UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

The Role of Chronic Exposure to Ambient Air Pollution in Dementia and Cognitive Aging

A dissertation submitted in partial satisfactory of the
requirements for the degree of Doctor of Philosophy

in

Public Health (Epidemiology)

by

Sindana Devayani Ilango

Committee in charge:

University of California San Diego

Professor Tarik Benmarhnia, Chair
Professor Harvey Checkoway
Professor Linda McEvoy

San Diego State University

Professor Linda Gallo
Professor Caroline Thompson

2020

The Dissertation of Sindana Devayani Ilango is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California San Diego

San Diego State University

2020

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to recognize the support of my advisor and dissertation committee chair, Tarik Benmarhnia. His mentorship, enthusiasm, and unwavering commitment contributed to an invaluable training experience. He has shaped the epidemiologist I am today.

I would like to thank Harvey Checkoway, Linda McEvoy, Linda Gallo, and Caroline Thompson, my dissertation committee members, for their insightful feedback on my research and guidance through this program.

I am grateful to my classmates for their meaningful friendship and camaraderie on this journey.

Finally, my family. Thank you for being my biggest champions. Where I am today has truly been a team effort.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Ilango, Sindana D; Mortamais, Marion; Gutierrez, Laura-Anne; Berr, Claudine; Benmarhnia, Tarik. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in the International Journal of Epidemiology 2020. Ilango Sindana D; Chen, Hong; Hystad, Perry; van Donkelaar, Aaron; Kwong, Jeffrey C; Tu, Karen; Martin, Randall V; Benmarhnia Tarik. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, has been submitted for publication of the material as it may appear in the Journal of Alzheimer's Disease. Ilango, Sindana D; Gonzalez, Kevin; Gallo, Linda; Allison, Matthew A; Cai, Jianwen; Isasi, Carmen R; Hosgood, Dean H; Vasquez, Priscilla M; Zeng, Donglin; Mortamais, Marion; Gonzalez, Hector M; Benmarhnia, Tarik. The dissertation author was the primary investigator and author of this paper.

Vita

Education

2020 Doctor of Philosophy, University of California San Diego & San Diego State University

2014 Master of Public Health, University of California Los Angeles

2011 Bachelor of Arts, University of California Berkeley

Publications

Ilango SD, Shaffer RM. Air Pollution, Cardiovascular Disease, and Dementia. JAMA Neurology [In Press]

Sun Y, **Ilango SD**, Schwarz L, Qang Q, Chen JC, Lawrence JM, Wu J, Benmarhnia T (2020) Examining the joint effects of heatwaves, air pollution, and green space on the risk of preterm birth in California. Environ. Res. Lett. [In Press]

Aguilera R, Hansen K, Gershunov A, **Ilango SD**, Sheridan P, Benmarhnia T (2020) Respiratory hospitalizations and wildfire smoke: a spatiotemporal analysis of an extreme firestorm in San Diego County, California. Environmental Epidemiology, 4(5), e114.

Ilango SD, McElroy S, Schwarz L (2020) Recommendations for epidemiologic studies of aging populations in a changing climate. Int J Public Health. doi:10.1007/s00038-020-01453-9

Chen H, Burnett R, Bai L, Kwong JC, Crouse D, Lavine E, Goldberg M, Copes R, Benmarhnia T, **Ilango SD**, Van Donkellar A, Hystad P (2020) Residential greenness and cardiovascular disease incidence, readmission, and mortality. Environmental Health Perspectives 128, 087005.

Aguilera R, Gershunov A, **Ilango SD**, Guzman-Morales J, Benmarhnia T (2020) Santa Ana winds of Southern California impact PM2. 5 with and without smoke from wildfires. GeoHealth 4, e2019GH000225.

Ilango SD, Chen H, Hystad P, van Donkelaar A, Kwong JC, Tu K, Martin R V, Benmarhnia T (2020) The role of cardiovascular disease in the relationship between air pollution and incident dementia: a population-based cohort study. Int. J. Epidemiol. 49, 36–44.

Leibel S, Nguyen M, Brick W, Parker J, **Ilango S**, Aguilera R, Gershunov A, Benmarhnia T (2020) Increase in Pediatric Respiratory Visits Associated with Santa Ana Wind–Driven Wildfire Smoke and PM2. 5 Levels in San Diego County. Ann. Am. Thorac. Soc. 17, 313–320.

Ilango SD, Weaver M, Sheridan P, Schwarz L, Clemesha RES, Bruckner T, Basu R, Gershunov A, Benmarhnia T (2020) Extreme heat episodes and risk of preterm birth in California, 2005–2013. Environ. Int. 137, 105541.

Schwarz L, Malig B, Guzman-Morales J, Guirguis K, **Ilango SD**, Sheridan P, Gershunov A, Basu R, Benmarhnia T (2020) The health burden fall, winter and spring extreme heat events in the in Southern California and contribution of Santa Ana Winds. Environ. Res. Lett. 15, 54017.

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

Sheridan P, **Ilango S**, Bruckner TA, Wang Q, Basu R, Benmarhnia T (2019) Ambient Fine Particulate Matter and Preterm Birth in California: Identification of Critical Exposure Windows. Am. J. Epidemiol. 188, 1608–1615.

Gallagher LG, **Ilango S**, Wundes A, Stobbe GA, Turk KW, Franklin GM, Linet MS, Freedman DM, Alexander BH, Checkoway H (2019) Lifetime exposure to ultraviolet radiation and the risk of multiple sclerosis in the US radiologic technologists cohort study. Mult. Scler. J. 25, 1162–1169.

Ilango SD, McEvoy LK, Laughlin GA, Bergstrom J, Barrett-Connor E, Kritz-Silverstein D (2019) Pregnancy history and cognitive aging among older women: the Rancho Bernardo Study. Menopause 26, 750–757.

Park HK, **Ilango S**, Charriez CM, Checkoway H, Riley D, Standaert DG, Bordelon Y, Shprecher DR, Reich SG, Hall D (2018) Lifetime exposure to estrogen and progressive supranuclear palsy: Environmental and Genetic PSP study. Mov. Disord. 33, 468–472.

Checkoway H, **Ilango S**, Li W, Ray RM, Tanner CM, Hu S, Wang X, Nielsen S, Gao DL, Thomas DB (2018) Occupational exposures and parkinsonism among Shanghai women textile workers. Am. J. Ind. Med. 61, 886–892.

Abstract of the Dissertation

The Role of Chronic Exposure to Ambient Air Pollution in Dementia and Cognitive Aging

by

Sindana Devayani Ilango

Doctor of Philosophy in Public Health (Epidemiology)

University of California San Diego, 2020
San Diego State University, 2020

Professor Tarik Benmarhnia, Chair

As average life-expectancy increases worldwide, there is a heightened public health concern about impaired cognitive function and dementia with advancing age. There are an estimated 10 million new cases of dementia each year and the worldwide prevalence is expected to triple by 2050. Without current effective treatment, research has expanded to identify modifiable risk factors for dementia to promote healthy cognitive aging.

Air pollution is a unique modifiable risk factor as levels can be shaped by individual behaviors and population-level environmental policies and regulations. Ambient air pollution is a mixture of particulate matter, gases, and other organic and metallic components. Emerging evidence suggests chronic exposure may affect diseases of the central nervous system, including dementia and cognitive impairment. It is biologically plausible that increased exposure to air pollution can produce a neuroinflammatory response, which in turn can result in structural changes in the brain. Although epidemiologic studies have observed a link among some populations, the causal pathway between air pollution and dementia remains unclear and these studies are subject to specific methodological challenges.

The first chapter of this dissertation reviews the epidemiological evidence regarding the relationship between dementia, cognitive impairment, and air pollution and outlines some methodological challenges my research addresses. The second chapter demonstrates different methods to account for competing events in a cohort of older adults in France and provides recommendations for future research in studies of air pollution and dementia. The third chapter examines the causal pathway between air pollution and dementia, evaluating cardiovascular disease as a potential intermediate in a population-based cohort in Ontario, Canada. The fourth chapter expands the generalizability of the current research by examining the relationship between air pollution and cognitive impairment in a US cohort of Hispanic/Latino adults, an understudied ethnic group with well-documented disparities in both air pollution exposure and chronic health outcomes. The final chapter of this dissertation summarizes key findings and highlights future directions to advance epidemiologic research of air pollution and cognitive aging.

1. Introduction

1.1. Overview of Dementia

Dementia is a chronic and progressive condition where higher cortical functions, such as memory, orientation, comprehension, and reasoning are impaired and affect daily living.[1] The progression varies by individual, but symptoms are often characterized by three stages: 1) early stage, where the affected individual becomes more forgetful, loses track of time and location, and has difficulty making decisions and carrying out complex tasks; 2) middle stage, where there is increasing difficulty with communication, decision-making, and observed changes in behavior; and 3) late stage with near total loss of independence and physical symptoms such as loss of bladder control and incontinence.[1] The effect of dementia on mortality is complex, as the interplay between co-occurring conditions typically account for death. However, a secular trend of a greater proportion of death certificates reporting dementia as cause of death has been observed in the United States in recent years.[2]

There are many forms of dementia, each with a unique pathology. The most common forms are Alzheimer's disease dementia (AD), accounting for 60-70% of cases[1], and vascular dementia (VaD). AD is diagnosed when a gradual decline in cognitive function affects daily functioning and the impairment is independent of other underlying health conditions. It is progressive in nature and primarily affects older adults. VaD can present like AD or have an abrupt onset but is the result of injuries to vessels that supply blood to the brain, depriving brain cells of oxygen. Currently, there is no diagnostic test to distinguish VaD from other dementias; clinicians use medical history (e.g., history of a previous stroke) to inform clinical diagnosis. These two forms of dementia comprise most of the dementia cases worldwide and often co-exist.

As the proportion of older adults in the world grows, there is increasing concern about dementia and dementia-related outcomes. Worldwide, approximately 50 million individuals are affected with dementia and there are an estimated 10 million new cases each year.[1] In 2018, an estimated 5.5 million Americans, about one in 10 people, aged 65 and older are living with Alzheimer's dementia in the United States (US).[3] Similarly, approximately 9% of adults 65 years and older in Canada, and adults 70 years and older in France are living with dementia.[4,5]

Race/ethnic differences in dementia prevalence exist may be due to disparities that affect dementia diagnosis. Dementia is largely underdiagnosed, especially among minority populations with barriers to accessible health care, reluctance toward screening tests, and perceptions surrounding impaired cognitive function.[6–8] Studies have suggested race/ethnic disparities in dementia incidence and diagnosis are largely driven by social and behavioral factors such as socioeconomic status, psychosocial pathways, behavioral norms, and vascular health.[9,10] The few studies of dementia disparities among Hispanics/Latinos suggest that Hispanics/Latinos have a slightly increased risk of Alzheimer's and related dementias than similarly aged older whites.[9,10] Studying early signs of dementia, such as cognitive impairment among race/ethnic minorities, such as Hispanics/Latinos, may improve understanding of dementia among minority populations.

1.2. Overview of Cognitive Function

Dementia is diagnosed when deficits in two or more cognitive domains affect daily living. The domains used for diagnosis include memory, executive function, visuospatial ability, and language.[11] A prodromal stage of noticeable decline in cognitive function frequently precedes dementia, although not all decline results in dementia. Thus, a broader understanding of cognitive impairment can help identify risk factors of early stage dementia with aims of disease prevention and healthier aging. As with dementia, there are no objective diagnostic criteria to identify

individuals with cognitive impairment. Rather, hundreds of standardized instruments have been developed to assess different domains of cognitive function and determine cognitive level. For example, episodic memory is commonly tested by asking patients to recall word lists. Visuospatial ability can be tested by requesting an individual to draw a clock face depicting a specific time. Many of these instruments were developed in predominantly non-Hispanic White populations and may not fully account for sociocultural differences that may affect test performance in diverse populations.[12] Thus, studying cognitive function with multiple measures validated in specific race/ethnic minority groups can improve our understanding of different aspects of cognitive level among vulnerable populations.

1.3. Modifiable Risk Factors of Dementia and Cognitive Impairment

Understanding modifiable risk factors of dementia and dementia-related outcomes is a public health priority, especially as there is no current treatment available to reverse the course of these conditions. Identifying highly prevalent modifiable exposures associated with dementia and cognitive impairment can maximize health benefits at a population level. Substantial epidemiological literature suggests the following factors affect risk of dementia: low education, midlife hearing loss, obesity, hypertension, late-life depression, smoking, physical inactivity, diabetes, and social isolation.[13] An estimated 35% of dementia cases could be prevented by modifying these factors.[13] Although many of these factors co-occur[14], this estimated population attributable fraction highlights the importance of understanding modifiable risk factors affected by health and behaviors across the life course. A marked reduction in dementia can be observed with behavioral and lifestyle interventions.

Recently, another highly prevalent, modifiable risk factor that may be related to dementia and cognitive level has been suggested. This risk factor is ambient air pollution and it has been

found to be associated with dementia and dementia-related outcomes in recent studies.[15,16] Ambient air pollution is ubiquitous; everyone is exposed, and some populations are exposed at high levels. The negative health impacts of exposure to air pollution are well-documented and are particularly pronounced among older adults.[17,18] These adverse health effects occur at exposure levels even below the national standards, suggesting further reduction of air pollutants may improve population health.[19] Thus, further exploration into the relationship between ambient air pollution and dementia and dementia-related outcomes is of heightened importance.

1.4. Air Pollution and the Brain: Overview and Biologic Plausibility

Exposure to ambient air pollution has a range of health implications. In 2015, an estimated 4 million premature deaths, globally, were attributable to air pollution exposure.[20] Ambient air pollution is a complex mixture of particulate matter (PM), gaseous pollutants (e.g., nitrogen dioxide (NO_2), ozone (O_3)), persistent organic metals, and heavy metals (Table 1.1).

Table 1.1: Sources of ambient air pollutants

Pollutant	Common Sources
Fine particulate matter ($PM_{2.5}$)	- Traffic - Wildfire smoke - Industrial emissions - Dust
Nitrogen Dioxide (NO_2)	- Burning of fossil fuels (e.g., traffic)
Ozone (O_3)	- Secondary pollutant; product of chemical reaction of nitrogen oxides (e.g., NO_2), volatile organic compounds, and ultraviolet radiation

Air pollution exposure can be classified as acute or chronic and each exposure period affects the human body through different mechanisms. Acute refers to exposure during a short time frame (e.g., hourly, daily or weekly) and is commonly due to a specific extreme event (e.g.,

wildfire) or meteorological conditions (e.g. atmospheric inversions). Alternatively, chronic refers to long-term exposure over months or several years. Chronic exposure to ambient air pollution will be the focus of this dissertation research. Chronic exposure to air pollution can result in oxidative stress to the human body and increased inflammatory responses which can subsequently impact the pathophysiological processes of major chronic diseases and exacerbate existing health conditions.[21]

Recently, there has been emerging evidence about exposure to air pollution affecting diseases of the central nervous system, including dementia.[15,22] Neuroimaging and biologic studies provide supporting evidence that air pollution may affect the brain.[22–29] Animal studies have also shown that fine particulate matter ($PM_{2.5}$) can move from the nose via the olfactory nerve and into the brain and nitrogen dioxide (NO_2) can impair synapses and induce neuronal damage.[23,24] A study comparing human post-mortem brain tissues in Manchester, UK, and Mexico found an abundance of nanoparticles from an external source (i.e., air pollution) in brain tissues from those who lived in more polluted areas.[25] Neuroimaging studies have similarly observed an adverse effect of $PM_{2.5}$ and traffic-related pollutants on brain volume[26,27], suggesting that air pollution may influence structural brain changes resulting in worse health outcomes. These findings can be explained by a "neuroinflammation hypothesis", where innate immune cells and microglia are affected by air pollution-induced central nervous system disruptions, directly and indirectly impacting risk of neurodegenerative diseases in later life.[24] Pollutants can directly reach the brain through the nasal pathway or through systemic circulation by crossing the blood brain barrier and trigger neuroinflammation.[22,24,28] Alternatively, pollutants can produce inflammation in other organs and tissues (e.g., cardiovascular systems) that can indirectly affect the central nervous

system and result in microvascular brain damage, impairments in cognitive function, and dementia.[24,29]

Air pollution and climate change are closely interlinked. The increase in global temperatures and increased frequency and severity of extreme weather events will greatly impact population exposure to air pollution.[30] For example, more frequent and persistent wildfires will result in increased wildfire smoke, contributing to increased concentrations of $PM_{2.5}$.[31] Adverse health impacts attributable to climate change are also expected to increase.[32] Thus, identifying and understanding adverse health impacts of increased exposure to air pollution is a public health priority. Furthermore, reducing air pollution emissions with climate change mitigation policies will have multiple co-benefits.[33]

The following sections review recent population studies examining the relationships between air pollution, dementia, and cognitive impairment:

1.5. Review of Epidemiologic Studies on Air Pollution and Dementia

Fourteen epidemiologic studies have published research on air pollution and incident dementia.[34–47] These studies were conducted in Canada[38,42,43], US[34,46,47], Taiwan[36,39,41,45], and Sweden[37,44]. With the exception of one case-control study[39], alll studies examined this relationship in cohort studies using data from administrative databases[35,36,38,43,47–49] or from ongoing prospective cohort studies[34,37,40,44]. These cohort studies used time to dementia diagnosis as the primary outcome of interest to estimate hazard ratios. Dementia diagnoses were either obtained from a combination of administrative databases[35,36,38,41–43,45,47] or from clinical examinations[34,37,40,44]. The majority of these studies focused on PM[34–36,38–40,42,44–47]; other air pollutants studied included nitrogen oxides[35–38,40–42,44], ozone[35,36,39,40,42,45], carbon monoxide[36,41], and indicators of traffic-related air pollution[43]. Air pollution exposure was assigned to individuals

using measurements from fixed site monitoring stations[36,41,45,46] or modeled estimates[34,35,38,40,42,44,47]. Exposure time windows studied ranged from annual to 15-year exposure windows. All studies found positive associations between increased exposure to air pollution and incident dementia.

