0) |
| Lazarevic et al. (2015) | Unspecified-Prevalence;<br>Questionnaire | • Self-report of diabetes diagnosis within the previous 3 years. | NO2 | • LUR model of NO2 using 3-year mean annual levels (2 years before the survey and during survey year)<br>• Distance to major road<br>Distance to minor roads | Mean (range)<br>3-year mean NO2 (ppb): 5.7 (2.4-11.3 |
| Coogan et al. (2016) | T2DM-Incidence;<br>Questionnaire | • Self-report of physician-diagnosed diabetes at age 30 or older. | NO2, Ozone | • LUR model of NO2 using annual levels for 2000-2010 at the block group level (56 cities).<br>Dispersion model of NO2 levels for 2000-2010 (27 cities). | Mean (SD)<br>NO2 at baseline (ppb):<br>• LUR Model (56 cities): 18.6 (6.5)<br>Dispersion Model (27 cities): 19.2 (5.5) |
| Hansen et al. (2016) | Unspecified-Incidence;<br>Health databases | NDR inclusion with the following:<br>• Hospital discharge (ICD-10 or ICD-9; or<br>• Chiropody; or<br>• Five blood glucose readings within one year; or<br>• Two blood glucose readings per year for 5 consecutive years; or<br>• Purchase of diabetes medication within 6 months.<br>Also, nurses who had either (ii) or (iv) as the sole inclusion criteria were not considered diabetic in the study. | NO2, PM10, PM2.5, NOx | Danish AirGIS human exposure modeling system:<br>• 5-year mean PM2.5, PM10, NO2, and NOx (1990-1995)<br>24-year mean NO2 and NOx | Mean (SD)<br>Annual air pollution at baseline address (μg/m$^3$)<br>• PM2.5: 18.1 (2.8)<br>• PM10: 21.7 (2.9)<br>• NO2: 12.5 (7.9)<br>NOx: 18.4 (22.7) |

Table 2 2: Exposure and outcome (cont.)

| Author | Diabetes type and source | Summary of outcome definition | Pollutants | Exposure definition | Exposure summary |
|---|---|---|---|---|---|
| C. Clark et al. (2017) | Unspecified-Incidence;<br>Health databases | Using the ICD-9 and ICD-10 codes for diabetes:<br>• One hospitalization for diabetes; or<br>• Two physician diagnosis of diabetes; or<br>• Two health care provider visits for diabetes within 1-year. | NO2, BC, NOx, PM2.5 | LUR model of NO2, NO, PM2.5, BC using 5-year monthly average levels in 2003. | Mean (IQR)<br>Average air exposure at residential address<br>• NO2 (μg/m$^3$): 32.1 (8.4)<br>• NO (μg/m$^3$): 32.0 (13.13)<br>• PM2:5 (μg/m$^3$): 4.1 (1.6)<br>Black carbon($10^{-5/}$m): 1.5 (0.9) |
| Li et al. (2017) | Unspecified-Prevalence;<br>Questionnaire | • Self-report of physician-diagnosed diabetes; or<br>• Taking diabetes medication as determined by two physicians | UFP | | Mean (SD)<br>Annual average particle number concentrations of UFP ($10^3$/cm3)<br>• Diabetes (Yes): 20 (6.6)<br>Diabetes (No): 21 (6.4) |
| Honda et al. (2017) | Unspecified-Prevalence;<br>Questionnaire, blood test | • HbA1c ≥ 6.5%; or<br>• Self-report of taking anti-diabetic medication. | NO2, PM10 | • PM2.5 levels obtained using Spatio-temporal generalized additive mixed models (GAMMS) for 1-5 year mean levels from 1999 to 2007 on a 6-km grid.<br>NO2 levels obtained using the nearest AQS monitor within an 80-km radius for 1-5 year mean levels. | Mean (SD)<br>• PM2.5 (μg/m$^3$) 10.4 (3.0)<br>NO2 (ppb) 13.7 (6.6) |
| Strak et al. (2017) | Unspecified-Prevalence;<br>Questionnaire, Health databases | • Self-report of physician-diagnosed diabetes; or<br>• Diabetes medication prescription | NO2, PM2.5 absorbance (BC), PM10, PM2.5, PM10 − 2.5, NOx | LUR model of NO2, PM2.5, BC, PM10, PM2.5, PM10 − 2.5, and NOx using annual average levels in 2009. | Mean (SD)<br>• PM10 (μg/m3): 24.76 (1.11)<br>• PM2.5 (μg/m3): 16.72 (0.69)<br>• PM10 − 2.5 (μg/m3): 8.30 (0.75)<br>• Absorbance ($10^{-5}$/m): 1.28 (0.22)<br>NO2 (μg/m3): 23.88 (6.06) |
| Eze et al. (2017) | Unspecified-Incidence;<br>Questionnaire, blood test | • Self-report of physician-diagnosed diabetes; or<br>• Self-report of taking anti-diabetic medication; or<br>• HbA1c ≥ 6.5%. | NO2, PM2.5 | • Dispersion model of NO2 and PM2.5 using annual mean levels in 2001.<br>• LUR model as above<br>Hybrid model as above | Median (IQR)<br>PM2.5 (μg/m3)<br>• Incident diabetes: 15.2 (4.5)<br>• No Incident diabetes: 14.6 (3.5)<br>NO2 (μg/m3)<br>• Incident diabetes: 20.4 (15)<br>No Incident diabetes: 21.1 (15.4) |
| O'Donovan et al. (2017) | T2DM-Prevalence;<br>Questionnaire, blood test | • Fasting glucose ≥7.0 mmol·L$^{-1}$; or<br>• Two-hour glucose ≥11.0 mmol·L$^{-1}$ | NO2, PM10, PM2.5 | DEFRA Pollution Climate Mapping (PCM) model using 3-year annual average levels of NO2, PM10, and PM2.5 on a 1x1 km grid. | Mean (SD)<br>• NO2 (μg/m3): 21.4 (5.8)<br>• PM2.5 (μg/m3): 12.0 (0.8)<br>PM10 (μg/m3): 16.4 (1.0) |

Table 2 2: Exposure and outcome (cont.)

| Author | Diabetes type and source | Summary of outcome definition | Pollutants | Exposure definition | Exposure summary |
|---|---|---|---|---|---|
| Yang et al. (2018) | T2DM-Prevalence; Questionnaire, blood test | • Fasting glucose ≥ 7·0 mmol/L; or<br>• 2-h glucose ≥ 11·1 mmol/L; or<br>• Intake of antidiabetic medication. | NO2, PM10, PM2.5, PM1, SO2, O3 | • Air monitoring stations within 1-km distance using 3-year (2006-08) average levels of NO2, PM10, SO2, and O3.<br>The spatial statistical model of PM1 and PM2.5 levels during (2006-08) | Mean (SD)<br>• PM1 (μg/m³): 66·0 (10·7)<br>• PM2·5 (μg/m³): 82·0 (14·8)<br>• PM10 (μg/m³): 123·1 (14·6)<br>• SO2 (μg/m³): 54·4 (14·3)<br>• NO2 (μg/m³): 35·3 (4·5)<br>O3 (μg/m³): 49·4 (14·1) |
| Orioli et al. (2018) | Unspecified-Prevalence; Questionnaire | • Self-report of physician-diagnosed diabetes. | NO2, PM10, PM2.5, O3 | AMS-MINNI national integrated dispersion model using 4-year annual average NO2, PM10, PM2.5, and O3 levels for the years 2003, 2005, 2007, and 2010. | Mean (SD)<br>• PM10 (μg/m³): 16.9 (7,4)<br>• PM2.5 (μg/m³): 15.9 (7.1)<br>• NO2 (μg/m³): 15.9 (11.3)<br>O3 (μg/m³): 103.2 (5.1) |
| Riant et al. (2018) | Unspecified-Prevalence; Questionnaire, blood test | • Intake of antidiabetic medication; or<br>• HbA1c ≥6.5%; or<br>• Fasting blood glucose level ≥1.26 g/L; or<br>• Non-fasting blood glucose level ≥2 g/L. | NO2, PM10, SO2 | • Dispersion model for NO2 and PM10 using annual mean concentrations between 2010 and 2013 in Lille, and between 2012 and 2013 in Dunkirk.<br>The dispersion model (like the above) for SO2 was available only for Dunkirk. | Median ($25^{th}$-$75^{th}$ percentile)<br>NO2 (μg/m3)<br>• Lille: 25.96 [22.52; 28.59]<br>• Dunkirk: 20.35 [18.23;21.46]<br>PM10 (μg/m3)<br>• Lille: 26.96 [25.65; 28.18]<br>• Dunkirk: 26.54 [25.85; 27.19]<br>SO2 (μg/m3)<br>• Lille: _ _<br>Dunkirk: 3.07 [2.15; 3.89] |
| Bai et al. (2018) | Unspecified-Incidence;<br>Health databases | Using a health database with ICD-9 and ICD-10 definitions:<br>• Hospital admission with a diagnosis of diabetes; or<br>• Two physician claims over 2 years. | NO2, UFP | LUR model of NO2 and UFP using 3-year moving averages of estimates of concentrations beginning from 1996. | Mean (SD)<br>• UFP (Count/$cm^3$): 28,383.1 (9,090.9)<br>• PM2.5 (μg/m3): 10.7 (1.6)<br>NO2 (ppb): 21.4 (3.5) |
| Shin et al. (2019) | Unspecified-Prevalence; Questionnaire | • Self-report of physician-diagnosed diabetes. | NO2, PM10, CO, SO2, O3 | Air monitoring stations of NO2, PM10, CO, SO2, and O3 using 10-year average concentrations during 2003-2012. | Mean (SD)<br>• PM10 (mg/$m^3$): 52.7 (8.6)<br>• SO2 (ppb): 5.6 (1.7)<br>• NO2 (ppb): 24.2 (7.9)<br>• CO (10 ppm): 5.7 (1.3)<br>O3 (ppb): 23.4 (4.5) |
| F. Liu et al. (2019) | T2DM-Prevalence; Questionnaire, blood test | • Self-report of type 2 diabetes diagnosis; or<br>• Intake of antidiabetic medication; or<br>• Fasting glucose ≥ 7·0 mmol/L | NO2, PM1, PM2.5 | Spatiotemporal model of NO2, PM1, and PM2.5 using 3-year average concentrations. | Mean (SD)<br>• PM1 (μg/m3): 57.4 (2.7)<br>• PM2.5 (μg/m3): 73.4 (2.6)<br>NO2 (μg/m3): 39.9 (3.6) |
| Howell et al. (2019) | Unspecified-Prevalence; Health databases | Using a health database:<br>• Hospital admission with a diagnosis of diabetes; or<br>• Two physician claims over 2 years. | NO2 | LUR model of NO2 using annual average concentration predicted for 2006. | Mean (SD)<br>NO2 (ppb): 18.0 (5.3) |

Table 2-3: Effect measure included in the analysis

| Citation | Pollutant | Outcome | Model | Gender | reported | Converted |
|---|---|---|---|---|---|---|
| Brook et al. 2008 | NO2 | Prevalence | LUR | All | 1.01 (0.98, 1.05) Per 1 ppb | 1.08 (0.90, 1.29) Per 10 ug/m$^3$ |
| Brook et al. 2008 | NO2 | Prevalence | LUR | Female | 1.04 (1.00, 1.08) Per 1 ppb | 1.23 (1.00, 1.51) Per 10 ug/m$^3$ |
| Brook et al. 2008 | NO2 | Prevalence | LUR | Male | 0.99 (0.95, 1.03) Per 1 ppb | 0.95 (0.76, 1.17) Per 10 ug/m$^3$ |
| Kramer et al. 2010 | NO2 | Incidence | Emission | Female | 1.15 (1.04, 1.27) Per 15 ug/m$^3$ | 1.1 (1.03, 1.170) Per 10 ug/m$^3$ |
| Kramer et al. 2010 | NO2 | Incidence | LUR | Female | 1.42 (1.16, 1.73) Per 15 ug/m$^3$ | 1.26 (1.10, 1.44) Per 10 ug/m$^3$ |
| Kramer et al. 2010 | NO2 | Incidence | Air monitor | Female | 1.34 (1.02, 1.76) Per 15 ug/m$^3$ | 1.22 (1.01, 1.46) Per 10 ug/m$^3$ |
| Anderson et al. 2011 | NO2 | Incidence | AirGIS | All | 1.04 (1.00, 1.08) Per 4.9 ug/m$^3$ | 1.08 (1.00, 1.17) Per 10 ug/m$^3$ |
| Anderson et al. 2011 | NO2 | Incidence | AirGIS | Female | 1.07 (1.01, 1.13) Per 4.9 ug/m$^3$ | 1.15 (1.02, 1.28) Per 10 ug/m$^3$ |
| Anderson et al. 2011 | NO2 | Incidence | AirGIS | Male | 1.01 (0.97, 1.07) Per 4.9 ug/m$^3$ | 1.02 (0.94, 1.15) Per 10 ug/m$^3$ |
| Eze et al. 2014 | NO2 | Prevalence | Hybrid | All | 1.21 (1.04, 1.4) Per 10 ug/m$^3$ | 1.21 (1.04, 1.40) Per 10 ug/m$^3$ |
| Eze et al. 2014 | NO2 | Prevalence | Hybrid | Female | 1.11 (0.91, 1.36) Per 10 ug/m$^3$ | 1.11 (0.91, 1.36) Per 10 ug/m$^3$ |
| Eze et al. 2014 | NO2 | Prevalence | Hybrid | Male | 1.25 (1.06, 1.48) Per 10 ug/m$^3$ | 1.25 (1.06, 1.48) Per 10 ug/m$^3$ |
| Lazarevic et al. 2015 | NO2 | Prevalence | LUR | Female | 1.04 (0.90, 1.20) Per 3.7 ppb | 1.06 (0.86, 1.30) Per 10 ug/m$^3$ |
| Coogan et al. 2016 | NO2 | Incidence | Dispersion | Female | 0.85 (0.71, 1.02) Per 9.7 ppb | 0.91 (0.83, 1.01) Per 10 ug/m$^3$ |
| Coogan et al. 2016 | NO2 | Incidence | LUR | Female | 0.88 (0.79, 0.98) Per 9.7 ppb | 0.93 (0.88, 0.99) Per 10 ug/m$^3$ |
| Hansen et al. 2016 | NO2 | Incidence | AirGIS | Female | 1.05 (0.98, 1.12) Per 7.53 ug/m$^3$ | 1.07 (0.97, 1.16) Per 10 ug/m$^3$ |
| Clark et al. 2017 | NO2 | Incidence | LUR | All | 1.00 (0.98, 1.02) Per 8.4 ug/m$^3$ | 1.00 (0.99, 1.01) Per 10 ug/m$^3$ |
| Eze et al. 2017 | NO2 | Incidence | Hybrid | All | 0.92 (0.67, 1.26) Per 15 ug/m$^3$ | 0.95 (0.77, 1.17) Per 10 ug/m$^3$ |
| Honda et al. 2017 | NO2 | Prevalence | Air monitor | All | 1.22 (1.07, 1.39) Per 8.3 ppb | 1.14 (1.04, 1.23) Per 10 ug/m$^3$ |
| O'Donovan et al. 2017 | NO2 | Prevalence | DEFRA-PCM | All | 0.91 (0.72, 1.16) Per 10 ug/m$^3$ | 0.91 (0.72, 1.16) Per 10 ug/m$^3$ |
| Renzi et al. 2017 | NO2 | Incidence | LUR | All | 1.00 (0.988, 1.01) Per 10ug/m$^3$ | 1.00 (0.99, 1.01) Per 10 ug/m$^3$ |
| Renzi et al. 2017 | NO2 | Incidence | LUR | Female | 1.00 (0.992, 1.01) Per 10ug/m$^3$ | 1.00 (0.99, 1.01) Per 10 ug/m$^3$ |
| Renzi et al. 2017 | NO2 | Incidence | LUR | Male | 1.00 (0.993, 1.01) Per 10ug/m$^3$ | 1.00 (0.99, 1.01) Per 10 ug/m$^3$ |
| Renzi et al. 2017 | NO2 | Prevalence | LUR | All | 1.01 (1.002, 1.02) Per 10ug/m$^3$ | 1.01 (1.00, 1.02) Per 10 ug/m$^3$ |
| Strak et al. 2017 | NO2 | Prevalence | LUR | All | 1.07 (1.05, 1.09) Per 7.76 ug/m$^3$ | 1.09 (1.06, 1.12) Per 10 ug/m$^3$ |
| Bai et al. 2018 | NO2 | Incidence | LUR | All | 1.06 (1.05, 1.07) Per 4 ppb | 1.08 (1.07, 1.09) Per 10 ug/m$^3$ |
| Bai et al. 2018 | NO2 | Incidence | LUR | Female | 1.07 (1.05, 1.08) Per 4 ppb | 1.09 (1.07, 1.11) Per 10 ug/m$^3$ |
| Bai et al. 2018 | NO2 | Incidence | LUR | Male | 1.05 (1.03, 1.06) Per 4 ppb | 1.07 (1.04, 1.08) Per 10 ug/m$^3$ |
| Orioli et al. 2018 | NO2 | Prevalence | Dispersion | All | 1.03 (1.01, 1.05) Per 10 ug/m$^3$ | 1.03 (1.01, 1.05) Per 10 ug/m$^3$ |
| Orioli et al. 2018 | NO2 | Prevalence | Dispersion | Female | 1.00 (0.99, 1.02) Per 10 ug/m$^3$ | 1 .00(0.99, 1.02) Per 10 ug/m$^3$ |
| Orioli et al. 2018 | NO2 | Prevalence | Dispersion | Male | 1.06 (1.04, 1.08) Per 10 ug/m$^3$ | 1.06 (1.04, 1.08) Per 10 ug/m$^3$ |
| Riant et al. 2018 | NO2 | Prevalence | Dispersion | All | 1.06 (0.9, 1.25) Per 5 ug/m$^3$ | 1.12 (0.81, 1.56) Per 10 ug/m$^3$ |
| Yang et al. 2018 | NO2 | Prevalence | Air monitor | All | 1.22 (1.12, 1.33) Per 9 ug/m$^3$ | 1.25 (1.13, 1.37) Per 10 ug/m$^3$ |
| Yang et al. 2018 | NO2 | Prevalence | Air monitor | Female | 1.10 (0.94, 1.3) Per 9 ug/m$^3$ | 1.11 (0.93, 1.34) Per 10 ug/m$^3$ |
| Yang et al. 2018 | NO2 | Prevalence | Air monitor | Male | 1.28 (1.11, 1.47) Per 9 ug/m$^3$ | 1.32 (1.12, 1.53) Per 10 ug/m$^3$ |
| Howell et al. 2019 | NO2 | Prevalence | LUR | All | 1.16 (1.14, 1.17) Per 10 ppb | 1.08 (1.07, 1.09) Per 10 ug/m$^3$ |
| Liu et al. 2019 | NO2 | Prevalence | Satellite | All | 1.05 (1.039, 1.06) Per 1 ug/m$^3$ | 1.30 (1.23, 1.37) Per 10 ug/m$^3$ |
| Liu et al. 2019 | NO2 | Prevalence | Satellite | Female | 1.04 (1.026, 1.05) Per 1 ug/m$^3$ | 1.47 (1.29, 1.66) Per 10 ug/m$^3$ |
| Liu et al. 2019 | NO2 | Prevalence | Satellite | Male | 1.07 (1.052, 1.09) Per 1 ug/m$^3$ | 1.95 (1.66, 2.30) Per 10 ug/m$^3$ |
| Shin et al. 2019 | NO2 | Prevalence | Air monitor | Female | 1.19 (1.07, 1.33) Per 13.6 ppb | 1.07 (1.03, 1.12) Per 10 ug/m$^3$ |
| Shin et al. 2019 | NO2 | Prevalence | Air monitor | Female | 1.19 (1.07, 1.33) Per 13.6 ppb | 1.07 (1.03, 1.12) Per 10 ug/m$^3$ |
| Shin et al. 2019 | NO2 | Prevalence | Air monitor | Male | 1.12 (1.02, 1.22) Per 13.6 ppb | 1.05 (1.01, 1.08) Per 10 ug/m$^3$ |
| Shin et al. 2019 | NO2 | Prevalence | Air monitor | Male | 1.12 (1.02, 1.22) Per 13.6 ppb | 1.05 (1.01, 1.08) Per 10 ug/m$^3$ |
| Li et al. 2017 | UFP | Prevalence | TAA-PNC | All | 0.71 (0.46, 1.10) Per 10112 count/cm$^3$ | 0.71 (0.46, 1.10) Per 10$^4$count/cm$^3$ |
| Bai et al. 2018 | UFP | Incidence | LUR | All | 1.06 (1.05, 1.08) Per 9948 count/cm$^3$ | 1.06 (1.05, 1.08) Per 10$^4$ count/cm$^3$ |
| Kramer et al. 2010 | BC | Incidence | LUR | Female | 1.27 (1.09, 1.48) Per 0.39 (10$^{-5}$/m) | 1.85 (1.25, 2.73) Per 10$^{-5}$/m |
| Clark et al. 2017 | BC | Incidence | LUR | All | 1.03 (1.01, 1.04) Per 0.90 (10$^{-5}$/m) | 1.03 (1.01, 1.04) Per 10$^{-5}$/m |
| Renzi et al. 2017 | BC | Incidence | LUR | All | 1.00 (0.981, 1.02) Per 1.0 (10$^{-5}$/m) | 1.00 (0.98, 1.02) Per 10$^{-5}$/m |
| Strak et al. 2017 | BC | Prevalence | LUR | All | 1.04 (1.02, 1.06) Per 0.24 (10$^{-5}$/m) | 1.18 (1.09, 1.27) Per 10$^{-5}$/m |

Table 2-4: Quality of evidence

| Quality factor | Rating | Basis |
|---|---|---|
| Downgrade | | |
| Risk of bias across studies | 0 | No evidence of substantial risk of bias across included studies |
| Indirectness | 0 | The studies assessed the population, exposure, and outcome of interest |
| Inconsistency | 0 | Except for three studies (Coogan et al., 2016; Egger et al., 1997; Eze et al., 2017; O'Donovan et al., 2017), study results were generally consistent in direction with the summary effect with varying degrees of magnitudes. |
| Imprecision | 0 | The CI of the pooled effect for exposure to NO2 and DM was narrow. |
| Publication Bias | 0 | No evidence of publication bias. |
| Upgrade | | |
| Large magnitude of effect | 0 | The effect estimate was not large |
| Dose-response | 0 | Several studies reported dose-response curves, but the evidence was not compelling enough to change the rating. |
| Confounding minimizes effect | 0 | No evidence was found that residual confounding would reduce the effect estimate. |
| Overall quality of evidence | Moderate | Moderate. The initial rating for human studies was moderate with no downgrade/upgrade of the rating |
| Summary of findings from the meta-analysis | NA | There is a positive association between the risk of DM and exposure to $NO_2$ |
| Summary of qualitative findings | NA | Dijkema et al. (not included in the quantitative analysis) showed a positive association between NO2 exposure and DM. |
| Strength of considerations | | |
| Quality of body of evidence | NA | Moderate |
| The direction of the effect estimate | NA | Risk of DM increased with increasing exposure o $NO_2$ |
| Confidence in the effect estimate | NA | It is unlikely that a new study would have an effect estimate that would make the results null. |
| Other compelling attributes of the data that may influence certainty | NA | None |
| Overall strength of evidence | Sufficient | We conclude that there is a positive association between NO2 exposure and risk of DM with sufficient evidence while reasonably ruling out chance, bias, and confounding as an explanation. |

Figure 2-1: Flow chart

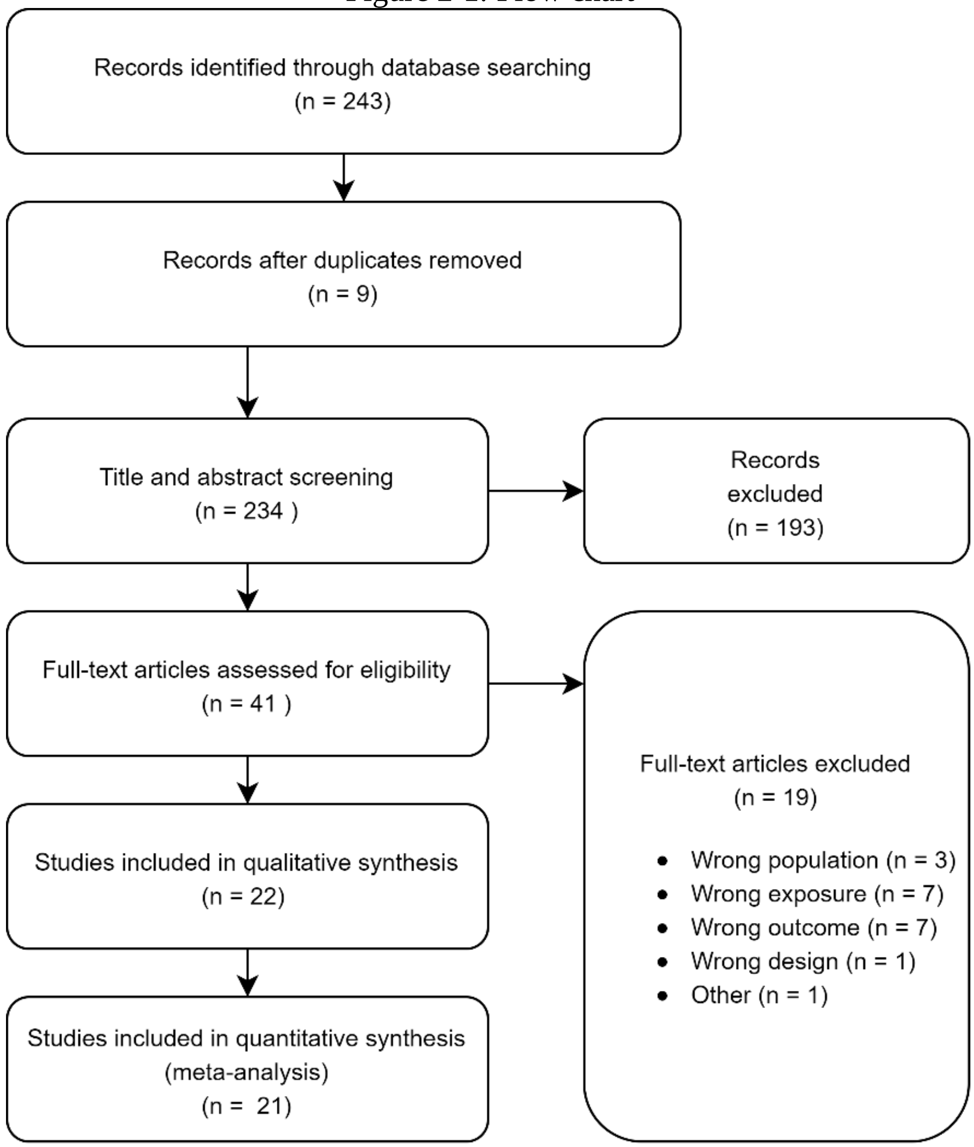


Figure 2-2: Risk of bias

| | Recruitment | Blinding | Exposure assessment | Confounding | Incomplete outcome data | Selective outcome reporting | Conflict of interest | Other sources of bias |
|---|---|---|---|---|---|---|---|---|
| Brook et al. (2008b) | L | L | L | L | H | L | L | L |
| Krämer et al. (2010) | L | L | L | L | L | L | L | L |
| Dijkema et al. (2011) | L | L | L | L | L | L | L | L |
| Andersen et al. (2012b) | L | L | L | L | L | L | L | L |
| Eze et al. (2014a) | L | L | L | L | L | L | L | L |
| Lazarevic et al. (2015) | L | L | L | L | H | L | L | L |
| Coogan et al. (2016) | H | H | L | L | H | L | L | L |
| Hansen et al. (2016) | L | L | L | L | H | L | L | L |
| Renzi et al. (2018) | L | L | L | L | L | L | L | L |
| C. Clark et al. (2017) | L | L | L | H | H | L | L | L |
| Li et al. (2017) | L | L | L | L | H | L | L | L |
| Honda et al. (2017) | L | L | H | L | L | L | L | L |
| Strak et al. (2012) | L | L | L | L | H | L | L | L |
| Eze et al. (2017) | L | L | L | L | L | L | L | L |
| O'Donovan et al. (2017) | L | L | L | L | L | L | L | L |
| Yang et al. (2018) | L | L | H | L | L | L | L | L |
| Orioli et al. (2018) | L | L | U | L | H | L | L | L |
| Riant et al. (2018) | L | L | L | L | L | L | L | L |
| Bai et al. (2018) | L | L | L | L | H | L | L | L |
| Shin et al. (2019) | L | L | H | L | H | L | L | L |
| Howell et al. (2019) | L | L | L | L | H | L | L | L |
| F. Liu et al. (2019) | L | L | L | L | L | L | L | L |