Some studies acknowledged a potential bias resulting in an underestimation of the true effect, due to selective attrition out of the study.[37,39] This selection bias can arise because of competing events, related to increased air pollution exposure, which preclude an individual from experiencing the outcome of interest. By selecting individuals who have been diagnosed with dementia, the analysis is inherently restricted to individuals who have survived through competing events, such as death. In the presence of unmeasured confounding between the competing outcomes (likely) and a link between exposure and competing event (also likely), a selection bias is induced.[50,51] Several approaches have been proposed to account for selection bias and competing events.[51–53] Despite this, the reviewed literature either failed to account for death in their analyses or censored deaths under the strong assumption that censoring occurred independently of exposure. Although all studies discussed the adverse health effects of air pollution, only two studies explicitly accounted for death as a competing event as a sensitivity analysis.[40,44]

The study of air pollution and dementia is a relatively new area of research. Limited understanding of the causal mechanisms exists. There are suggestions that air pollution may affect health outcomes, such as cardiovascular disease, which may in turn affect risk of dementia. While several studies acknowledged this, potential mechanisms have not been formally evaluated. Studies that examined the potential mediating role of comorbidities did so by comparing estimates from regression models with and without cardiovascular disease and other comorbidities and found minimal differences.[35,37,41,42,46] Two studies examined effect measure modification determined by

the interaction between air pollution and pre-existing comorbidities (e.g., stroke, heart failure, diabetes, and hypertension) and found suggestions of modification of some comorbidities on the multiplicative scale.[42,44] Study participants with heart disease had a stronger relationship between $PM_{2.5}$ and dementia in the Swedish cohort but not in the population cohort in Ontario, Canada.[42,44] These two studies reported conflicting evidence of the potentially modifying role of stroke.[42,44] Recent methodological developments in causal inference that decompose the mechanism into direct and indirect effects can be applied to investigate research questions on causal mediation.[54]

1.6. Review of Epidemiologic Studies on Air Pollution and Cognitive Function

Ten cross-sectional studies of air pollution and cognitive function among older adults have been conducted within cohorts in U.S.[55–60], China[61], and Germany [62–64]. There are several domains that capture cognitive function, each can be assessed by one of several validated instruments. In the reviewed literature, the outcome of interest was cognitive function measured at a single time point. Indicators of global cognitive function were evaluated from specific assessments of global cognitive function (e.g., Mini-Mental State Exam or Six-Item Screener) or created from combining domain-specific scores. The primary cognitive domains evaluated were memory, executive function, orientation, abstraction, and global cognitive function. These assessments were administered in-person at clinical evaluations or over the telephone. Scores were evaluated as raw scores, dichotomized into cognitively impaired or not, or standardized. The air pollutants studied were PM[55–58,63,64], ozone [57,58], nitrogen oxides[58,62,64], and combined measurements of air pollution[61] or traffic exposure[62–64]. Chronic exposure was measured over one to 15-year exposure windows using data from monitoring stations[55–61,63], satellites[56], or modeled estimates[62,64].

With the exception of one study[56], all found at least one notable adverse association between increased exposure to air pollution and cognitive function. However, it is important to

note that many associations were tested when examining batteries of cognitive assessments with multiple measurements of air pollution; no consistent pattern with specific associations was observed. For example, increased exposure to particulate matter was found to be associated with worse verbal learning[58], abstraction[62], working memory[55], and orientation[55], and mild cognitive impairment. Indicators of air pollution, such as gross air quality indices and traffic exposure were also found to be associated with worse cognitive outcome. [61–64]

There is limited generalizability with respect to race/ethnicity in these cross-sectional studies of air pollution and cognitive function. Most studies were conducted in predominantly white populations in the U.S. or in Germany. Only two studies included individuals who identified as Hispanic ethnicity, although this ethnic group made up less than 15% of the study population.[58,59]

1.7. Methodological Challenges and Proposed Solutions in Studying Air Pollution and Neurocognitive Outcomes

There are several unique methodological challenges in studying neurocognitive outcomes among older adults. Over the past two decades, developments in causal inference methods have been introduced to the field of epidemiology to answer causal questions from observational studies by employing a mathematical framework with assumptions required to identify a causal effect. Many of these methods have been adapted from social sciences, economics, mathematics, and computer science. The following methodological issues are central to the motivation for this dissertation work: 1) informed censoring due to death, 2) limited understanding of the causal pathway, and 3) limited generalizability to Hispanic populations in the current evidence base.

1.7.1. Informed Censoring due to Death

Since dementia is typically diagnosed at older ages, events more common in later stages of life may prevent a dementia diagnosis. As previously introduced, these events are known as competing events as they "compete" with the outcome of interest. An example of a competing event in longitudinal studies of older adults is death. In most of the reviewed literature, death was considered a censored event in analyses. This analytical decision implicitly assumes death is independent of the exposure. However, this may be a poor assumption as seen by substantial evidence linking air pollution to adverse health outcomes and premature death.[65] A large proportion of the study population could have died from air-pollution related causes (e.g., cardiovascular disease, respiratory conditions) before living long enough to be diagnosed with dementia. This restriction could result in a study of healthier individuals who are resilient to the adverse effects of air pollution and threaten internal validity. In other words, this selection of participants could result in a false comparison between the lower exposed groups and a healthier, more resilient, group of highly exposed individuals. This is ultimately an example of selection bias. Failure to account for informed censoring due to the competing event, death, may underestimate dementia incidence and bias estimated effects of exposures toward the null, as the two comparison groups (low exposure versus highly exposed) are more similar in the study setting than in the target population (Figure 1.1).

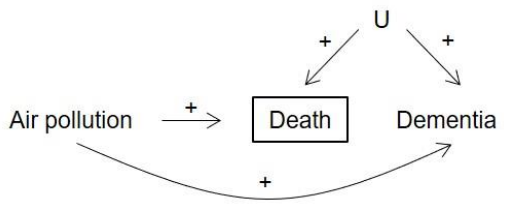

Figure 1.1: Selection bias due to informed censoring due to death

Several approaches to account for competing events exist. However, these approaches are underutilized perhaps due to the lack of clarity in target inference and estimand.[53,66] None of the previous studies on air pollution and dementia have considered censoring in a way that accounts the dependence between air pollution and censored event. Therefore, this dissertation will demonstrate and compare different approaches to account for competing events using a case study of air pollution and dementia, with death as a competing event.

1.7.2. Causal Pathways

Despite growing evidence suggesting a link between increased exposure to air pollution and dementia incidence, the exact mechanism is not well understood, as even our understanding of the progression of dementia is limited. One way to investigate causal pathways is to decompose the relationship between air pollution and dementia into its direct and indirect effects through an intermediate of interest. Cardiovascular disease and events (CVD) is an intermediate of interest (Figure 1.2) as air pollution has been consistently related to the increased risk of incidence, complications, and mortality from CVD.[67]

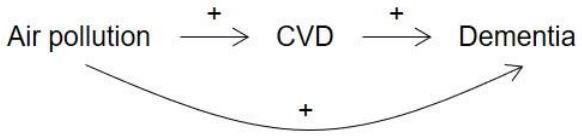

Figure 1.2: Hypothesized direct and indirect effect of air pollution on dementia

Furthermore, growing evidence has linked CVD to impaired cognitive function and it often co-occurs with dementia.[7,68] Previous studies have acknowledged that indirect pathways through CVD and other comorbidities that may explain the observed relationship, however, this hypothesis was examined by taking the difference between effect sizes with and without adjustment for

CVD.[35,37,41,42,46] This traditional difference approach is unable to quantify a causal estimand of the indirect effect or provide a causal interpretation without identification assumptions or additional assumptions about unmeasured confounding between the 1) exposure and outcome, 2) exposure and mediator, 3) and mediator and outcome. Additionally, 4) there should be no mediator-outcome confounder affected by the exposure.[54] Recent developments in formal methods to decompose a total effect into its direct and indirect effect have been developed which can provide a causal interpretation.[54] Applying these methods to disentangle the pathway from air pollution and dementia can offer insights into whether there is an indirect path through cardiovascular events which can improve understanding of the biological mechanism and prioritize intervention efforts. Examining the mediating role of CVD can offer insights into the etiology and biological mechanisms between air pollution and neurocognitive outcomes and is examined in this dissertation.

1.7.3. Generalizability

While there is a growing body of literature examining the relationship between air pollution and dementia and dementia-related outcomes, the majority of studies are typically conducted in areas of predominantly white populations (e.g., Canada, US, United Kingdom, Sweden, Germany). Representativeness is not a requirement of epidemiologic studies.[69] However, there are many instances where studying subpopulations is important for identifying particularly vulnerable groups. Although air pollution is ubiquitous, it is important to understand the relationship in minority race/ethnic groups, as some groups are more likely to live in neighborhoods with higher levels of air pollution from closer proximity to highly trafficked roads and industrial operations.

The relationship between air pollution and neurocognitive outcomes among Hispanics/Latinos is understudied. Disparities in both air pollution exposure and cognitive

functioning between older Hispanics and non-Hispanic Whites are well documented.[70] Only two previous studies of the reviewed literature included Hispanics/Latinos, albeit a small (<15%) proportion of the full study population, in analyses.[57,59] Thus, in this dissertation the association of air pollution with cognitive function in Hispanics/Latinos is examined.

1.7.4. Introduction to Causal Inference and Assumptions

Traditional statistical regression models and hypothesis testing aim at examining associations between variables. However, many researchers are interested in inferring causal associations, for example, rather than examining the association between smoking and lung cancer, an epidemiologist might wonder does smoking *cause* lung cancer? Developments in causal inference methods have been borne out of fields of computer science, economics, and social sciences and expand upon traditional statistical techniques by introducing a specific mathematical notation and making explicit assumptions required to make causal inferences.

The counterfactual framework will be introduced to define the causal effect. An individual causal effect is observed when treatment A has a causal effect on an individual's outcome Y, specifically, when exposure a is set to 1, $Y^{a=1}$ does not equal the potential outcome if a is set to 0, $Y^{a=0}$. $Y^{a=1}$ and $Y^{a=0}$ are counterfactual outcomes – a theoretical concept that describes the potential outcome of the individual had the same individual received a different treatment. This is a theoretical concept because it is impossible to directly observe a contrast between an individual's observed, or factual, outcome with his/her unobserved, or counterfactual, outcome[71]; an individual can only experience one version of a treatment and subsequent outcome. To overcome this, we can turn to examining average causal effects in a population to assess causal associations. Using the same notation, an average causal effect of treatment A on outcome Y is observed if $\Pr[Y^{a=1} = 1] \neq \Pr[Y^{a=0} = 1]$ in the population of interest. Under the counterfactual framework for causal

inference, there are three assumptions required to approximate a counterfactual comparison group to identify causal associations: 1) Exchangeability, 2) Positivity, and 3) Consistency. Each of these are described below:

Exchangeability requires that the groups being compared, for example, the exposed and unexposed groups, are identical with respect to all confounders such that the same outcome will be observed had the exposed group been unexposed. Formally, $\Pr[Y^a = 1|A = 1] = \Pr[Y^a = 1|A = 0]$ for both $a = 0$ and $a = 1$.This is equivalent to independence between the counterfactual outcome and the observed treatment ($Y^a \coprod A$ for all a). Covariate balance is required to achieve exchangeability between comparison groups.

Positivity ensures that individuals in every stratum of covariates have a non-zero probability of being in both the exposed and unexposed groups, or $\Pr[A = a|L = l] > 0$ for all values l with $\Pr[L = l] \neq 0$ in the population of interest. The covariates L are the covariates required to achieve exchangeability. Without this condition, it is impossible to assess the conditional effect of A on outcome Y. For example, if an investigator is interested in the conditional effect of treatment A on mortality adjusting for sex, there must be individuals in the dataset with all combinations of stratum: treatment A, without treatment A, death, survival, male, and female. If there were no males in the study, the investigator would be unable to make an inference about the causal effect in a population with both males and females.[72]

Consistency is the condition that requires the same outcome $Y^a = Y$ for every $A = a$. This can be achieved with specific, well-defined treatments or exposures of interest, that can be related to a specific intervention. For example, diet or exercise is a more consistent and

defined treatment than obesity.[73] Furthermore, an individual's potential outcome is not affected by other individual's exposure to treatment (Stable Unit Treatment Value Assumption).

Under the counterfactual framework and the assumptions of exchangeability, positivity, and consistency, methods have been developed for observational studies to identify causal associations.

1.7.5. Proposed Solutions to Methodological Challenges

In this dissertation, we expand on the current literature of air pollution and neurocognitive outcomes by addressing some of the unique methodological challenges common in studies of air pollution and aging related health outcomes. First, multiple approaches to account for competing events will be demonstrated to highlight the differences in target inference and estimand in approaches common in survival analyses. Next, the causal pathway between air pollution and incident dementia will be decomposed into its direct and indirect effect through cardiovascular disease by employing a formal causal mediation analysis. Finally, the relationship between air pollution and cognitive impairment will be examined in a cohort of Hispanic/Latino adults to identify if the relationship exists in this race/ethnic group. Assumptions required to identify a causal effect will be made throughout this dissertation.

1.8. Specific Aims

In this dissertation, I examined the relationship between chronic exposure to ambient air pollution and incident dementia and cognitive function. The following aims are addressed:

Aim 1: To demonstrate three common approaches to account for competing events in a study of air pollution and dementia, considering death as a competing event

Aim 2: To perform a causal mediation analysis to decompose the total effect of chronic exposure to air pollution on incident dementia into its direct and indirect effects through incident cardiovascular disease

Aim 3: To examine the effect of chronic exposure to air pollution on cognitive function in a well-established cohort of Hispanics/Latinos in the US, the Hispanic Community Health Study/Study of Latinos

This research studied populations of middle to older-aged adults in France, Canada, and in the United States. The case study used for Aim 1 draws from an ongoing prospective cohort study of older adults France (Three-City Cohort). A retrospective cohort of older adults living in Ontario, Canada was created from linkage of several existing health, environmental, and census data sources and was used for Aim 2. Finally, a prospective cohort study of Hispanics/Latinos living in San Diego, California, was used for Aim 3.

Findings from these aims provide valuable insight into the relationship between air pollution and neurocognitive outcomes by addressing some of the current challenges in dementia research by asking inferential questions and applying causal inference methods.

2. Competing events in studies of air pollution and dementia: A comparison of multiple approaches to account for informed censoring due to death

2.1. Abstract

One challenge in studying the effect of chronic exposure to air pollution on aging-related outcomes is the presence of informed censoring due to competing events such as death. In studies of older adults, premature death may preclude or "compete with" an outcome of interest. Different approaches to account for informed censoring have been described in epidemiologic literature but are underutilized in air pollution studies and the target estimand is typically poorly defined. The first step to answer a causal question is to define a target estimand. Then the investigator can design and implement an identification strategy to generate an estimator from a mathematical function (e.g., regression model) that takes data as input and produces an estimate (e.g., regression coefficient) to approximate the target causal estimand. In this study, we demonstrated three approaches to account for competing risks, each with unique assumptions and targeted estimands. First, a controlled direct effect of air pollution on dementia, not mediated by the competing event, was estimated with a cause-specific hazard ratio. Next, a Fine and Gray approach was applied to estimate the subdistribution hazard ratio. Finally, a weighting scheme was applied with inverse probability weights to correct for time-varying informed censoring and estimate a weighted cause-specific hazard ratio. Each approach is defined under the counterfactual framework and demonstrated in a case study evaluating the effect of air pollution on incident dementia in a cohort in France, considering death as a competing event. We provide recommendations for future investigators considering competing events in survival analyses of air pollution and dementia.

2.2. Introduction

In longitudinal studies of older adults, competing events are a common phenomenon where an event may preclude an individual from experiencing the outcome of interest. For example, consider a study of stroke among older adults. Some study participants will die over the study period without developing stroke and consequently drop out of the study at time of death. Death from cancer in a stroke-free individual would be considered a competing event, since investigators do not know if that individual would have experienced a stroke had they completed the study. Here, we can think of death from cancer as an event that "competes" with stroke.

The presence of competing events can result in an inadvertent selection that threatens the internal validity of a study depending notably on the structural relationship between the exposure, outcome, and competing event (Figure 2.1). If an adverse exposure and competing event are dependent (i.e. the exposure increases the probability of competing event), individuals who remain at risk of experiencing the outcome of interest will be systematically different than the population of diagnosed cases in the target population we are truly interested in. For example, a simulation study demonstrated that the observed protective relationship between smoking and malignant melanoma may be due to competing events related to smoking (e.g., death due to lung cancer, heart disease, or COPD).[50]

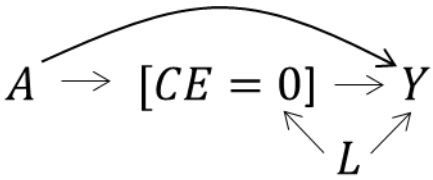

Figure 2.1: Collider stratification bias due to conditioning on the competing event where A=exposure, CE=competing event, Y=outcome, L=vector of unmeasured covariates that are common causes of CE and Y

Despite the pervasiveness of competing events in studies of older adults, there is limited discussion of accounting for them. Several approaches have been proposed to account for competing events and have been described by a number of reviews.[52,66,74–77] However, there is slow uptake, possibly because of unclear target estimands from each approach. Young et al. recently argued that a unifying framework for estimands common in survival analyses may help clarify analysis decisions.[53] Briefly, the estimation of the *cause-specific hazard* estimates the risk under the elimination of competing events. This is equivalent to the controlled direct effect of the exposure on the outcome, not mediated through the competing event. It is typically estimated by censoring individuals when they experience the competing event. The *subdistribution hazard* coincides with the cumulative incidence function and estimates the risk in the presence of competing events; competing events are not eliminated in this setting. Rather, the *subdistribution hazard* is equivalent to the total effect of exposure on outcome by considering individuals with the competing event as at-risk for the outcome for the remainder of the study period. Instead of including the competing event as part of the censoring definition as seen in the estimation of the cause-specific hazard, individuals with the competing event are considered "cured" from the primary outcome. As an extension to the cause-specific hazard, weighting methods can be applied in conjunction with hazard functions to further account for specific dependencies that are otherwise not considered. In this tutorial, traditional approaches estimating the *cause-specific hazard* and *subdistribition hazard* will be compared with a weighted cause-specific hazard ratio, an approach recently formalized by Young et al. that uses inverse probability weights and applies the potential outcomes framework.[53] Each approach accounts for competing events, however there are differences in the target estimand and inference that are often overlooked.

The focus of this paper is to compare these three approaches to account for competing events in survival analyses. As a case study, we will use the relationship between air pollution and dementia as an example. In recent years, there has been growing evidence suggesting a link between increased exposure to ambient air pollution and increased risk of dementia.[34–37,39–43,45,46,49,78] Competing events are common in longitudinal studies of older adults, where study participants are vulnerable to several comorbidities that could result in failure to complete the study. Increased exposure to air pollution exacerbates many of these comorbidities (e.g., respiratory disease, cardiovascular disease, and lung cancer) and shorten life expectancy.[65] Air pollution alone contributed to 8.7% of global mortality, or 4.9 million deaths, in 2017.[65] Older adults are particularly vulnerable to climate stressors, including poor air quality.[17] The adverse effects of air pollution disproportionately affect adults at the same ages at which they may be enrolled in dementia studies (e.g., ≥ 65 years). Thus, considering competing events should be carefully considered in studies of air pollution and dementia.