Figure 2-3: $NO_2$ and diabetes

| Study | OR | 95% CI | Weight (fixed) | Weight (Random) |
|---|---|---|---|---|
| Brook et al. 2008 | 1.08 | [0.89; 1.28] | 0.1% | 1.6% |
| Kramer et al. 2010 | 1.26 | [1.09; 1.43] | 0.1% | 2.1% |
| Andersen et al. 2011 | 1.08 | [1.00; 1.17] | 0.3% | 4.7% |
| Eze et al. 2014 | 1.21 | [1.03; 1.39] | 0.1% | 1.9% |
| Lazarevic et al. 2015 | 1.06 | [0.84; 1.28] | 0.0% | 1.4% |
| Coogan et al. 2016 | 0.93 | [0.88; 0.99] | 0.7% | 6.2% |
| Hansen et al. 2016 | 1.07 | [0.97; 1.16] | 0.2% | 4.3% |
| Renzi et al. 2017 | 1.00 | [0.99; 1.01] | 23.6% | 8.1% |
| Clark et al. 2017 | 1.00 | [0.99; 1.01] | 13.3% | 8.0% |
| Honda et al. 2017 | 1.14 | [1.04; 1.23] | 0.2% | 4.2% |
| Strak et al. 2017 | 1.09 | [1.06; 1.12] | 3.1% | 7.6% |
| O'Donovan et al. 2017 | 0.91 | [0.69; 1.13] | 0.0% | 1.4% |
| Yang et al. 2018 | 1.25 | [1.13; 1.37] | 0.1% | 3.3% |
| Orioli et al. 2018 | 1.03 | [1.01; 1.05] | 5.3% | 7.8% |
| Riant et al. 2018 | 1.12 | [0.75; 1.50] | 0.0% | 0.5% |
| Bai et al. 2018 | 1.08 | [1.07; 1.09] | 11.6% | 8.0% |
| Shin et al. 2019 - M | 1.05 | [1.01; 1.08] | 1.6% | 7.2% |
| Shin et al. 2019 - F | 1.07 | [1.02; 1.12] | 1.0% | 6.7% |
| Liu et al. 2019 | 1.30 | [1.22; 1.37] | 0.4% | 5.3% |
| Howell et al. 2019 | 1.08 | [1.07; 1.09] | 38.2% | 8.1% |
| Eze et al. 2017 | 0.95 | [0.75; 1.15] | 0.1% | 1.6% |
| **Fixed effect model** | **1.05** | **[1.04; 1.05]** | **100.0%** | -- |
| **Random effects model** | **1.07** | **[1.04; 1.10]** | -- | **100.0%** |
| **Prediction interval** | | **[0.96; 1.18]** | | |

Heterogeneity: $I^2 = 95\%$, $\tau^2 = 0.0025$, $p < 0.01$

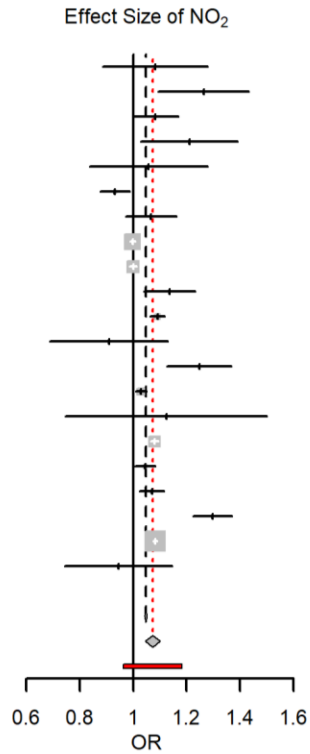


Figure 2-4: $NO_2$ and diabetes by study design

| Study and subgroup | OR | 95% CI | Weight (fixed) | Weight (Random) |
|---|---|---|---|---|
| Incidence | | | | |
| Kramer et al. 2010 | 1.26 | [1.09; 1.43] | 0.1% | 1.7% |
| Andersen et al. 2011 | 1.08 | [1.00; 1.17] | 0.2% | 4.1% |
| Coogan et al. 2016 | 0.93 | [0.88; 0.99] | 0.5% | 5.7% |
| Hansen et al. 2016 | 1.07 | [0.97; 1.16] | 0.2% | 3.7% |
| Renzi et al. 2017 | 1.00 | [0.99; 1.01] | 17.1% | 7.8% |
| Clark et al. 2017 | 1.00 | [0.99; 1.01] | 9.6% | 7.7% |
| Bai et al. 2018 | 1.08 | [1.07; 1.09] | 8.4% | 7.7% |
| Eze et al. 2017 | 0.95 | [0.75; 1.15] | 0.0% | 1.3% |
| Fixed effect model | 1.02 | [1.01; 1.02] | 36.1% | -- |
| Random effects model | 1.03 | [0.99; 1.07] | -- | 39.8% |
| Heterogeneity: $I^2 = 95\%$, $\tau^2 = 0.0021$, $p < 0.01$ | | | | |
| Prevalence | | | | |
| Brook et al. 2008 | 1.08 | [0.89; 1.28] | 0.0% | 1.4% |
| Eze et al. 2014 | 1.21 | [1.03; 1.39] | 0.0% | 1.6% |
| Lazarevic et al. 2015 | 1.06 | [0.84; 1.28] | 0.0% | 1.1% |
| Honda et al. 2017 | 1.14 | [1.04; 1.23] | 0.2% | 3.7% |
| Strak et al. 2017 | 1.09 | [1.06; 1.12] | 2.2% | 7.2% |
| O'Donovan et al. 2017 | 0.91 | [0.69; 1.13] | 0.0% | 1.1% |
| Yang et al. 2018 | 1.25 | [1.13; 1.37] | 0.1% | 2.8% |
| Orioli et al. 2018 | 1.03 | [1.01; 1.05] | 3.9% | 7.5% |
| Riant et al. 2018 | 1.12 | [0.75; 1.50] | 0.0% | 0.4% |
| Shin et al. 2019 - M | 1.05 | [1.01; 1.08] | 1.2% | 6.7% |
| Shin et al. 2019 - F | 1.07 | [1.02; 1.12] | 0.7% | 6.2% |
| Liu et al. 2019 | 1.30 | [1.22; 1.37] | 0.3% | 4.8% |
| Howell et al. 2019 | 1.08 | [1.07; 1.09] | 27.7% | 7.8% |
| Renzi et al. 2017 | 1.01 | [1.00; 1.02] | 27.4% | 7.8% |
| Fixed effect model | 1.05 | [1.04; 1.05] | 63.9% | -- |
| Random effects model | 1.09 | [1.06; 1.13] | -- | 60.2% |
| Heterogeneity: $I^2 = 95\%$, $\tau^2 = 0.0025$, $p < 0.01$ | | | | |
| **Fixed effect model** | **1.04** | **[1.03; 1.04]** | **100.0%** | **--** |
| **Random effects model** | **1.07** | **[1.04; 1.09]** | **--** | **100.0%** |
| Heterogeneity: $I^2 = 95\%$, $\tau^2 = 0.0021$, $p < 0.01$ | | | | |
| Residual heterogeneity: $I^2 = 95\%$, $p < 0.01$ | | | | |

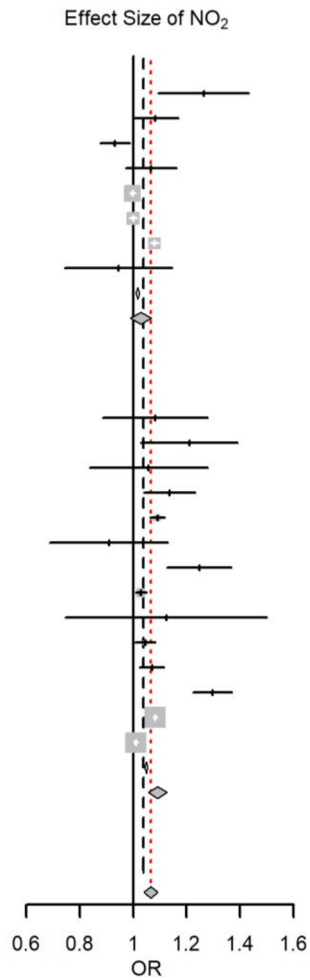


Figure 2-5: $NO_2$ and diabetes by gender

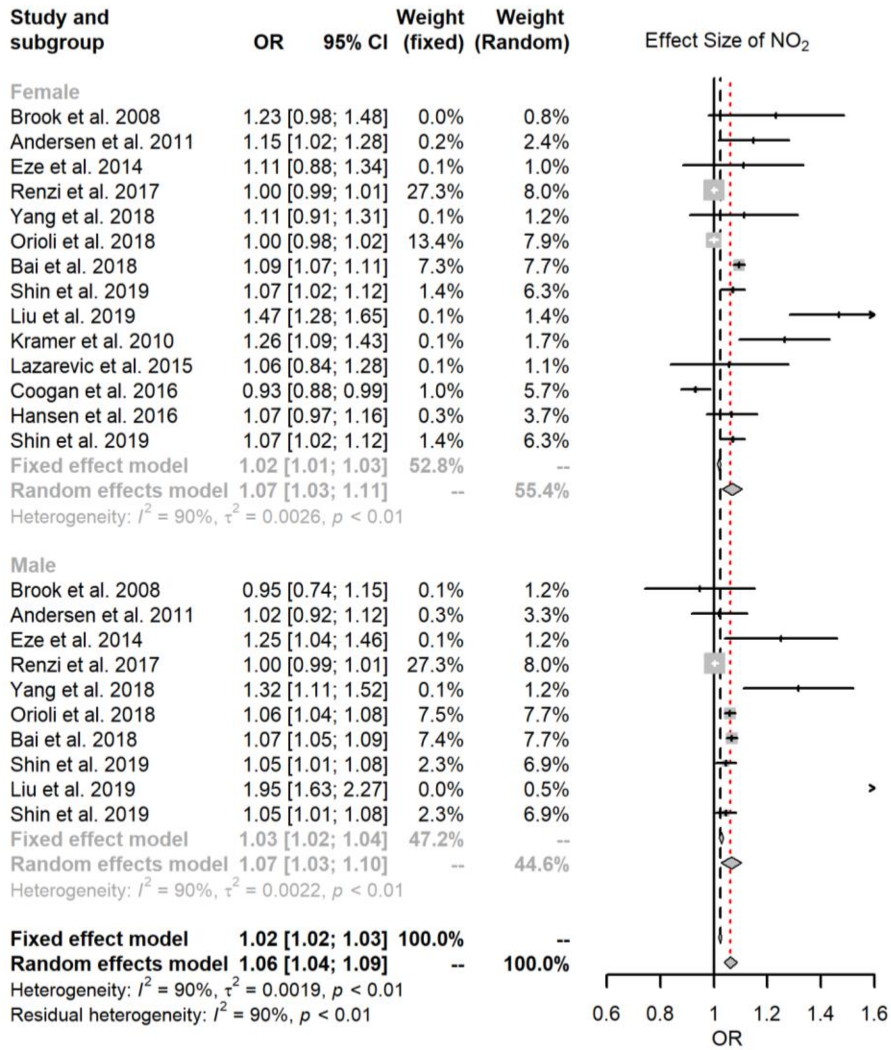


Figure 2-6: $NO_2$ and diabetes by minimum age of inclusion

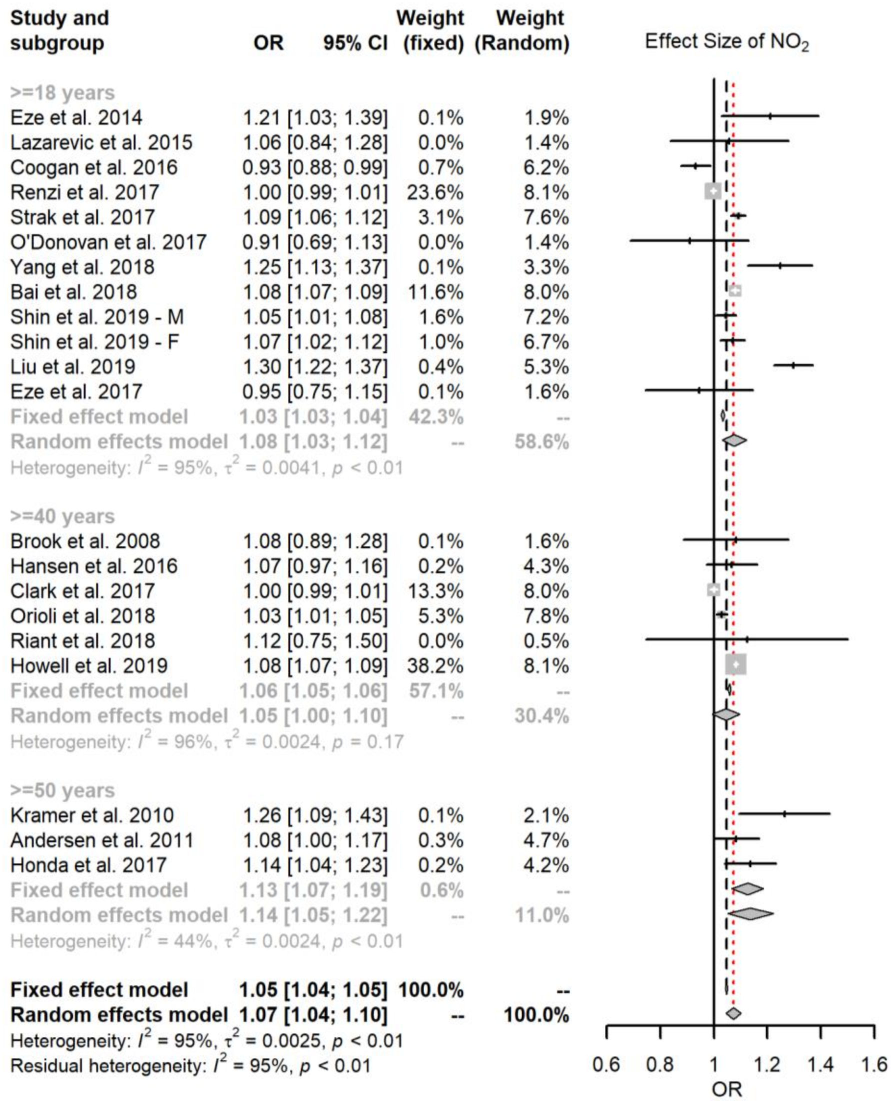


Figure 2-7: $NO_2$ and diabetes by location

| Study and subgroup | OR | 95% CI | Weight (fixed) | Weight (Random) |
|---|---|---|---|---|
| America | | | | |
| Brook et al. 2008 | 1.08 | [0.89; 1.28] | 0.1% | 1.6% |
| Coogan et al. 2016 | 0.93 | [0.88; 0.99] | 0.7% | 6.2% |
| Clark et al. 2017 | 1.00 | [0.99; 1.01] | 13.3% | 8.0% |
| Honda et al. 2017 | 1.14 | [1.04; 1.23] | 0.2% | 4.2% |
| Bai et al. 2018 | 1.08 | [1.07; 1.09] | 11.6% | 8.0% |
| Howell et al. 2019 | 1.08 | [1.07; 1.09] | 38.2% | 8.1% |
| Fixed effect model | 1.06 | [1.06; 1.07] | 64.0% | -- |
| Random effects model | 1.04 | [1.00; 1.09] | -- | 36.2% |
| Heterogeneity: $I^2$ = 97%, $\tau^2$ = 0.0022, $p$ < 0.01 | | | | |
| Asia | | | | |
| Yang et al. 2018 | 1.25 | [1.13; 1.37] | 0.1% | 3.3% |
| Shin et al. 2019 - M | 1.05 | [1.01; 1.08] | 1.6% | 7.2% |
| Shin et al. 2019 - F | 1.07 | [1.02; 1.12] | 1.0% | 6.7% |
| Liu et al. 2019 | 1.30 | [1.22; 1.37] | 0.4% | 5.3% |
| Fixed effect model | 1.10 | [1.07; 1.12] | 3.2% | -- |
| Random effects model | 1.16 | [1.05; 1.27] | -- | 22.5% |
| Heterogeneity: $I^2$ = 93%, $\tau^2$ = 0.0115, $p$ < 0.01 | | | | |
| Australia | | | | |
| Lazarevic et al. 2015 | 1.06 | [0.84; 1.28] | 0.0% | 1.4% |
| Fixed effect model | 1.06 | [0.84; 1.28] | 0.0% | -- |
| Random effects model | 1.06 | [0.84; 1.28] | -- | 1.4% |
| Heterogeneity: not applicable | | | | |
| Europe | | | | |
| Kramer et al. 2010 | 1.26 | [1.09; 1.43] | 0.1% | 2.1% |
| Andersen et al. 2011 | 1.08 | [1.00; 1.17] | 0.3% | 4.7% |
| Eze et al. 2014 | 1.21 | [1.03; 1.39] | 0.1% | 1.9% |
| Hansen et al. 2016 | 1.07 | [0.97; 1.16] | 0.2% | 4.3% |
| Renzi et al. 2017 | 1.00 | [0.99; 1.01] | 23.6% | 8.1% |
| Strak et al. 2017 | 1.09 | [1.06; 1.12] | 3.1% | 7.6% |
| O'Donovan et al. 2017 | 0.91 | [0.69; 1.13] | 0.0% | 1.4% |
| Orioli et al. 2018 | 1.03 | [1.01; 1.05] | 5.3% | 7.8% |
| Riant et al. 2018 | 1.12 | [0.75; 1.50] | 0.0% | 0.5% |
| Eze et al. 2017 | 0.95 | [0.75; 1.15] | 0.1% | 1.6% |
| Fixed effect model | 1.01 | [1.01; 1.02] | 32.7% | -- |
| Random effects model | 1.06 | [1.02; 1.10] | -- | 39.9% |
| Heterogeneity: $I^2$ = 86%, $\tau^2$ = 0.0021, $p$ < 0.01 | | | | |
| **Fixed effect model** | **1.05** | **[1.04; 1.05]** | **100.0%** | **--** |
| **Random effects model** | **1.07** | **[1.04; 1.10]** | **--** | **100.0%** |
| Heterogeneity: $I^2$ = 95%, $\tau^2$ = 0.0025, $p$ < 0.01 | | | | |
| Residual heterogeneity: $I^2$ = 93%, $p$ < 0.01 | | | | |

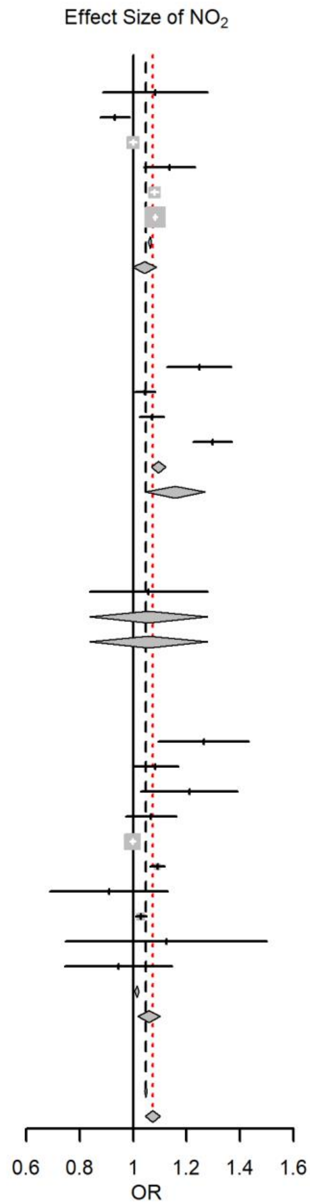


Figure 2-8: $NO_2$ and diabetes by exposure model

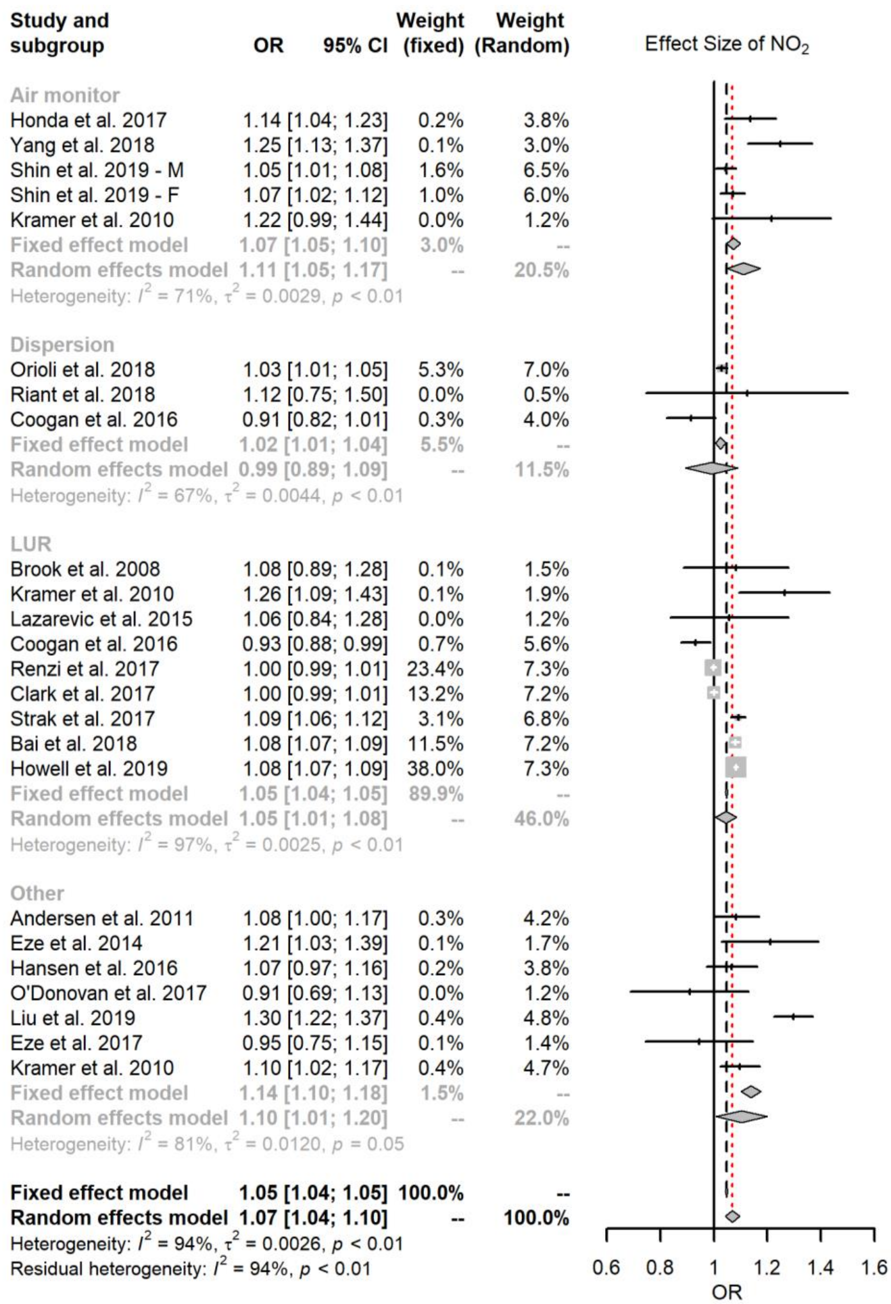


Figure 2-9: $NO_2$ and diabetes by outcome definition

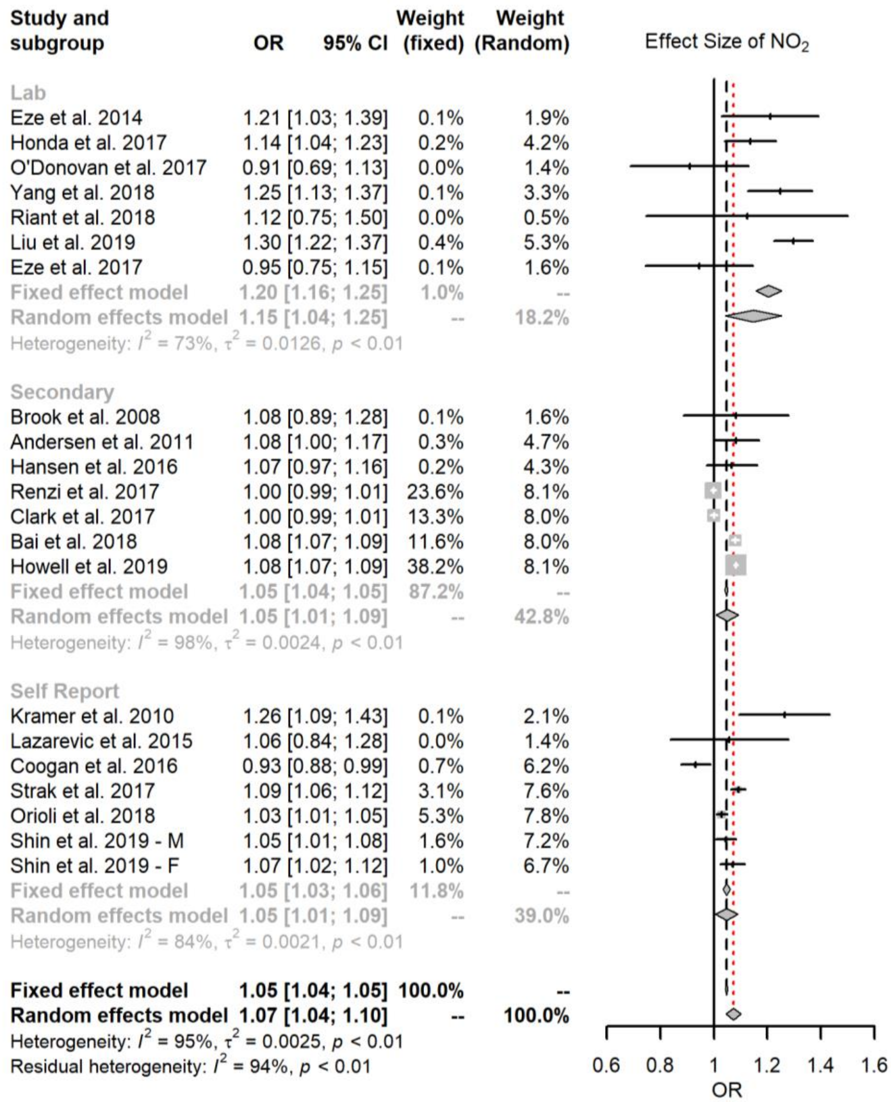


Figure 2-10: Black carbon and diabetes

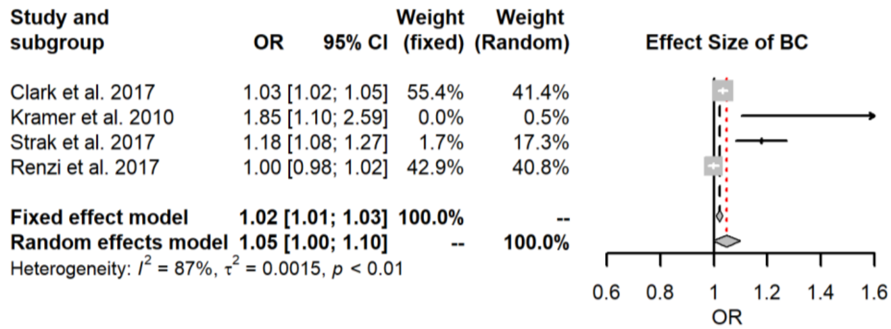



Figure 2-11: Ultrafine particles and diabetes

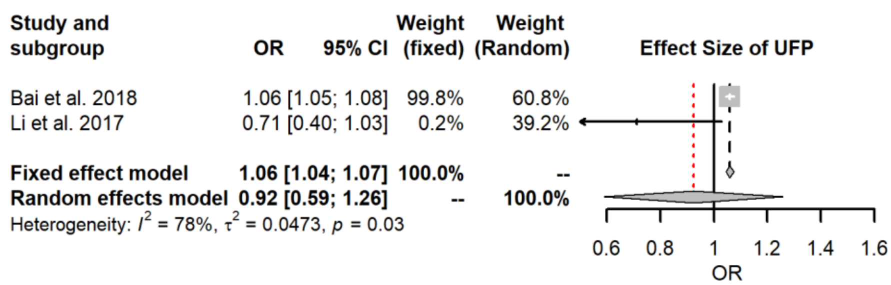


Figure 2-12: Funnel plot

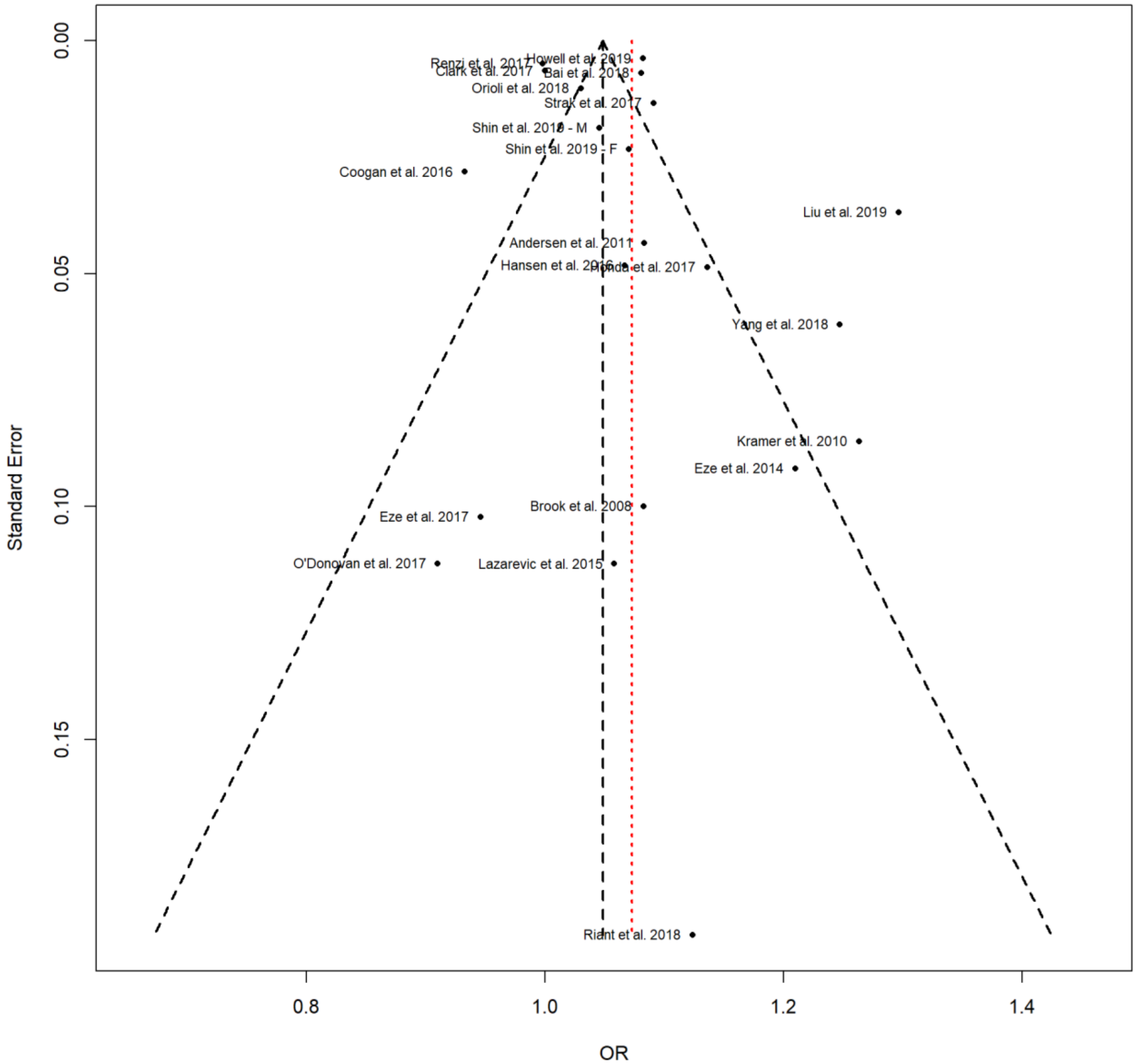


Figure 2-13: Adjustment factors

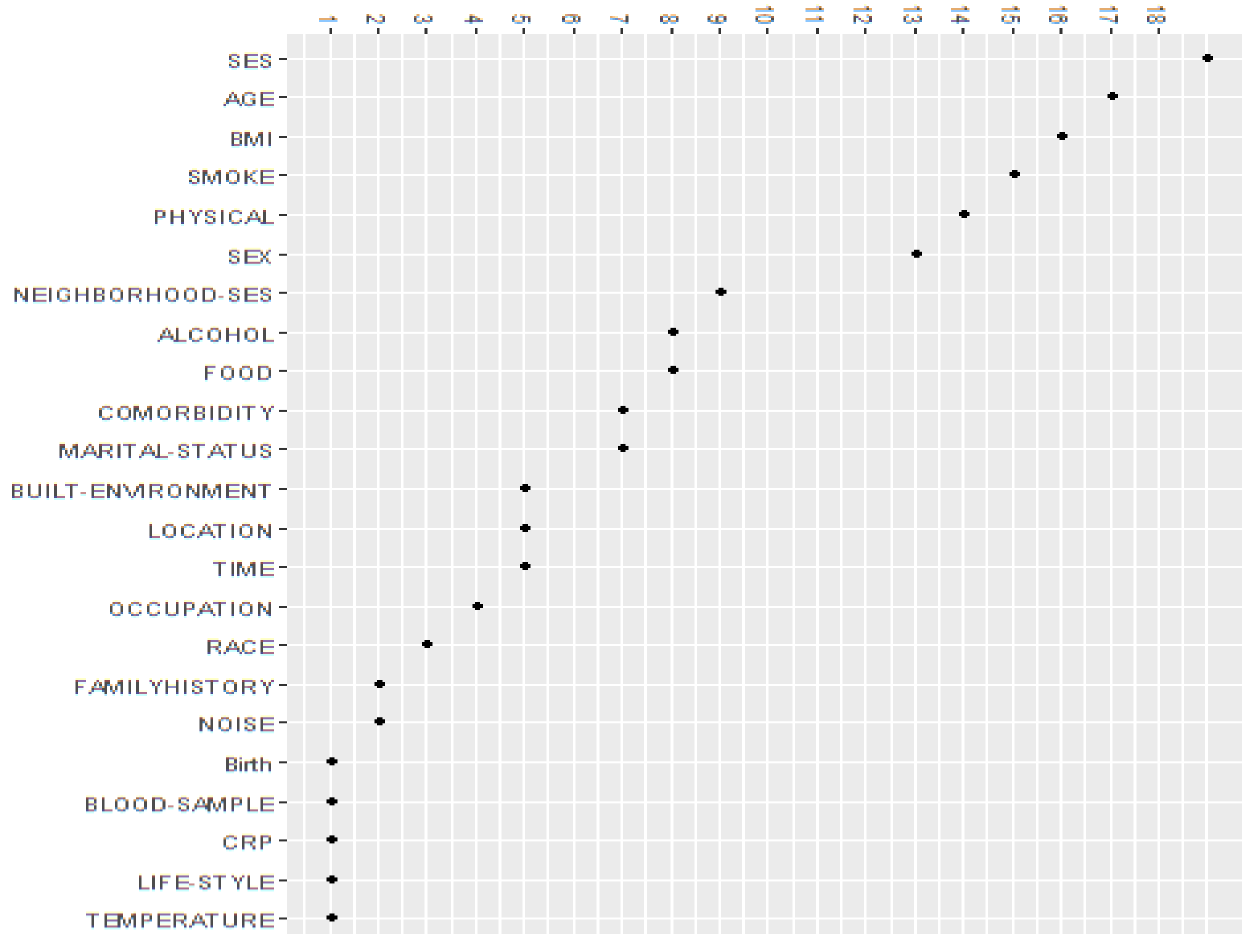
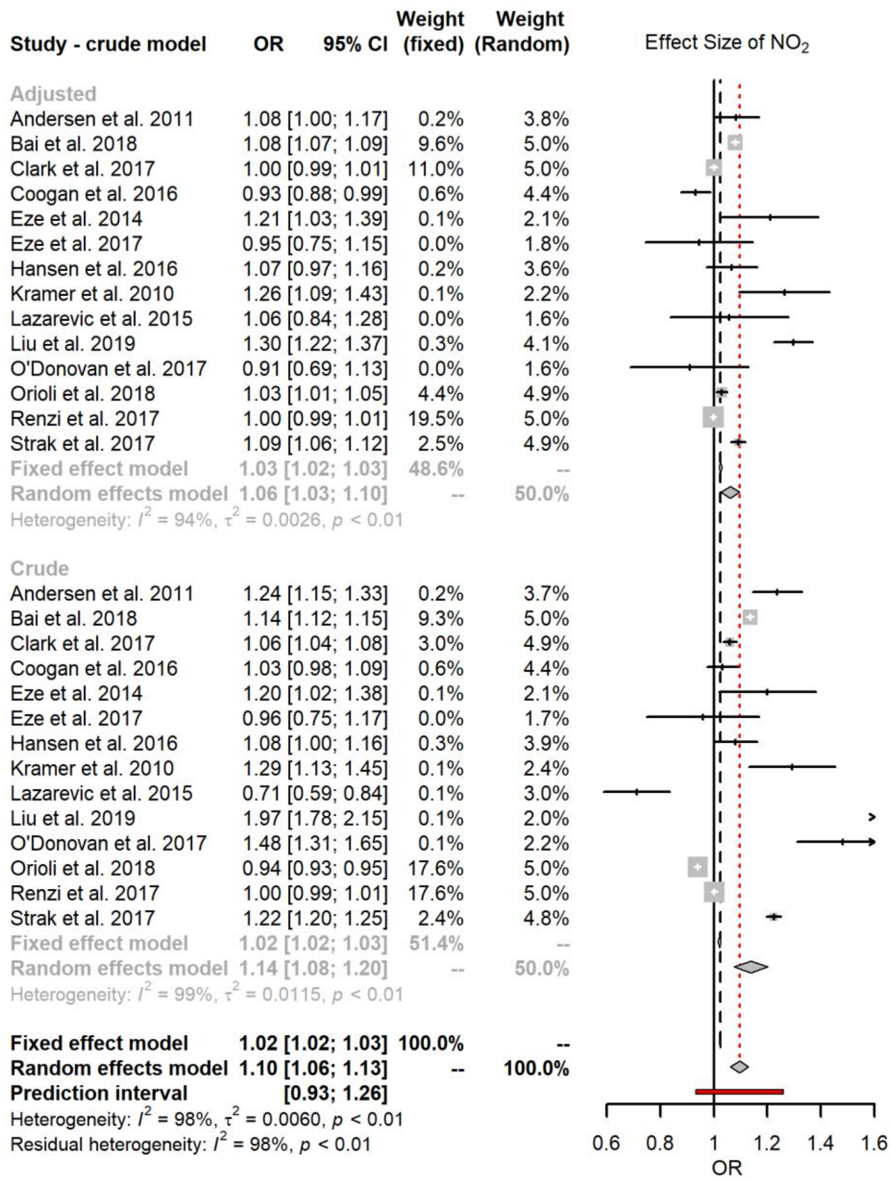


Figure 2-14: $NO_2$ and diabetes crude vs adjusted



# 3. BURDEN OF DISEASE ASSESSMENT

## 3.1 Introduction

Air pollution is a growing contributor to the global burden of disease. An estimated 4.2 million premature deaths are due to ambient air pollution in 2015 (Forouzanfar et al., 2016; Prüss-Üstün et al., 2016). Moreover, air pollution is a leading risk factor for multiple non-communicable diseases including cardiovascular, respiratory, renal, and other diseases (Landrigan et al., 2018). Recent toxicological and epidemiological evidence link air pollution exposure to the development of diabetes mellitus (Table 1-1) (Eze et al., 2015; Wang et al., 2014). There have been several calls from leading health professionals to examine the burden air pollution had on multiple health outcomes including diabetes, and to examine the health disparities associated with such burden (Landrigan et al., 2018). Recent advances in air pollution modeling techniques have made air pollution measurement at fine geographical levels possible. The availability of such measurements makes it possible for public health professionals to model the burden of air pollution on large scales by combining air pollution measurements with multiple publicly available data. This study aims to quantify the burden of diabetes among adults due to air pollution exposure in the United States and compare the burden across several geographic levels.

## 3.2 Methods

### 3.2.1 Study Area and Timeline

We combined data from multiple sources for the 2010 contiguous United States (48 states plus the District of Columbia) at the finest geographical level available. Data included the 2010 US Census, air pollution concentration using a land-use regression model, diabetes incidence rate at the county level, and concentration-response functions for the incident and prevalent diabetes due to $NO_2$ exposure.

### 3.2.2 Census Data

The United States 2010 decennial census data was obtained from the National Historical Geographic Information System (NHGIS) (Manson et al., 2019). The NHGIS provides easy access to US census data. Population counts for adults >= 18 years of age and race and ethnicity were obtained at the census block levels. A census block level is the finest geographical level used to tabulate census data. Census blocks are not consistent in size and are usually defined by physically visible boundaries like roads, rivers, railroads, or land lots (US Census Bureau, 1994). Census blocks are designated as either urban or rural based on population thresholds, nonresidential lad use, and distance from other urban areas. Urban blocks are subdivided into two categories based on population size; urbanized areas (>= 50,000 people) and urban clusters (>= 2,500 to <50,000 people) (Ratcliffe et al., 2016).

### 3.2.3 Diabetes Incidence and Prevalence

County-level diabetes prevalence and incidence rates were readily available and obtained from the United States Diabetes Surveillance System (USDSS) (CDC, 2017b). The USDSS uses data from the Behavioral Risk Factor Surveillance System (BRFSS) to calculate population estimates. The BRFSS is a continuous state-based telephone-based health survey of the adult population (CDC, 2009). The following questions were considered:

- "Has a doctor ever told you that you have diabetes?" – if "yes" and the respondent was not "pregnant" the respondent is considered to have "diagnosed diabetes".
- "How old were you when you were told you have diabetes?" – if the respondent has “diagnosed diabetes” and the difference between their age at the time of the survey and the

age of diagnosis was less than one year the respondent is considered a "newly diagnosed case", however, if the time difference is between one and two years the respondent is weighted as "half a newly diagnosed case".

The precision of estimates was increased using three years of data, the year before and after. For example, the 2010 estimates used data from 2009, 2010, and 2011 to increase precision. Bayesian multilevel modeling techniques for small area estimates were used to obtain county-specific rates. The model makes a county estimate by borrowing BRFSS data from other counties (Barker et al., 2013; Cadwell et al., 2010; Malec et al., 1997; Rao, 2003).

### 3.2.4 Exposure Assessment Model

Exposure levels were assigned as a function of the annual average $NO_2$ concentration at each census block. $NO_2$ concentrations were obtained using a satellite-based land-use regression model (Bechle et al., 2015). The model predicts concentrations at unmeasured areas by combining monthly average readings of $NO_2$ using Environmental Protection Agency (EPA) air quality monitors, remote sensing (satellite) readings, and geographical information systems (GIS) covariates (e.g. major roads, elevation, impervious surfaces, and forests). The developers validated the model using hold-out cross-validation to test the predictive power for $NO_2$ at unmeasured locations (Bechle et al., 2015). The validation showed an ($R^2 = 0.82$) which is relatively good compared to similar LUR models (Beelen et al., 2009b; Hystad et al., 2011; Novotny et al., 2011; Vienneau et al., 2013)

### 3.2.5 Concentration-Response Function

Concentration-response functions (CRF) of incidence and prevalence were obtained from the meta-analytic effect size of the systematic review. The CRF is a measure of how each unit change in $NO_2$ exposure translates into a change in prevalence or incidence of diabetes mellitus among exposed individuals.

### 3.2.6 The Burden of Disease Model

We calculated the attributable number of incident and prevalent cases of diabetes due to $NO_2$ exposure by combining census data, $NO_2$ concentration, diabetes incidence, and prevalence rates, and the concentration-response function. The attributable number of incidence cases ($AC_{IR}$) for each census block is calculated by multiplying the attributable fraction ($AF_b$) with the incident cases ($IC_b$) within a census block.

$$\mathrm{AC_{IR}} = \sum_{\mathrm{i=1}}^{\mathrm{b}} (\mathrm{AF_b} * \mathrm{IC_b})$$

The attributable fraction ($AF_b$) is the relative risk rate difference for each exposure increase in a unit of $NO_2$ ($RR_{diff}$). The $IC_b$ is estimated by multiplying the diabetes mellitus incidence rate, with the number of the at-risk population within a census block. The at-risk population is the number of individuals who either don't have the outcome of interests (diabetes) or were incident cases and is estimated by subtracting the prevalent cases from the total adult population within a census block ($Adult_b$).

$$\mathrm{AC_{IR}} = \sum_{\mathrm{i=1}}^{\mathrm{b}} \left[ \frac{(\mathrm{RR\ diff_b} - 1)}{\mathrm{RR\ diff_b}} * \mathrm{IR} * (\mathrm{Adult_b} - (\mathrm{Adult_b} * \mathrm{PR})) \right]$$

The attributable number of prevalent cases ($AC_{PR}$) for each census block is calculated by multiplying the $RR_{diff}$ with prevalent diabetes cases ($PC_b$) within a census block.

$$AC_{PR} = \sum_{i=1}^{b} (AF_b * PC_b)$$

$$= \sum_{i=1}^{b} \left[ \frac{(RR\ diff_b - 1)}{RR\ diff_b} * PR * \mathrm{Adult_b} \right]$$

The relative risk rate difference is the difference in risk for each unit increase in exposure.

$$RR\ diff_b = e^{\left(\left(\frac{ln(RR)}{RR_u}\right)*NO_{2_b}\right)}$$

Where

$b$ = Represents populated census blocks.

$RR_u$= Exposure unit for the concentration-response (per 4 µg/m$^3$).

$AF_b$= Attributable fraction of diabetes due to $NO_2$ exposure in census block $b$

$NO_{2_b}$= Mean concentration of $NO_2$ in census block $b$.

### 3.2.7 Alternative Scenarios

We modeled and compared the number of attributable cases using an alternative scenario in which $NO_2$ concentrations at any given census block did not exceed the lowest $NO_2$ concentration detected for each corresponding living location category it lies within. This was achieved by replacing $NO_2$ concentrations for each census block with the corresponding concentration:

- Rural areas = 1.48 µg/m$^3$
- Urban clusters = 1.57 µg/m$^3$
- Urbanized areas = 2.59 µg/m$^3$

## 3.3 Results

### 3.3.1 Census Data

There was a total of 6,182,882 populated census blocks in the contiguous US in 2010, of which (58%) were designated as urban areas. The total number of adults was 223,953,591 (73%

of the total adult population in the US). By living location, more than 80% of adults lived in an urban area. A summary of the population characteristics is summarized in Table 3-1.

### 3.3.2 $NO_2$ Concentration

The mean $NO_2$ concentration across populated blocks in the US was 13.2 μg/m$^3$ ranging between 1.5 μg/m$^3$ to 58.3 μg/m$^3$ (Table 3-2). Urban designated blocks had a higher average air pollution concentration then rural blocks. The state with the highest and lowest mean $NO_2$ concentrations was District of Columbia (26.3 μg/m$^3$) and South Dakota (5.2 μg/m$^3$), respectively (Table A 1).

### 3.3.3 Diabetes Prevalent and Incident Cases

Using the county diabetes prevalence and incidence rates, the estimated total number of diabetes prevalent and incident cases among adults was 21,299,056 and 1,938,813 respectively (Table 3-1). More than 75% of both prevalent and incident cases lived in an urban designated census block. A summary of the total diabetes cases by the state is provided in (Table A 1).

### 3.3.4 Attributable Number of Diabetes Cases

The total number of diabetes prevalent and incident cases attributable to air pollution exposure (and fraction) among adults were estimated to be around 5,978,048 (28.1%) and 213,641 (11%) respectively (Table 3-3). The state with the highest attributable prevalent and incident cases was California with 2,106,691 and 197,425 cases respectively (Table A 1). The state with the highest attributable fraction of prevalent and incident cases was the District of Columbia (43.5% and 17.8%) respectively, while the state with the lowest levels were South Dakota (13% and 4.7%). Figure 3-1provides a summary of the distribution of attributable fraction by census block across each state.

### 3.3.5 Attributable Number of Diabetes Cases by Living Location

By living area, the total number of prevalent and incident attributable cases in urban designated areas was 5,212,792(87%) and 188,464 (88%) respectively (Table 3-3). The attributable fractions were highest in blocks designated as urbanized areas (32.8% and 13.0%) compared to urban clusters and rural areas for both the prevalent and incident cases respectively. Table A 2 provides a summary of the attributable fraction of prevalent and incident cases by living locations across each state.

### 3.3.6 Alternative Scenarios

Table 3-4 presents a summary of the change in the number and original estimates using air pollution concentrations reduced to the lowest modeled concentration among each living location. The total reduction was 83% for prevalent cases and 85% for incident cases.

## 3.4 Discussion

### 3.4.1 Main Results

In this study, we modeled the burden of diabetes due to exposure to air pollution across the contingent US using $NO_2$ concentrations at the census block level, diabetes prevalence and incidence rates at the county level, and concentration-response functions derived from a metanalysis. Based on the model we estimated that a large proportion of diabetes cases among adults in the US can be attributable to air pollution exposure. Overall, we found that the total number of attributable cases of prevalent and incident diabetes due to air pollution exposure reached 5,978,048 and 213,641, respectively, among the adult US population (Table 3-3). Adults living in census blocks designated as urbanized areas had a higher attributable fraction compared to other census designations for both prevalent and incident cases. This can be explained by higher $NO_2$ concentrations found in urbanized areas compared to other census designations (Table 3-2). We present a summary of our findings across states (Table A 1-A4 & Figure 3-1). Finally, reducing air pollution levels to the lowest detectable levels may have the potential to reduce the attributable number of cases by 89% (Table 3-4).

### 3.4.2 Comparison with Similar Studies

Bowe et al. (2018a) examined the burden of diabetes due to $PM_{2.5}$ exposure. The study estimated the number of attributable incident cases globally was at 3.2 million (2.2-3.8) while the three largest countries in terms of total cases (in the thousands) were China with 600.3 (447·2–757·3), India 590·5 (447·0–737·1) followed by the US 149·5 (85·2–210·3). The number of attributable cases per population count varied across countries. Pakistan had an ABD per 100 000 population of 58·8 (44·1–74·3), followed by the US 46·3 (26·4–65·1), and India with 44·9 (34·0–56·0). In comparison to our study, Bowe et al. (2018a) examined $PM_{2.5}$ as the exposure of interest using satellite-based data while we examine $NO_2$. , Bowe et al. (2018a) used a theoretical minimum risk exposure level (TMREL) in which exposure values between the minimum and fifth percentile of the exposure distribution did not contribute to the risk of developing diabetes, while we assumed all exposure levels attributed to the risk of developing diabetes.

### 3.4.3 Strengths and Sources of Error

We used a satellite-based model with a relatively high predictive power at unmeasured locations (Bechle et al., 2015). The model provides very fine local level exposure estimates. Despite the high precision of air pollution concentration at the residential location, this model had several limitations described next. First, we used $NO_2$ as a marker of exposure. However, air pollution exposure occurs as a mixture of pollutants (Leaderer et al., 1993). $NO_2$ is a more specific marker of urban sources of air pollution (i.e. vehicle emissions) compared to other pollutants like PM. Studies examining the concentrations of pollutants near motorways show that $NO_2$ levels decline with distance to roadside while $PM_{10}$ and $PM_{2.5}$ do not show a concentration gradient (Roorda-Knape et al., 1998). Second, the LUR model measures air pollutions levels as an indirect method of exposure as opposed to a direct method (i.e. internal dose or personal measurements) (Ott, 1982). However, we believe this method is feasible in our study for the following reasons: First, a direct measurement becomes infeasible to apply to large populations as the cost of would outweigh the benefit. Second, indoor and outdoor air pollution mixtures differ. The main sources of indoor $NO_2$ levels are smoking and gas stoves while the main sources of outdoor $NO_2$ levels in an urban setting are combustion, and in the absence of indoor sources the major source of $NO_2$ levels are outdoor sources (Monn, 2001). Third, although the model estimates exposure of individuals at residential locations, it does not consider spatiotemporal variation (i.e. exposure at work or during grocery shopping). However, the concentration-response function used in the analysis was derived from studies that examined exposure at the residential location (Table 2-2). Fourth, studies using longitudinal repeated measurements found a stronger correlation between outdoor and personal values compared to a single measurement (Monn, 2001). Finally, our exposure model does not consider indoor sources of $NO_2$, studies showed that outdoor values were strongly associated with mortality and morbidity indicating that the health effects are more likely form outdoor sources (Dockery et al., 1993; Schwartz et al., 1996).