Despite this, only two studies in the reviewed literature on air pollution and dementia discuss and account for competing events.[40,44] Both considered mortality as a competing event in sensitivity analyses and found minimal differences after accounting for competing events. The first accounted for competing events by estimating the subdistribution hazard ratio[40] and the second did not specify their method[44]. Some studies censored study participants at the time-of-death in analyses,[35,37,42,43,45,46] an analytical decision that implicitly accounts for competing events by estimating a cause-specific hazard ratio. However, this was not specified. Furthermore, these studies applied Cox proportional hazard models to generate estimates, which assumes independence between exposure and censored event.[79] This is likely violated as indicated by substantial evidence linking air pollution to adverse health outcomes and premature death.[65]

Failure to account for competing events in settings where a dependency between exposure and competing event exists may attenuate or even produce a spurious protective result.[50] In dementia research, targeting the true causal effect can effectively inform policies and interventions to ensure healthy aging.

Despite the literature documenting the importance and interpretation of estimands that account for competing events in survival analyses, in practice there is underutilization and often misinterpretation of results.[77,79] Previous reviews and tutorials on methods to account for competing events have provided both conceptual[52,66,75–77] and technical[53,79] details about the estimation of cause-specific and subdistribution hazards. However, there is limited literature on the conceptual application of these methods and weighting procedures to account for competing events. We believe a non-technical discussion describing each approach and outlining specific analytic recommendations can improve understanding and promote the proper application of methods.

In this tutorial, we demonstrate three approaches to account for competing events. Traditional approaches estimating *cause-specific* and *subdistribution hazards* will be compared with estimation of a *weighted cause-specific hazard*. Each approach will be defined under the counterfactual framework proposed by Young et al.[53] and demonstrated in a case study evaluating the effect of air pollution on incident dementia, considering death as a competing event in a cohort of older adults in France. Finally, we will compare approaches and provide recommendations for future investigators considering competing events in survival analyses of air pollution and dementia.

2.3. Methods

Key definitions

We begin this tutorial by defining key terminology used in survival analyses. Statistical analyses where information on the timing of events (e.g., date of dementia diagnosis) is available to the investigator is referred to as analyses in failure-time settings, time-to-event analyses, or survival analyses. Survival analyses are advantageous in estimating associations in the presence of time-varying confounding and covariates. The timescale, or time axis, refers to the unit of time (e.g., calendar years, age, study duration) used to quantify how long study participants are followed. The risk set defines the individuals considered at risk for the event of interest. Several events can occur to a study participant. There is the event of interest, censored events, and competing events. The event of interest is the outcome under study. Censored events occur when the investigator no longer considers an individual at risk for the event of interest and can be due to loss-to-follow up or study completion. Competing events are events that may preclude an individual from experiencing the outcome of interest and can be considered as censored events in specific analyses. A hazard function is a function that describes the instantaneous rate of an event in a population and is commonly modeled with Cox proportional hazards models to estimate a hazard ratio. The hazard ratio represents the relative change in hazard function associated with a 1-unit increase in a continuous covariate. Young et al., suggests that a unifying, formal framework for estimands in survival analyses could clarify analysis decisions and interpretations. Thus, in this tutorial, we apply their translations of statistical estimands to potential outcome notation and include directed acyclic graphs (DAGs).[53]

Naïve approaches where competing events are not considered

In studies where competing events are not considered, individuals are followed until they experience the outcome of interest or are censored and no longer contribute to risk sets. If

competing events are not considered, it is implicitly assumed they do not exist. Individuals who truly experience a competing event (this fact is unrealized to the investigator) remain in the risk set until their next study visit, at which point they are censored. In naïve analyses using Cox proportional hazard models, we assume proportional hazards and independence between the censored event and exposure. The estimand of interest is the marginal hazard ratio, where all competing events have been eliminated. Under the counterfactual framework, the contrast between marginal hazards with a binary exposure can be defined with the notation:

$$\Pr\left[Y_{t+1}^{a=1,\,\overline{ce}=\overline{0}} = 1 \mid Y_t^{a=1,\overline{ce}=\overline{0}} = 0\right] \text{ vs. } \Pr\left[Y_{t+1}^{a=0,\,\overline{ce}=\overline{0}} = 1 \mid Y_t^{a=0,\overline{ce}=\overline{0}} = 0\right]$$

Where Y=outcome, a=exposure, ce=competing event, and t=time. The overline above the variables represents all subsequent events. For example, $\overline{ce} = \overline{0}$ means all competing events and subsequent competing events are set to 0 by a hypothetical intervention that eliminates competing events.

Cause-specific hazard ratio: the direct effect of exposure on outcome, not mediated by competing event

In contrast to naïve approaches that assume competing events are eliminated, the cause-specific hazard function represents an instantaneous rate of a specific event among individuals who are free of any event (primary outcome and competing event) at time t. Individuals are censored at the time of the competing event and subsequently removed from all risk sets. This essentially conditions on the competing event by restricting the study sample to individuals who are free of competing events.

For example, suppose an individual enters the study at age 65 and dies, dementia-free, at age 70. In a survival analysis considering age as the time scale, this individual would be considered

at-risk for dementia from ages 65 to 70 and then censored. After 70, this individual no longer contributes to risk sets for older dementia cases.

A Cox proportional hazard model can be used to estimate a cause-specific hazard ratio that represents the relative change in the cause-specific hazard associated with a 1-unit change in a continuous covariate. When competing events do not exist, this hazard ratio coincides with the incidence. However, in the presence of competing events, this equivalence does not hold. Rather, the cause-specific hazard ratio can be interpreted as a *rate* of occurrence of the event of interest, among study participants who are event-free at that time. Additionally, the Cox proportional hazard model assumes independence between exposure and censored events, which may be violated if the exposure increases the probability of the competing event. In this scenario, alternative methods should be explored.

In estimating the cause-specific hazard ratio, the target estimand is the contrast between cause-specific hazards and is equivalent to the controlled direct effect of exposure on the outcome, not mediated through the competing event (Figure 2.2). Under the counterfactual framework, it is assumed that competing events are conditioned on by restriction to the population without competing events. The contrast between cause-specific hazards with a binary exposure A can be defined with the notation:

$$\Pr[Y_{t+1}^{a=1} = 1 | \, CE_{t+1}^{a=1} = Y_t^{a=1} = 0] \text{ vs. } \Pr[Y_{t+1}^{a=0} = 1 | \, CE_{t+1}^{a=0} = Y_t^{a=0} = 0]$$

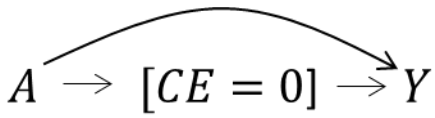

Figure 2.2: Effect of A on Y conditioned (by restriction) on competing events, where A=exposure, CE=competing event, and Y=outcome

Subdistribution hazard ratio: the total effect of exposure on outcome

Rather than estimating the direct effect, an investigator may be interested in the total effect of exposure on outcome, considering all competing events. Fine and Grey proposed a method in 1999 involving estimation of the subdistribution hazard which is commonly used in medical research to account for competing events.[80] In this approach, individuals who experience a competing event during the study period are included in all risk sets. By definition, this person cannot truly experience the outcome of interest because they experienced the competing event. However, since we do not know whether they would have experienced dementia had they not died, the subdistribution hazard and resulting subdistribution hazard ratio estimates the total effect of exposure on outcome during the follow up period. By maintaining individuals who have experienced the competing event in the risk-set, we inflate the number of people at-risk for dementia.

For example, consider the same individual who entered the study at age 65 and was censored at age 70 due to death in the estimation of the cause-specific hazard ratio. To estimate the subdistribution hazard ratio, this individual would continue to contribute to risk sets until the end of the study period, even after their death at age 70.

A Cox proportional hazard model can then be used to estimate a subdistribution hazard ratio that represents the relative change in the subdistribution hazard associated with a 1-unit change in a covariate. In contrast to the cause-specific hazard ratio, estimating subdistribution hazards does not require any assumption about the independence between exposure and events since the competing event is not censored. Another advantage of the subdistribution hazard is it coincides with an estimate of the cumulative incidence function. The direction of the subdistribution hazard coincides with the direction of the risk ratio.[79]

There is no formal way to include competing events in DAGs, however it has been proposed to consider the outcome as a composite outcome, including both the competing event and primary outcome (Figure 2.3).[81] The estimand of interest is the contrast of subdistribution hazards which is an estimate of the total effect comprising of the direct effect of exposure on outcome and the indirect effect of exposure on outcome mediated by competing events. The potential outcomes notation for this estimand is defined by: $\Pr[Y_{t+1}^{a=1} = 1 | Y_t^{a=1} = 0]$ vs. $\Pr[Y_{t+1}^{a=0} = 1 | Y_t^{a=0} = 0]$.

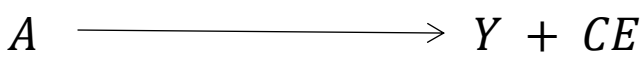

Figure 2.3: Effect of A on Y and CE, where *A*=exposure and *Y*=primary outcome, and *CE*=competing event

Weighted cause-specific hazard ratio: an application of inverse probability of censoring weights

The final approach in this tutorial is an application of inverse probability of censoring weights to account for informed censoring due to competing events. If there is dependence between the exposure and competing events, the application of a Cox proportional hazard model to estimate the traditional cause-specific hazard ratio is not appropriate and will result in a biased estimate if there is unmeasured confounding between the competing event and outcome (Figure 2.1). In the example of air pollution and dementia, the dependency between air pollution and death will result in an underestimation of the true effect. Alternatively, a weighted approach can be applied. Inverse probability weights are created and applied to the baseline population to create a pseudopopulation where there are no dependencies between the competing event and the exposure of interest.[82] In the context of competing events, these weights are applied to account for the dependency between exposure and competing event. To estimate weights, logistic regression models can be used to

predict the probability of survival to the next follow-up period, conditional on factors that may predict death (e.g., demographics and socioeconomic characteristics, health conditions, lifestyle behaviors). The inverse of the predicted probabilities from the logistic regression models are then used to weight each subject in the analysis to create a pseudopopulation at each time point by taking the product of the time-specific weights. This essentially upweights uncensored individuals who have a higher probability of experiencing the competing event and down-weights individuals who have a lower probability of being censored, based on the characteristics included in the logistic regression model. This weighting attempts to correct for bias due to informed censoring due to death by accounting for competing events in the weighting procedure. The estimand of interest for this approach is the contrast between weighted cause-specific hazards. This is similar to the cause-specific approach where competing events are eliminated, however by using a marginal structural model we remove a potential selection bias due competing events and allow for a more flexible application of the Cox proportional hazard model (Figure 2.4). This weighted hazard ratio can be represented by the following potential outcomes notation:

$$\Pr[Y_{t+1}^{a=1} = 1 | \, CE_{t+1}^{a=1} = Y_t^{a=1} = 0] \text{ vs. } \Pr[Y_{t+1}^{a=0} = 1 | \, CE_{t+1}^{a=0} = Y_t^{a=0} = 0]$$

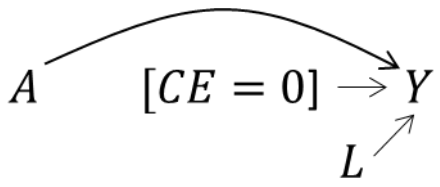

Figure 2.4: Effect of A on Y in pseudopopulation where paths $A \rightarrow CE$ and $L \rightarrow CE$ are removed; A=exposure, CE=competing event, Y=outcome, L= covariates that affect Y

Case-Study: A comparison of approaches to account for competing events in Three-City Cohort

Study Population

Data from the Three-City (3C) Cohort was used to demonstrate three approaches to account for death as a competing event in a study of air pollution and dementia. 3C is a cohort of 9,294 adults 65 years and older from electoral rolls in three cities in France (Bordeaux, Dijon, and Montpellier). This cohort was designed with aims of understanding risk factors of dementia and cognitive impairment and has been previously described.[83] Study participants attended a baseline visit in 1999-2001 and subsequent follow-up visits every 2-3 years for face-to-face interviews and clinical evaluations. Demographic and socioeconomic characteristics, health conditions, and dementia status was updated at each follow-up visit. Individuals were included in this analysis if they were free of dementia at the baseline visit and had available residential information. Finally, death was assessed through medical, hospital, and death records for individuals who dropped out of the study after the baseline visit. The final analysis included 8,314 study participants (Figure 2.5).

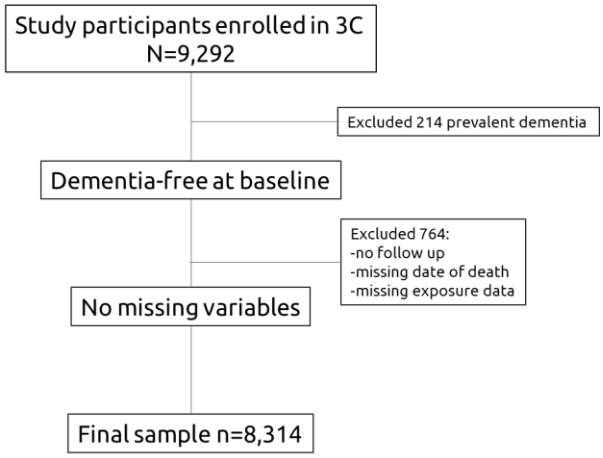

Figure 2.5: Description of study population from the Three-City Cohort

Long-term exposure to air pollution was assigned to each participant. Briefly, individuals were assigned a 10-year average exposure before study enrollment based on residential address reported at study visits. Air pollution estimates for the years 2005-2015 were generated using land-use regression models using air pollution concentration data from EuroAirnet monitoring sites and land use characteristics, population density, road characteristics, and topography. Models to generate air pollution estimates in France have been previously validated.[84] For air pollution estimates 1999-2004, pollutant concentrations were estimated with the CHIMERE chemistry transport model.[85]

Statistical Analysis

Study participants were followed over *t* time (measured in days), from study baseline to the end of their follow-up period, determined by a dementia diagnosis or censored event. Dementia cases were assigned a 10-year average exposure to air pollution ($PM_{2.5}$) before the baseline study visit.

We first applied a naïve approach where competing events are not considered. Next, we applied an approach that estimates the direct effect of air pollution on dementia not mediated by the competing event (cause-specific hazard ratio). We then applied an approach that estimates the total effect of air pollution that considers competing events (subdistribution hazard ratio). Finally, we apply a weighted extension of the cause-specific hazard ratio using inverse probability of censoring weights. Apart from the naïve approach, all methods uniquely account for the competing event, death. To estimate the effect of air pollution on time-to-dementia, hazard ratios were estimated with multi-level Cox proportional hazards models with calendar time as the time scale. We also estimated the effect of air pollution on time-to-competing event by estimating the death-

specific hazard ratio. For each approach, we included age at baseline study visit, sex, and education as potential confounders and the study site as a fixed effect to account for spatial clustering.

2.4. Results

The final study sample included 8,314 older adults enrolled in 3C. Study participants had a mean age of 74 years at baseline. Approximately 10% of individuals developed dementia over the study period and 34% died by the end of the study. The average $PM_{2.5}$ exposure for participants, 10 years before study baseline, was 20.5 ug/ml (IQR=2.2). Descriptive statistics comparing individuals who developed dementia or experienced the competing event are included in (Table 2.1). Air pollution increased the risk of death before dementia (HR=1.27 95% CI: 1.06, 1.53).

Table 2.1: Description of Three-City Cohort by event type

Characteristic	Mean (SD) or Freq (%)	
	Dementia (N=947)	Competing Event (N=1,841)
Age at baseline, years	76.9 (5.4)	76.8 (6.0)
Sex		
Male	327 (34.5)	977 (53.1)
Female	620 (65.6)	864 (46.9)
Education		
Low	418 (44.2)	596 (32.4)
Middle	223 (23.6)	569 (30.9)
High	304 (32.2)	676 (36.7)
Center		
Bordeaux	324 (34.2)	448 (26.5)
Dijon	448 (47.3)	924 (50.2)
Montpellier	175 (18.5)	429 (23.3)
$PM_{2.5}$, median (IQR)	28.5 (IQR=2.0)	28.4 (IQR=1.8)

Results from analyses using the different approaches are summarized in Table 2.2. The naïve analysis where competing events are not considered yielded a hazard ratio of 1.03 (95% CI: 0.80, 1.34) per 5ug/ml increase in $PM_{2.5}$. Censoring individuals at the time of the competing event yielded a cause-specific hazard ratio of 1.02 (95% CI: 0.79, 1.32). Maintaining individuals who

experienced the competing event in the risk set produced a subdistribution hazard ratio of 0.99 (0.76, 1.28). Finally, the weighted hazard ratio that accounts for informed censoring due to competing events produced a hazard ratio of 1.10 (0.87, 1.37).

Table 2.2: Hazard ratios estimated using different approaches to account for competing events

Approach	HR (95% CI)	How is competing event considered?
Naïve	1.03 (0.80, 1.34)	Not considered.
Cause-specific	1.02 (0.79, 1.32)	Censored at time of competing event.
Subdistribution	0.99 (0.76, 1.28)	Individuals with competing event remain in risk set.
IP weighted	1.10 (0.87, 1.37)	Individuals weighted by the inverse conditional probability of experiencing the competing event.

2.5. Discussion

This tutorial demonstrates three approaches to account for competing events in studies of older adults. We consider a case study of air pollution and dementia among older adults in France, accounting for death as a competing event by estimating a cause-specific hazard ratio, a subdistribution hazard ratio, and a weighted cause-specific hazard ratio. After applying the three approaches, we found that the cause-specific hazard model, an estimate of the controlled direct effect, aligned with a naïve model, where competing events were ignored. The subdistribution hazard model generated an attenuated hazard ratio. In contrast, the weighted method generated a weighted cause-specific hazard ratio greater than all other estimates. Although the effect sizes are relatively small and imprecise, the pattern in direction is explained by the unique and distinct target inferences for each approach.