### 3.4.4 Concentration Response Function

We used concentration-response functions derived from meta-analytic methods where multiple studies exploring the risk of developing diabetes due to exposure air pollution are pooled together to produce a single effect estimate. Although pooled effect estimates have higher precision, one limitation is that it also aggregates existing bias across included studies (Rothman et al., 2008). Our estimates assume a causative association between exposure to air pollution and diabetes based on a positive concentration-response function,  however, several limitations exist. First, the pooled estimate suffered from a high level of heterogeneity. However, the pooled estimates remained positive across the various strata in the analysis. Second, while the included studies controlled for important confounders of diabetes there remains the possibility of residual confounding which could explain the association. Third, interaction with other pollutants was not considered in all the included studies which also might explain the association. Fourth, the definition of diabetes varied across studies. For example, some studies defined diabetes as self-reported while others used lab-based methods. Finally, the exposure assessment methods varied across included studies, for example, some studies used LUR models, others used dispersion models and air monitors.

### 3.4.5 Incidence and Prevalence Rate

A Bayesian multilevel model was used to estimate the incidence and prevalence rates at the county level. The model had a couple of limitations, though. First, the data used to estimate the county level incidence and prevalence rates were obtained from the BRFSS.which is designed to provide state-level health estimates since not all county estimates (CDC, 2009). The Bayesian model makes indirect estimates by allowing the effect of age, gender, and race/ethnicity of prevalence and incidence to vary by county (Barker et al., 2013; Cadwell et al., 2010). Although we used modeled rates as opposed to direct estimates, we believe the modeled estimates are good estimates that are validated by comparing the modeled to direct estimates of available counties. Secondly, stratified estimates (i.e. age or race/ethnicity) were not possible since sample sizes from the BRFSS at the county level do not support finer level estimates.

Third, the burden estimates are for one year of available data. These estimates might have considerable variation due to changing incidence and prevalence rates and air pollution levels by year. Finally, the incidence and prevalence rates are based on self-reported diabetes and do not account for undiagnosed diabetes. Undiagnosed diabetes could be up to 34% of all diabetes cases in the US (Demmer et al., 2013).

### 3.5 Summary and Conclusion

In summary, our study quantified the burden of diabetes due to air pollution exposure in the United States by combining census data with $NO_2$ concentrations obtained at the census block level from a satellite-based land-use regression model, diabetes prevalence and incidence rates at the county level, and a concentration-response function from a meta-analysis. The study contributes to the limited number of literature estimating the burden of diabetes due to air pollution and answers a call from leading health professionals in this regard. We found that around 28% and 11% of diabetes cases among adults in the United States may be attributable to air pollution exposure. A reduction in air pollution concentrations to the lowest measured concentrations by living location may considerably reduce the number of attributable cases by up to 89%.

Table 3-1: Census data

| | ADULT (%) | | $CASES_{PR}$ (%) | | $CASES_{IR}$ (%) | |
|---|---|---|---|---|---|---|
| Total | 223,953,591 | | 21,299,056 | | 1,938,813 | |
| Rural | 43,927,049 | 19.6% | 4,682,345 | 22.0% | 420,371 | 21.7% |
| Urban cluster | 20,901,097 | 9.3% | 2,153,411 | 10.1% | 194,655 | 10.0% |
| Urbanized area | 159,125,445 | 71.1% | 14,463,299 | 67.9% | 1,323,788 | 68.3% |

Table 3-2: Air pollution summary

| | Mean | Min | first | Median | third | Max |
|---|---|---|---|---|---|---|
| Total | 13.2 | 1.5 | 7.9 | 11.4 | 16.6 | 58.3 |
| Rural | 8.0 | 1.5 | 6.0 | 7.8 | 9.8 | 37.7 |
| Urban cluster | 12.0 | 1.6 | 9.6 | 11.9 | 14.2 | 35.6 |
| Urbanized area | 18.4 | 2.6 | 13.0 | 17.0 | 22.1 | 58.3 |
| <$20,000 | 16.1 | 2.0 | 10.4 | 14.9 | 20.1 | 56.8 |
| $20,000 to <$35,000 | 13.2 | 1.6 | 8.1 | 11.7 | 16.7 | 58.3 |
| $35,000 to <$50,000 | 11.8 | 1.5 | 7.0 | 10.0 | 14.5 | 58.0 |
| $50,000 to <$75,000 | 12.8 | 1.6 | 7.6 | 10.8 | 15.7 | 55.7 |
| >=$75,000 | 16.5 | 2.1 | 10.9 | 14.9 | 20.6 | 55.5 |
| Not defined | 16.0 | 1.8 | 9.2 | 13.6 | 20.2 | 56.3 |
| African American | 16.9 | 1.8 | 10.1 | 15.7 | 22.0 | 56.2 |
| Asian | 22.5 | 2.0 | 14.7 | 21.3 | 29.4 | 55.2 |
| Hispanic | 18.6 | 1.6 | 10.8 | 16.1 | 23.9 | 58.3 |
| Other | 11.5 | 1.6 | 6.9 | 9.5 | 14.1 | 56.7 |
| White | 12.3 | 1.5 | 7.7 | 10.8 | 15.3 | 56.3 |

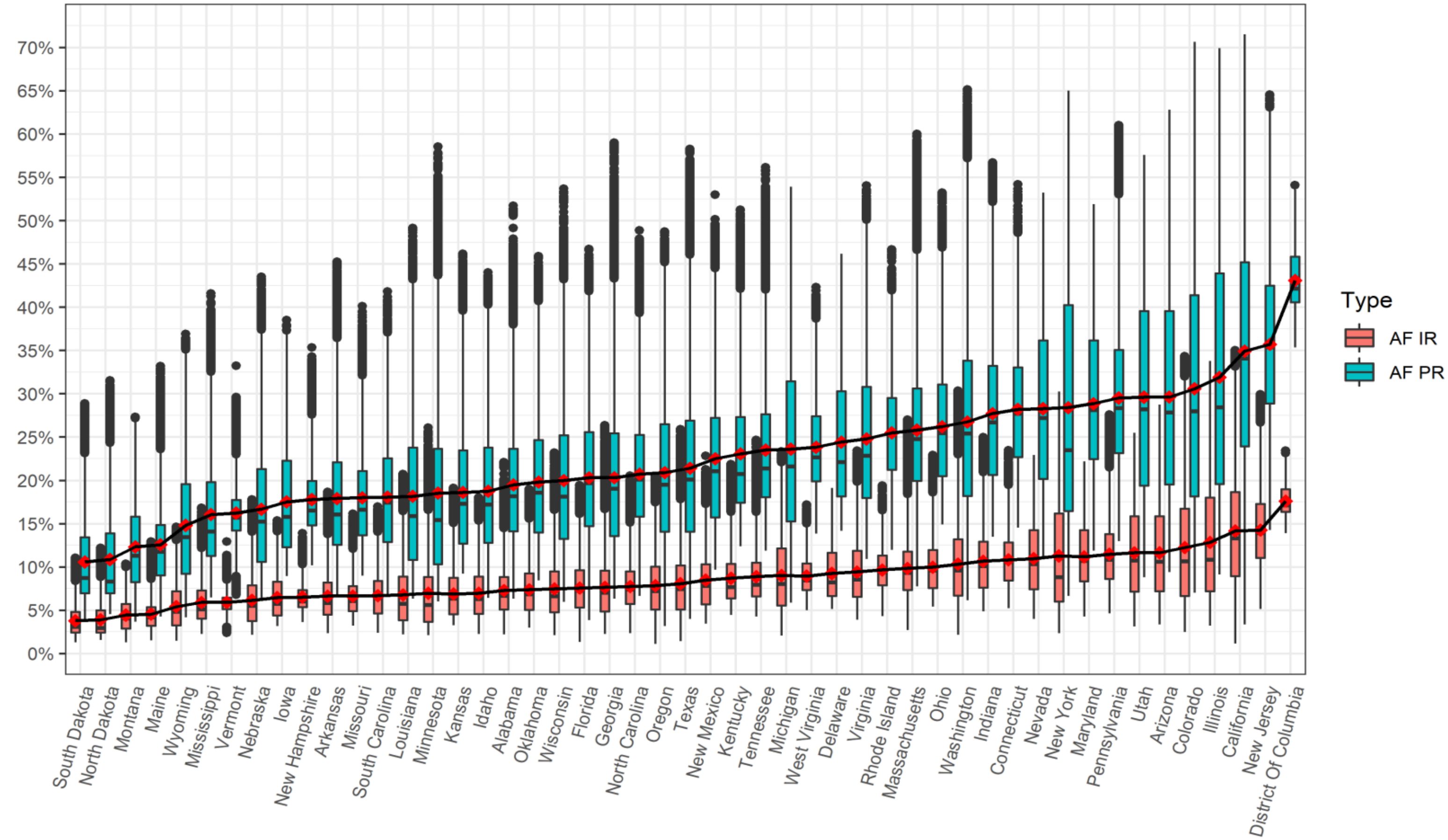
Table 3-3: Burden estimates

| | $AC_{PR}$ | | $AF_{PR}$ | $AC_{IR}$ | | $AF_{IR}$ |
|---|---|---|---|---|---|---|
| Total | 5,978,048 | (% of total) | 28.1% | 213,641 | (% of total) | 11.0% |
| Rural | 765,256 | 13% | 16.3% | 25,177 | 12% | 6.0% |
| Urban cluster | 466,119 | 8% | 21.6% | 15,743 | 7% | 8.1% |
| Urbanized area | 4,746,673 | 79% | 32.8% | 172,721 | 81% | 13.0% |

Table 3-4: Alternative scenario estimates

| | $AC_{PR}$ (change %) | | $AC_{IR}$ (change %) | |
|---|---|---|---|---|
| Total | 1,003,937 | -83% | 31,927 | -85% |
| Rural | 146,718 | -81% | 4,565 | -82% |
| Urban cluster | 71,527 | -85% | 2,242 | -86% |
| Urbanized area | 785,692 | -83% | 25,120 | -85% |

Figure 3-1: Attributable fraction by state



# 4. HEALTH DISPARITY

## 4.1 Introduction

Air pollution is a risk factor for multiple non-communicable diseases and all-cause mortality (Forouzanfar et al., 2016; Landrigan et al., 2018; Prüss-Üstün et al., 2016). Communities of lower and middle socioeconomic status are disproportionally burdened by air pollution (WHO, 2016a). In the US, lower socioeconomic communities are disproportionally exposed to air pollution and health disparities among social strata are known to exist (C. Clark et al., 2017). However, whether the disproportionate air pollution exposure in the US is relevant to public health needs to be examined further. This study aims to explore the health disparities across racial and income strata associated with the burden of diabetes due to air pollution exposure and create easily accessible interactive tools to visualize and explore the burden of disease in the United States.

## 4.2 Methods

We estimated the burden of diabetes due to air pollution for the contiguous United States for the year 2010 using the following data sets; a) decennial census data at the block and block groups level, b) air pollution concentration at the block level, c) diabetes incidence rates at the county level, and d) concentration-response function from a pooled effect estimate of longitudinal studies.

### 4.2.1 Census Data

The 2010 decennial census data for the United States were obtained from the National Historical Geographic Information System (NHGIS) (Manson et al., 2019) which provides easy access to US census data. We categorized census blocks by race/ethnicity as either white, African American, Asian, Hispanic, or other, based on the predominant race and ethnicity residing within the census block. For example, if the largest number of individuals residing within a census block were Hispanic, we would categorize the census block as predominantly Hispanic. If more than one predominant race (i.e. equal number of individuals) resided within the census block we assigned the block as an “other” category.

Median household income was available at the census block group level (one level higher than the census block). Each census block was assigned the median household income of its respective block group. We stratified the income groups using two methods, a) dollar amount and b) centiles of the income distribution. The first methods used the following categories; <\$20,000, \$20,000 to <\$35,000, \$35,000 to <\$50,000, \$50,000 to <\$75,000 and ≥\$75,000 (L. P. Clark et al., 2014). The second method was defined by dividing the median household income into ten equal categories (centiles) based on the income distribution at a) the national level and b) the county level. Census block groups with missing income data were assigned as “unknown”.

Census blocks are designated as either urban or rural based on population thresholds, nonresidential lad use, and distance from other urban areas. Urban blocks are subdivided into two categories based on population size; urbanized areas (>= 50,000 people) and urban clusters (>= 2,500 to <50,000 people)  (Ratcliffe et al., 2016).

### 4.2.2 Air Pollution Concentration

Air pollution concentrations of $NO_2$ for each census block were obtained using a land-use regression model developed by Bechle et al. (2015). The model predicts air pollution concentrations by incorporates data from satellite-based readings, EPA air quality monitors, and GIS covariates including major roads, elevation, impervious surfaces, and forests. The model has a relatively good validation compared to similar models of $R^2 = 0.82$ (Beelen et al., 2009b; Hystad et al., 2011; Novotny et al., 2011; Vienneau et al., 2013).

### 4.2.3 Concentration-Response Function and Diabetes Incidence Rate

We used a concentration-response function from a pooled effect estimate of longitudinal studies examining the risk of exposure to air pollution in the form of $NO_2$ and incident diabetes mellitus among adults. The OR was 1.02 per 10 ug/m$^3$ increase in $NO_2$ exposure. Diabetes incidence rates for 2010 by county was obtained from the USDSS (Barker et al., 2013; Cadwell et al., 2010; CDC, 2017a; Malec et al., 1997; Rao, 2003).

### 4.2.4 The Burden of Disease Model

We estimated the attributable fraction of incident diabetes cases due to $NO_2$ exposure across each census block by combining the previous data sets (census, $NO_2$ concentration, CRF, and diabetes incidence rates) using a burden of disease model described previously (see 3.2.6 The Burden of Disease Model). To examine the health disparities, we compared the attributable fraction of diabetes cases due to air pollution across median household income, predominant race, and living location.

### 4.2.5 Interactive Tools

Lookup tables and maps summarizing the burden across each county were created. The table presents the total population, adults, incident cases, overall attributable fraction, the attributable fraction stratified by race and income. The maps present the attributable fraction by county.

## 4.3 Results

### 4.3.1 Census

Of the 223,953,591 adults living in the US in 2010, 75.1% were white, 12.6% were Hispanic, 9.6% where African American, and 2.1% were Asian. The largest percent lived within a block with a median income of "$50,000 to <$75,000" (Table 4-1). Population counts by centiles of income are summarized in (Table 4-1).

### 4.3.2 Burden

The number of incident attributable cases was highest among census blocks with a predominantly white race (146,414 cases), and attributable fraction among predominantly Asians (17.8%) (Table 4-1). By median household income, the highest attributable cases were among "$50,000 to <$75,000", and the attributable fractions among "<$20,000" (12.7%). Table 4-1 summarizes incident cases, attributable cases, and fractions and by centiles.

The distribution of attributable fraction by race showed that predominantly Asian blocks had the highest mean value (Figure 4-1). Mean values by race were more variable in urbanized areas compared to rural and urban clusters (Figure 4-2). By median household income, census block showed a U shaped distribution, higher in the "<$20,000" and ">=$75,000" groups (Figure 4-3). The U shaped distribution was more apparent in urbanized areas, compared to rural and urban clusters (Figure 4-4). Predominantly Asian blocks had the highest mean value regardless of income level (Figure 4-5).

Centiles using the national distribution of income showed that across living locations higher centile groups had a larger mean value in rural areas, while in urbanized areas there was a U shaped distribution. When allowing the income distribution to vary across counties the burden becomes higher in lower centiles groups within urban clusters and urbanized areas (Figure 4-6).

By race, the mean values of centile groups showed a U shaped with an upward trend. When allowing the centile distribution to vary across counties the trend becomes downward for predominantly African American, Asian, and Hispanic blocks (Figure 4-7). Table A 4 provides a summary of the attributable fraction of prevalent and incident cases by race across each state and Table A 2 provides a summary of the attributable fraction of prevalent and incident cases by median household income across each state

### 4.3.3 Interactive Tools

Using the data, we developed two tools 1) an interactive map by county, and 2) an interactive lookup table by county (see attached HTML files). The interactive map shows each county within a color scheme from green (lower attributable fractions of incident diabetes cases) to dark red (higher attributable fraction). When hovering over a county with the mouse information for the county is presented including the name of the county, state, total adult population, mean $NO_2$ concentration, the estimated attributable diabetes incident cases due to $NO_2$ exposure, and attributable fraction. The interactive lookup table presents each row as a county with the following columns: the state, county name, the total population within the county, total adult population, the estimated number of incident diabetes cases, the estimated attributable number of incident cases, the attributable fraction for the county, and the attributable fraction stratified by predominant race within a census block, and the attributable fraction of census block groups by median household income. Empty cells indicate that the county does not have a corresponding stratum (for example some counties do not have census blocks with some types of predominant race or median household income group). The lookup table also supports multiple features including a search bar, the ability to reorder rows by specific columns, the ability to copy and transfer the data to CSV, excel, pdf, or print. Using the interactive tools we can see that the burden of the disease becomes higher among counties around urbanized areas (northeast regions around New, York, norther regions around Illinois, southwestern in California around Los Angeles) while the burden was lower rural areas (mid-US), we can also hover over any county with the mouse and look up the counties by name using the lookup table for more information within the county. Table A 5 & A6 were built using the interactive lookup table and show the top 20 counties in terms of the attributable number of incident cases and their fractions.

## 4.4 Discussion

### 4.4.1 Health Disparity

We examined the burden of diabetes due to air pollution in the US across racial and income strata. We found the burden was highest among blocks with a predominantly Asian population and lowest income groups. The burden was higher among predominantly Asian blocks across each income group. The burden was more variable within urbanized areas compared to rural and urban clusters. We examined the health disparity within counties of lower vs higher income groups by dividing the income distribution within a county into centiles and found that lower-income centiles had a larger burden compared to higher centiles of income among predominantly non-white blocks. Bowe et al. (2018a) examined the burden of diabetes due to air pollution exposure in the form of $PM_{2.5}$ and found substantial variability among geographies with low-income and low-to-mid-income countries having a higher burden. We examined the burden of diabetes due to exposure to $NO_2$ within the US and found that the burden was highest among the lowest income groups and higher among minority populations.

Our study had several limitations. First, our concentration-response function did not vary by race or income group. Second, when grouping by predominant race for a census we are misclassifying individuals living within a block of a different race. Third, the resolution of air pollution data is a limitation since it does not take me into effect temporal or spatial variation of the individuals exposed nor does it measure indoor air pollution levels.

### 4.4.2 Interactive Tools

We developed an online interactive map and a lookup table to visualize and explore the burden of diabetes due to air pollution at the county level. The interactive map illustrates by color intensity the attributable fraction of diabetes cases due to air pollution exposure with the green color being low and darker red color being a higher value. The map also provides county-level information including the county name, state, the total population of adults, the estimated attributable number of cases, the estimated attributable fraction, and the mean $NO_2$ concentration of all included census blocks within the county. This information is presented when hovering over the county by mouse. We also presented static maps to visualize the attributable fraction of counties stratified by median household income and predominant race (Figure A 1- A3). The interactive lookup table provides more detailed information for each county and includes the state, county name, total population, total adult population, total estimated incident cases, attributable number of incident cases, the attributable fraction of incident cases, and the attributable fraction stratified by predominant race and median household income. The table provides the ability to search each column through the search bar, re-ordering any column by descending or ascending order, and to copy and export the data within the table. For example, Table A 5 & A6 were extracted from the lookup tables and show the burden by county ordered by attributable cases and fractions respectively.

## 4.5 Summary and Conclusion

In summary, we found that the burden of diabetes due to air pollution varied across and race and income levels within the US, the variability was more prominent in urbanized areas. we also developed and made publicly available easy access interactive tools for interested researchers, health professionals, and the general public.

Table 4-1: Health Disparity

| | Category | Adults | Incident Cases | ACIR | AFIR |
|---|---|---|---|---|---|
| Race | White | 168,087,741 | 1,461,053 | 146,414 | 10.0% |
| | African American | 21,446,944 | 205,290 | 25,979 | 12.7% |
| | Asian | 4,621,359 | 34,390 | 6,115 | 17.8% |
| | Hispanic | 28,214,484 | 222,904 | 33,741 | 15.1% |
| | Other | 1,583,063 | 15,177 | 1,391 | 9.2% |
| Income Group | <$20,000 | 7,694,460 | 71,692 | 9,135 | 12.7% |
| | $20,000 to <$35,000 | 36,737,336 | 341,799 | 37,944 | 11.1% |
| | $35,000 to <$50,000 | 57,649,537 | 520,385 | 53,932 | 10.4% |
| | $50,000 to <$75,000 | 69,371,589 | 592,241 | 63,586 | 10.7% |
| | >=$75,000 | 51,758,988 | 406,266 | 48,419 | 11.9% |
| | Not defined | 741,681 | 6,430 | 625 | 9.7% |
| Centile by National Distribution | 1st | 17,565,690 | 164,083 | 19,909 | 12.1% |
| | 2nd | 19,479,462 | 181,240 | 19,937 | 11.0% |
| | 3rd | 20,578,249 | 189,492 | 19,902 | 10.5% |
| | 4th | 21,313,834 | 193,003 | 20,041 | 10.4% |
| | 5th | 22,067,756 | 196,589 | 20,209 | 10.3% |
| | 6th | 22,910,490 | 200,212 | 21,079 | 10.5% |
| | 7th | 23,800,891 | 203,857 | 21,751 | 10.7% |
| | 8th | 24,944,059 | 207,396 | 22,843 | 11.0% |
| | 9th | 25,717,092 | 206,729 | 23,882 | 11.6% |
| | 10th | 24,738,905 | 188,971 | 23,367 | 12.4% |
| Centile by County Distribution | 1st | 26,221,869 | 223,043 | 27,942 | 12.5% |
| | 2nd | 22,876,194 | 194,860 | 23,939 | 12.3% |
| | 3rd | 22,155,299 | 190,457 | 22,577 | 11.9% |
| | 4th | 21,732,490 | 187,719 | 21,620 | 11.5% |
| | 5th | 21,577,946 | 187,070 | 20,740 | 11.1% |
| | 6th | 21,476,600 | 186,495 | 20,073 | 10.8% |
| | 7th | 21,580,627 | 187,864 | 19,628 | 10.4% |
| | 8th | 21,671,843 | 189,458 | 19,080 | 10.1% |
| | 9th | 22,293,006 | 195,649 | 19,126 | 9.8% |
| | 10th | 21,625,520 | 189,764 | 18,292 | 9.6% |

Figure 4-1: Attributable fraction by predominant race

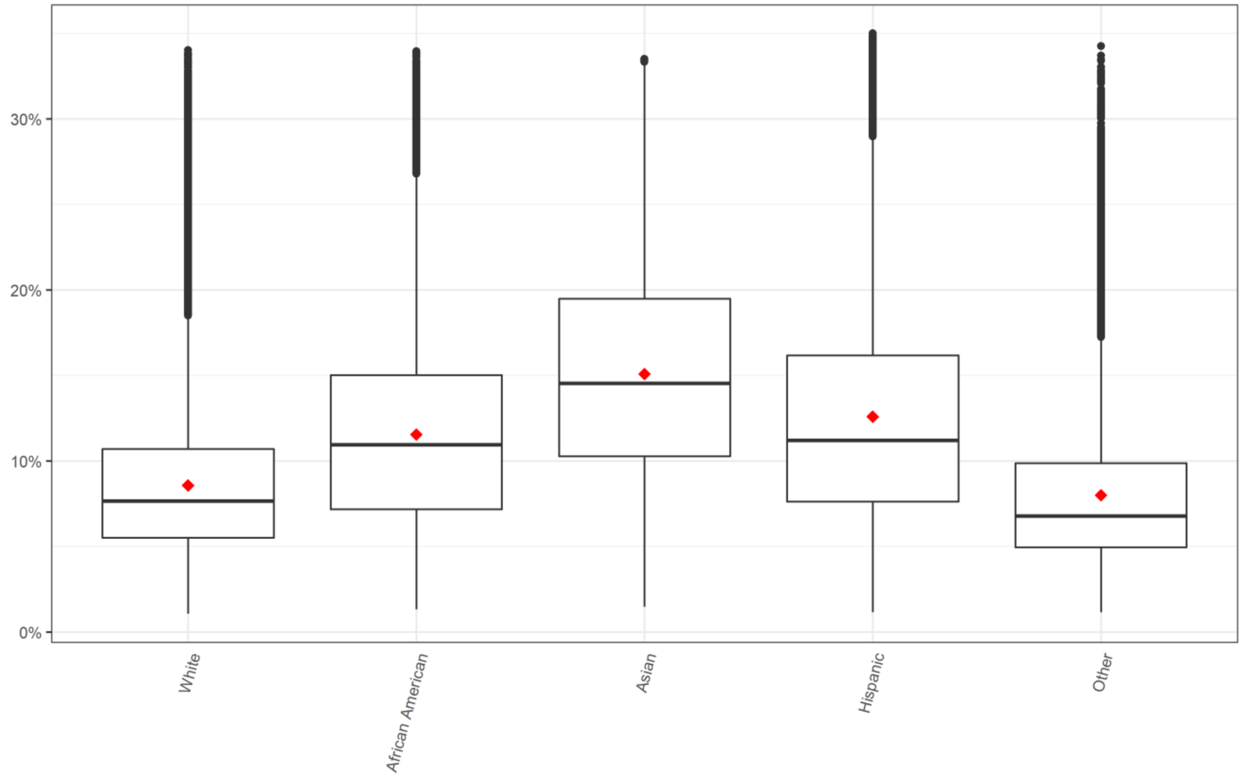



Figure 4-2: Attributable fraction of predominant race by living location

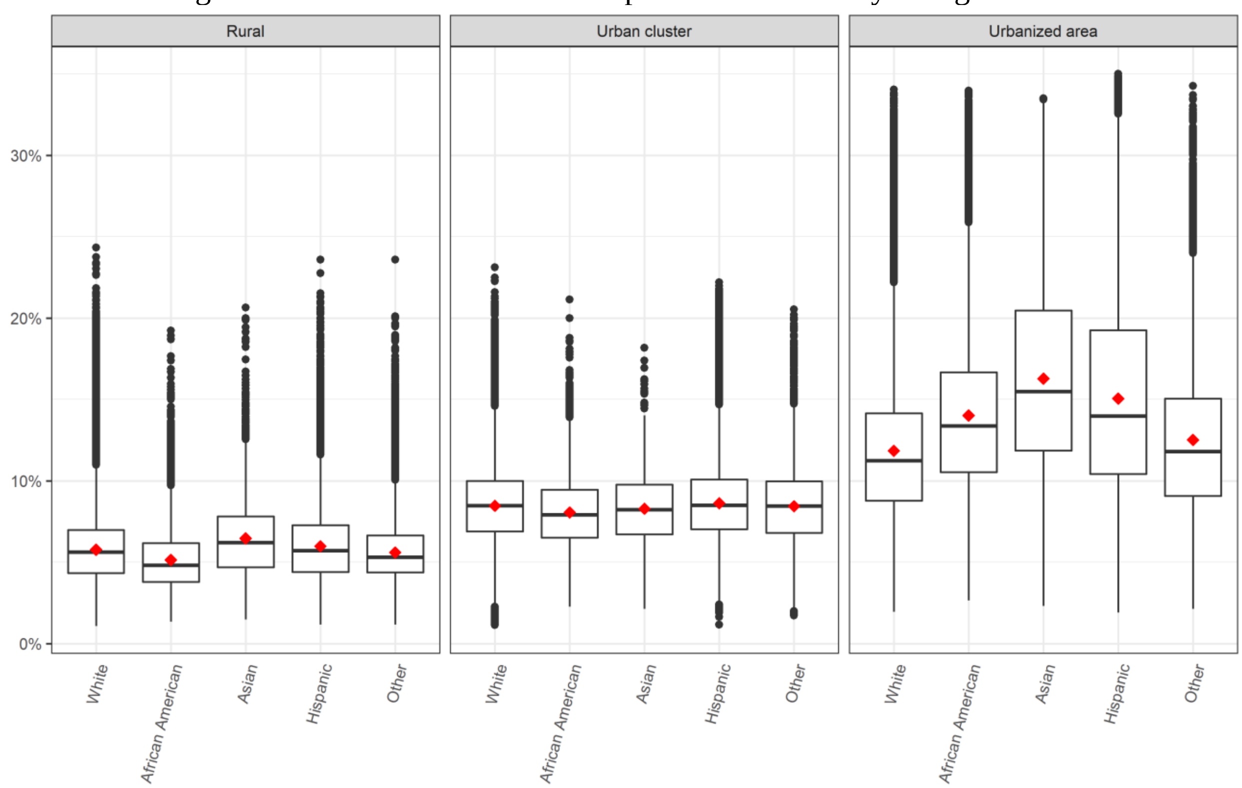


Figure 4-3: Attributable fraction by income

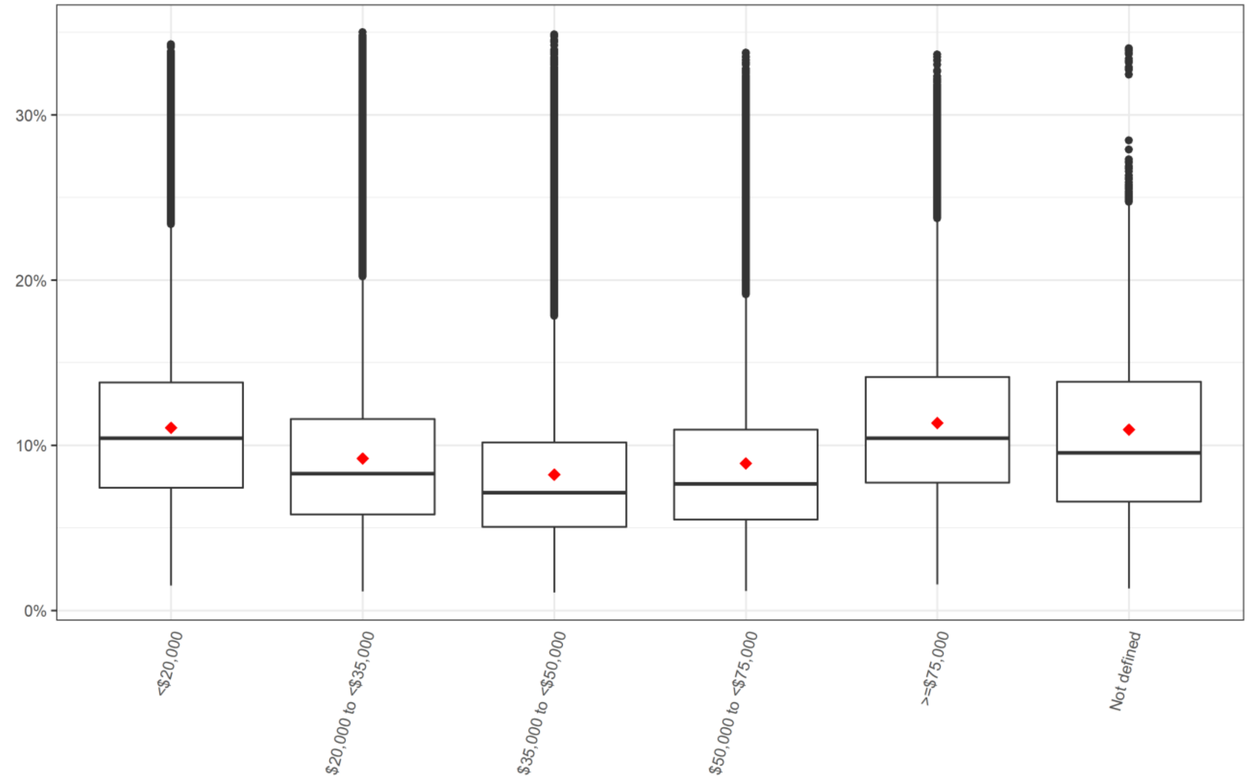



Figure 4-4: Attributable fraction of income by living location

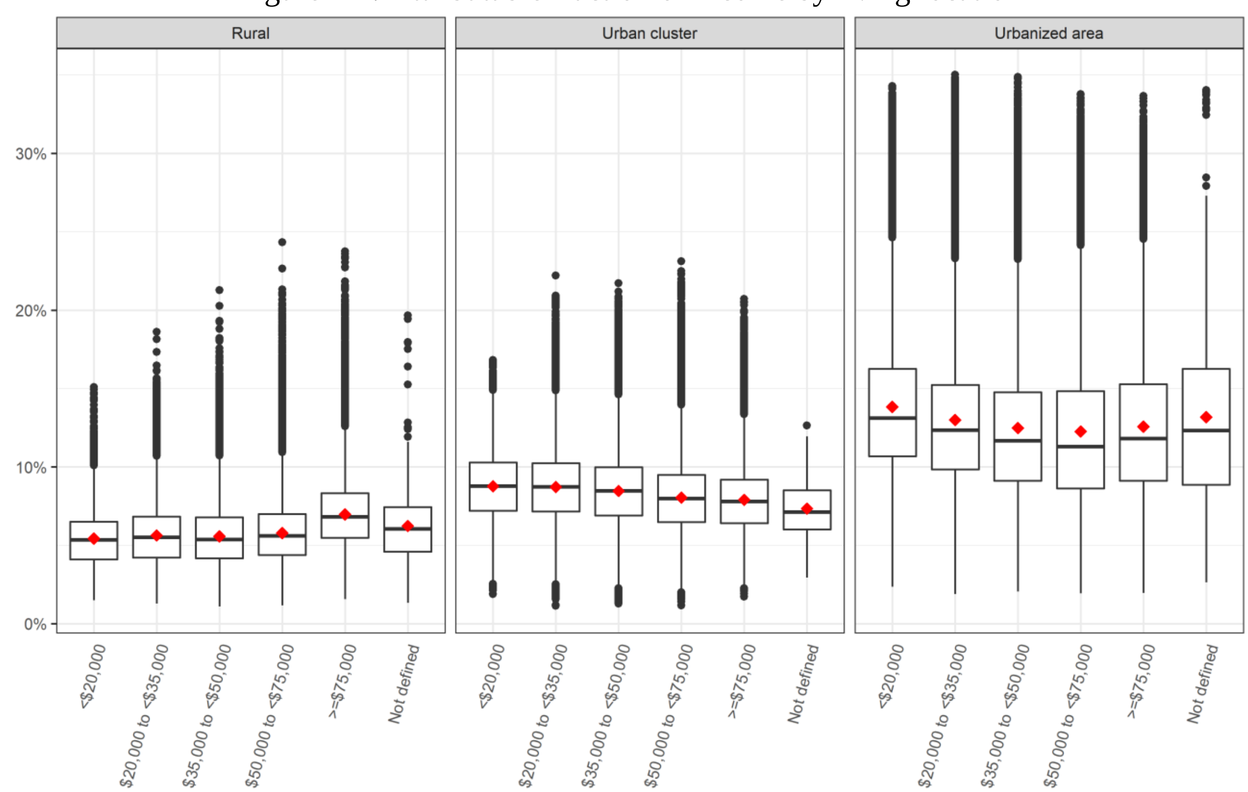


Figure 4-5: Attributable fraction of income by predominant race

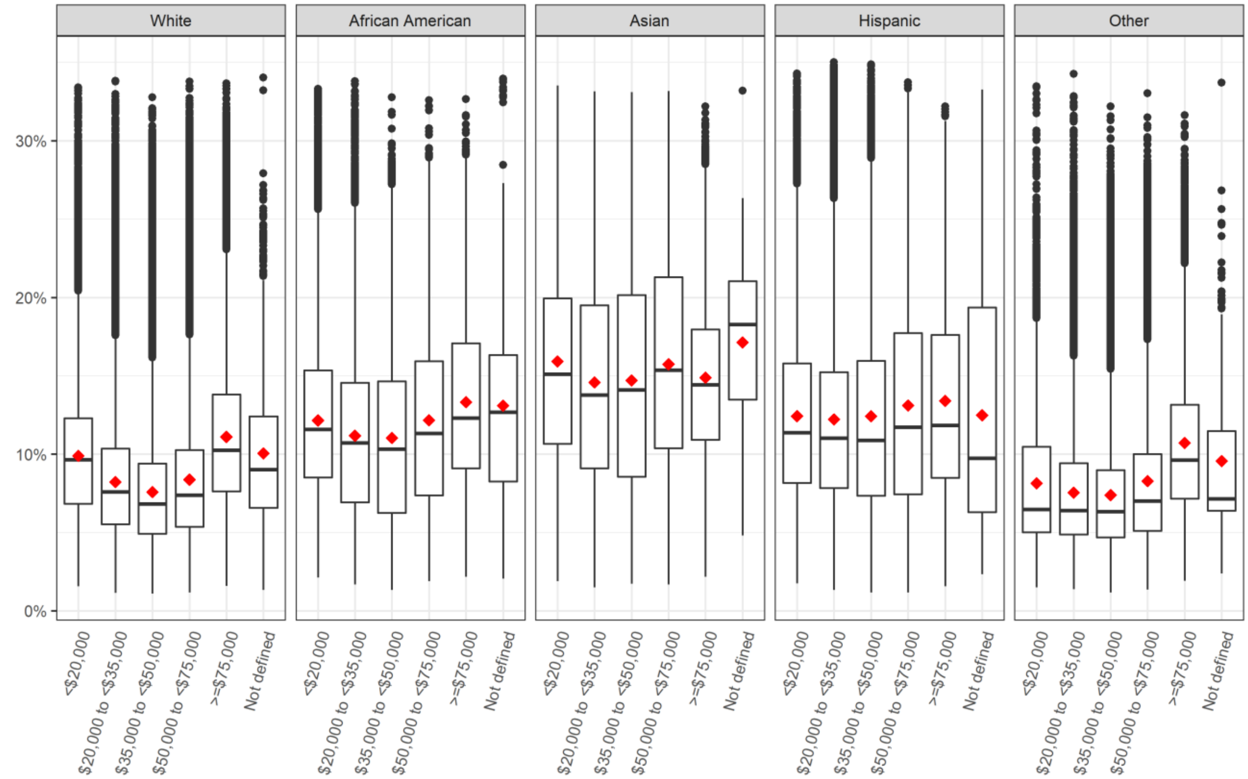
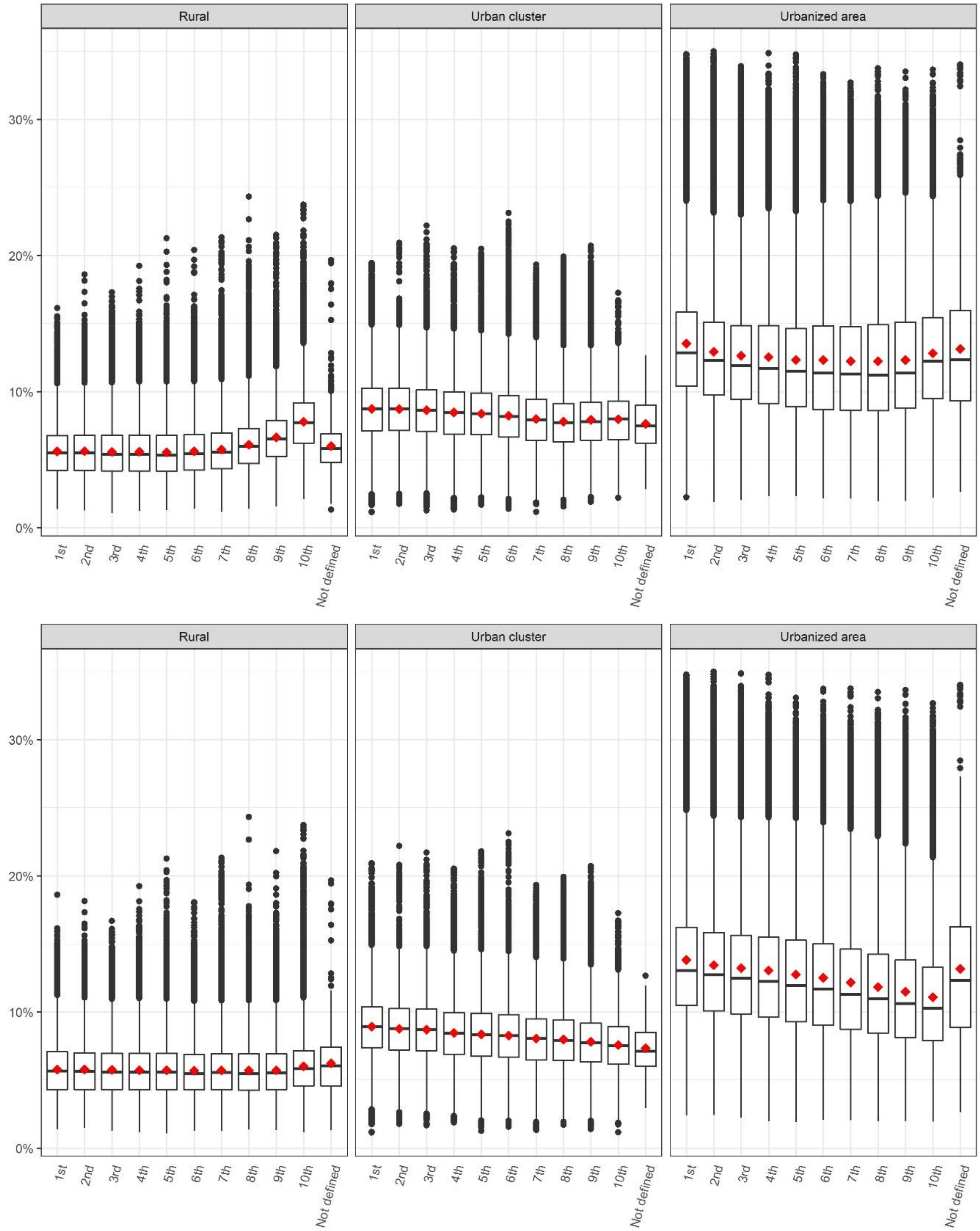
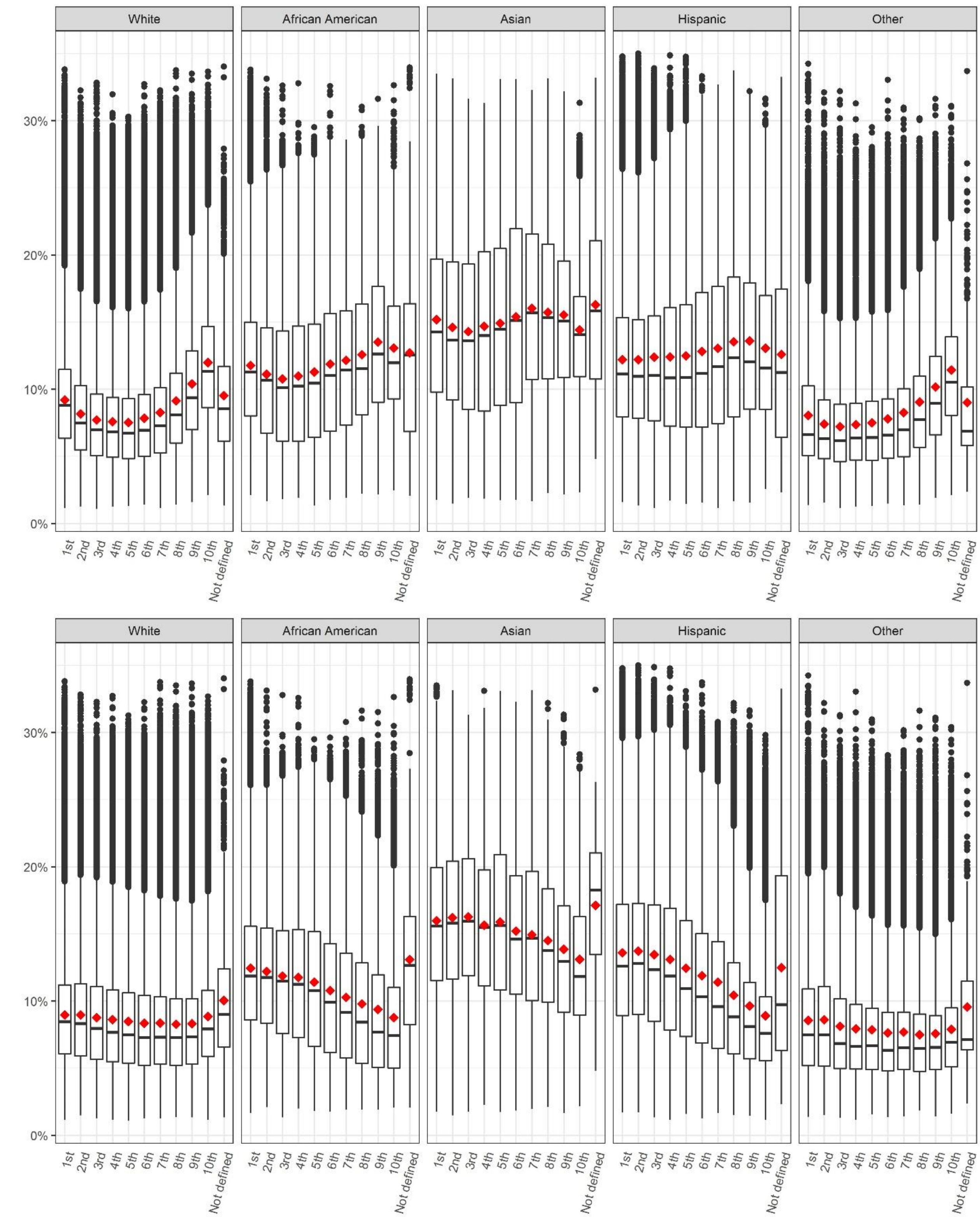


Figure 4-6: Attributable fraction of centile of income by living location




*The upper figure shows centile groups based on the national distribution of income while the lower figure shows centile groups based on the county-level distribution of income*

Figure 4-7: Attributable fraction of centiles of income by predominant race




*The upper figure shows centile groups based on the national distribution of income while the lower figure shows centile groups based on the county-level distribution of incom*

# 5. SUMMARY

The purpose of the dissertation was to assess whether exposure to air pollution increases the risk of developing diabetes mellitus among adults, quantify the burden of diabetes due to air pollution in the United States, and explore the health disparities associated with the burden. We first conducted a systematic review and meta-analysis of studies examine the exposure to air pollution and the risk of developing diabetes mellitus. Secondly, we conducted a burden assessment of diabetes mellitus due to air pollution exposure by combing several data sets for the census, air pollution, prevalence, and incidence rates. Third, we explored how the burden of the disease varies across geographical and social strata including state, county, urban vs rural, predominant race,  and income. Fourth, we created interactive tools to visualize and lookup the burden data at the county level. A summary of our findings is presented in the following sections.

## 5.1 Systematic Review and Meta-Analysis

Air pollution is an emerging global health risk that has been linked to an increase in all-cause mortality and as a cause of multiple non-communicable diseases including cardiovascular diseases, respiratory diseases, neurological and developmental among others. There have been several studies exploring the link between air pollution and the risk of development of air pollution including toxicological and epidemiological studies. Previously published systematic reviews have shown a positive association between air pollution exposure and the risk of developing diabetes mellitus. However, the number of include studies was small and the results showed a wide confidence interval. The number of studies published studies since then has increased, and we aimed to conduct a systematic review and meta-analysis to update the current state of knowledge.

We conducted a review of the literature for studies examining exposure to air pollution in the form of either $NO_2$, BC, or UFP and the risk of developing diabetes mellitus among adults. We included 21 studies in the quantitative analysis. Of these, 20 studies examined the association between exposure to $NO_2$ and the risk of developing diabetes mellitus, 4 examining the exposure to BC, and 2 examined the exposure to UFP. We have concluded that there is sufficient evidence of an association between exposure to $NO_2$ and risk of diabetes among adults based on a moderate quality of evidence, an effect estimate with a positive direction, a pooled effect with a narrow confidence interval with a direction of effect that is unlikely to reverse or reach the null value with an addition of a new study, and a consistent direction of effect estimates among smaller studies. We were not able to reach a similar conclusion for the other pollutants BC and UFP because of the limited number of studies for each.

## 5.2 The Burden of Disease Assessment and Health Disparity

Air pollution is increasingly being recognized as a leading contributor to the global health burden in terms of mortality and morbidity. There have been recent calls to explore and quantify the health burden air pollution is having on society. Therefore, with the advent of technology that measures air pollution at a very fine level, we aimed to quantify the burden of disease from diabetes due to air pollution exposure in the United States utilizing several publicly available data sets and to compare the health disparity in burden among different social strata while creating publicly available and easily accessible interactive tools to visualize and explore the burden of disease.

We joined census data at the census block level, $NO_2$ concentrations from a satellite-based land-use regression model, county-level prevalence and incidence rates, and concentration-response functions from poled effect estimates using a meta-analysis study. We found that many diabetes cases across the United States may be attributable to air pollution exposure and that the burden varied across states, counties, urban vs rural, predominant race within a census block, and median household income. Using the joined data, we were able to create an interactive map and lookup table that easily accessible. The interactive maps help visualize the distribution of burden geographically across counties, while the lookup table can be used hand in hand with the maps to search for more information regarding a specific county. The data is also easily accessible to the general public with the ability to transfer the data in multiple commonly used formats.

## BIBLIOGRAPHY

ADA. (2003). Peripheral arterial disease in people with diabetes. *Diabetes care, 26*(12), 3333-3341.

ADA. (2018). Economic costs of diabetes in the US in 2017. *Diabetes care, 41*(5), 917-928.

ADA. (2019a). 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2019. *Diabetes care, 42*(Supplement 1), S13-S28.

ADA. (2019b). Classification and diagnosis of diabetes: standards of medical care in diabetes—2019. *Diabetes care, 42*(Supplement 1), S13-S28.

Agardh, Emilie, Allebeck, Peter, Hallqvist, Johan, Moradi, Tahereh, & Sidorchuk, Anna. (2011). Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *International journal of epidemiology, 40*(3), 804-818.

Allred, Elizabeth N, Bleecker, Eugene R, Chaitman, Bernard R, Dahms, Thomas E, Gottlieb, Sidney O, Hackney, Jack D, . . . Warren, Jane. (1989). Short-term effects of carbon monoxide exposure on the exercise performance of subjects with coronary artery disease. *New England Journal of Medicine*, *321*(21), 1426-1432.

Andersen, Zorana J, Raaschou-Nielsen, Ole, Ketzel, Matthias, Jensen, Steen S, Hvidberg, Martin, Loft, Steffen, . . . Sorensen, Mette. (2012a). Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes care, 35*(1), 92-98.

Andersen, Zorana J, Raaschou-Nielsen, Ole, Ketzel, Matthias, Jensen, Steen S, Hvidberg, Martin, Loft, Steffen, . . . Sørensen, Mette. (2012b). Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes care, 35*(1), 92-98.

Anderson, HR, Favarato, Graziella, & Atkinson, Richard W. (2011). Long-term exposure to outdoor air pollution and the prevalence of asthma: meta-analysis of multi-community prevalence studies. Air Qual Atmos Health 2013; 6: 57–68. Nishimura KK, Galanter JM, Roth LA, et al. Early life air pollution and asthma risk in minority children: the GALA II & SAGE II studies. *Am J Respir Crit Care Med, 188*, 309-318.

Anderson, HR, Favarato, Graziella, & Atkinson, Richard W. (2013). Long-term exposure to air pollution and the incidence of asthma: meta-analysis of cohort studies. *Air Quality, Atmosphere & Health, 6*(1), 47-56.

Anderson, Ryan J, Freedland, Kenneth E, Clouse, Ray E, & Lustman, Patrick J. (2001). The prevalence of comorbid depression in adults with diabetes: a meta-analysis. *Diabetes care, 24*(6), 1069-1078.

Atkinson, Mark A, Eisenbarth, George S, & Michels, Aaron W. (2014). Type 1 diabetes. *The Lancet, 383*(9911), 69-82.

Bai, Li, Chen, Hong, Hatzopoulou, Marianne, Jerrett, Michael, Kwong, Jeffrey C, Burnett, Richard T, . . . Van Ryswyk, Keith. (2018). Exposure to ambient ultrafine particles and nitrogen dioxide and incident hypertension and diabetes. *Epidemiology, 29*(3), 323-332.

Baker, Michael K, Simpson, Kylie, Lloyd, Bradley, Bauman, Adrian E, & Singh, Maria A Fiatarone. (2011). Behavioral strategies in diabetes prevention programs: a systematic review of randomized controlled trials. *Diabetes research and clinical practice, 91*(1), 1-12.

Balshem, Howard, Helfand, Mark, Schünemann, Holger J, Oxman, Andrew D, Kunz, Regina, Brozek, Jan, . . . Norris, Susan. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of clinical epidemiology, 64*(4), 401-406.

Balti, Eric V, Echouffo-Tcheugui, Justin B, Yako, Yandiswa Y, & Kengne, Andre P. (2014). Air pollution and risk of type 2 diabetes mellitus: a systematic review and meta-analysis. *Diabetes research and clinical practice, 106*(2), 161-172.

Barker, Lawrence E, Thompson, Theodore J, Kirtland, Karen A, Boyle, James P, Geiss, Linda S, McCauley, Mary M, & Albright, Ann L. (2013). Bayesian small area estimates of diabetes incidence by United States county, 2009. *Journal of data science: JDS, 11*(1), 269.

Bateson, Thomas F, & Schwartz, Joel. (2004). Who is sensitive to the effects of particulate air pollution on mortality?: a case-crossover analysis of effect modifiers. *Epidemiology, 15*(2), 143-149.

Baulig, Augustin, Poirault, Jean-Jacques, Ausset, Patrick, Schins, Roel, Shi, Tingming, Baralle, Delphine, . . . Baeza-Squiban, Armelle. (2004). Physicochemical characteristics and biological activities of seasonal atmospheric particulate matter sampling in two locations of Paris. *Environmental science & technology, 38*(22), 5985-5992.

Bechle, Matthew J, Millet, Dylan B, & Marshall, Julian D. (2015). National spatiotemporal exposure surface for NO2: monthly scaling of a satellite-derived land-use regression, 2000–2010. *Environmental science & technology, 49*(20), 12297-12305.

Beelen, Rob, Hoek, Gerard, Houthuijs, Danny, van den Brandt, Piet A, Goldbohm, R Alexandra, Fischer, Paul, . . . Brunekreef, Bert. (2009a). The joint association of air pollution and noise from road traffic with cardiovascular mortality in a cohort study. *Occupational and Environmental Medicine, 66*(4), 243-250.

Beelen, Rob, Hoek, Gerard, Pebesma, Edzer, Vienneau, Danielle, de Hoogh, Kees, & Briggs, David J. (2009b). Mapping of background air pollution at a fine spatial scale across the European Union. *Science of the Total Environment, 407*(6), 1852-1867.