Cause-specific hazard ratios are an intuitive estimate of the controlled direct effect between exposure and outcome and are useful for answering etiologic research questions (e.g., how does increased exposure to air pollution directly affect risk of dementia among individuals without the

competing event?). However, in a situation where increased exposure increases probability of the competing event, collider bias may result in an attenuated cause-specific hazard ratio (Figure 2.1).[50] Similarly, a subdistribution hazard ratio will be attenuated, and in this case reverse directions, when the increased exposure increases risk of the competing event. This is due to the increased proportion of highly exposed individuals who die during the study period and are considered "at-risk" for dementia in analyses. While this inflated denominator seems counterintuitive, it is in fact an estimate of the total effect of air pollution on dementia, including the indirect path through the competing event and coincides with what is observed in the real world, where air pollution has a stronger effect on death then dementia. The target estimand of interest depends on the truth the investigator is seeking to make an inference about. If the investigator is interested in the total effect observed in the real world, that accounts for the relationship between the exposure and competing event, the contrast between subdistribution hazards is appropriate. However, careful attention to the relationship between exposure and competing event, covariate distribution across event types, and sample size are necessary for proper interpretation. For example, in the extreme scenario where the exposure has a causal effect on the competing event in most participants, the subdistribution hazard ratio will appear protective. This is what is observed in the real world, as the exposure has a strong relationship with the competing event and the study is underpowered to detect an effect on the primary outcome. Previous discussions on the interpretation of subdistribution hazard models suggest that the hazards are sensitive to the dependencies between covariates and sample size and require careful interpretation of estimates.[77,86–88]

Now, suppose the investigator is interested in the etiologic research questions. Then, the contrast between cause-specific hazards is appropriate. However, careful attention to potential

collider bias and appropriate modeling strategy is necessary. If there is a potential link between exposure and competing event, then we recommend the use of the weighted approach to estimate the weighted cause-specific hazard to answer etiologic questions. In the context of air pollution and dementia, the evidence base suggesting a link between air pollution and dementia is growing and further etiologic research is necessary before targeted intervention strategies. Specifically, understanding the extent to which air pollution affects risk of dementia can allow for future implementation of policy changes. For this reason, the ability to estimate an unbiased direct effect is the first step toward translating findings to inform environmental regulation and health policy.

The application of weighting approaches that account for dependencies between air pollution and competing events will improve the internal validity of estimates. This is because the weighting procedure creates a pseudopopulation that removes biasing pathways due to known and measured covariates. For example, in our 3C case study, air pollution increased the risk of death. Thus, the weighted pseudopopulation upweights individuals in the study who have an increased probability of death. Weights were created by estimating the inverse probability of the competing event, conditioned on air pollution exposure, demographic and socioeconomic characteristics, and health conditions. Individuals with a lower probability of death, determined by these same characteristics, are assigned smaller weights. Weights were updated for each year of follow-up to account for the differing relationships between the covariates and competing events over time. This approach accounts for competing events by "randomizing" the competing event in the pseudopopulation and the investigator can proceed with a weighted Cox proportional hazard model to estimate a weighted cause-specific effect that accounts for the competing event in the weighting procedure. There are some limitations to this weighting approach. The creation of the pseudopopulation hinges on a strong understanding of the causal relationships and selection

mechanisms and measured variables on the biasing pathways. Furthermore, one may be hesitant to apply a weighting scheme that essentially "upweights" dead participants and creates an immortal study population.[89] However, if we seek to target a causal interpretation, we suggest that such weighting approach is appropriate.[90]

The weighting approach is contrary to the traditional cause-specific approach where competing events are considered by restricting the study sample to the population who had not experienced the competing event (i.e., censoring competing events). It is important to acknowledge that this traditional approach constitutes a well-suited approach to deal with etiologic questions but may be subject to a selection bias and if so, may not fit the assumptions required to estimate a hazard ratio from a Cox proportional hazard model. The weighting approach also contrasts with the subdistribution hazard approach, which is an estimate of the total effect. As suggested by Rudolph et al, for research questions related to policy evaluation, estimating the subdistribution hazard ratio may be an appropriate approach, as estimating the total effect mirrors what we would observe in the real world where both the outcome and competing event occur.[66]

We recommend the following strategies when implementing a survival analyses of air pollution and dementia, where competing events are of concern:

1. Clarify the question of interest. Is the investigator interested in the direct effect of exposure on outcome, not mediated by the competing event? Or, the total effect that includes pathways through the competing event.

2. Think through the causal mechanisms through which the covariates and selection are dependent (can present DAG in supplementary material).

3. Identify the target inference and estimand of interest (cause-specific or subdistribution hazard ratio?).

4. Describe the competing event by defining how it is classified (e.g., deaths among dementia-free individuals before the last follow-up visit).

5. Present the distribution of exposure of interest and covariates for primary outcome and competing events in supplementary material (e.g., Table 2.1).

6. Present different approaches to account for competing events in supplementary material (e.g., Table 2.2).

We believe implementing these steps will help clarify the target estimand when considering competing events and will result in clarity about the expected direction of bias after accounting for competing events in analyses. This will make estimates more comparable across studies to help triangulate evidence about air pollution and dementia. These guidelines can be extended to settings outside of air pollution and dementia. Further developments have been made to extend consideration of competing events to deal with confounding between competing events and primary outcome, time-varying covariates, delayed entry into study, and missing data.[81,91–93]

In summary, in this tutorial we demonstrated three approaches to account for competing events, using an example of air pollution and dementia, with death as a competing event. We recommend careful thought about the research question and target inference. In general, we recommend IP weighting approaches to account for competing events to answer etiologic questions. However, the application of any approach requires careful understanding of assumptions and interpretation and should be communicated in research.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Ilango, Sindana D; Mortamais, Marion; Gutierrez, Laura-Anne; Berr, Claudine; Benmarhnia, Tarik. The dissertation author was the primary investigator and author of this paper.

3. The role of cardiovascular disease in the relationship between air pollution and incident dementia: a population-based cohort study

Sindana D. Ilango, Hong Chen, Perry Hystad, Aaron van Donkellar, Jeffrey C Kwong, Karen Tu, Randall V Martin, Tarik Benmarhnia

3.1. Abstract

Background: Evidence suggests a link between air pollution and dementia. Cardiovascular disease (CVD) may be a potential determinant of dementia. This motivated us to quantify the contribution of CVD to the association between air pollution and dementia.

Methods: A cohort of Canadian-born residents of Ontario, who participated in the 1996–2003 Canadian Community Health Surveys, was followed through 2013 or until dementia diagnosis. Exposure to nitrogen dioxide (NO_2) and fine particulate matter ($PM_{2.5}$) was estimated with a 3-year average and 5-year lag before dementia diagnosis. Incident CVD was evaluated as a mediator. We used multi-level Cox proportional and Aalen additive hazard regression models, adjusting for individual- and neighborhood-level risk factors to estimate associations with NO_2 and $PM_{2.5}$. We estimated the total, direct and indirect effects of air pollution on dementia through cardiovascular disease.

Results: This study included 34,391 older adults. At baseline, the mean age of this cohort was 59 years. The risk of dementia was moderately higher among those more exposed to NO_2 (hazard ratio (HR) 1.10, 95% confidence interval (CI) 0.99–1.19; and 100 additional cases per 100,000 [standard error (SE) $<100x10^{-5}$]) and $PM_{2.5}$ (HR 1.29, 95% CI 0.99–1.64; 200 additional cases per 100,000 [SE $100x10^{-5}$]) after adjusting for covariates; however, these estimates are imprecise. A greater proportion of the relationship between $PM_{2.5}$ and dementia was mediated through CVD than NO_2 for both scales.

Conclusions: These results suggest some of the association between air pollution and dementia is mediated through CVD, indicating that improving cardiovascular health may prevent dementia in areas with higher exposure to air pollution.

3.2. Introduction

Ambient air pollution has a range of acute and chronic health implications. In 2015, an estimated 4 million premature deaths globally were attributable to exposure to air pollution.[20] Ambient air pollution is a complex mixture of particulate matter (PM), gaseous pollutants (e.g., nitrogen dioxide [NO_2]), persistent organic pollutants, and heavy metals. Long-term exposure can result in oxidative stress and increased inflammation which can impact the pathophysiological processes of major chronic diseases and exacerbate existing health conditions.[21]

Recent evidence suggests air pollution affects diseases of the central nervous system, including dementia.[22] Insight into the etiology and prevention of these diseases and neurodegeneration is valuable, especially with increasing concern with aging populations and the current absence of a cure. The worldwide prevalence of dementia is expected to increase sharply, from 44 million people with dementia in 2013 to an estimated 135 million in 2050.[94]

A systematic review of 18 epidemiologic studies published through 2014 on air pollution and dementia and cognitive function found most studies identified at least one notable association between increased exposure to air pollution and worse cognitive outcome.[15] All previous studies examining the relationship between long-term exposure to ambient air pollution and dementia suggest an association.[34,37,39,41–43,45,46] For example, traffic-related air pollutants and close proximity to heavy traffic roads were found to be associated with increased risk of dementia in the UK, Sweden and Ontario, Canada.[37,42,43,95] Evidence relating air pollution to cognitive

performance is less consistent. Associations between air pollution and measures of cognitive function and decline were found in studies of older US adults[55,59,96], in select urban cities in Germany[64], but not found in a studies of older adults in Los Angeles[58] or the UK[97].

Despite growing evidence linking air pollution to dementia[34,37,39,41–43,45,46], the exact mechanism is not well understood. Emerging literature suggests oxidative stress, neuroinflammation, and cardiovascular disease (CVD) to be contributing factors.[22,24,98] Knowledge of the causal pathway can provide valuable insight into disease etiology and pathophysiology and inform where medical and public health interventions can be most effectively applied to reduce disease burden. Identifying *modifiable* intermediates along a causal pathway are of particular interest in public health research because they offer opportunities to intervene at a population level and can have lasting effects.

In this study, we hypothesize that CVD is on the causal pathway between air pollution and dementia. Air pollution has been consistently related to increased risk of incidence, complications, and mortality from CVD.[67,99–101] Furthermore, growing evidence has linked CVD and its risk factors to impaired cognitive function and often co-occur with dementia.[7,68]

To date, no study has investigated the extent to which CVD plays an intermediate role in the association between air pollution, specifically $PM_{2.5}$ and NO_2, and incident dementia. We aimed to disentangle this relationship on both relative and absolute scales that allow for describing the strength of the effect and quantifying the potential public health benefits, respectively. Dissecting this relationship can provide important insight into the causal mechanism of dementia, identify vulnerable populations, and prioritize public health efforts to reduce the burden of dementia. We thus performed a causal mediation analysis in a large population-based cohort in Ontario, Canada.

3.3. Methods

Study Design and Population

We conducted a population-based cohort study of older adults in Ontario, Canada. Eligible participants included Canadian-born Ontario residents who participated in the 1996-1997 cycle of the National Population Health Survey (NPHS) and the 2000/2001, 2003, and 2005 cycles of the Canadian Community Health Survey (CCHS). NPHS and CCHS are population-based surveys administered across Canada that collect information about health status, health care utilization, and health determinants, covering approximately 98% of the Canadian population aged 12 years or older with a response rate of about 80%.[102] Data from these surveys have been widely used for public health surveillance and research.

Participants were included in this study if they lived in Ontario for at least 5 years and were 45 years or older at the date of survey (i.e. study baseline). This allowed us to measure prior cumulative exposure to air pollution in Ontario and capture older individuals who are at higher risk of developing dementia.

To study the potentially mediating role of CVD, the exposure must accumulate before the mediator, which must subsequently occur prior to the outcome. We achieved this by using the year of survey completion to define time periods, to ensure proper temporality. Briefly, we estimated chronic exposure to air pollution before baseline survey completion. Then, CVD was assessed during a window after the baseline survey. Finally, to ascertain dementia, individuals were followed from the end of CVD follow-up through 2013 or until incident dementia diagnosis. We *a priori* considered this cohort structure to incorporate a hypothesized lag before the effects of air pollution can present itself in neurocognitive outcomes. The Research Ethics Board of Sunnybrook Health Sciences Centre, Toronto, Canada approved the study.

Exposure Assessment

Chronic exposure to ambient air pollutants, specifically fine particulate matter ($PM_{2.5}$) and nitrogen dioxide (NO_2), was the exposure of interest for this study. We used previously estimated mean measurements of $PM_{2.5}$ and NO_2 at a spatial resolution of about 1x1km for each year between 1993 and 2013 to calculate these exposures. Details are included in the supplementary material.

To estimate chronic exposure to ambient air pollutants, running averages of pollutant measurements over the three years leading up to the time of baseline survey completion were calculated (Figure 3.1). For example, if a participant completed the survey at the end of 1996, the participant's chronic exposure was estimated as an average of pollutant measurement from 1994, 1995, and 1996. This three-year running average was estimated for both $PM_{2.5}$ and NO_2 separately. These estimates were updated as the follow-up continued through the study period.

Outcome Assessment

Cases of incident dementia were defined as having at least one of the following three criteria (see supplementary material for further details):

1) at least one hospital admission with a diagnosis of dementia [see list in supplementary material], or

2) at least three physician claims over a two-year period, or

3) a prescription relating to dementia

We ascertained this information using data linkage to population-based health administrative databases. The databases included: hospital discharge records from the Canadian Institute for Health Information's Discharge Abstracts Database, physician claims from the Ontario Health Insurance Plan database, and prescription claims from the Ontario Drug Benefits database. These

datasets were linked using unique encoded identifiers and analyzed at ICES. The province of Ontario has a single-payer, universal health care system offered through the provincial government; virtually all Ontario residents are covered through this system and are included in these registries. The algorithm for identifying dementia cases has been validated with medical chart review and has a sensitivity of 79% and specificity of 99%.[103]

Cohort members were followed for incident dementia from five years after completion of the baseline health survey through 2013; this took into account a hypothesized five-year lag period for air pollution to have an effect on dementia. For example, the same individual who completed the survey at the end of 1996 would be followed for dementia from 2001 through 2013. We chose 5 years as the lag period because this was the greatest length we could statistically account for, given the size of this cohort. Individuals diagnosed with dementia during the lag period were not included in the analysis.

Mediator Assessment

We examined incident cardiovascular events as a potential mediator between exposure to air pollution and dementia. After excluding individuals with prevalent cardiovascular events at the time of survey completion, cohort members were followed for first cardiovascular event for up to five years, beginning at the year of baseline health survey. Cardiovascular events were defined as hospital admissions or medical procedures for: coronary heart disease, stroke, arrhythmia, and congestive heart failure. We obtained information about these events from hospital discharge abstracts, medical procedure codes, and the Ontario Congestive Heart Failure Database (see supplementary material for details).[104]

The five-year mediator follow-up occurred immediately after the interval during which air pollution exposure was measured; it coincided with the five-year lag between air pollution and

dementia follow-up. For instance, if the cohort member completed the health survey in 1996, he/she would be followed for cardiovascular events from 1996 to 2001 and then followed for dementia. Figure 3.1 describes the timing of outcome, exposure, and mediator follow-up periods.

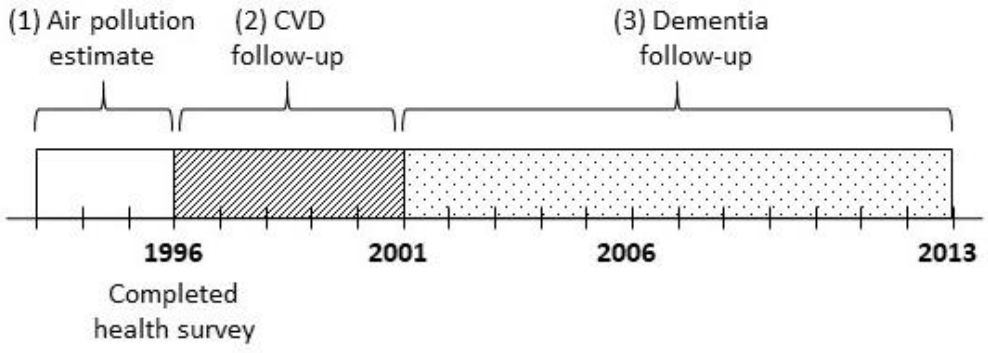

Figure 3.1: Schematic of cohort follow-up periods. Example exposure, mediator, and outcome follow-up periods for an individual who completed a health survey in 1996. (1) three-year air pollution measurement (exposure); (2) five-year follow-up for CVD (mediator); (3) dementia follow-up through 2013 (outcome)

Covariates

We selected *a priori* potential confounders to include as covariates in the model. We ascertained demographic and health behavior information from the health surveys. The details of all covariates can be found in the supplementary material.

Statistical Analysis

The study cohort was described with means (SD) and frequencies (%) for all variables of interest. We then estimated the association between air pollution and dementia and air pollution and cardiovascular events, separately. We generated a Cox proportional hazards model and an Aalen additive hazards model for each of these relationships to estimate hazard ratios (HR) and 95% confidence intervals (CI) and parameter estimates (β) and standard errors (SE) for every 5 ppb and 10 ug/m^3 increase in NO_2 and $PM_{2.5}$, respectively. In these eight models, we accounted

for spatial clustering by incorporating two-level clustering with census neighborhood nested within census division as a random effect. We additionally ran minimally adjusted models for each pollutant. Final adjusted models included covariates for individual age, sex, education, marital status, income quintile, smoking status, body-mass index, physical activity, rural residence, and northern region, and neighborhood-level percentages of recent immigrants, income quintile, unemployment, and less than high school education.

Causal Mediation Analysis

The primary objective of this paper was to perform a formal causal mediation analysis to decompose the total effect of air pollution on incident dementia into its natural direct and natural indirect effect through cardiovascular events. We used the 2-stage regression method for mediation analysis for survival data under (1) Cox proportional and (2) Aalen additive hazard models.[54,105,106] Survival analysis in epidemiology most frequently employs Cox proportional hazard models. A newly developed, alternative, and more flexible approach to mediation analysis with survival data involves Aalen additive hazard models.[107] This approach offers additional flexibility by not requiring proportional hazards and is a straightforward and intuitive way of interpreting effect sizes are extended to absolute number of events.[106]

We made the following assumptions on both models: no unmeasured confounding and no mediator-outcome confounder affected by the exposure itself. We also *a priori* assumed no exposure-mediator interaction thus, we expect the controlled direct effect (CDE) to coincide with the natural direct effect (NDE) and interpreted them interchangeably.[105] For multiplicative and additive scales, the NDE compares the dementia incident risk or additional cases for each unit increase in exposure to air pollution, controlling for the CVD pathway and all other covariates. The natural indirect effect (NIE) represents the change in dementia risk or additional cases when

exposure to ambient air pollution is held constant while CVD risk changes in response to one unit increase in air pollution exposure.