Beelen, Rob, Hoek, Gerard, van Den Brandt, Piet A, Goldbohm, R Alexandra, Fischer, Paul, Schouten, Leo J, . . . Brunekreef, Bert. (2007). Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). *Environmental health perspectives, 116*(2), 196-202.

Biessels, Geert Jan, Staekenborg, Salka, Brunner, Eric, Brayne, Carol, & Scheltens, Philip. (2006). Risk of dementia in diabetes mellitus: a systematic review. *The Lancet Neurology, 5*(1), 64-74.

Borenstein, Michael, Hedges, Larry V, Higgins, Julian PT, & Rothstein, Hannah R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research synthesis methods, 1*(2), 97-111.

Boulton, Andrew JM, Vinik, Arthur I, Arezzo, Joseph C, Bril, Vera, Feldman, Eva L, Freeman, Roy, . . . Ziegler, Dan. (2005). Diabetic neuropathies: a statement by the American Diabetes Association. *Diabetes care, 28*(4), 956-962.

Bowe, Benjamin, Xie, Yan, Li, Tingting, Yan, Yan, Xian, Hong, & Al-Aly, Ziyad. (2018a). The 2016 global and national burden of diabetes mellitus attributable to PM2· 5 air pollution. *The Lancet Planetary Health, 2*(7), e301-e312.

Bowe, Benjamin, Xie, Yan, Li, Tingting, Yan, Yan, Xian, Hong, & Al-Aly, Ziyad. (2018b). Particulate matter air pollution and the risk of incident CKD and progression to ESRD. *Journal of the American Society of Nephrology, 29*(1), 218-230.

Brancati, Frederick L, Whelton, Paul K, Randall, Bryan L, Neaton, James D, Stamler, Jeremiah, & Klag, Michael J. (1997). Risk of end-stage renal disease in diabetes mellitus: a prospective cohort study of men screened for MRFIT. *Jama, 278*(23), 2069-2074.

Brauer, Michael, Hoek, Gerard, Van Vliet, Patricia, Meliefste, Kees, Fischer, Paul H, Wijga, Alet, . . . Kerkhof, Marjan. (2002). Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. *American journal of respiratory and critical care medicine, 166*(8), 1092-1098.

Brook, Robert D, Franklin, Barry, Cascio, Wayne, Hong, Yuling, Howard, George, Lipsett, Michael, . . . Smith Jr, Sidney C. (2004). Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. *Circulation, 109*(21), 2655-2671.

Brook, Robert D, Jerrett, Michael, Brook, Jeffrey R, Bard, Robert L, & Finkelstein, Murray M. (2008a). The relationship between diabetes mellitus and traffic-related air pollution. *Journal of occupational and environmental medicine, 50*(1), 32-38.

Brook, Robert D, Jerrett, Michael, Brook, Jeffrey R, Bard, Robert L, & Finkelstein, Murray M. (2008b). The relationship between diabetes mellitus and traffic-related air pollution. *Journal of occupational and environmental medicine, 50*(1), 32-38.

Cadwell, Betsy L, Thompson, Theodore J, Boyle, James P, & Barker, Lawrence E. (2010). Bayesian small area estimates of diabetes prevalence by US county, 2005. *J Data Sci, 8*(1), 173-188.

Carroll, Raymond J, Ruppert, David, Stefanski, Leonard A, & Crainiceanu, Ciprian M. (2006). *Measurement error in nonlinear models: a modern perspective*: CRC press.

CDC. (2009). Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2019. In.

CDC. (2017a). National diabetes statistics report, 2017. *Atlanta, GA: Centers for Disease Control and Prevention, US Department of Health and Human Services, 20*.

CDC. (2017b). National diabetes statistics report, 2017. *Atlanta, GA: Centers for Disease Control and Prevention*.

CDC. (2020). National diabetes statistics report, 2020. *Atlanta, GA: Centers for*.

Cho, NH, Shaw, JE, Karuranga, S, Huang, Y, da Rocha Fernandes, JD, Ohlrogge, AW, & Malanda, B. (2018). IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. *Diabetes research and clinical practice, 138*, 271-281.

Clark, Charlotte, Sbihi, Hind, Tamburic, Lillian, Brauer, Michael, Frank, Lawrence D, & Davies, Hugh W. (2017). Association of long-term exposure to transportation noise and traffic-related air pollution with the incidence of diabetes: a prospective cohort study. *Environmental health perspectives, 125*(8), 087025.

Clark, Lara P, Millet, Dylan B, & Marshall, Julian D. (2014). National patterns in environmental injustice and inequality: outdoor NO 2 air pollution in the United States. *PloS one, 9*(4), e94431.

Coogan, Patricia F, White, Laura F, Yu, Jeffrey, Burnett, Richard T, Marshall, Julian D, Seto, Edmund, . . . Jerrett, Michael. (2016). Long term exposure to NO2 and diabetes incidence in the Black Women's Health Study. *Environmental research, 148*, 360-366.

Dabelea, Dana, Mayer-Davis, Elizabeth J, Saydah, Sharon, Imperatore, Giuseppina, Linder, Barbara, Divers, Jasmin, . . . Crume, Tessa. (2014). Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *Jama, 311*(17), 1778-1786.

Daneman, Denis. (2006). Type 1 diabetes. *The Lancet, 367*(9513), 847-858.

Demmer, Ryan T, Zuk, Aleksandra M, Rosenbaum, Michael, & Desvarieux, Moïse. (2013). Prevalence of diagnosed and undiagnosed type 2 diabetes mellitus among US adolescents: results from the continuous NHANES, 1999–2010. *American Journal of Epidemiology, 178*(7), 1106-1113.

DerSimonian, Rebecca, & Laird, Nan. (1986). Meta-analysis in clinical trials. *Controlled clinical trials, 7*(3), 177-188.

Dijkema, Marieke BA, Mallant, Sanne F, Gehring, Ulrike, van den Hurk, Katja, Alssema, Marjan, van Strien, Rob T, . . . Hoek, Gerard. (2011). Long-term exposure to traffic-related air pollution and type 2 diabetes prevalence in a cross-sectional screening-study in the Netherlands. *Environmental Health, 10*(1), 76.

Dockery, Douglas W, Pope, C Arden, Xu, Xiping, Spengler, John D, Ware, James H, Fay, Martha E, . . . Speizer, Frank E. (1993). An association between air pollution and mortality in six US cities. *New England Journal of Medicine, 329*(24), 1753-1759.

Egger, Matthias, Smith, George Davey, Schneider, Martin, & Minder, Christoph. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj, 315*(7109), 629-634.

Environmental Protection Agency. (2011). *The Benefits and Costs of the Clean Air Act from 1990 to 2020: Final Report, Rev. A*: US Environmental Protection Agency, Office of Air and Radiation.

Eze, Ikenna C, Foraster, Maria, Schaffner, Emmanuel, Vienneau, Danielle, Héritier, Harris, Rudzik, Franziska, . . . von Eckardstein, Arnold. (2017). Long-term exposure to transportation noise and air pollution in relation to incident diabetes in the SAPALDIA study. *International journal of epidemiology, 46*(4), 1115-1125.

Eze, Ikenna C, Hemkens, Lars G, Bucher, Heiner C, Hoffmann, Barbara, Schindler, Christian, Kunzli, Nino, . . . Probst-Hensch, Nicole M. (2015). Association between ambient air pollution and diabetes mellitus in Europe and North America: systematic review and meta-analysis. *Environmental health perspectives, 123*(5), 381-389.

Eze, Ikenna C, Schaffner, Emmanuel, Fischer, Evelyn, Schikowski, Tamara, Adam, Martin, Imboden, Medea, . . . Künzli, Nino. (2014a). Long-term air pollution exposure and diabetes in a population-based Swiss cohort. *Environment international, 70*, 95-105.

Eze, Ikenna C, Schaffner, Emmanuel, Fischer, Evelyn, Schikowski, Tamara, Adam, Martin, Imboden, Medea, . . . Probst-Hensch, Nicole. (2014b). Long-term air pollution exposure and diabetes in a population-based Swiss cohort. *Environment international, 70*, 95-105.

Flegal, Katherine M, Keyl, Penelope M, & Nieto, F Javier. (1991). Differential misclassification arising from nondifferential errors in exposure measurement. *American Journal of Epidemiology, 134*(10), 1233-1246.

Fong, Donald S, Aiello, Lloyd, Gardner, Thomas W, King, George L, Blankenship, George, Cavallerano, Jerry D, . . . Klein, Ronald. (2004). Retinopathy in diabetes. *Diabetes care, 27*(suppl 1), s84-s87.

Fonken, Laura K, Xu, Xiaohua, Weil, Zachary M, Chen, Guohua, Sun, Qinghua, Rajagopalan, Sanjay, & Nelson, Randy J. (2011). Air pollution impairs cognition, provokes depressivelike behaviors and alters hippocampal cytokine expression and morphology. *Molecular psychiatry, 16*(10), 987.

Forouzanfar, Mohammad H, Afshin, Ashkan, Alexander, Lily T, Anderson, H Ross, Bhutta, Zulfiqar A, Biryukov, Stan, . . . Charlson, Fiona J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet, 388*(10053), 1659-1724.

Fox, Caroline S, Coady, Sean, Sorlie, Paul D, Levy, Daniel, Meigs, James B, D'Agostino, Ralph B, . . . Savage, Peter J. (2004). Trends in cardiovascular complications of diabetes. *Jama, 292*(20), 2495-2499.

Gowers, Alison M, Cullinan, Paul, Ayres, Jon G, Anderson, H Ross, Strachan, David P, Holgate, Stephen T, . . . Maynard, Robert L. (2012). Does outdoor air pollution induce new cases of asthma? Biological plausibility and evidence; a review. *Respirology, 17*(6), 887-898.

Greenland, Sander. (1987). Quantitative methods in the review of epidemiologic literature. *Epidemiologic reviews, 9*(1), 1-30.

Hajat, Anjum, Hsia, Charlene, & O'Neill, Marie S. (2015). Socioeconomic disparities and air pollution exposure: a global review. *Current environmental health reports, 2*(4), 440-450.

Hansen, Anne Busch, Ravnskjær, Line, Loft, Steffen, Andersen, Klaus Kaae, Bräuner, Elvira Vaclavik, Baastrup, Rikke, . . . Brandt, Jørgen. (2016). Long-term exposure to fine particulate matter and incidence of diabetes in the Danish Nurse Cohort. *Environment international, 91*, 243-250.

Health Effects Institute. (2010). *Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects*: Health Effects Institute.

Honda, Trenton, Pun, Vivian C, Manjourides, Justin, & Suh, Helen. (2017). Associations between long-term exposure to air pollution, glycosylated hemoglobin and diabetes. *International journal of hygiene and environmental health, 220*(7), 1124-1132.

Howell, Nicholas A, Tu, Jack V, Moineddin, Rahim, Chen, Hong, Chu, Anna, Hystad, Perry, & Booth, Gillian L. (2019). Interaction between neighborhood walkability and traffic-related air pollution on hypertension and diabetes: the CANHEART cohort. *Environment international, 132*, 104799.

Howells, Lara, Musaddaq, Besma, McKay, Ailsa J, & Majeed, Azeem. (2016). Clinical impact of lifestyle interventions for the prevention of diabetes: an overview of systematic reviews. *BMJ open, 6*(12), e013806.

Hystad, Perry, Setton, Eleanor, Cervantes, Alejandro, Poplawski, Karla, Deschenes, Steeve, Brauer, Michael, . . . Jerrett, Michael. (2011). Creating national air pollution models for population exposure assessment in Canada. *Environmental health perspectives, 119*(8), 1123-1129.

IARC. (2006). *IARC monographs on the evaluation of carcinogenic risks to humans* (Vol. 86): World Health Organization.

Johnson, Paula I, Sutton, Patrice, Atchley, Dylan S, Koustas, Erica, Lam, Juleen, Sen, Saunak, . . . Woodruff, Tracey J. (2014). The Navigation Guide—evidence-based medicine meets environmental health: systematic review of human evidence for PFOA effects on fetal growth. *Environmental health perspectives, 122*(10), 1028-1039.

Khreis, Haneen, Kelly, Charlotte, Tate, James, Parslow, Roger, Lucas, Karen, & Nieuwenhuijsen, Mark. (2017). Exposure to traffic-related air pollution and risk of development of childhood asthma: a systematic review and meta-analysis. *Environment international, 100*, 1-31.

Kim, SJ, Hong, YP, Lew, WJ, Yang, SC, & Lee, EG. (1995). Incidence of pulmonary tuberculosis among diabetics. *Tubercle and lung disease, 76*(6), 529-533.

Koike, Eiko, & Kobayashi, Takahiro. (2006). Chemical and biological oxidative effects of carbon black nanoparticles. *Chemosphere, 65*(6), 946-951.

Kramer, Ursula, Herder, Christian, Sugiri, Dorothea, Strassburger, Klaus, Schikowski, Tamara, Ranft, Ulrich, & Rathmann, Wolfgang. (2010). Traffic-related air pollution and incident type 2 diabetes: results from the SALIA cohort study. *Environmental health perspectives, 118*(9), 1273-1279.

Krämer, Ursula, Herder, Christian, Sugiri, Dorothea, Strassburger, Klaus, Schikowski, Tamara, Ranft, Ulrich, & Rathmann, Wolfgang. (2010). Traffic-related air pollution and incident type 2 diabetes: results from the SALIA cohort study. *Environmental health perspectives, 118*(9), 1273-1279.

Landrigan, Philip J, Fuller, Richard, Acosta, Nereus JR, Adeyi, Olusoji, Arnold, Robert, Baldé, Abdoulaye Bibi, . . . Breysse, Patrick N. (2018). The Lancet Commission on pollution and health. *The Lancet, 391*(10119), 462-512.

Lazarevic, Nina, Dobson, Annette J, Barnett, Adrian G, & Knibbs, Luke D. (2015). Long-term ambient air pollution exposure and self-reported morbidity in the Australian Longitudinal Study on Women's Health: a cross-sectional study. *BMJ open, 5*(10), e008714.

Leaderer, Brian P, Lioy, Paul J, & Spengler, John D. (1993). Assessing exposures to inhaled complex mixtures. *Environmental health perspectives, 101*(suppl 4), 167-177.

Lelieveld, Jos, Evans, John S, Fnais, Mohammed, Giannadaki, Despina, & Pozzer, Andrea. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature, 525*(7569), 367-371.

Leslie, R David, Palmer, Jerry, Schloot, Nanette C, & Lernmark, Ake. (2016). Diabetes at the crossroads: relevance of disease classification to pathophysiology and treatment. *Diabetologia, 59*(1), 13-20.

Li, Yu, Lane, Kevin J, Corlin, Laura, Patton, Allison P, Durant, John L, Thanikachalam, Mohan, . . . Brugge, Doug. (2017). Association of long-term near-highway exposure to ultrafine particles with cardiovascular diseases, diabetes and hypertension. *International journal of environmental research and public health, 14*(5), 461.

Liu, Cuiqing, Bai, Yuntao, Xu, Xiaohua, Sun, Lixian, Wang, Aixia, Wang, Tse-Yao, . . . Harkema, Jack. (2014). Exaggerated effects of particulate matter air pollution in genetic type II diabetes mellitus. *Particle and fibre toxicology, 11*(1), 27.

Liu, Feifei, Guo, Yuming, Liu, Yisi, Chen, Gongbo, Wang, Yuxin, Xue, Xiaowei, . . . Hou, Yitan. (2019). Associations of long-term exposure to PM1, PM2. 5, NO2 with type 2 diabetes mellitus prevalence and fasting blood glucose levels in Chinese rural populations. *Environment international, 133*, 105213.

Lotufo, Paulo A, Gaziano, J Michael, Chae, Claudia U, Ajani, Umed A, Moreno-John, Gina, Buring, Julie E, & Manson, JoAnn E. (2001). Diabetes and all-cause and coronary heart disease mortality among US male physicians. *Archives of Internal Medicine, 161*(2), 242-247.

Lund, Amie K, Knuckles, Travis L, Obot Akata, Chrys, Shohet, Ralph, McDonald, Jacob D, Gigliotti, Andrew, . . . Campen, Matthew J. (2006). Gasoline exhaust emissions induce vascular remodeling pathways involved in atherosclerosis. *Toxicological sciences, 95*(2), 485-494.

Lundberg, V, Stegmayr, B, Asplund, K, Eliasson, M, & Huhtasaari, F. (1997). Diabetes as a risk factor for myocardial infarction: population and gender perspectives. *Journal of internal medicine, 241*(6), 485-492.

Malec, Donald, Sedransk, J, Moriarity, Christopher L, & LeClere, Felicia B. (1997). Small area inference for binary variables in the National Health Interview Survey. *Journal of the American Statistical Association, 92*(439), 815-826.

Manson, S, Schroeder, J, Van Riper, D, & Ruggles, S. (2019). IPUMS National Historicla Geographical Information System: Version 14.0 [database]. IPUMS. *Institute for Social Research and Data Innovation. University of Minnesota*.

Marshall, Julian D, Swor, Kathryn R, & Nguyen, Nam P. (2014). Prioritizing environmental justice and equality: diesel emissions in Southern California. *Environmental science & technology, 48*(7), 4063-4068.

Miettinen, Heikki, Lehto, Seppo, Salomaa, Veikko, Mähönen, Markku, Niemelä, Matti, Haffner, Steven M, . . . Group, FINMONICA Myocardial Infarction Register Study. (1998). Impact of diabetes on mortality after the first myocardial infarction. *Diabetes care, 21*(1), 69-75.

Mills, NL, Tornqvist, H, Robinson, SD, Darnley, K, Gonzales, M, Boon, NA, . . . Sandstrom, T. (2005). *Diesel exhaust inhalation causes vascular dysfunction and impaired endogenous fibrinolysis: An explanation for the increased cardiovascular mortality associated with air pollution.* Paper presented at the Journal of the American College of Cardiology.

Monn, Christian. (2001). Exposure assessment of air pollutants: a review on spatial heterogeneity and indoor/outdoor/personal exposure to suspended particulate matter, nitrogen dioxide and ozone. *Atmospheric Environment, 35*(1), 1-32.

Murabito, Joanne M, D'Agostino, Ralph B, Silbershatz, Halit, & Wilson, Peter WF. (1997). Intermittent claudication: a risk profile from the Framingham Heart Study. *Circulation, 96*(1), 44-49.

Novotny, Eric V, Bechle, Matthew J, Millet, Dylan B, & Marshall, Julian D. (2011). National satellite-based land-use regression: NO2 in the United States. *Environmental science & technology, 45*(10), 4407-4414.

O'Donovan, Gary, Chudasama, Yogini, Grocock, Samuel, Leigh, Roland, Dalton, Alice M, Gray, Laura J, . . . Henson, Joe. (2017). The association between air pollution and type 2 diabetes in a large cross-sectional study in Leicester: The CHAMPIONS Study. *Environment international, 104*, 41-47.

Olefsky, Jerrold M, & Glass, Christopher K. (2010). Macrophages, inflammation, and insulin resistance. *Annual review of physiology, 72*, 219-246.

Orioli, Riccardo, Cremona, Giuseppe, Ciancarella, Luisella, & Solimini, Angelo G. (2018). Association between PM10, PM2. 5, NO2, O3 and self-reported diabetes in Italy: A cross-sectional, ecological study. *PloS one, 13*(1).

Ott, Wayne R. (1982). Concepts of human exposure to air pollution. *Environment international, 7*(3), 179-196.

Ouzzani, Mourad, Hammady, Hossam, Fedorowicz, Zbys, & Elmagarmid, Ahmed. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews, 5*(1), 210.

Prüss-Üstün, Annette, Wolf, Jennyfer, Corvalán, Carlos, Bos, Robert, & Neira, Maria. (2016). *Preventing disease through healthy environments: a global assessment of the burden of disease from environmental risks*: World Health Organization.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rao, JNK. (2003). Small area estimation. Ho oken (NJ): John ile & Sons. In: Inc.

Ratcliffe, Michael, Burd, Charlynn, Holder, Kelly, & Fields, Alison. (2016). Defining rural at the US Census Bureau. *American community survey and geography brief, 1*, 8.

Renzi, Matteo, Cerza, Francesco, Gariazzo, Claudio, Agabiti, Nera, Cascini, Silvia, Di Domenicantonio, Riccardo, . . . Cesaroni, Giulia. (2018). Air pollution and occurrence of type 2 diabetes in a large cohort study. *Environment international, 112*, 68-76.

Riant, Margaux, Meirhaeghe, Aline, Giovannelli, Jonathan, Occelli, Florent, Havet, Anais, Cuny, Damien, . . . Dauchet, Luc. (2018). Associations between long-term exposure to air pollution, glycosylated hemoglobin, fasting blood glucose and diabetes mellitus in northern France. *Environment international, 120*, 121-129.

Roorda-Knape, Mirjam C, Janssen, Nicole AH, De Hartog, Jeroen J, Van Vliet, Patricia HN, Harssema, Hendrik, & Brunekreef, Bert. (1998). Air pollution from traffic in city districts near major motorways. *Atmospheric Environment, 32*(11), 1921-1930.

Rothman, Kenneth J, Greenland, Sander, & Lash, Timothy L. (2008). *Modern epidemiology*: Lippincott Williams & Wilkins.

Sawaya, George F, Guirguis-Blake, Janelle, LeFevre, Michael, Harris, Russell, & Petitti, Diana. (2007). Update on the methods of the US Preventive Services Task Force: estimating certainty and magnitude of net benefit. *Annals of Internal Medicine, 147*(12), 871-875.

Schwartz, Joel, Dockery, Douglas W, & Neas, Lucas M. (1996). Is daily mortality associated specifically with fine particles? *Journal of the Air & Waste Management Association, 46*(10), 927-939.

Schwarzer, Guido, & Schwarzer, Maintainer Guido. (2012). Package 'meta'. *The R Foundation for Statistical Computing, 9*.

Shin, Jinyoung, Choi, Jaekyung, & Kim, Kyoung Jin. (2019). Association between long-term exposure of ambient air pollutants and cardiometabolic diseases: A 2012 Korean Community Health Survey. *Nutrition, Metabolism and Cardiovascular Diseases, 29*(2), 144-151.

Shoelson, Steven E, Lee, Jongsoon, & Goldfine, Allison B. (2006). Inflammation and insulin resistance. *The Journal of clinical investigation, 116*(7), 1793-1801.

Sies, Helmut. (1997). Oxidative stress: oxidants and antioxidants. *Experimental Physiology: Translation and Integration, 82*(2), 291-295.

Stevenson, Catherine R, Critchley, Julia A, Forouhi, Nita G, Roglic, Gojka, Williams, Brian G, Dye, Christopher, & Unwin, Nigel C. (2007). Diabetes and the risk of tuberculosis: a neglected threat to public health? *Chronic illness, 3*(3), 228-245.

Strak, Maciej, Janssen, Nicole AH, Godri, Krystal J, Gosens, Ilse, Mudway, Ian S, Cassee, Flemming R, . . . Brunekreef, Bert. (2012). Respiratory health effects of airborne particulate matter: the role of particle size, composition, and oxidative potential—the RAPTES project. *Environmental health perspectives, 120*(8), 1183-1189.

Strak, Maciej, Janssen, Nicole, Beelen, Rob, Schmitz, Oliver, Vaartjes, Ilonca, Karssenberg, Derek, . . . Brunekreef, Bert. (2017). Long-term exposure to particulate matter, NO2 and the oxidative potential of particulates and diabetes prevalence in a large national health survey. *Environment international, 108*, 228-236.

Sun, Qinghua, Yue, Peibin, Deiuliis, Jeffrey A, Lumeng, Carey N, Kampfrath, Thomas, Mikolaj, Michael B, . . . Parthasarathy, Sampath. (2009). Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation, 119*(4).

US Census Bureau. (1994). Geographic areas reference manual. In: US Department of Commerce Washington, DC.

Vazquez, Gabriela, Duval, Sue, Jacobs Jr, David R, & Silventoinen, Karri. (2007). Comparison of body mass index, waist circumference, and waist/hip ratio in predicting incident diabetes: a meta-analysis. *Epidemiologic reviews, 29*(1), 115-128.

Vienneau, Danielle, De Hoogh, Kees, Bechle, Matthew J, Beelen, Rob, Van Donkelaar, Aaron, Martin, Randall V, . . . Marshall, Julian D. (2013). Western European land use regression incorporating satellite-and ground-based measurements of NO2 and PM10. *Environmental science & technology, 47*(23), 13555-13564.

Vincent, Renaud, Kumarathasan, Premkumari, Goegan, Patrick, Bjarnason, Stephen G, Guénette, Josée, Bérubé, Denis, . . . Miller, Frederick J. (2001). *Inhalation toxicology of urban ambient particulate matter: acute cardiovascular effects in rats*: Health Effects Institute Boston, MA.

Viswanathan, Meera, Ansari, Mohammed T, Berkman, Nancy D, Chang, Stephanie, Hartling, Lisa, McPheeters, Melissa, . . . Tsertsvadze, Alexander. (2012). Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In *Methods guide* *for effectiveness and comparative effectiveness reviews [Internet]*: Agency for Healthcare Research and Quality (US).

Wang, Bin, Xu, Donghua, Jing, Zhaohai, Liu, Dawei, Yan, Shengli, & Wang, Yangang. (2014). Effect of long-term exposure to air pollution on type 2 diabetes mellitus risk: a systemic review and meta-analysis of cohort studies. *European journal of endocrinology, 171*(5), R173-182.

Weyer, Christian, Bogardus, Clifton, Mott, David M, & Pratley, Richard E. (1999). The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. *The Journal of clinical investigation, 104*(6), 787-794.

WHO. (2016a). Ambient air pollution: A global assessment of exposure and burden of disease.

WHO. (2016b). Global report on diabetes. Geneva: World Health Organization; 2016. In.

WHO. (2019). *Classification of diabetes mellitus*. Geneva: WHO.

Wolfsdorf, Joseph, Craig, Maria E, Daneman, Denis, Dunger, David, Edge, Julie, Lee, Warren, . . . Hanas, Ragnar. (2009). Diabetic ketoacidosis in children and adolescents with diabetes. *Pediatr Diabetes, 10*(Suppl 12), 118-133.

Yang, Bo-Yi, Qian, Zhengmin Min, Li, Shanshan, Chen, Gongbo, Bloom, Michael S, Elliott, Michael, . . . Wang, Si-Quan. (2018). Ambient air pollution in relation to diabetes and glucose-homoeostasis markers in China: a cross-sectional study with findings from the 33 Communities Chinese Health Study. *The Lancet Planetary Health, 2*(2), e64-e73.

Ying, Zhekang, Xu, Xiaohua, Bai, Yuntao, Zhong, Jixin, Chen, Minjie, Liang, Yijia, . . . Sun, Qinghua. (2013). Long-term exposure to concentrated ambient PM2. 5 increases mouse blood pressure through abnormal activation of the sympathetic nervous system: a role for hypothalamic inflammation. *Environmental Health Perspectives, 122*(1), 79-86.

You, Wen-Peng, & Henneberg, Maciej. (2016). Type 1 diabetes prevalence increasing globally and regionally: the role of natural selection and life expectancy at birth. *BMJ Open Diabetes Research and Care, 4*(1), e000161.

Zheng, Ze, Xu, Xiaohua, Zhang, Xuebao, Wang, Aixia, Zhang, Chunbin, Hüttemann, Maik, . . . Sun, Qinghua. (2013). Exposure to ambient particulate matter induces a NASH-like phenotype and impairs hepatic glucose metabolism in an animal model. *Journal of hepatology*, *58*(1), 148-154.