For each hazard model, we fit two multi-level, mixed effects regression models to estimate the total effect (TE), NDE, and NIE. Both model types were considered to assess mediation on multiplicative and additive scales. Further details about estimating these quantities are described in the supplementary material. With these estimates we calculated the estimated proportion of the total effect of ambient air pollution on dementia mediated through CVD. TEs, NDEs, and NIEs and their 95% CIs were computed using bootstrapping procedures (250 replications). We conducted various sensitivity analyses to assess the robustness of our findings (see details in the supplementary material). All data management and statistical analyses were conducted using RStudio Version 1.1.423 with the extension packages *coxme* and *timereg*.[108,109]

3.4. Results

This study included 34 391 older adults from Ontario, Canada, who contributed a total of 366 208 person-years. Approximately 7% (n=2559) of individuals developed dementia during this period. At baseline, the mean age of this cohort was ~60 years. Fifty-eight percent were female and about half of the population (55%) attended some or completed college. About one-third of the study population had no history of smoking.

Table 3.1: Baseline characteristics of study population in Ontario, Canada (n=34,391)

Characteristic	Mean (SD) or n (%)
Demographics	
Age at entry (years)	60.19 (10.56)
Sex	
Male	14,555 (42%)
Female	19,836 (58%)
Education	
Less than high school	9,055 (26%)
High school diploma	6,462 (19%)
Some college or more	18,874 (55%)
Income	
Lowest	758 (2%)
Low-middle	1856 (5%)
Middle	5,117 (15%)
Middle-high	8,083 (24%)
Highest	6,645 (19%)
Unknown	11,932 (35%)
Marital Status	
Married	21,175 (62%)
Single	2827 (8%)
Separated, widowed, divorced	10373 (30%)
Unknown	16 (<1%)
Health Information	
Physical activity	
Active	7,346 (21%)
Moderate	8,792 (26%)
Not active	18,253 (53%)
Smoking status	
Never smoker	9,976 (29%)
Former smoker	16,441 (48%)
Current smoker	7,974 (23%)
Weight status	
Underweight	500 (2%)
Normal	13,505 (39%)
Overweight	13,455 (39%)
Obese	6,931 (20%)
Pre-existing comorbidity	
Diabetes	3,262 (10%)
Hypertension	12,026 (35%)
Traumatic brain injury	1,279 (4%)
Physician density per 1,000	1.49 (1.17)
Geography	
Northern latitude	5989 (17%)
Missing	6 (<1%)
Rural residence	11,243 (33%)
Area-level risk factors	
Percentage of recent immigrants	1.87 (2.61)
Percentage unemployed	6.71 (1.93)
Percentage under high school	28.50 (5.74)

The three-year cumulative exposures to NO_2 and $PM_{2.5}$ five years before dementia follow-up were 10.4 ppb (range: 2.2-54.4 ppb; IQR: 7.6 ppb) for NO_2 and 8.6 µg/m3 (range: 0.8-35.2; µg/m3 IQR: 4.7 µg/m3) for $PM_{2.5}$.

For every 5 ppb and 10 $\mu g/m^3$ unit increase in cumulative exposure to NO_2 and $PM_{2.5}$, there was a positive association with the incidence of dementia, with fully adjusted HRs of 1.10 (95%CI 0.99-1.19) and 1.29 (95%CI 0.99-1.64), respectively (Table 3.2). Additive models indicate that for each unit increase, 100 (SE $<100x10^{-5}$) and 200 (SE $100x10^{-5}$) additional cases of dementia per 100 000 per year are diagnosed for NO_2 and $PM_{2.5}$, respectively (Table 3.2). See Table 3.5 for minimally adjusted estimates.

Associations between exposure to NO_2 and $PM_{2.5}$ and CVD were also detected, with HRs of 1.01 (95%CI 0.96-1.06) and 1.08 (95%CI 0.94-1.24), although these estimates have wide confidence intervals and are imprecise (Table 3.2). For each unit increase, 100 (SE $<100x10^{-5}$) and 300 (SE $100x10^{-5}$) additional CVD cases per 100 000 individuals per year were diagnosed for NO_2 and $PM_{2.5}$, respectively (Table 3.2).

Table 3.2: Associations between air pollutant, dementia, and cardiovascular disease

	Cox PH Model HR[a] (95% CI)	Aalen model Estimate[a] (SE)
NO_2 [b]		
Dementia [c]	1.10 (0.99, 1.19)	$100x10^{-5}$ ($<100x10^{-5}$)
Cardiovascular disease	1.01 (0.96, 1.06)	$100x10^{-5}$ ($<100x10^{-5}$)
$PM_{2.5}$ [b]		
Dementia [c]	1.29 (0.99, 1.64)	$200x10^{-5}$ ($100x10^{-5}$)
Cardiovascular disease	1.08 (0.94, 1.24)	$300x10^{-5}$ ($100x10^{-5}$)

HR= hazard ratio; CI= confidence interval; SE=standard error
[a] Adjusted for age, sex, education, marital status, income quintile, smoking status, body-mass index, physical activity, rural residence, and northern region; area level: recent immigrants, unemployment, and education
[b] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3
[c] Total effect obtained from product method

Our mediation analysis shows that the effect of air pollution on dementia may be partially mediated through cardiovascular events for both scales (Table 3.3, Table 3.4). On the multiplicative scale, we observed an indirect effect HR of 1.01 (95%CI 0.98-1.03) for NO_2 and 1.06 (95%CI 0.99-1.12) for $PM_{2.5}$ mediated through CVD. These translate to approximately 9% of the effect of NO_2 on dementia and 21% of the effect of $PM_{2.5}$ on dementia being mediated through cardiovascular events in this study population. In additive models, for NO_2, we observe approximately 1.5 (95% CI 1.0-2.6) additional cases per 100 000 per year can be attributed to the pathway through CVD. For $PM_{2.5}$, approximately 4.2 (95% CI 2.9-6.9) additional cases per 100,000 per year can be attributed to the pathway through CVD. These translate to approximately 2% and 4% of the pathway from NO_2 and $PM_{2.5}$ can be attributed to the pathway through CVD. For both relative and absolute scales, a greater proportion of the effect of $PM_{2.5}$ on dementia may be mediated through CVD than for NO_2. It is important to note that measures of proportion mediated should be interpreted as a qualitative measure and are imprecise due to the wide confidence intervals of our indirect effects. Our conclusions did not change appreciably with sensitivity analyses (see supplementary material).

Table 3.3: Total, controlled direct, and natural indirect effects of ambient air pollutant through cardiovascular disease (Cox proportional hazards model)

Pollutant	Total effect HR[a] (95% CI)	Natural direct effect HR[b] (95% CI)	Natural indirect effect HR[b] (95% CI)
NO_2[c]	1.10 (0.99 – 1.19)	1.09 (1.00 – 1.18)	1.01 (0.98 – 1.03)
$PM_{2.5}$[c]	1.29 (0.99 – 1.64)	1.22 (0.95 – 1.56)	1.06 (0.99 – 1.12)

HR=hazard ratio; CI=confidence interval

[a] Total effect obtained from product method

[b] Adjusted for age, sex, education, marital status, income quintile, smoking status, body-mass index, physical activity, rural residence, and northern region; area level: recent immigrants, unemployment, and education

[c] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

Table 3.4: Total, controlled direct, and natural indirect effects of ambient air pollutant through cardiovascular disease (Aalen additive hazards model)

Pollutant	Total effect Estimate[a] (95% CI)	Natural direct effect Estimate[b] (95% CI)	Natural indirect effect Estimate[b] (95% CI)
NO_2[c]	$100x10^{-5}$ ($1.2x10^{-5}$, $100x10^{-5}$)	$100x10^{-5}$ (<$100x10^{-5}$, $100x10^{-5}$)	$1.45x10^{-5}$ ($1x10^{-5}$, $2.6x10^{-5}$)
$PM_{2.5}$[c]	$100x10^{-5}$ ($3.6x10^{-5}$, $300x10^{-5}$)	$100x10^{-5}$ (<$100x10^{-5}$, $300x10^{-5}$)	$4.20x10^{-5}$ ($2.9x10^{-5}$, $6.9x10^{-5}$)

CI=confidence interval

[a] Total effect obtained from product method

[b] Adjusted for age, sex, education, marital status, income quintile, smoking status, body-mass index, physical activity, rural residence, and northern region; area level: recent immigrants, unemployment, and education

[c] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

3.5. Discussion

Using a population-based cohort of 34,391 individuals in Ontario, Canada, we decomposed the total effect of exposure to ambient air pollutants, specifically $PM_{2.5}$ and NO_2, on incident dementia into its respective direct and indirect effects through CVD on multiplicative and additive scales. We found an increased risk of dementia among those with higher exposure to NO_2 (HR 1.10, β $100x10^{-5}$) and $PM_{2.5}$ (HR 1.15, β $200x10^{-5}$). We found some evidence of an indirect effect through CVD for both pollutants, with incident CVD mediating more of the relationship between $PM_{2.5}$ and dementia than the relationship between NO_2 and dementia. These effects are observed in a region with pollutant concentrations that are among the lowest in the world.

This total effect is in line with previous studies. Two recent systematic reviews of air pollution and cognitive functioning and dementia reported that the majority of reviewed studies found positive associations between higher exposure to air pollution ($PM_{2.5}$ or living in a high traffic area) and worse cognitive aging and dementia.[15,16]

Such findings are biologically plausible and are supported by neuroimaging and biologic studies. Pathology studies have shown that dementia is the result of a combination of neurodegenerative and vascular lesions, suggesting commonalities in the mechanisms of dementia and vascular disease.[54,105] Taking this into account with the substantial literature linking air

pollution with cardiovascular and cerebrovascular risk factors and disease provides motivation to investigate CVD as an intermediate on the pathway.[67,99–101]

Understanding dementia disease etiology and the role of CVD in dementia can provide invaluable insight toward prevention strategies. Since well-defined interventions to prevent CVD exist, targeted efforts to improve cardiovascular health may be beneficial to dementia prevention in areas with increased exposure to air pollution. For instance, cardiovascular health programs, screening, and access to CVD healthcare can be prioritized in highly polluted areas to not only improve CVD outcomes but also potentially reduce the risk of dementia.

Many methodological challenges exist when studying dementia at the population level.[110] First, our findings are subject to selective attrition due to mortality. While this would likely result in underestimating the effect sizes, we acknowledge the possibility that a portion of study participants died from air pollution-related causes (e.g., CVD, respiratory conditions) before living long enough to develop dementia. Thus, those who developed dementia could be a healthier group of participants who were less vulnerable to detrimental air pollution effects. We attempted to address this concern by examining the influence of potential risk factors for competing risks in our models.

Second, our classification of dementia is limited to diagnosed cases. This can be a concern when cases are ascertained from a data source that doesn't definitively capture dementia information.[110] For example, using the hospital discharge data alone underestimates dementia cases because the diagnosis and management of dementia does not require hospital admission. We addressed this by using three sources (i.e., hospital discharge records, physician claims, and prescription claims) to identify cases in a population with universal access to healthcare. It is also possible that there might be inconsistencies in diagnoses, because the diagnosis often depends on

caregiver's concern and access to care. Additionally, individuals with CVD conditions may interact more frequently with the healthcare system and therefore increase their probability of being diagnosed earlier for dementia. However, since our study population is restricted to individuals who have completed a health survey, individuals in our study are more likely to be health-conscious and are likely of higher socioeconomic status. While this may limit the generalizability of our findings, we believe studying dementia in this population minimizes potential outcome misclassification, and thus improves the validity of our findings. Similarly, we were unable to account for undiagnosed CVD cases.

Next, the results from this mediation analysis rely on the assumptions of no unmeasured confounding and no mediator-outcome confounder affected by the exposure.[54] To check this assumption, we conducted sensitivity analyses to assess the extent of confounding by measured risk factors and health status variables (e.g., diabetes, hypertension, traumatic brain injury) that may violate the assumption of no unmeasured mediator-outcome confounding and found no appreciable change in our observed indirect effect sizes. We also assumed no interaction between air pollution and CVD, allowing the CDE and NDE to be interpreted similarly.

This study has major strengths including its large size, ability to ascertain incident cardiovascular events and incident dementia from validated sources, availability of individual socioeconomic and health behavior information, and analytic approach. We encourage future studies to replicate our methods in established cohort studies where participants have routine evaluations for dementia and cognitive decline.

A formal causal mediation analysis is a valuable tool to disentangle complex relationships and provides insight toward understanding disease etiology and identifying specific prevention efforts. Our results suggest an increased risk of dementia among individuals with higher

cumulative exposure to air pollution, which were partially mediated through CVD. Our study identified and highlighted two modifiable risk factors, ambient air pollution and CVD, for dementia. Intervening on one or both have the potential to significantly reduce the burden of dementia.

3.6. Appendix

Ascertainment of Long-term Exposure to $PM_{2.5}$ and NO_2

Assessment of Ambient Concentrations of $PM_{2.5}$ and NO_2

Estimates of ground-level concentrations of $PM_{2.5}$ were derived by relating satellite retrievals of aerosol optical depth, a measure of light extinction by aerosols in the total atmospheric column, to ground-level $PM_{2.5}$ using the temporally and spatially varying relationship simulated by a global atmospheric chemistry transport model (GEOS-Chem CTM).[111] Ground-level observations of $PM_{2.5}$ were then incorporated using a geographically weighted regression with predictors that included information on urban land cover, elevation, and aerosol composition at a spatial resolution of 1×1km for each year between 1998 and 2012.[112] Covering all North America below 70^{o}N, which includes all of Ontario, these annual estimates of $PM_{2.5}$ have been shown to closely agree with out of sample ground measurements at fixed-site monitoring stations across North America (R^2=0.82, n=1440). Similar $PM_{2.5}$ estimates have been used to determine the associations of $PM_{2.5}$ with cardiorespiratory mortality and morbidity and the global disease burden attributable to air pollution.[113–117]

To derive exposure to NO_2, we made use of a national land-use regression (LUR) model developed using measurements of NO_2 at the fixed-site stations of Environment Canada's National Air Pollution Surveillance Network.[118] Briefly, this model was constructed by regressing observed

annual mean concentrations of NO_2 in Canada against an array of predictors (*e.g.*, satellite estimates of NO_2, area of industrial land use, road length) to capture background and regional variations of NO_2.[119] The estimates were then augmented by incorporating local-scale variations of NO_2 due to vehicle emissions by applying spatially-varying multipliers that represented distance-decay gradient in NO_2.[119] The LUR model explained 73% of the variability in annual 2006 measurements of NO_2, with a root mean square error of 2.9 parts per billion.[119] The LUR-derived NO_2 estimates have been used to estimate the effect of traffic-related air pollution on mortality[7] and adverse birth outcomes in Canada.[120]

Temporal Calibration of Exposure Surfaces of $PM_{2.5}$ and NO_2

Because the exposure surface of NO_2 was derived for 2006 (the mid-point of the study period), to incorporate changes in ambient concentrations of NO_2 over time, we conducted yearly calibrations of NO_2 exposure surface by scaling it with a ratio between the average concentrations of NO_2 at all fixed-site monitors across Ontario in a given year and that in 2006, thus producing annual mean estimates of NO_2 between 1994 (seven years before cohort inception to allow for lagging exposure) and 2013. This approach assumes that the spatial pattern in NO_2 did not change appreciably during this period. This is a reasonable assumption because despite decreasing concentrations over time, the degree of change in NO_2 concentrations and thus the spatial difference remained stable in Ontario.

In addition, we have shown previously that areas in Ontario with higher concentrations of $PM_{2.5}$ have retained their spatial ranking from 1996 to 2010 and that variability in longer-term exposure to $PM_{2.5}$ is primarily spatial rather than temporal.[114] Because annual estimates of $PM_{2.5}$ were not available before 1998 and after 2012, we extrapolated $PM_{2.5}$ estimates in 1998 annually to 1994-1997, and to 2013, using $PM_{2.5}$ estimates in 2012.

We further verified long-term stability in the spatial patterns of annual mean concentrations of NO_2 during the period of 2003 to 2014. In doing this, we compiled historical data on the monitoring of NO_2 from Environment Canada's National Air Pollution Surveillance (NAPS) network. We considered the start year of 2003 because many new fixed-site stations were added after 2003. We excluded fixed-site monitors that were located outside Ontario and that were operated for less than half of this period, leaving sufficient data to derive annual mean concentrations for 20 cities in Ontario, such as Toronto, Hamilton, and Ottawa.

Using the monitoring data, we estimated long-term average concentrations of NO_2 for each city over the period 2003-2014. Using the annual mean concentrations of NO_2 from the 20 cities, we further estimated the total variance of NO_2 across the 20 cities and throughout the period 2003-2014. In addition, we estimated the variance of NO_2 due to temporal variability from 2003 to 2014. This was done by calculating mean exposure averaged across the 20 cities for each year and then estimating the variance of the annual averages over time. The total variance was 16.3 $(ppb)^2$ while the temporal variance was 5.4 $(ppb)^2$. Thus, 67% of the total variation in the concentrations of NO_2 among the 20 cities was associated with spatial variability and only 33% with variation over time. This result suggests that variability in the concentrations of NO_2 in Ontario is primarily spatial in nature and not temporal.

The representativeness of NO_2 measurements derived using a land-use regression model for longer-term spatial contrast in NO_2 has been reported in several previous studies.[121–123] For example, in the European Study of Cohorts for Air Pollution Effects (ESCAPE), Eeftens and coworkers assessed the relationship between NO_2 measured in 1999 and NO_2 measured in 2007 at 40 locations in the Netherlands.[122] They found strong correlation among NO_2 measurements between the two periods (coefficient of determination, R^2=86% or r=0.92), indicating that the areas

with higher NO_2 concentrations in earlier periods were likely to retain their spatial ranking. Long-term stability in the spatial patterns of NO_2 has also been shown in a study conducted in Rome, Italy, and another study conducted in Vancouver, Canada.[121,123] We therefore expect that the spatial contrast in NO_2 from our land-use regression model provides reasonable estimates of longer-term spatial exposure to NO_2 in Ontario.

To further evaluate the robustness of our results, we conducted another sensitivity analysis by deriving station-specific temporal scaling factors and applied these spatially-varying scaling factors to cohort members living near these fixed-site stations. Because only 13 stations were operated continuously between 1994 and 2013, this analysis was restricted to those living within 50 km from these stations.

Ascertainment of Dementia

Cases of incident dementia were defined as having at least one of the following three criteria:

1) at least one hospital admission with a diagnosis of dementia [International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9CM) diagnostic code 46.1, 290.0-290.4, 294, 331.0, 331.1, 331.5, 331.82 or ICD-10 code F00-F03, G30 after 2002], or

2) at least three physician claims (code 290, 331) over a two-year period, or

3) a prescription relating to dementia (e.g., donepezil, galantamine)

Ascertainment of Cardiovascular Events

Cardiovascular events were defined as experiencing one or more of the following: coronary heart disease, stroke, arrhythmia, and congestive heart failure. Each condition was ascertained from the following information:

Coronary Heart Disease: hospital discharge data using the most responsible or secondary diagnostic code (ICD-9: 410-414; ICD10: I20-I25) in conjunction with medical procedure codes for percutaneous coronary intervention (PCI) and coronary artery bypass grafting (CABG) before 2002: PCI (4802, 4803, 4809) and CABG 481; and after 2002: PCI (1IJ50, 1IJ57GQxx) and CABG 1IJ76

Arrhythmia: hospital discharge database with most responsible or secondary diagnostic code (ICD-9: 427; ICD-10: I47, I48, I49, I460, I469, R001)

Stroke: hospital discharge database with most responsible or secondary diagnostic code (ICD-9: 430, 431, 434, 436 and ICD-10: I60, I61, I63.0, I63.1, I63.2, I63.3, I63.4, I63.5, I63.7, I63.8, I63.9, I64, H34)

Congestive Heart Failure: Ontario Congestive Heart Failure (CHF) Database which uses data from CIHI discharge abstract database, physician service claims from the OHIP database, and emergency department records from National Ambulatory Care Reporting System (NACRS).[104] CHF was defined as one hospital admission with a CHF diagnosis or an OHIP claim/ emergency department record with a CHF diagnosis followed within two years by either 3 a second OHIP claim/NACRS record or a hospital admission with a CHF diagnosis (ICD-9: 428; ICD-10: I500, I501, I509).

Ascertainment of Comorbidities

Information on prevalent depression was obtained from the health survey taken by the participant. People with diabetes, hypertension, or heart failure were identified using validated databases of all residents diagnosed with these conditions in Ontario.[124–126] The presence of the diagnosis of a specific disease between 1991 and 2001 was defined as the presence of that comorbidity.

Covariates

We considered potential confounders between each of the relationships of interest (i.e., ambient air pollution and dementia, CVD and dementia, and air pollution and CVD).

The covariates we used were: age of cohort member (years), sex (male/female), education (less than high school, high school graduate, some college or more), income quintile, marital status (single; married, common law, or living with partner; separated, widowed or divorced), rural residence (yes/no), smoking status (never, current, former), body mass index (BMI) category (underweight, normal weight, overweight, obese) and physical activity (active, moderate, not active). We also created dichotomous variables using the individual health region to classify individuals residing in more northern latitudes (yes/no) and in the Greater Toronto Area (yes/no). Two of the 14 health regions in Ontario were considered as northern Ontario (North East and North West). We also considered additional covariates including prevalent conditions (yes/no) (diabetes, hypertension, and traumatic brain injury) in sensitivity analyses.

We used neighborhood-level covariates from the Canadian Census to account for clustering of socioeconomic status by neighborhood and census division. Estimates from the Census are released every five years; we ascertained aggregate estimates from the Census completed in the year closest to the year of follow-up to be in line with the time-varying estimates of air pollution

exposure in the analysis. Income quintile was aggregated at the neighborhood level and percent of recent immigrants, percent of unemployment, and percent of population with less than a high school education were aggregated at the census division level. For sensitivity analyses, we also included physician density per 1,000 at the municipal or city level using the ICES physician database.

Causal Mediation Analysis Methods

To account for the time to event outcome and mediator, we fit (1) Cox proportional and (2) Aalen additive hazard regression models.[54,106]

We fit the following multi-level, mixed-effects models, with age as the time scale, to estimate the natural direct effect of air pollution on incident dementia:

$$\lambda(t|A,M,C) = \lambda_0(t)\exp(\theta_1 A + \theta_2 M + \theta_3 C) \quad \text{(Cox proportional hazard model)}$$

$$\lambda(t|A,M,C) = \lambda_0(t) + \lambda_1 A + \lambda_2 M + \lambda_3 C \quad \text{(Aalen additive hazard model)}$$

Here, A represents ambient air pollution, M represents incident cardiovascular event, C represents the set of covariates described above, and $\lambda_0(t)$ represents the baseline dementia hazard at age t for an individual exposed to the lowest unit of chronic ambient air pollution, conditioning on the mediator and set of covariates. The controlled direct effects are then estimated by using $\exp(\theta_1)$ and λ_1 for Cox proportional and Aalen additive hazard models, respectively.

We then fit a second multi-level, mixed-effects model, with age as the time scale, as our mediator model to estimate CVD risk:

$$\lambda(t|A,C) = \lambda_0(t)\exp(\beta_1 A + \beta_2 C) \quad \text{(Cox proportional hazard model)}$$

$$\lambda(t|A,,C) = \lambda_0(t) + \lambda'_1 A + \lambda'_2 C \quad \text{(Aalen additive hazard model)}$$

Here, A represents ambient air pollution, C represents the set of covariates described above, and $\lambda_0(t)$ represents the baseline CVD hazard at age t for an individual exposed to the lowest unit of chronic ambient air pollution conditioning on covariates. The natural indirect effect is estimated to be the exponential product of β_1 and θ_2 for the Cox model and the product of λ'_1 and λ_1 for the Aalen model to simultaneously account for the effect of air pollution on CVD risk and the effect of CVD risk on dementia.[54,106]

We calculated the estimated proportion of the total effect of ambient air pollution on dementia mediated through CVD with the following expressions: $\frac{NDE*(NIE-1)}{NDE*NIE-1}$ for Cox proportional hazard models and $\frac{NIE}{\text{NDE+NIE}}$ for Aalen additive models.[54,106]

Sensitivity analysis, methods

First, we considered minimally adjusted models and included age, sex, education, and income as covariates. Next, we considered additional comorbidities (further adjusting for diabetes, hypertension, and traumatic brain injury) and physician density in our outcome model to address potential mediator-outcome confounding. Finally, we estimated natural direct and indirect effects accounting for potential exposure-mediator interaction in Cox proportional hazard models using the following expressions:

$$NDE^{PH} = \frac{\exp(\gamma_1 a)\{1+\exp(\gamma_2 + y_3 a + \beta_0 + \beta_1 a^* + \beta_2' c)\}}{\exp(\gamma_1 a^*)\{1+\exp(\gamma_2 + y_3 a^* + \beta_0 + \beta_1 a^* + \beta_2' c)\}}$$

$$NIE^{PH} = \frac{\{1+\exp(\beta_0 + \beta_1 a^* + \beta_2' c)\}\{1+\exp(\gamma_2 + y_3 a + \beta_0 + \beta_1 a + \beta_2' c)\}}{\{1+\exp(\beta_0 + \beta_1 a + \beta_2' c)\}\{1+\exp(\gamma_2 + y_3 a + \beta_0 + \beta_1 a^* + \beta_2' c)\}}$$

These expressions are from extensions of the previously described Cox proportional hazard outcome and mediator models and include an exposure-mediator interaction term, λ_3.[127] To

simplify these expressions, we assumed $a - a^* = 1$, $\beta_0 = 0$, and $c = 0$ for binary variables and standardized continuous variables.

Sensitivity Analyses, results

We found no appreciable difference between minimally adjusted and fully adjusted total effect models (Table 3.5). Including comorbidities and physician density as additional covariates in total effect models had no effect or slightly attenuated effect sizes (Table 3.6, Table 3.7). Finally, accounting for potential exposure-mediator interaction attenuated the estimated natural indirect effect and proportion mediated through CVD for both pollutants (Table 3.8).

Supplementary Tables

Table 3.5: Minimally adjusted total effect between air pollutant and dementia

Pollutant	HR[a] (95% CI)	Estimate[a] (SE)
NO_2[b]	1.09 (1.04-1.13)	$61.3x10^{-5}$ ($13x10^{-5}$)
$PM_{2.5}$[b]	1.28 (1.07-1.53)	$200x10^{-5}$ ($<100x10^{-5}$)

HR=hazard ratio; CI=confidence interval; SE=standard error
[a] Adjusted for age, sex, education, income
[b] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

Table 3.6: Total effect of ambient air pollutant with further adjustment for comorbidities

Pollutant	Cox PH model Total effect HR[a,b] (95% CI)	Aalen model Total effect Estimate[a,b] (SE)
NO_2[c]	1.08 (1.00 – 1.16)	$73.1x10^{-5}$ ($22.8x10^{-5}$)
$PM_{2.5}$[c]	1.15 (0.93 – 1.42)	$200x10^{-5}$ ($100x10^{-5}$)

HR=hazard ratio; CI=confidence interval; SE=standard error
[a] Adjusted for age, sex, education, income, marital status, smoking status, weight status, physical activity, rural residence, northern region, diabetes, hypertension, traumatic brain injury; area level: percent recent immigrants, unemployed, less than high school education
[b] Total effect from a total effect model (not product method)
[c] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

Table 3.7: Total effect of ambient air pollutant with further adjustment for physician density

Pollutant	Cox PH model Total effect HR[a,b] (95% CI)	Aalen model Total effect Estimate[a,b] (SE)
NO_2[c]	1.08 (1.00-1.16)	$70.1x10^{-5}$ ($23.9x10^{-5}$)
$PM_{2.5}$[c]	1.15 (0.93-1.42)	$200x10^{-5}$ ($100x10^{-5}$)

HR=hazard ratio; CI=confidence interval; SE=standard error
[a] Adjusted for age, sex, education, income, marital status, smoking status, weight status, physical activity, rural residence, northern region, physician density; area level: percent recent immigrants, unemployed, less than high school education
[b] Total effect from a total effect model (not product method)
[c] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

Table 3.8: Natural direct and indirect effects of ambient air pollutant through cardiovascular disease accounting for exposure-mediator interaction (Cox proportional hazards model)

Pollutant	Natural direct effect HR[a] (95% CI)	Natural indirect effect HR[a] (95% CI)	Proportion mediated
NO_2[b]	1.14 (1.03 – 1.24)	1.00 (1.00 – 1.01)	1%
$PM_{2.5}$[b]	1.39 (1.01 – 1.89)	1.01 (1.00 – 1.03)	5%

HR=hazard ratio; CI=confidence interval

[a] Adjusted for age, sex, education, marital status, income quintile, smoking status, body-mass index, physical activity, rural residence, and northern region; area level: recent immigrants, unemployment, and education

[b] NO_2 per 5 ppb, $PM_{2.5}$ per 10 ug/m^3

Chapter 3, in full, is a reprint of the material as it appears in the International Journal of Epidemiology 2020. Ilango Sindana D; Chen, Hong; Hystad, Perry; van Donkelaar, Aaron; Kwong, Jeffrey C, Tu, Karen; Martin, Randall V; Benmarhnia Tarik. The dissertation author was the primary investigator and author of this paper.

4. Long-term exposure to ambient air pollution and cognitive function among Hispanics/Latinos in San Diego, California

Sindana D. Ilango, Kevin Gonzalez, Linda Gallo, Matthew A. Allison, Jianwen Cai, Carmen R. Isasi, H. Dean Hosgood, Priscilla M Vasquez, Donglin Zeng, Marion Mortamais, Hector Gonzalez, Tarik Benmarhnia

4.1. Abstract

Background: Hispanic/Latinos in the US are more likely to live in neighborhoods with greater exposure to air pollution and are projected to have the largest increase in dementia among race/ethnic minority groups.

Objective: We examined the associations of air pollution with performance on cognitive function tests in Hispanics/Latinos.

Methods: We used data from the San Diego site of the Hispanic Community Health Study/Study of Latinos, an ongoing cohort of Hispanics/Latinos. This analysis focused on individuals ≥45 years of age who completed a neurocognitive battery examining overall mental status, verbal learning, memory, verbal fluency, and executive function (n=2,089). Air pollution ($PM_{2.5}$ and O_3) before study baseline was assigned to participants' zip code. Logistic and linear regression were used to estimate the association of air pollution on overall mental status and domain-specific standardized test scores. Models accounted for complex survey design, demographic, and socioeconomic characteristics.

Results: We found that for every 10-ug/m^3 increase in $PM_{2.5}$, verbal fluency worsened (β: -0.21 [95% CI: -0.68, 0.25]). For every 10-ppb increase in O_3, verbal fluency and executive function

worsened (β: -0.19 [95% CI: -0.34, -0.03]; β: -0.01 [95% CI: -0.01, 0.09], respectively). We did not identify any detrimental effect of pollutants on other domains.

Conclusion: Although we found suggestions that air pollution may impact verbal learning and executive function, we observed no consistent or precise evidence to suggest an adverse impact of air pollution on cognitive level among this cohort of Hispanics/Latinos.

4.2. Introduction

As average life-expectancy increases in the US and worldwide, there is a heightened public health concern about impaired cognitive function with advancing age. Previous literature indicates that race/ethnic minorities are at higher risk for age-related cognitive dysfunction compared to non-Hispanic Whites.[9,128] The Hispanic/Latino population has been projected to have the largest increase in Alzheimer's Disease and related dementias (ADRD) among other race/ethnic minority groups over the next four decades.[128] Cognitive impairment frequently precedes dementia. Thus, improved understanding of cognitive impairment and its determinants can reveal important insights into dementia prevention. Furthermore, understanding these determinants in Hispanic/Latino populations can inform targeted intervention strategies. In this study, we evaluate air pollution as a potential determinant of cognitive function.

Identifying ubiquitous and modifiable risk factors, such as ambient air pollution, is of key interest because they are highly prevalent, affect all populations, and can have multiple benefits if effectively intervened upon.[129,130] Ambient air pollution is a mixture of particulate matter and gaseous pollutants. Emerging evidence suggests that increased exposure to air pollution is associated with cognitive impairment and dementia among older adults.[15,16] Air pollutants including fine particulate matter ($PM_{2.5}$) and ozone (O_3) can impact the brain through both direct and indirect pathways.[24] First, pollutants can directly reach the brain through the nasal pathway or

through systemic circulation by crossing the blood brain barrier and trigger neuroinflammation.[22,24,28] Secondly, pollutants can produce inflammation in other organs and tissues (e.g., cardiovascular systems) that can indirectly affect the central nervous system.[22,24,29]

Unlike previous studies that have been conducted in predominantly white populations,[55–59] this study investigates the association between air pollution and cognitive function in a Hispanic/Latino population. Examining this relationship in diverse race/ethnic populations, specifically in Hispanics/Latinos, is crucial given the expected relative increase in ADRD prevalence over the next 40 years.[128] Furthermore, there is a strong race/ethnic disparity in neighborhood environments.[131] In California, Hispanics/Latinos and Blacks are more likely to live in socioeconomically segregated communities that are disproportionally exposed to higher levels of ambient air pollution than White, Asian/Pacific Islander, and Native American populations.[132] These disparities highlight the need to examine the effect of air pollution on cognitive function among race/ethnic groups to inform the need to promote more equitable access to healthier neighborhoods. Currently, there is limited information available on the effect of air pollution on cognitive function among Hispanics/Latinos, the largest ethnic group in California.[58,59,132] Thus, we aimed to examine associations between long-term exposure to air pollution and cognitive function in the San Diego metropolitan area using data from a representative cohort of community-dwelling Hispanic/Latino adults.

4.3. Methods

Description of Study Participants

This study used data from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL), an ongoing, prospective cohort study of 16,415 community-dwelling Hispanic/Latino adults from four U.S. sites: Bronx, New York; Chicago, Illinois; Miami-Dade,

Florida; and San Diego, California. Details of this study have been previously published.[133] Briefly, adult participants aged 18-74 years were recruited for the study, with a two-stage probability sampling of households. Participants attended a clinic visit in 2008-2011 which included a clinical examination and questionnaire about demographic background and health conditions. Middle- and older-aged adults aged 45 years and older further underwent neurocognitive testing using a battery of five tests (see details below). The present study focuses on individuals who completed neurocognitive testing and reside in the San Diego field center catchment area. This analysis includes middle-aged adults to capture any early indication of dementia. The study population is predominantly Mexican heritage (immigrants, second and third generations). Out of 4,086 participants in the San Diego site, n=1,658 were younger than 45 and were excluded from the analyses. An additional n=339 participants were excluded for missing data. The final study included 2,089 participants. Differences in age, sex, Mexican heritage, and education level between those included and excluded in the study are presented in Table 4.3. This study received approval from the Institutional Review Board of the participating study site.

Cognitive Function

The primary outcome for this study was cognitive score, measured by performance on a neurocognitive battery of five tests at a single time point. Examinations were administered at the clinic visit by study staff trained and supervised by doctorate-level, licensed, clinical psychologists. The battery examined domains associated with aging and included a screener for overall mental status and tests of episodic verbal learning, memory, verbal fluency, and executive function.

Detailed information about the test battery and their application within HCHS/SOL has been previously published.[134] Briefly, the test included the Six-Item-Screener (SIS), Brief-Spanish

English Verbal Learning Test (B-SEVLT Sum and Recall), Word Fluency (WF), and Digit Symbol Substitution test (DSST). The SIS is a mental status test derived from the Mini-Mental State Examination that includes three-item recall and three-item temporal orientation probes.[135,136] Scores range from 1 to 6, with higher scores indicating better performance. The measure of overall mental status was dichotomized in line with previous work in this cohort, with normal mental status defined as SIS > 4.[134] The B-SEVLT is an episodic learning and memory test. Participants were presented with and asked to recall a list of 15 common words immediately (B-SEVLT Sum) and after an interference task and short delay (B-SEVLT Recall). The total score for each is the number of correctly recalled words across three trials.[137] WF is a phonemic verbal fluency test. Participants were asked to recite as many words beginning with the letters "F" and "A" within 60 seconds. The total score is the number of correctly generated words during this time span.[134] The DSST is a mental processing speed and executive function measure. Participants were asked to encode symbols to numbers. The total score is the number of correctly coded symbols.[138] Domain-specific test scores were transformed into z-scores using the means and standard deviations of the measures in the study population, to allow comparability of results across the different tests.

Air Pollution Measurement

The air pollutants $PM_{2.5}$ (ug/m^3) and O_3 (ppb) were the primary exposures for this study, specifically, a four-year average of daily estimates at the zip-code level. Neighborhood exposure to $PM_{2.5}$ and O_3 was estimated using 24-hour daily means and 8-hour daily maximums, respectively, sampled and analyzed by the US Environmental Protection Agency Air Quality System. Measured concentrations within a 20 km radius of each population-weighted centroid were used for interpolation.[139] Values were estimated using an inverse distance weighting each point of interest; this gives greater importance to values reported by monitoring stations closer to

the point of interest than monitoring stations farther away in distance.[45,140] We assumed participants resided in the same zip code and estimated long-term exposure with a running average of daily estimates the four years before study entry. ArcMap10.3 was used to generate air pollution estimates.