## APPENDIX – A

## SUPPLEMENTARY TABLES AND FIGURES

Table A 1: Burden of disease by state

| STATE | NO2 Mean | ADULT | $CASES_{PR}$ | $AC_{PR}$ | $AF_{PR}$ | $CASES_{IR}$ | $AC_{IR}$ | $AF_{IR}$ |
|---|---|---|---|---|---|---|---|---|
| Alabama | 10.3 | 3,503,424 | 443,339 | 86,596 | 19.5% | 38,756 | 2,814 | 7.3% |
| Arizona | 17.0 | 4,572,376 | 390,685 | 129,933 | 33.3% | 36,922 | 4,883 | 13.2% |
| Arkansas | 9.3 | 2,119,988 | 243,761 | 46,237 | 19.0% | 22,613 | 1,600 | 7.1% |
| California | 21.1 | 26,801,914 | 2,106,691 | 832,573 | 39.5% | 197,425 | 31,964 | 16.2% |
| Colorado | 18.1 | 3,664,504 | 214,637 | 80,000 | 37.3% | 20,459 | 3,145 | 15.4% |
| Connecticut | 15.6 | 2,658,321 | 220,773 | 64,946 | 29.4% | 19,424 | 2,204 | 11.3% |
| Delaware | 13.2 | 664,131 | 64,592 | 16,378 | 25.4% | 5,181 | 506 | 9.8% |
| D.C. | 26.3 | 478,003 | 38,718 | 16,854 | 43.5% | 3,163 | 564 | 17.8% |
| Florida | 10.7 | 14,288,320 | 1,486,399 | 306,616 | 20.6% | 133,841 | 10,365 | 7.7% |
| Georgia | 10.8 | 6,906,024 | 724,122 | 166,470 | 23.0% | 64,319 | 5,640 | 8.8% |
| Idaho | 9.8 | 1,092,301 | 95,665 | 20,214 | 21.1% | 8,263 | 656 | 7.9% |
| Illinois | 19.0 | 9,334,110 | 842,843 | 325,042 | 38.6% | 76,793 | 12,182 | 15.9% |
| Indiana | 15.4 | 4,677,220 | 492,231 | 146,023 | 29.7% | 43,746 | 5,053 | 11.6% |
| Iowa | 9.1 | 2,225,845 | 192,282 | 39,784 | 20.7% | 18,241 | 1,412 | 7.7% |
| Kansas | 9.7 | 2,042,474 | 192,709 | 43,070 | 22.3% | 18,243 | 1,536 | 8.4% |
| Kentucky | 12.4 | 3,193,163 | 381,375 | 93,953 | 24.6% | 35,375 | 3,321 | 9.4% |
| Louisiana | 9.6 | 3,279,135 | 390,125 | 77,112 | 19.8% | 36,238 | 2,688 | 7.4% |
| Maine | 6.3 | 1,017,402 | 97,604 | 12,967 | 13.3% | 9,242 | 446 | 4.8% |
| Maryland | 16.1 | 4,256,926 | 412,495 | 129,744 | 31.5% | 38,432 | 4,733 | 12.3% |
| Massachusetts | 14.1 | 4,926,486 | 437,751 | 122,232 | 27.9% | 35,723 | 3,856 | 10.8% |
| Michigan | 12.9 | 7,234,755 | 776,164 | 211,546 | 27.3% | 69,533 | 7,322 | 10.5% |
| Minnesota | 9.9 | 3,872,714 | 288,585 | 70,870 | 24.6% | 23,699 | 2,254 | 9.5% |
| Mississippi | 8.3 | 2,117,802 | 280,115 | 47,356 | 16.9% | 25,994 | 1,620 | 6.2% |
| Missouri | 9.3 | 4,387,516 | 443,897 | 90,215 | 20.3% | 41,632 | 3,148 | 7.6% |
| Montana | 6.2 | 738,379 | 55,437 | 7,662 | 13.8% | 4,847 | 243 | 5.0% |
| Nebraska | 8.6 | 1,313,869 | 111,293 | 24,050 | 21.6% | 9,404 | 767 | 8.2% |
| Nevada | 15.9 | 1,964,223 | 163,182 | 52,555 | 32.2% | 16,337 | 2,069 | 12.7% |
| N. Hampshire | 9.1 | 990,668 | 85,619 | 15,850 | 18.5% | 7,803 | 531 | 6.8% |
| New Jersey | 21.0 | 6,500,690 | 575,430 | 218,150 | 37.9% | 51,439 | 7,918 | 15.4% |
| New Mexico | 12.1 | 1,479,338 | 111,017 | 27,823 | 25.1% | 10,734 | 1,031 | 9.6% |
| New York | 16.6 | 14,480,591 | 1,342,342 | 527,222 | 39.3% | 114,397 | 18,835 | 16.5% |
| North Carolina | 11.0 | 6,976,803 | 739,441 | 155,795 | 21.1% | 68,506 | 5,422 | 7.9% |
| North Dakota | 5.4 | 500,656 | 41,260 | 6,142 | 14.9% | 3,845 | 211 | 5.5% |
| Ohio | 14.3 | 8,469,378 | 917,587 | 258,341 | 28.2% | 82,852 | 8,951 | 10.8% |
| Oklahoma | 10.4 | 2,709,741 | 305,974 | 66,966 | 21.9% | 28,907 | 2,382 | 8.2% |
| Oregon | 11.1 | 2,858,891 | 240,685 | 54,276 | 22.6% | 24,651 | 2,114 | 8.6% |
| Pennsylvania | 16.6 | 9,522,989 | 960,661 | 308,994 | 32.2% | 89,978 | 11,319 | 12.6% |
| Rhode Island | 13.8 | 790,809 | 66,597 | 17,368 | 26.1% | 6,155 | 610 | 9.9% |
| South Carolina | 9.4 | 3,400,939 | 401,904 | 73,780 | 18.4% | 35,471 | 2,414 | 6.8% |
| South Dakota | 5.2 | 587,440 | 51,862 | 6,758 | 13.0% | 4,450 | 210 | 4.7% |
| Tennessee | 12.7 | 4,669,984 | 536,163 | 135,392 | 25.3% | 49,115 | 4,721 | 9.6% |
| Texas | 11.5 | 17,523,847 | 1,604,168 | 379,517 | 23.7% | 145,816 | 13,155 | 9.0% |
| Utah | 17.0 | 1,801,348 | 123,060 | 42,373 | 34.4% | 11,456 | 1,583 | 13.8% |
| Vermont | 8.3 | 475,486 | 34,909 | 5,912 | 16.9% | 2,857 | 177 | 6.2% |
| Virginia | 13.5 | 5,917,339 | 574,533 | 157,289 | 27.4% | 53,592 | 5,678 | 10.6% |
| Washington | 14.9 | 4,954,645 | 422,099 | 122,946 | 29.1% | 43,383 | 4,938 | 11.4% |
| West Virginia | 12.7 | 1,413,781 | 182,729 | 44,038 | 24.1% | 15,434 | 1,400 | 9.1% |
| Wisconsin | 10.6 | 4,184,790 | 362,316 | 86,932 | 24.0% | 31,128 | 2,867 | 9.2% |
| Wyoming | 7.6 | 412,113 | 31,228 | 5,186 | 16.6% | 2,999 | 184 | 6.1% |

Table A 2: Attributable fractions by state and living location

| | $AF_{PR}$ | | | $AF_{IR}$ | | |
|---|---|---|---|---|---|---|
| STATE | Rural | Urban cluster | Urbanized area | Rural | Urban cluster | Urbanized area |
| Alabama | 15% | 19% | 24% | 5% | 7% | 9% |
| Arizona | 18% | 23% | 37% | 7% | 9% | 15% |
| Arkansas | 14% | 21% | 25% | 5% | 8% | 9% |
| California | 20% | 24% | 42% | 7% | 9% | 17% |
| Colorado | 19% | 23% | 43% | 7% | 9% | 18% |
| Connecticut | 21% | 25% | 31% | 8% | 9% | 12% |
| Delaware | 18% | 19% | 29% | 7% | 7% | 11% |
| District of Columbia | N/A | N/A | 44% | N/A | N/A | 18% |
| Florida | 11% | 15% | 22% | 4% | 5% | 8% |
| Georgia | 15% | 18% | 28% | 5% | 7% | 11% |
| Idaho | 14% | 22% | 25% | 5% | 8% | 10% |
| Illinois | 18% | 25% | 43% | 7% | 10% | 18% |
| Indiana | 21% | 28% | 34% | 8% | 11% | 13% |
| Iowa | 15% | 22% | 25% | 6% | 8% | 9% |
| Kansas | 16% | 23% | 26% | 6% | 8% | 10% |
| Kentucky | 18% | 25% | 32% | 7% | 9% | 12% |
| Louisiana | 11% | 16% | 25% | 4% | 6% | 9% |
| Maine | 11% | 16% | 19% | 4% | 6% | 7% |
| Maryland | 22% | 21% | 34% | 8% | 8% | 13% |
| Massachusetts | 18% | 20% | 29% | 6% | 7% | 11% |
| Michigan | 16% | 20% | 32% | 6% | 7% | 13% |
| Minnesota | 13% | 20% | 32% | 5% | 7% | 13% |
| Mississippi | 13% | 21% | 22% | 5% | 8% | 8% |
| Missouri | 15% | 18% | 24% | 5% | 6% | 9% |
| Montana | 10% | 16% | 18% | 4% | 6% | 6% |
| Nebraska | 14% | 21% | 26% | 5% | 8% | 10% |
| Nevada | 18% | 22% | 34% | 7% | 8% | 14% |
| New Hampshire | 15% | 20% | 21% | 6% | 7% | 8% |
| New Jersey | 24% | 27% | 39% | 9% | 10% | 16% |
| New Mexico | 18% | 23% | 31% | 6% | 9% | 12% |
| New York | 16% | 21% | 44% | 6% | 8% | 19% |
| North Carolina | 16% | 20% | 25% | 6% | 8% | 10% |
| North Dakota | 10% | 16% | 20% | 4% | 6% | 7% |
| Ohio | 21% | 26% | 31% | 8% | 10% | 12% |
| Oklahoma | 16% | 22% | 27% | 6% | 8% | 10% |
| Oregon | 14% | 19% | 27% | 5% | 7% | 10% |
| Pennsylvania | 22% | 26% | 36% | 8% | 10% | 14% |
| Rhode Island | 18% | 21% | 27% | 6% | 8% | 10% |
| South Carolina | 14% | 19% | 21% | 5% | 7% | 8% |
| South Dakota | 9% | 15% | 17% | 3% | 6% | 6% |
| Tennessee | 19% | 24% | 30% | 7% | 9% | 12% |
| Texas | 13% | 19% | 27% | 5% | 7% | 10% |
| Utah | 18% | 23% | 38% | 7% | 9% | 15% |
| Vermont | 15% | 21% | 20% | 5% | 8% | 7% |
| Virginia | 19% | 25% | 32% | 7% | 9% | 12% |
| Washington | 18% | 22% | 33% | 7% | 8% | 13% |
| West Virginia | 21% | 26% | 28% | 8% | 10% | 11% |
| Wisconsin | 15% | 22% | 30% | 5% | 8% | 12% |
| Wyoming | 11% | 18% | 22% | 4% | 7% | 8% |

Table A 3: Attributable fraction by state and median income

| STATE | $AF_{PR}$ <\$20,000 | $AF_{PR}$ \$20,000 to <\$35,000 | $AF_{PR}$ \$35,000 to <\$50,000 | $AF_{PR}$ \$50,000 to <\$75,000 | $AF_{PR}$ >=\$75,000 | $AF_{IR}$ <\$20,000 | $AF_{IR}$ \$20,000 to <\$35,000 | $AF_{IR}$ \$35,000 to <\$50,000 | $AF_{IR}$ \$50,000 to <\$75,000 | $AF_{IR}$ >=\$75,000 |
|---|---|---|---|---|---|---|---|---|---|---|
| Alabama | 23% | 20% | 18% | 19% | 21% | 9% | 7% | 7% | 7% | 8% |
| Arizona | 33% | 35% | 33% | 33% | 32% | 13% | 14% | 13% | 13% | 13% |
| Arkansas | 25% | 19% | 18% | 18% | 20% | 9% | 7% | 7% | 7% | 8% |
| California | 45% | 43% | 41% | 40% | 37% | 19% | 18% | 17% | 16% | 15% |
| Colorado | 43% | 42% | 38% | 36% | 36% | 18% | 18% | 16% | 15% | 14% |
| Connecticut | 35% | 34% | 33% | 29% | 27% | 14% | 13% | 13% | 11% | 10% |
| Delaware | 36% | 30% | 24% | 24% | 27% | 14% | 12% | 9% | 9% | 10% |
| D.C. | 45% | 42% | 44% | 43% | 44% | 18% | 17% | 18% | 18% | 18% |
| Florida | 27% | 23% | 21% | 19% | 19% | 10% | 9% | 8% | 7% | 7% |
| Georgia | 24% | 22% | 21% | 23% | 26% | 9% | 8% | 8% | 9% | 10% |
| Idaho | 24% | 24% | 21% | 20% | 22% | 9% | 9% | 8% | 7% | 8% |
| Illinois | 42% | 40% | 38% | 38% | 39% | 18% | 17% | 16% | 15% | 16% |
| Indiana | 37% | 34% | 30% | 27% | 29% | 15% | 13% | 12% | 10% | 11% |
| Iowa | 28% | 24% | 21% | 19% | 20% | 11% | 9% | 8% | 7% | 7% |
| Kansas | 28% | 25% | 22% | 21% | 22% | 11% | 9% | 8% | 8% | 8% |
| Kentucky | 26% | 24% | 24% | 25% | 27% | 10% | 9% | 9% | 10% | 10% |
| Louisiana | 24% | 20% | 18% | 19% | 21% | 9% | 8% | 7% | 7% | 8% |
| Maine | 20% | 14% | 12% | 13% | 14% | 8% | 5% | 5% | 5% | 5% |
| Maryland | 40% | 35% | 34% | 32% | 30% | 16% | 14% | 13% | 12% | 11% |
| Massachusetts | 37% | 33% | 31% | 27% | 26% | 15% | 13% | 12% | 11% | 10% |
| Michigan | 34% | 30% | 25% | 26% | 28% | 14% | 12% | 10% | 10% | 11% |
| Minnesota | 36% | 27% | 22% | 23% | 27% | 15% | 11% | 9% | 9% | 10% |
| Mississippi | 20% | 17% | 16% | 16% | 19% | 8% | 6% | 6% | 6% | 7% |
| Missouri | 24% | 21% | 19% | 20% | 21% | 9% | 8% | 7% | 7% | 8% |
| Montana | 18% | 16% | 13% | 13% | 12% | 7% | 6% | 5% | 5% | 4% |
| Nebraska | 30% | 24% | 21% | 20% | 21% | 12% | 9% | 8% | 8% | 8% |
| Nevada | 43% | 38% | 35% | 31% | 28% | 17% | 15% | 14% | 12% | 11% |
| N. Hampshire | 22% | 23% | 19% | 18% | 17% | 8% | 9% | 7% | 7% | 6% |
| New Jersey | 45% | 42% | 41% | 38% | 36% | 19% | 17% | 17% | 16% | 14% |
| New Mexico | 26% | 25% | 25% | 25% | 25% | 10% | 10% | 10% | 10% | 10% |
| New York | 46% | 43% | 37% | 37% | 41% | 20% | 18% | 16% | 15% | 17% |
| N. Carolina | 25% | 21% | 20% | 21% | 23% | 9% | 8% | 8% | 8% | 9% |
| North Dakota | 22% | 17% | 15% | 13% | 14% | 8% | 6% | 5% | 5% | 5% |
| Ohio | 34% | 31% | 28% | 27% | 27% | 13% | 12% | 11% | 10% | 10% |
| Oklahoma | 26% | 23% | 21% | 21% | 22% | 10% | 9% | 8% | 8% | 8% |
| Oregon | 27% | 23% | 22% | 22% | 24% | 10% | 9% | 8% | 8% | 9% |
| Pennsylvania | 42% | 36% | 30% | 31% | 32% | 17% | 14% | 12% | 12% | 12% |
| Rhode Island | 33% | 30% | 28% | 25% | 22% | 13% | 12% | 11% | 10% | 8% |
| South Carolina | 22% | 19% | 18% | 18% | 18% | 8% | 7% | 7% | 7% | 7% |
| South Dakota | 14% | 14% | 13% | 12% | 13% | 5% | 5% | 5% | 4% | 5% |
| Tennessee | 33% | 26% | 24% | 25% | 26% | 13% | 10% | 9% | 9% | 10% |
| Texas | 26% | 25% | 23% | 22% | 24% | 10% | 10% | 9% | 8% | 9% |
| Utah | 43% | 39% | 35% | 33% | 34% | 18% | 16% | 14% | 13% | 14% |
| Vermont | 24% | 21% | 17% | 16% | 15% | 9% | 8% | 6% | 6% | 6% |
| Virginia | 31% | 26% | 26% | 27% | 29% | 12% | 10% | 10% | 10% | 11% |
| Washington | 34% | 30% | 29% | 28% | 30% | 14% | 12% | 11% | 11% | 12% |
| West Virginia | 28% | 23% | 24% | 25% | 25% | 11% | 9% | 9% | 9% | 10% |
| Wisconsin | 35% | 29% | 23% | 22% | 24% | 14% | 11% | 9% | 8% | 9% |
| Wyoming | 28% | 21% | 17% | 15% | 16% | 11% | 8% | 6% | 6% | 6% |

Table A 4: Attributable fraction by state, race, and ethnicity

| | $AF_{PR}$ | | | | | $AF_{IR}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| STATE | African American | Asian | Hispanic | Other | White | African American | Asian | Hispanic | Other | White |
| Alabama | 22% | 19% | 24% | 19% | 19% | 8% | 7% | 9% | 7% | 7% |
| Arizona | 41% | 43% | 38% | 20% | 32% | 17% | 18% | 16% | 7% | 13% |
| Arkansas | 24% | 22% | 23% | 20% | 18% | 9% | 8% | 9% | 8% | 7% |
| California | 47% | 44% | 44% | 36% | 35% | 20% | 18% | 19% | 15% | 14% |
| Colorado | 51% | 48% | 45% | 36% | 36% | 22% | 20% | 19% | 15% | 15% |
| Connecticut | 35% | 34% | 36% | 33% | 28% | 14% | 13% | 14% | 13% | 11% |
| Delaware | 29% | 31% | 28% | 27% | 25% | 11% | 12% | 11% | 10% | 9% |
| D.C. | 43% | 47% | 43% | 46% | 44% | 18% | 20% | 18% | 19% | 18% |
| Florida | 25% | 22% | 26% | 22% | 19% | 10% | 8% | 10% | 8% | 7% |
| Georgia | 26% | 33% | 29% | 23% | 21% | 10% | 13% | 12% | 9% | 8% |
| Idaho | 28% | 22% | 22% | 17% | 21% | 11% | 8% | 8% | 6% | 8% |
| Illinois | 49% | 48% | 50% | 40% | 35% | 21% | 20% | 22% | 17% | 14% |
| Indiana | 40% | 32% | 40% | 35% | 29% | 16% | 13% | 17% | 14% | 11% |
| Iowa | 29% | 21% | 25% | 21% | 21% | 11% | 8% | 9% | 8% | 8% |
| Kansas | 30% | 24% | 29% | 22% | 22% | 12% | 9% | 11% | 8% | 8% |
| Kentucky | 35% | 33% | 31% | 27% | 24% | 14% | 13% | 12% | 10% | 9% |
| Louisiana | 23% | 25% | 27% | 17% | 18% | 9% | 10% | 11% | 6% | 7% |
| Maine | 24% | 16% | 11% | 9% | 13% | 9% | 6% | 4% | 3% | 5% |
| Maryland | 35% | 32% | 37% | 32% | 29% | 14% | 13% | 15% | 13% | 11% |
| Massachusetts | 36% | 36% | 35% | 32% | 27% | 14% | 15% | 14% | 13% | 10% |
| Michigan | 38% | 34% | 36% | 24% | 26% | 15% | 13% | 14% | 9% | 10% |
| Minnesota | 40% | 37% | 33% | 17% | 24% | 16% | 15% | 13% | 6% | 9% |
| Mississippi | 19% | 18% | 18% | 16% | 16% | 7% | 7% | 7% | 6% | 6% |
| Missouri | 27% | 26% | 24% | 21% | 20% | 10% | 10% | 9% | 8% | 7% |
| Montana | 12% | 13% | 14% | 11% | 14% | 4% | 5% | 5% | 4% | 5% |
| Nebraska | 32% | 30% | 26% | 19% | 21% | 12% | 12% | 10% | 7% | 8% |
| Nevada | 33% | 32% | 39% | 27% | 30% | 13% | 13% | 16% | 11% | 12% |
| N Hampshire | 24% | 21% | 28% | 17% | 19% | 9% | 8% | 11% | 6% | 7% |
| New Jersey | 43% | 43% | 46% | 40% | 35% | 18% | 18% | 19% | 16% | 14% |
| New Mexico | 23% | 28% | 25% | 23% | 25% | 9% | 11% | 10% | 9% | 10% |
| New York | 49% | 54% | 52% | 42% | 34% | 21% | 23% | 23% | 18% | 14% |
| N. Carolina | 23% | 26% | 24% | 16% | 21% | 9% | 10% | 9% | 6% | 8% |
| North Dakota | 16% | 20% | 12% | 9% | 15% | 6% | 8% | 4% | 3% | 6% |
| Ohio | 35% | 34% | 34% | 31% | 27% | 14% | 13% | 13% | 12% | 11% |
| Oklahoma | 28% | 25% | 29% | 19% | 22% | 11% | 10% | 11% | 7% | 8% |
| Oregon | 35% | 27% | 24% | 19% | 22% | 14% | 10% | 9% | 7% | 9% |
| Pennsylvania | 45% | 45% | 42% | 37% | 31% | 18% | 19% | 17% | 15% | 12% |
| Rhode Island | 33% | 35% | 32% | 32% | 25% | 13% | 14% | 12% | 13% | 10% |
| South Carolina | 18% | 23% | 22% | 18% | 18% | 7% | 9% | 8% | 7% | 7% |
| South Dakota | 22% | 16% | 15% | 9% | 13% | 8% | 6% | 5% | 3% | 5% |
| Tennessee | 34% | 29% | 32% | 28% | 24% | 13% | 11% | 13% | 11% | 9% |
| Texas | 28% | 28% | 26% | 23% | 22% | 11% | 11% | 10% | 9% | 8% |
| Utah | 45% | 40% | 43% | 23% | 34% | 19% | 17% | 18% | 9% | 14% |
| Vermont | 16% | 18% | 14% | 17% | 17% | 6% | 6% | 5% | 6% | 6% |
| Virginia | 30% | 35% | 36% | 27% | 27% | 12% | 14% | 14% | 11% | 10% |
| Washington | 44% | 41% | 29% | 24% | 29% | 18% | 17% | 11% | 10% | 11% |
| West Virginia | 26% | 28% | 27% | 25% | 24% | 10% | 11% | 10% | 9% | 9% |
| Wisconsin | 38% | 29% | 36% | 17% | 23% | 15% | 11% | 14% | 6% | 9% |
| Wyoming | 19% | 20% | 19% | 10% | 17% | 7% | 7% | 7% | 4% | 6% |

Table A 5: Top 20 counties ordered by attributable number of incident cases

| | | | | $AF_{IR}$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **STATE** | **COUNTY** | **CASES** | **AC** | **All** | **African American** | **Asian** | **Hispanic** | **Other** | **White** | **$20,000 to <$35,000** | **$35,000 to <$50,000** | **$50,000 to <$75,000** | **>=$75,000** | **<$20,000** |
| California | Los Angeles County | 48,540 | 11,313 | 23% | 24% | 25% | 25% | 22% | 21% | 25% | 24% | 23% | 21% | 27% |
| Illinois | Cook County | 31,256 | 7,086 | 23% | 23% | 23% | 24% | 23% | 22% | 24% | 23% | 22% | 21% | 25% |
| New York | Kings County | 16,862 | 3,745 | 22% | 23% | 21% | 23% | 23% | 21% | 22% | 22% | 22% | 22% | 23% |
| Arizona | Maricopa County | 20,350 | 3,321 | 16% | 18% | 19% | 19% | 15% | 15% | 19% | 18% | 16% | 14% | 20% |
| New York | Queens County | 13,509 | 3,226 | 24% | 23% | 25% | 25% | 24% | 23% | 24% | 24% | 24% | 23% | 21% |
| California | Orange County | 15,759 | 2,744 | 17% | 22% | 19% | 20% | 19% | 16% | 19% | 19% | 18% | 16% | 19% |
| California | San Diego County | 16,618 | 2,643 | 16% | 19% | 16% | 18% | 15% | 15% | 19% | 17% | 16% | 14% | 19% |
| Texas | Harris County | 20,558 | 2,460 | 12% | 13% | 12% | 13% | 12% | 11% | 14% | 13% | 11% | 10% | 15% |
| California | San Bernardino County | 11,413 | 2,335 | 20% | 20% | 22% | 22% | 17% | 17% | 20% | 20% | 20% | 21% | 22% |
| Pennsylvania | Philadelphia County | 11,452 | 2,332 | 20% | 20% | 22% | 21% | 21% | 20% | 21% | 20% | 19% | 20% | 21% |
| Michigan | Wayne County | 15,257 | 2,262 | 15% | 17% | 15% | 17% | 15% | 14% | 16% | 15% | 14% | 12% | 17% |
| New York | New York County | 8,272 | 2,183 | 26% | 26% | 27% | 26% | 25% | 27% | 26% | 26% | 26% | 27% | 26% |
| New York | Bronx County | 8,948 | 2,102 | 23% | 23% | 24% | 24% | 24% | 22% | 24% | 23% | 23% | 22% | 24% |
| California | Riverside County | 14,074 | 1,998 | 14% | 16% | 18% | 15% | 14% | 13% | 14% | 14% | 15% | 14% | 16% |
| Texas | Dallas County | 13,926 | 1,923 | 14% | 13% | 13% | 14% | 13% | 14% | 14% | 14% | 13% | 14% | 14% |
| Florida | Miami-Dade County | 15,090 | 1,672 | 11% | 13% | 9% | 11% | 10% | 9% | 12% | 11% | 11% | 9% | 13% |
| Washington | King County | 10,134 | 1,642 | 16% | 19% | 18% | 17% | 17% | 16% | 19% | 18% | 17% | 15% | 23% |
| Nevada | Clark County | 12,014 | 1,635 | 14% | 13% | 13% | 16% | 13% | 13% | 17% | 15% | 13% | 11% | 18% |
| California | Santa Clara County | 9,300 | 1,441 | 15% | 18% | 16% | 16% | 16% | 15% | 16% | 17% | 16% | 15% | 19% |
| New York | Nassau County | 6,992 | 1,391 | 20% | 21% | 21% | 21% | 21% | 20% | 20% | 20% | 20% | 20% | 21% |

Table A 6: Top 20 counties by the attributable fraction of incident cases

| | | | | Attributable fraction of incident cases | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| STATE | COUNTY | CASES | AC | All | African American | Asian | Hispanic | Other | White | $20,000 to <$35,000 | $35,000 to <$50,000 | $50,000 to <$75,000 | >=$75,000 | <$20,000 |
| New York | New York County | 8,272 | 2,183 | 26% | 26% | 27% | 26% | 25% | 27% | 26% | 26% | 26% | 27% | 26% |
| New York | Queens County | 13,509 | 3,226 | 24% | 23% | 25% | 25% | 24% | 23% | 24% | 24% | 24% | 23% | 21% |
| New Jersey | Hudson County | 3,645 | 858 | 24% | 24% | 24% | 24% | 23% | 22% | 24% | 24% | 23% | 23% | 24% |
| Colorado | Denver County | 2,621 | 640 | 24% | 23% | 24% | 24% | 24% | 25% | 25% | 24% | 24% | 24% | 25% |
| California | Los Angeles County | 48,540 | 11,313 | 23% | 24% | 25% | 25% | 22% | 21% | 25% | 24% | 23% | 21% | 27% |
| Illinois | Cook County | 31,256 | 7,086 | 23% | 23% | 23% | 24% | 23% | 22% | 24% | 23% | 22% | 21% | 25% |
| New York | Bronx County | 8,948 | 2,102 | 23% | 23% | 24% | 24% | 24% | 22% | 24% | 23% | 23% | 22% | 24% |
| New York | Kings County | 16,862 | 3,745 | 22% | 23% | 21% | 23% | 23% | 21% | 22% | 22% | 22% | 22% | 23% |
| California | San Bernardino County | 11,413 | 2,335 | 20% | 20% | 22% | 22% | 17% | 17% | 20% | 20% | 20% | 21% | 22% |
| Pennsylvania | Philadelphia County | 11,452 | 2,332 | 20% | 20% | 22% | 21% | 21% | 20% | 21% | 20% | 19% | 20% | 21% |
| New York | Nassau County | 6,992 | 1,391 | 20% | 21% | 21% | 21% | 21% | 20% | 20% | 20% | 20% | 20% | 21% |
| Colorado | Arapahoe County | 2,398 | 486 | 20% | 23% | 20% | 23% | 22% | 20% | 23% | 22% | 21% | 18% | 23% |
| Colorado | Adams County | 2,014 | 407 | 20% | 23% | 22% | 23% | 21% | 19% | 23% | 23% | 20% | 17% | 24% |
| New Jersey | Essex County | 4,870 | 934 | 19% | 20% | 19% | 21% | 20% | 17% | 21% | 20% | 20% | 17% | 21% |
| New York | Richmond County | 2,964 | 550 | 19% | 18% | 20% | 18% | 18% | 19% | 19% | 18% | 19% | 19% | 18% |
| New Jersey | Union County | 2,914 | 544 | 19% | 19% | 19% | 20% | 19% | 18% | 20% | 20% | 19% | 17% | 20% |
| Colorado | Jefferson County | 2,087 | 396 | 19% | 16% | 19% | 23% | 22% | 19% | 23% | 21% | 20% | 16% | 22% |
| Virginia | Arlington County | 913 | 172 | 19% | 19% | 19% | 17% | 19% | 19% | 16% | 18% | 19% | 19% | 16% |
| Illinois | DuPage County | 4,556 | 813 | 18% | 17% | 17% | 18% | 18% | 18% | 19% | 18% | 18% | 18% | 16% |
| Utah | Salt Lake County | 4,505 | 810 | 18% | 20% | 18% | 20% | 19% | 18% | 21% | 20% | 18% | 16% | 21% |

Figure A 1: Attributable fraction of diabetes due to $NO_2$

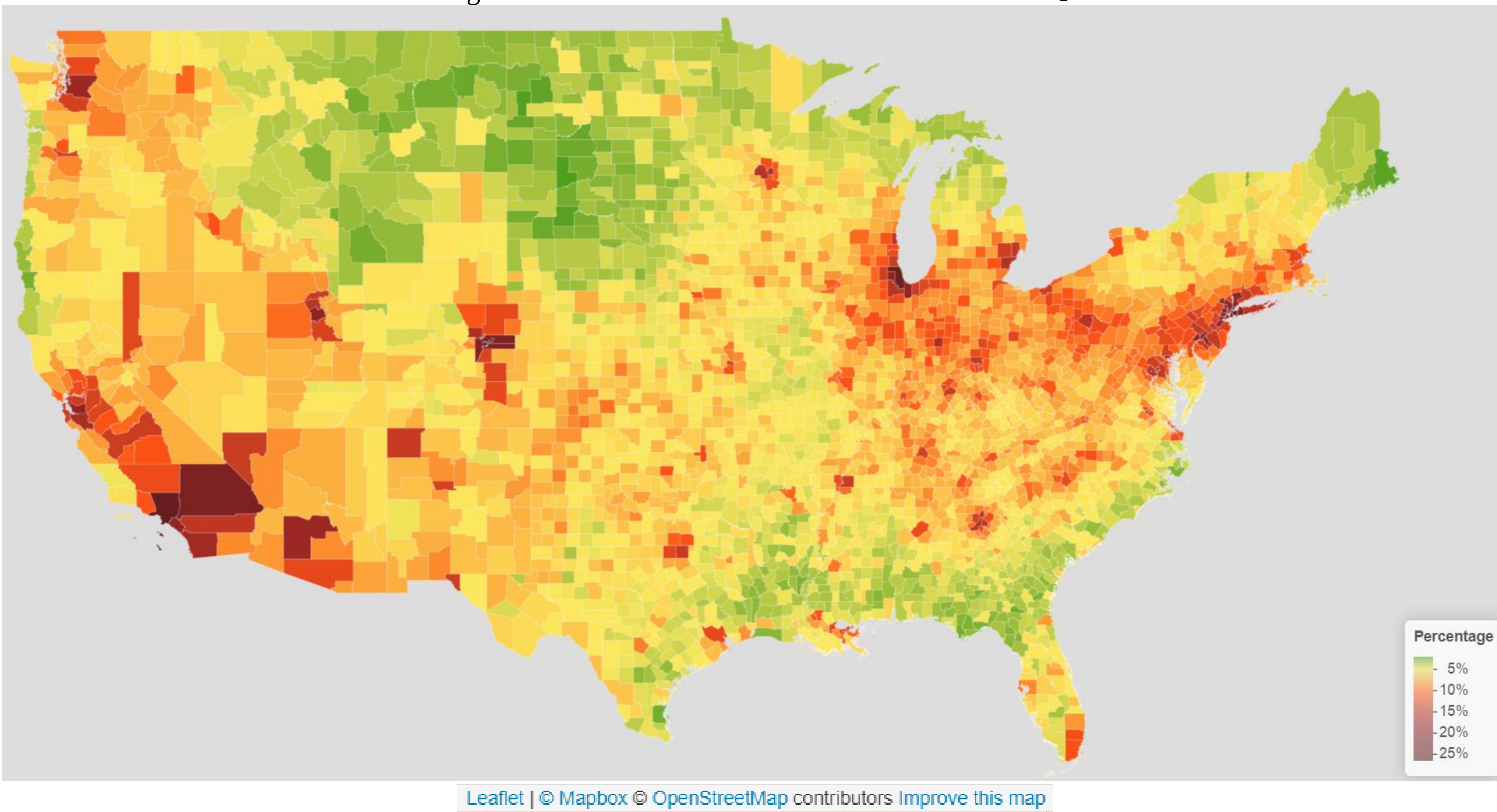
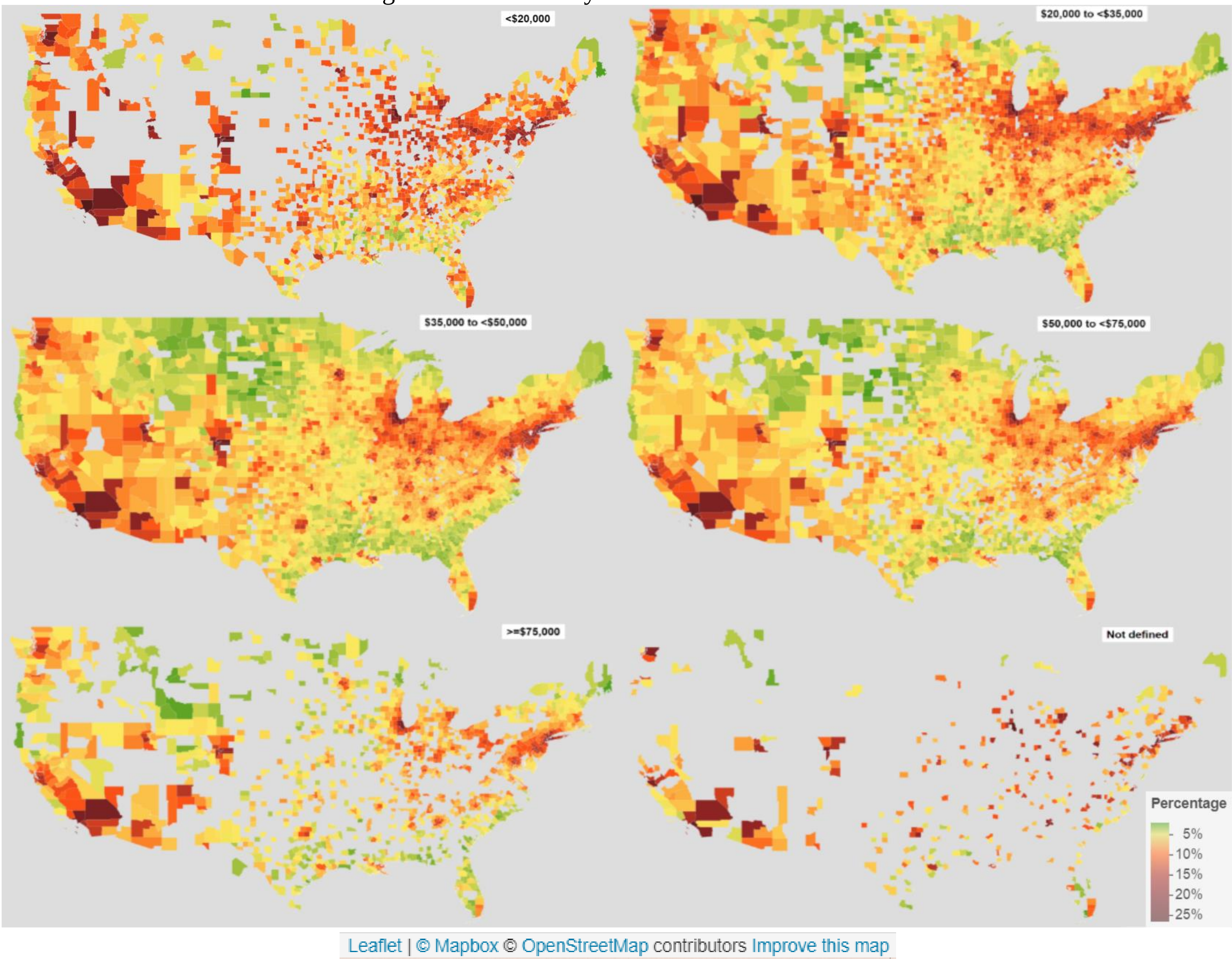
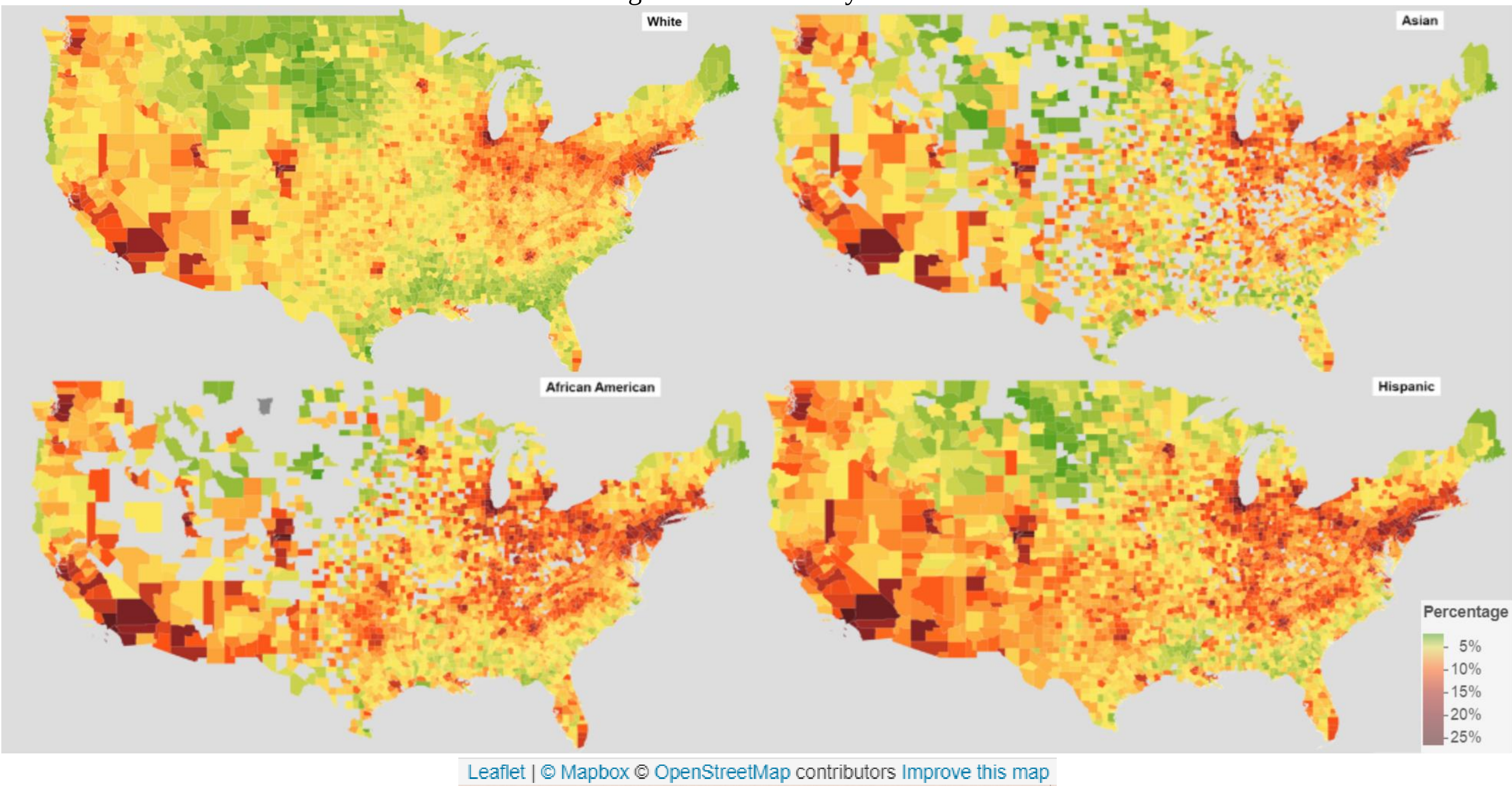


Figure A 2: Burden by median household income




*Counties with empty spaces do not have census blocks with the corresponding median household income

Figure A 3: Burden by race




*Counties with empty spaces do not have census blocks with the corresponding predominant race

APPENDIX – B

SEARCH STRATEGY

**EMBASE Ovid**

Searched on Oct-31-2019 [39 cards]

1. diabet*.ti,ab.
2. exp diabetes mellitus/
3. 1 or 2
4. exp cohort analysis/ or exp longitudinal study/ or exp prospective study/ or exp follow up/ or cohort$.tw. or exp case control study/ or (case$ and control$).tw.
5. (cohort or longitudinal or prospective or retrospective).ti,ab,kw.
6. (cross-sectional or prevalence or transversal).ti,ab,kw.
7. ((case* adj5 control*) or (case adj3 comparison*) or control group*).ti,ab,kw.
8. incidence.ti,ab,kw.
9. or/4-8
10. exp nitrous oxide/
11. (nitrogen ?oxide or NOx or NO2).ti,ab.
12. exp black carbon/
13. (black carbon or carbon black or soot).ti,ab.
14. exp ultrafine particles/
15. (ultrafine or Ultra fine particles or ultrafine particles or ultra fine particulate or ultrafine particulate or UFP or UFPS).ti,ab,kf.
16. or/10-15
17. 3 and 9 and 16
18. limit 17 to english language

**Medline Ovid**

Searched on Oct-31-2019 [206 cards]

1. exp Diabetes Mellitus/
2. diabet*.ti,ab.
3. or/1-2
4. epidemiological methods/

5. limit 4 to yr=1971-1988
6. exp cohort studies/ or controlled clinical trial.pt. or exp case-control studies/ or (case adj2 control*).tw. or Epidemiologic Studies/ or Cohort studies/ or Longitudinal studies/ or Follow-up studies/ or Prospective studies/ or Retrospective studies/
7. (cohort or longitudinal or prospective or retrospective).ti,ab,kw.
8. Case-Control Studies/ or Control Groups/ or Matched-Pair Analysis/ or retrospective studies/
9. Cross-Sectional Studies/ or Prevalence/ or (cross-sectional or prevalence or transversal).ti,ab,kw.
10. Incidence/ or incidence.ti,ab,kw.
11. or/4-10
12. (nitrogen adj2 dioxide).ti,ab.
13. Nitrogen Dioxide/ or (nitrogen ?oxide or NOx or NO2).ti,ab,kf.
14. (black adj2 carbon).ti,ab.
15. Black Carbon/ or carbon black.mp. or soot.ti,ab,kf. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]
16. Ultra fine particles/ or (ultrafine or ultrafine particles or ultra fine particulate or ultrafine particulate or UFP or UFPS).ti,ab,kf.
17. or/12-16
18. 3 and 11 and 17
19. limit 18 to english language

**Transportation Research Information Services (TRIS) Database and the OECD's Joint Transport Research Centre's International Transport Research Documentation (ITRD) Database**

Searched on Oct 30 2019 [7 cards]

(diabetes) AND (no2 or nitrous oxide or nitrogen dioxide or black carbon or carbon black or soot or ultrafine particles or ultrafine particulate or ultra fine particles or ultra fine particulate or ultrafine)

**Risk of Bias From**

Instructions for Making Risk of Bias Determinations

[Note: These questions have been modified from previous applications of the Navigation Guide.]
***Please answer LOW RISK, UNCERTAIN, HIGH RISK, or NOT APPLICABLE and provide details/justification.*** *The following answers pertain to the risk assignment:*

- *"Yes"* → *"Low risk of bias."*
- *"Uncertain"*
- *"No"* → *"High risk of bias"*

1. Was the strategy for recruiting participants consistent across study groups?

***LOW risk of bias (i.e., answer: "YES"):***
Protocols for recruitment and inclusion/exclusion criteria were applied similarly across the study groups, and any one of the following:

- Study participants were recruited from the same population at the same time frame; or
- Study participants were not all recruited from the same population, but the proportions of participants from each population in each study group are uniform

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
Any of the following:

- Protocols for recruitment or inclusion/exclusion criteria were applied differently across study groups; or
- Study participants were recruited at different time frames; or
- Study participants were recruited from different populations and proportions of participants from each population in each study group are not uniform
- A differential loss to follow-up between groups
- Reported refusal/non-response is uniform between groups

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that participant selection is not an element of study design capable of introducing risk of bias in the study.

2. Was knowledge of the exposure adequately prevented during the study?

***LOW risk of bias (i.e., answer: "YES"):***
Any of the following:

- No blinding, but the review authors judge that the outcome and the outcome measurement, as well as the exposure and exposure measurement, are not likely to be influenced by lack of blinding (such as differential outcome assessment where the outcome is assessed using different measurement or estimation metrics across exposure groups, or differential exposure assessment where exposure is assessed using different measurement or estimation metrics across diagnostic or outcome groups); or

- Blinding of key study personnel was ensured, and it is unlikely that the blinding could have been broken; or
- Study personnel was not blinded, but exposure and outcome assessment was blinded and the non-blinding of others is unlikely to introduce bias. For example, investigators were effectively blinded to the exposure and/or outcome groups, or, if the exposure was measured by a separate entity and the outcome was obtained from a hospital record.

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
Any of the following:

- No blinding or incomplete blinding, and the outcome or outcome measurement or exposure and exposure measurement is likely to be influenced by lack of blinding (i.e., differential outcome or exposure assessment); or
- Blinding of key study personnel attempted, but likely that the blinding could have been broken introducing bias; or
- Study personnel was not blinded, and the non-blinding of others was likely to introduce bias.

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that blinding is not an element of study design capable of introducing a risk of bias in the study.

3. Were exposure assessment methods robust?

***The overall considerations include:***

1. What is the quality of the metric being used?
2. Has the metric been validated for the scenario for which it is being used?
3. Did the analysis account for prediction uncertainty?
4. How was missing data accounted for, and any data imputations incorporated?
5. Were sensitivity analyses performed?

***For exposure assessment models consider the following:***

1. Were the input data in the study suspected to systematically under-or over-estimate exposure?
2. What type of model was used (geostatistical interpolation, land-use regression, dispersion models, personal air sampling models, hybrid models, etc.)?
3. Were data on land use, topography, traffic, monitoring data, emission rates, etc. incorporated and justified by authors in their selection?
4. What were the spatial variation (e.g., distance from the source) and geographic/spatial accuracy (county, census tract, individual residence)?

5. What was the address completeness (e.g., only home address at one point in time, or more complete address history throughout pregnancy/postnatal life and other locations such as work)?
6. What was the space-time coverage of the model?
7. Were time-activity patterns accounted for?

***LOW risk of bias (i.e., answer: "Yes"):***
The reviewers judge that there is a low risk of exposure misclassification, i.e.:

- There is high confidence in the accuracy of the exposure assessment methods (i.e. tested for validity and reliability) in measuring the targeted exposure; or
- Less-established or less direct exposure measurements are validated against well-established or direct methods AND if applicable (e.g. for laboratory measurements), appropriate QA/QC for methods are described and are satisfactory, with at least three of the following items reported, or at least two of the following items reported plus evidence of satisfactory performance in a high-quality inter-laboratory comparison:
- a measure of repeatability;

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
The reviewers judge that there is a high risk of exposure misclassification and any one of the following:

- There is low confidence in the accuracy of the exposure assessment methods; or
- Less-established or less direct exposure measurements are not validated and are suspected to introduce bias that impacts the outcome assessment (example: participants are asked to report exposure status retrospectively, subject to recall bias); or
- Uncertain how exposure information was obtained; or:
  A) Monitoring: Information from databases or otherwise was gathered that indirectly assessed exposure without considering variables noted in the List of Considerations above, such as spatial variability, land-use regression, etc., or there is sufficient evidence that relevant factors from the List of Considerations above would imply a risk of bias in the exposure assessment.
  B) Modeling: the air pollution model used has been demonstrated not to pertain to area-based or person-based measures or has otherwise been previously demonstrated to be unable to describe air levels of exposure for assigning exposure in a research situation, or there is sufficient evidence that relevant factors from the List of Considerations above would imply a risk of bias in the exposure assessment.

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that exposure assessment methods are not capable of introducing a risk of bias in the study.
4. Was confounding adequately addressed?

The following are a list of confounders we considered "important" to the outcome of interest:

- Age
- Comorbidity ; may include any of the following; hypertension, dyslipidemia, myocardial infarction, or metabolic syndrome.
- Obesity; measured using either weight, BMI, hip to waist ratio or any other method indicating that the investigators controlled for obesity
- Family History
- Lifestyle variables; in the form of physical activity, exercise, or diet.
- Gender
- Socioeconomic status; ascertained using several metrics including the level of income, education, or neighborhood status.
- Smoking; includes active or passive smoking (also known as environmental smoking or secondhand smoking)

***LOW risk of bias (i.e., answer: "Yes"):***
The study accounted for (i.e., matched, stratified, multivariate analysis or otherwise statistically controlled for) four or more confounders or reported that potential confounders were evaluated and omitted because inclusion did not substantially affect the results.

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
The study accounted for less than four of the listed potential confounders.

5. Was incomplete outcome data adequately addressed?

***LOW risk of bias (i.e., answer: "Yes"):***
Participants were followed long enough to obtain outcome measurements; OR any one of the following:

- No missing outcome data; or
- Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to introduce bias); or
- Attrition or missing outcome data balanced in numbers across exposure groups, with similar reasons for missing data across groups; or
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a relevant impact on the intervention effect estimate; or
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a relevant impact on the observed effect size; or
- Missing data have been imputed using appropriate methods

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***

Participants were not followed long enough to obtain outcome measurements; OR any one of the following:

- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across exposure groups; or
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce biologically relevant bias in intervention effect estimate; or
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce biologically relevant bias in observed effect size; or
- Potentially inappropriate application of imputation.

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that incomplete outcome data is not capable of introducing a risk of bias in the study.

6. Does the study report appear to have been comprehensive in its outcome reporting?

***LOW risk of bias (i.e., answer: "Yes"):***
All the study's pre-specified (primary and secondary) outcomes outlined in the protocol, methods, abstract, and/or introduction that are of interest in the review have been reported in the pre-specified way.

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
Anyone of the following:

- Not all the study's pre-specified primary outcomes (as outlined in the protocol, methods, abstract, and/or introduction) have been reported; or
- One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified; or
- One or more reported primary outcomes were not pre-specified (unless a clear justification for their reporting is provided, such as an unexpected effect); or
- One or more outcomes of interest are reported incompletely

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that selective outcome reporting is not capable of introducing a risk of bias in the study.

7. Is the study free of financial conflict of interest in any of the exposures studied?

***LOW risk of bias (i.e., answer: "Yes"):***

The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples include the following:

- The funding source is limited to government, non-profit organizations, or academic grants funded by the government, foundations and/or non-profit organizations;
- Chemicals or other treatment used in the study were purchased from a supplier;
- Company affiliated staff are not mentioned in the acknowledgments section;
- Authors were not employees of a company with a financial interest in the outcome of the study;
- A company with a financial interest in the outcome of the study was not involved in the design, conduct, analysis, or reporting of the study and authors had complete access to the data;
- Study authors make a claim denying conflicts of interest;
- Study authors are unaffiliated with companies with a financial interest, and there is no reason to believe a conflict of interest exists;
- All study authors are affiliated with a government agency (are prohibited from involvement in projects for which there is a conflict of interest or an appearance of a conflict of interest).

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***
The study received support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples of support include:

- Research funds;
- Chemicals, equipment or testing provided at no cost;
- Writing services;
- Author/staff from the study was an employee or otherwise affiliated with a company that has a financial interest;
- Company limited author access to the data;
- The company was involved in the design, conduct, analysis, or reporting of the study;
- Study authors claim a conflict of interest

***NOT APPLICABLE (risk of bias domain is not applicable to study):***
There is evidence that conflicts of interest are not capable of introducing a risk of bias in the study.

8. Did the study appear to be free of other problems that could put it at risk of bias?

***LOW risk of bias (i.e., answer: "Yes"):***
The study appears to be free of other sources of bias.

***Uncertain:***
There is insufficient information to permit a judgment of 'LOW' or 'HIGH' risk of bias

***HIGH risk of bias (i.e., answer: "No"):***

There is at least one important risk of bias. For example, the study:

- Had a potential source of bias related to the specific study design used; or
- Stopped early due to some data-dependent process (including a formal-stopping rule); or
- The conduct of the study is affected by interim results (e.g. recruiting additional participants from a subgroup showing greater or lesser effect); or
- Has been claimed to have been fraudulent; or
- Had some other problem