Covariates

Demographic and socioeconomic characteristics were obtained from questionnaires administered at the clinic visit. The following covariates were considered in this analysis: Hispanic/Latino heritage (Mexican vs. not), age, sex (male or female), educational attainment (less than 12 years, 12 years or equivalent, or greater than 12 years) and household income (less than $20k, $20k-$40k, or more than $40k).

Statistical Analysis

The study population was described with means (SD) and frequencies (%) of demographic characteristics. All analyses were conducted for each pollutant, evaluating exposures as both continuous (per 10 ug/m^3 and 10 ppb for $PM_{2.5}$ and O_3, respectively) and categorical variables (cut points at the 5^{th}, 25^{th}, $50^{th,}$ 75^{th}, and 95^{th} percentiles) to explore potential non-linear dose-responses. The primary outcome of interest was cognitive score measured at a single timepoint. Since scores were transformed, results reflect a change in outcome relative to a standard deviation increase. We used logistic regression models for survey data to estimate the effect of air pollution on normal mental status function (SIS > 4). Linear regression models for survey data were generated to estimate the effect of air pollution exposure on performance on each domain-specific test. For all tests, high scores reflect better performance. Models were adjusted for Hispanic/Latino heritage, age, sex, education, and income; this is a minimal set of variables, decided a priori, that may confound the relationship between air pollution and cognitive function. In sensitivity analyses, we

stratified the domain-specific models by age group (<55 years and ≥ 55 years). All descriptive and regression models accounted for the complex sampling design of the HCSC/SOL, to improve generalizability to the target population. Survey weights were calibrated to 2010 US census characteristics by age, sex, and Hispanic/Latino heritage for each study site to account for nonresponse, oversampling of subpopulations, and spatial structure of participants. All analyses were executed using the survey function in Stata v. 16.

4.4. Results

The mean age of the target population was 55.3 (SD: 9.5) years. Approximately 55% of the population were female and 70% reported an annual household income of less than $40,000. The mean exposure concentration was 12.0 ug/m^3 (SD: 1.3, range: 10.2 to 13.3) and 50.1 ppb (SD: 4.1, range 39.4 to 52.6) for $PM_{2.5}$ and O_3, respectively.

Table 4.1: Characteristics of study population (unweighted n=2,089)

Characteristic	Unweighted N (%)
Age	
=< 54	1127 (52.72)
55-64	711 (30.88)
>=65	251 (16.60)
Sex	
Female	1381 (55.46)
Male	708 (44.54)
Hispanic/Latino heritage	
Mexican	1960 (93.84)
Not Mexican	129 (6.16)
Marital status	
Single	239 (10.49)
Married/Living with partner	1304 (63.27)
Separated/Divorced/Widower	546 (26.33)
Education	
Less than high school	887 (39.53)
High school or equivalent	420 (18.42)
Greater than high school	782 (42.05)
Income	
Less than $20k	893 (41.57)
>20k - <$40k	742 (31.00)
More than $40k	454 (27.43)

Approximately 90% of the study population were classified as normal cognitive function (n=1,859). On average, individuals had raw scores of SIS: 5.4 (SD: 0.03, range: 0 to 6), B-SEVLT Sum: 23.7 (SD: 0.2, range 3 to 40), B-SEVLT Recall: 8.9 (SD: 0.1, range: 0 to 15), WF: 20.8 (SD: 0.4, range: 1 to 45), and DSST: 39.3 (SD: 0.6, range: 0 to 81).

In regression analyses, increased exposure to $PM_{2.5}$ and O_3 was associated with lower odds of normal mental status (SIS > 4). For domain-specific measures, standardized performance on cognitive function exams was marginally worse for WF (β: -0.21 SD; 95% CI: -0.68, 0.25) but slightly better for B-SEVLT-Sum (β=0.89 SD; 95% CI: 0.42, 1.35), SEVLT-Recall (β=0.67 SD; 95% CI: 0.20, 1.13), and DSST (β=0.13 SD; 95% CI: -0.33, 0.60) for every 10 ug/m3 increase in $PM_{2.5}$. For every 10 ppb increase in O_3, we observed worse performance for WF (β= -0.19 SD;

95% CI: -0.34, -0.03) and DSST (β= -0.01 SD; 95% CI: -0.12, 0.09) and slightly better performance in standardized cognitive function scores for B-SEVLT Sum (β=0.12 SD; 95% CI: -0.01, 0.25), and SEVLT Recall (β=0.15 SD; 95% CI: 0.02, 0.29) (Table 4.2).

Table 4.2: Associations between air pollution and cognitive performance

Test	**$PM_{2.5}$**	**O_3**
	Overall Mental Status OR (95% CI)	
SIS>4	0.67 (0.05, 8.39)	0.69 (0.34, 1.44)
	Domain-Specific Standardized Scores β (95% CI)	
B-SEVLT-Sum	0.89 (0.42, 1.35)	0.12 (-0.01, 0.25)
B-SEVLT-Recall	0.67 (0.20, 1.13)	0.15 (0.02, 0.29)
WF	-0.21 (-0.68, 0.25)	-0.19 (-0.34, -0.03)
DSST	0.13 (-0.33, 0.60)	-0.01 (-0.12, 0.09)

*Models adjusted for age, sex, heritage, marital status, education, and income
*Abbreviations: $PM_{2.5}$ (per 10 ug/m3); O_3 (per 10 ppb); SIS: Six-Item-Screener; B-SEVLT: Brief Spanish English Verbal Learning Test; WF: Word Fluency; DSST: Digit Symbol Substitution Test. Higher scores indicate better performance for all tests.

For $PM_{2.5}$, we observed some indication of worsening performance on DSST for $PM_{2.5}$ concentrations greater than the 50^{th} percentile. We found no consistent pattern of worse cognitive performance with increasing exposure levels for either B-SEVLT test, WF, or overall mental status (Figure 4.1, Table 4.5). Similarly, we observed decreasing performance on DSST for higher O_3 exposure groups and no pattern of worse cognitive performance with higher exposure levels for B-SEVLT exams, WF, or overall mental status (Figure 4.1, Table 4.6).

In analyses stratified by age group, we found suggestions of a stronger effect of air pollution on worse performance on WF and DSST for $PM_{2.5}$, and WF for O_3 in the older age group (Table 4.7).

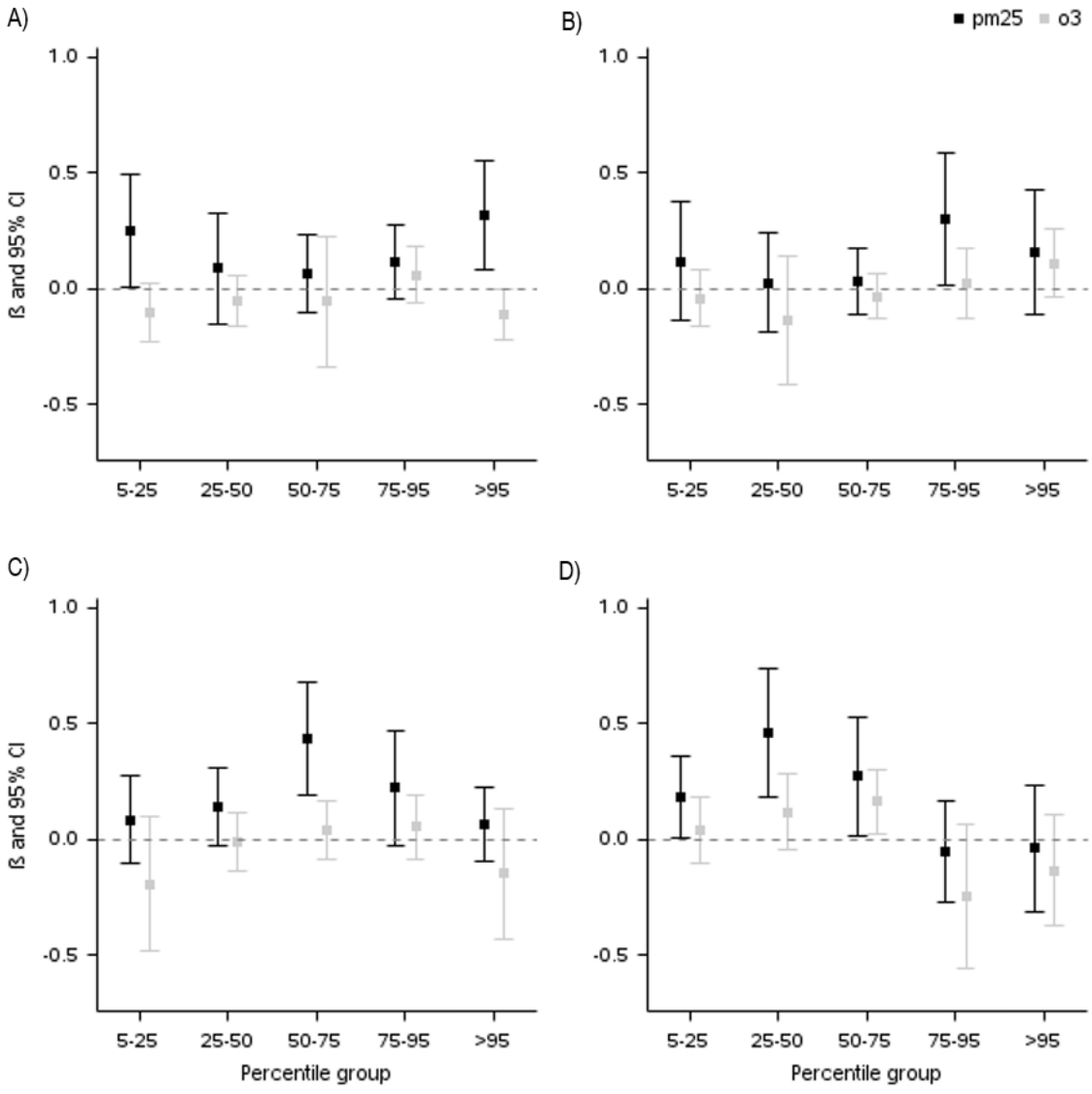

Figure 4.1: Domain-specific associations between standardized scores of cognitive function and air pollution percentile groups. Lowest exposure group (<5th percentile) is reference. A) Brief Spanish English Verbal Learning Test (B-SEVLT) - Sum; B) B-SEVLT-Recall; C) Word Fluency; D) Digit Symbol Substitution Test

4.5. Discussion

This study examined the effect of exposure to $PM_{2.5}$ and O_3 on performance in several domains of cognitive function in a cohort of middle- and older-aged Hispanic/Latino adults in San Diego, CA. We observed that increased exposure to air pollutants was marginally associated with lower mental status, poorer verbal fluency and worse executive function; however, these results

were not precise. We did not identify any detrimental influence of $PM_{2.5}$ and O_3 exposure on measures of learning or memory included in this study. Overall, we did not find consistent or robust evidence for a role air pollution exposure in worse cognitive function across the considered domains in this cohort of Hispanic/Latino adults.

In contrast with the present findings, previous studies have observed a link between particulate matter and worse verbal learning [58], abstraction [62], working memory [55], and orientation [55]. Previous studies have also found exposure to ozone may be associated with reduced performance in neurocognitive tests of short-term memory, attention, and perceptual function, and executive function.[57,58] Although it is biologically plausible that air pollution can produce a neuroinflammatory response resulting in structural brain changes[26,27], findings in the present study were imprecise as indicated by wide confidence intervals. This may be explained by the relatively young age of our cohort (mean age is 55.26 years), the selection of our participants, and the limited variability of air pollution exposure in our study population.

Potential explanations for our findings first include the study population which comprised middle- and older- aged adults to potentially capture the prodromal stages of cognitive decline that develops at earlier ages. This selection of middle-aged adults may attenuate the effect of air pollution on cognitive function because notable influences on cognitive function may not be observed until later in life. We found suggestions of a stronger effect among the older age group and we recommend future cohorts of Hispanics/Latinos that include a wider age range of older adults to replicate this work.

Second, it is possible that Hispanic/Latinos have higher cognitive resilience and are less vulnerable to the effects of air pollution on cognitive function. This resilience is observed in a recent study of air pollution and cognitive decline in New York, where less decline in global

cognitive function was observed among Hispanics compared to non-Hispanic White and Black participants, across $PM_{2.5}$ and other pollutants.[141]

Third, our study area focused on a relatively small geographic area within San Diego county. Despite applying survey weights to obtain a representative sample of the target population, we may not have had enough exposure heterogeneity to detect substantive effects. In this study population, exposure to fine particulate matter and ozone ranged from 10 to 13 ug/m^3 and 39 to 53 ppb, respectively. In contrast, a cohort study in Los Angeles, CA of the same pollutants had exposures ranging from approximately 6 to 29 ug/m^3 and 21 to 59 ppb.[58] We suggest future studies explore diverse geographical areas in order to capture various levels of air pollution exposure.

We acknowledge limitations in our study. Our participants were assigned individual exposure to pollutants based on residential zip codes reported at study baseline. These measurements do not account for residential movement and each zip code is assigned a uniform concentration. Furthermore, we assumed that exposure is fully experienced at the residential zip code. Exposure misclassification may also arise due to seasonal and traffic characteristics that are not accounted for in our air pollution estimates. While we acknowledge some misclassification due to residential movement varying time spent at home, and modeling method, we expect this misclassification to be unrelated to cognitive performance. Any bias due to this misclassification would underestimate estimates.

In addition, we examined cognitive function measured at one time-point. This is susceptible to unmeasured confounding due to socioeconomic and sociocultural characteristics that affect both neighborhood residence and performance on exam. We recommend longitudinal studies of cognitive decline to examine within-person changes in cognitive function to overcome this challenge.

In this study, we examined the effect of long-term exposure to air pollution on cognitive function in a well-characterized cohort of Hispanic/Latino adults in San Diego. We did not find consistent evidence of an association between increased exposure to $PM_{2.5}$ and O_3 and worse performance on cognitive function exams. Further work among ethnically diverse older study populations and with repeated measures of cognitive function are important to confirm or challenge these findings.

4.6. Appendix

Table 4.3: Comparison of age, sex, Hispanic/Latino, and education characteristics for those included (n=2,089) and excluded (n=339) from study due to missingness

Characteristic	Not Missing (%)	Missing (%)	Total
Age			
<54	52.52	46.24	51.61
55-64	30.88	33.29	31.23
65+	16.60	20.47	17.16
Sex			
Female	55.46	62.44	56.47
Male	44.54	37.56	43.53
Mexican Status			
Not Mexican	6.16	3.56	5.80
Mexican	93.84	96.44	94.20
Education			
Less than High School (HS)	39.53	42.76	39.97
HS or Equivalent	18.42	19.23	18.53
Greater than HS	42.05	38.01	41.50

Table 4.4: Unadjusted associations between air pollution and cognitive level

Test	$PM_{2.5}$	O_3
	Overall Mental Status OR (95% CI)	
SIS>4	0.73 (0.07, 7.30)	0.86 (0.42, 1.76)
	Domain-specific Standardized Scores β (95% CI)	
B-SEVLT-Sum	0.83 (0.34, 1.33)	0.16 (-0.02, 0.34)
B-SEVLT-Recall	0.61 (0.11, 1.10)	0.19 (-0.02, 0.40)
WF	-0.23 (-0.82, 0.35)	-0.15 (-0.35, 0.04)
DSST	0.18 (-0.52, 0.88)	0.11 (-0.15, 0.36)

*Abbreviations: PM25 (per 10 ug/m3); O3 (per 10 ppb); SIS: Six-Item-Screener; B-SEVLT: Brief Spanish English Verbal Learning Test; WF: Word Fluency; DSST: Digit Symbol Substitution Test. Higher scores indicate better performance for all tests.

Table 4.5: Associations between fine particulate matter percentile groups and cognitive level

Test	<5 [ref]	5-25	25-50	50-75	75-95	95+
			Overall Mental Status OR (95% CI)			
SIS > 4	--	2.01 (0.57, 7.10)	1.81 (0.61, 5.40)	0.95 (0.28, 3.22)	1.98 (0.58, 6.76)	2.85 (0.79, 10.21)
			Domain-Specific Standardized Scores β (95% CI)			
SEVLT-Sum	--	0.25 (0.01, 0.50)	0.32 (0.08, 0.55)	0.30 (0.02, 0.58)	0.44 (0.12, 0.68)	0.46 (0.18, 0.74)
SEVLT-Recall	--	0.09 (-0.15, 0.33)	0.12 (-0.13, 0.37)	0.16 (-0.11, 0.42)	0.22 (-0.02, 0.47)	0.27 (0.02, 0.53)
WF	--	0.07 (-0.10, 0.23)	0.03 (-0.19, 0.24)	0.09 (-0.10, 0.27)	0.07 (-0.09, 0.23)	-0.05 (-0.27, 0.17)
DSST	--	0.12 (-0.04, 0.28)	0.03 (-0.11, 0.18)	0.14 (-0.02, 0.31)	0.18 (0.01, 0.36)	-0.04 (-0.31, 0.23)

*Models adjusted for age, sex, heritage, marital status, education, and income
*Abbreviations: SIS: Six-Item-Screener; B-SEVLT: Brief Spanish English Verbal Learning Test; WF: Word Fluency; DSST: Digit Symbol Substitution Test. Higher scores indicate better performance for all tests.

Table 4.6: Associations between ozone percentile groups and cognitive level

Test	<5 [ref]	5-25	25-50	50-75	75-95	95+
			Overall Mental Status OR (95% CI)			
SIS > 4	--	0.13 (0.04, 0.49)	0.11 (0.03, 0.46)	0.13 (0.03, 0.47)	0.16 (0.04, 0.55)	0.24 (0.06, 1.06)
			Domain-Specific Standardized Scores β (95% CI)			
SEVLT-Sum	--	-0.10 (-0.23, 0.02)	-0.11 (-0.22, -0.001)	0.03 (-0.12, 0.18)	0.04 (-0.09, 0.17)	0.12 (-0.05, 0.29)
SEVLT-Recall	--	-0.05 (-0.16, 0.06)	-0.04 (-0.16, 0.08)	0.11 (-0.04, 0.26)	0.06 (-0.09, 0.20)	0.16 (0.03, 0.30)
WF	--	-0.05 (-0.34, 0.23)	-0.14 (-0.41, 0.14)	-0.19 (-0.48, 0.10)	-0.15 (-0.43, 0.13)	-0.25 (-0.55, 0.06)
DSST	--	0.06 (-0.06, 0.18)	-0.03 (-0.13, 0.06)	-0.01 (-0.13, 0.11)	0.04 (-0.10, 0.18)	-0.13 (-0.37, 0.11)

*Models adjusted for age, sex, heritage, marital status, education, and income
*Abbreviations: SIS: Six-Item-Screener; B-SEVLT: Brief Spanish English Verbal Learning Test; WF: Word Fluency; DSST: Digit Symbol Substitution Test. Higher scores indicate better performance for all tests.

Table 4.7: Associations between air pollution and cognitive performance, by age group

Test	$PM_{2.5}$	O_3
Ages 45 to 54 years (n=1127)		
Overall Mental Status OR (95% CI)		
SIS>4	1.63 (0.12, 23.29)	0.90 (0.43;1.97)
Domain-Specific Standardized Scores β (95% CI)		
B-SEVLT-Sum	1.26 (0.69, 1.83)	0.14 (-0.04, 0.32)
B-SEVLT-Recall	1.00 (0.44, 1.55)	0.12 (-0.05, 0.29)
WF	0.17 (-0.34, 0.68)	-0.02 (.0.22, 0.18)
DSST	0.25 (-0.25, 0.74)	-0.12 (-0.25, 0.01)
55 years and older (n=962)		
Overall Mental Status OR (95% CI)		
SIS>4	0.35 (0.01, 13.16)	0.71 (0.29, 1.68)
Domain-Specific Standardized Scores β (95% CI)		
B-SEVLT-Sum	0.35 (-0.28, 0.97)	0.08 (-0.07, 0.22)
B-SEVLT-Recall	0.15 (-0.50, 0.80)	0.15 (-0.05, 0.35)
WF	-0.68 (-1.37, -0.00)	-0.35 (-0.54, -0.16)
DSST	-0.07 (-0.93, 0.78)	0.07 (-0.09, 0.23)

*Models adjusted for age, sex, heritage, marital status, education, and income

*Abbreviations: $PM_{2.5}$ (per 10 ug/m3); O_3 (per 10 ppb); SIS: Six-Item-Screener; B-SEVLT: Brief Spanish English Verbal Learning Test; WF: Word Fluency; DSST: Digit Symbol Substitution Test. Higher scores indicate better performance for all tests.

Chapter 4, in full, has been submitted for publication of the material as it may appear in the Journal of Alzheimer's Disease. Ilango, Sindana D; Gonzalez, Kevin; Gallo, Linda; Allison, Matthew A; Cai, Jianwen; Isasi, Carmen R; Hosgood, Dean H; Vasquez, Priscilla M; Zeng, Donglin; Mortamais, Marion; Gonzalez, Hector M; Benmarhnia, Tarik. The dissertation author was the primary investigator and author of this paper.

5. Discussion

5.1. Summary of dissertation research

In the last decade, there has been a surge of epidemiologic studies investigating the effect of air pollution on dementia and dementia-related outcomes. The compelling evidence resulted in the inclusion of air pollution exposure in later life as a potentially modifiable risk factor in the 2020 Lancet Commission report on dementia.[142] In this report, an estimated 2.3 percent of dementia cases, worldwide, are attributable to living in urban areas with high air pollution exposure.[142] Cognitive impairment can be an early indicator of dementia, thus understanding the effect of air pollution on cognitive function may also improve our understanding of how air pollution impacts the brain. Although the evidence base on air pollution and dementia appears to be consistent and growing, there are some methodological concerns which are discussed in the literature but have not been adequately assessed. Addressing these concerns will improve the internal and external validity of findings which can help shape our understanding of disease etiology and effective intervention strategies.

The purpose of this dissertation was to examine the role of chronic exposure to air pollution on dementia and cognitive impairment. This research expands the current literature by addressing methodological challenges common in epidemiologic research of environmental exposures and aging-related outcomes. The first aim of this dissertation underscores the importance of considering competing events, an area that is typically not discussed or clearly accounted for in studies of air pollution and dementia. This study demonstrated multiple approaches by detailing the difference between the target inference, computation, and statistical estimand. This work expands on the previous reviews on competing events by applying a causal framework to explain

the differences between the compared approaches and providing specific recommendations to consider competing events in studies of air pollution and dementia.

The second aim of this dissertation examined the mediating role of cardiovascular disease in the causal pathway between air pollution and incident dementia. Several papers discussed cardiovascular disease as a potential intermediate to explain the association, but this had not been formally investigated. An adverse relationship between increased exposure to air pollution and incident dementia had been established in a population-based cohort in Ontario, Canada.[42] This paper expands on this work by applying a causal mediation analysis to decompose the total effect into its component natural direct and indirect effects through cardiovascular disease. We found that some of the effect was mediated by air pollution, with a greater proportion mediated for $PM_{2.5}$ (4% on the additive scale, 21% on the multiplicative scale) than NO_2 (2% on the additive scale, 9% on the multiplicative scale).

The final aim of this study addressed the lack of race/ethnic diversity and generalizability of the current research on air pollution and cognitive outcomes. In this aim, the relationship between air pollution and multiple domains of cognitive function was examined in a cohort of Hispanic/Latino adults in San Diego. Most of the previous research on air pollution, dementia, and dementia-related outcomes have been conducted in predominantly White populations. To expand the external validity of present findings it is important to study these same research questions in understudied, vulnerable populations. Furthermore, understanding the relationship in specific subpopulations can inform tailored intervention efforts. Hispanics/Latinos are a minority group that are projected to have the largest increase in Alzheimer's disease and related dementias by 2050.[128] Similarly, Hispanic/Latino adults are a minority group disproportionately exposed to

higher levels of air pollution.[132] In this study, we found limited evidence to suggest a relationship between air pollution and cognitive impairment among this cohort of Hispanic/Latino adults.

This dissertation advances the field of air pollution and dementia. It discusses a source of selection bias that is common but typically not well-addressed in studies of air pollution and dementia. The tutorial presented in Chapter 2 will increase the accessibility of methods to account for competing events and encourage stronger methodological discussions to strengthen the evidence base. Chapter 3 contains the first application of causal mediation to disentangle the relationship between air pollution and dementia. Causal mediation analyses are useful for two primary reasons. First, they are a tool that can be applied to investigate disease etiology through population studies. Second, they can inform prevention efforts by identifying intermediates or specific vulnerable populations where interventions would have effective reductions in disease. This work has motivated subsequent work answering similar research questions.[44] Finally, Chapter 4 is the first study of air pollution and cognitive function in a cohort of Hispanic/Latino adults, expanding the external generalizability of previous research on this topic.

5.2. The importance of understanding the relationship between air pollution and dementia

Understanding the effect of air pollution on dementia is an important area of research for several reasons. First, it improves our knowledge of dementia etiology. Biologic studies support a "neuroinflammation hypothesis", where innate immune cells and microglia are affected by air pollution-induced central nervous system disruptions, directly and indirectly impacting risk of neurodegenerative diseases in later life.[24] Identifying these pathways from a population-level lens can help narrow the research focus to understand dementia etiology.

Second, as there is currently no treatment to reverse the course of the disease, research focus has expanded to identify modifiable risk factors. Air pollution is a unique modifiable

exposure that can be modified by both individual behaviors and population-level policies and regulations. Mitigation strategies to reduce greenhouse emissions include economic policies (e.g., to incentivize fuel standards), physical policies (e.g., land-use policies), soft policies (e.g., advertising campaigns to promote lifestyle and behavioral changes), and knowledge policies (e.g., the support of research).[143] Implementation of such policies and regulations can have multiple co-benefits including improved physical and behavioral health outcomes and overall wellbeing.[143] Even if the magnitude of the observed effect sizes between air pollution and dementia is small, reducing air pollution by a small amount across a large population can have lasting reductions of dementia and improved population health.

Third, air pollution has differential impacts on population health. Health is closely interlinked with where an individual lives, as residential address determines access to health care, healthy food, green spaces, quality education, and air pollution levels. For example, Hispanics/Latinos are more likely to live in socioeconomically segregated communities that are disproportionally exposed to higher levels of ambient air pollution than White, Asian/Pacific Islander, and Native American populations.[132] Furthermore, Hispanics/Latinos are projected to have the largest increase in Alzheimer's Disease and related dementias (ADRD) among other minority groups over the next four decades.[128] These disparities highlight the need to examine the effect of air pollution on dementia and dementia-related outcomes among Hispanics/Latinos to promote equal access to healthy neighborhoods. Understanding the specific relationship in specific race/ethnic minority groups is valuable for identifying disproportionately vulnerable populations and for providing societal basis for health disparities.

5.3. Air pollution, climate change, and aging-related health outcomes

Research on air pollution and dementia can be extended to underline the importance of studying how climate change may impact the health of older adults. In the context of climate change, extreme weather conditions and events are becoming more frequent, more severe, and less predictable; all are linked to affecting air pollution levels and adverse health outcomes like heart disease, respiratory conditions, and premature mortality.[17,144] Older adults are less resilient to climate change because they are more likely than younger populations to have impaired physical function, comorbidities, compromised immune systems, and to be more socially isolated.[17] Expanding public health research to examine the effects of multiple aspects of climate change (e.g., ambient air pollution, extreme weather events) and aging-related health outcomes can benefit older adults and the greater population.

5.4. Recommendations for future work in studies of air pollution and dementia

This dissertation research is not without limitations. First, several assumptions about unmeasured confounding, causal identification, critical windows of exposure, and model specification are made throughout analyses. For the studies in this dissertation, the robustness of findings was tested with various sensitivity analyses. However, there are still gaps in this body of work where future research is recommended.

First, further work on disease etiology is recommended. In Chapter 3, we considered a broad definition of cardiovascular disease. This work can be expanded by investigating specific disease pathways by exploring multiple dependent mediators.[145,146] The application of causal mediation analysis in studies with available information on biomarkers and brain scans can immensely add to this field.

Second, selection bias is inherent in studies of older adults. Selection bias due to competing events was addressed in Chapter 2 of this dissertation. However, selection bias may also arise

from selective entry into the study, as individuals who are dementia free at baseline are likely a healthier group than those who have been excluded because of prevalent disease. Furthermore, there may be selective attrition from the study due to reasons apart from competing events. Additional weighting approaches, principal stratification, and multiple imputation can be applied to account for selection bias in studies of older adults.[147]

Next, misclassification of the exposure and outcome should be formally evaluated. Much of the reviewed literature and this dissertation research relied on residential addresses to determine exposure to air pollution. Residential movement and alternative methods to assign exposure is recommended to understand the effect of exposure misclassification. Air pollution is also considered at different critical windows (e.g., exposure before baseline, before diagnosis) which could be inferentially problematic given the secular trends in air pollution. For example, given the overall decline in $PM_{2.5}$ in North America[148], studies with rolling entry and exposure assigned before the baseline study visit may result in systematically less exposure for individuals who participated in the study during later periods. This can be accounted for by considering such secular trends with time-varying covariates and alternative exposure windows as sensitivity analyses. Research efforts identifying the effect of specific components of air pollutants are also recommended. For example, particulate matter is comprised of several components defined by size However, the composition is heterogenous across time and geographic regions and differentially impacts health.[149]

Similarly, outcome misclassification is largely dependent on socioeconomic status, access to health care, and race/ethnicity. The extent to which potential differential exposure or outcome misclassification affects results is poorly understood and can be formally evaluated with quantitative bias analyses. Studying cognitive function as an alternative outcome offers advantages

in understanding early phases of dementia, as cognitive impairment typically precedes dementia. As previously described, hundreds of assessments have been developed to measure cognitive function. Each test has its own distribution and range, and performance is often sensitive to socioeconomic and sociocultural factors as these assessments were originally developed for specific subpopulations.[110] Test performance is affected by these socioeconomic and sociocultural characteristics which in turn can influence place of residence, occupation, and ultimately, long term exposure to ambient air pollution. These characteristics are difficult to fully capture in standard adjustment of demographic variables (e.g., age, sex, race/ethnicity, education); there is likely unmeasured confounding when studying the relationship between air pollution and cognitive function measured at a single point in time. One method to deal with this unmeasured confounding is to study associations with cognitive decline, instead of cognitive level at a single measurement, by using repeated measurements of cognitive function as the outcome. Studies examining associations with within-person change in cognitive function over time will adjust for time-fixed confounders by design and future research of air pollution and cognitive decline is recommended.

Additionally, applying alternative study designs, such as quasi-experimental methods can help triangulate results and deal with residual confounding. For example, residential relocation can be considered as a natural experiment where air pollution exposure changes depending on where the individual relocated to. Dementia risk can then be compared for populations that moved to higher or lower levels of air pollution.

Further exploration of social disparities by expanding research to diverse cohorts and asking research questions about health inequity, access to health care, race/ethnic disparities, can improve the external validity of our findings and help us understand the specific relationships in particularly vulnerable populations. For example, causal mediation analyses can be extended to

decompose the total effect of race/ethnicity on dementia into its direct and indirect effect through air pollution to identify specific areas to intervene on.

Finally, the recent growth of epidemiologic research on air pollution and dementia research motivates a parallel, understudied area of research: the effect of climate change on older adults.[17,150] Research on other aspects of climate such as extreme heat, exposure to smoke from wildfires, noise, and greenspaces in relation to aging-related outcomes is recommended.

5.5. Concluding remarks

In conclusion, this dissertation offers a thorough epidemiologic examination of the effect of chronic exposure to air pollution on dementia and cognitive aging, by building upon the evidence base to account for competing events, identify causal pathways, expand the generalizability of previous work. This body of research advances the field of air pollution and cognitive outcomes by addressing some of the methodological issues common in epidemiologic studies of environmental exposures and aging-related outcomes and provides detailed recommendations to guide future research.

References

1. World Health Organization. *Dementia: A Public Health Priority*.; 2015. doi:ISBN 9789241564458.

2. Weuve J, Hebert LE, Scherr PA, Evans DA. Deaths in the United States among persons with Alzheimer's disease (2010-2050). *Alzheimer's Dement*. 2014;10(2):e40-e46. doi:10.1016/j.jalz.2014.01.004.

3. Hebert LE, Weuve J, Scherr PA, Evans DA. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology*. 2013;80(19):1778-1783. doi:10.1212/WNL.0b013e31828726f5.

4. The Alzheimer Society of Canada. *Prevalence and Monetary Costs of Dementia in Canada*.; 2016.

5. Anthony S, Pradier C, Chevrier R, Festraëts J, Tifratene K, Robert P. The French National Alzheimer database: a fast growing database for researchers and clinicians. *Dement Geriatr Cogn Disord*. 2014;38(5-6):271-280.

6. Ortiz F, Fitten LJ. Barriers to healthcare access for cognitively impaired older Hispanics. *Alzheimer Dis Assoc Disord*. 2000;14(3):141-150. doi:10.1097/00002093-200007000-00005.

7. Alzheimer's Association. 2018 Alzheimer's disease facts and figures includes a special report on the financial and personal benefits of early diagnosis. *Alzheimers Dement*. 2018;14(3):367-429. doi:10.1016/j.jalz.2016.03.001.

8. Williams CL, Tappen RM, Rosselli M, Keane F, Newlin K. Willingness to be screened and tested for cognitive impairment: Cross-cultural comparison. *Am J Alzheimers Dis Other Demen*. 2010;25(2):160-166. doi:10.1177/1533317509352333.

9. Mayeda ER, Glymour MM, Quesenberry CP, Whitmer RA. Inequalities in dementia incidence between six racial and ethnic groups over 14 years. *Alzheimer's Dement*. 2016;12(3):216-224.

10. Gurland BJ, Wilder DE, Lantigua R, et al. Rates of dementia in three ethnoracial groups. *Int J Geriatr Psychiatry*. 1999;14(6):481-493.

11. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005.

12. Burke SL, Naseh M, Rodriguez MJ, Burgess A, Loewenstein D. Dementia-related neuropsychological testing considerations in non-Hispanic White and Latino/Hispanic populations. *Psychol Neurosci*. 2019;12(2):144-168. doi:10.1037/pne0000163.

13. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. *Lancet*. 2017;390. doi:10.1016/S0140-6736(17)31363-6.

14. Kivimäki M, Singh-Manoux A. Prevention of dementia by targeting risk factors. *Lancet*.

2018;391(10130):1574-1575. doi:10.1016/S0140-6736(18)30578-6.

15. Power MC, Adar SD, Yanosky JD, Weuve J. Exposure to air pollution as a potential contributor to cognitive function, cognitive decline, brain imaging, and dementia: A systematic review of epidemiologic research. *Neurotoxicology*. 2016;56:235-253. doi:10.1016/j.neuro.2016.06.004.

16. Clifford A, Lang L, Chen R, Anstey KJ, Seaton A. Exposure to air pollution and cognitive functioning across the life course - A systematic literature review. *Environ Res*. 2016;147:383-398. doi:10.1016/j.envres.2016.01.018.

17. Gamble JL, Hurley BJ, Schultz PA, Jaglom WS, Krishnan N, Harris M. Climate change and older Americans: State of the science. *Environ Health Perspect*. 2013;121(1):15-22. doi:10.1289/ehp.1205223.

18. Wang L, Green FHY, Smiley-Jewell SM, Pinkerton KE. Susceptibility of the aging lung to environmental injury. In: *Seminars in Respiratory and Critical Care Medicine*. Vol 31. © Thieme Medical Publishers; 2010:539-553.

19. Bowe B, Xie Y, Yan Y, Al-Aly Z. Burden of Cause-Specific Mortality Associated With PM2.5 Air Pollution in the United States. *JAMA Netw open*. 2019;2(11):e1915834. doi:10.1001/jamanetworkopen.2019.15834.

20. Cohen AJ, Brauer M, Burnett R, et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. *Lancet*. 2017;389(10082):1907-1918. doi:10.1016/S0140-6736(17)30505-6.

21. Kampa M, Castanas E. Human health effects of air pollution. *Environ Pollut*. 2008;151(2):362-367. doi:10.1016/j.envpol.2007.06.012.

22. Block ML, Calderón-Garcidueñas L. Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends Neurosci*. 2009;32(9):506-516. doi:10.1016/j.tins.2009.05.009.

23. Li H, Xin X. Nitrogen dioxide (NO2) pollution as a potential risk factor for developing vascular dementia and its synaptic mechanisms. *Chemosphere*. 2013;92(1):52-58.

24. Jayaraj RL, Rodriguez EA, Wang Y, Block ML. Outdoor ambient air pollution and neurodegenerative diseases: the neuroinflammation hypothesis. *Curr Environ Heal reports*. 2017;4(2):166-179.

25. Maher BA, Ahmed IAM, Karloukovski V, et al. Magnetite pollution nanoparticles in the human brain. *Proc Natl Acad Sci*. 2016;113(39):10797-10801.

26. Chen J, Wang X, Wellenius GA, et al. Ambient air pollution and neurotoxicity on brain structure: evidence from women's health initiative memory study. *Ann Neurol*. 2015;78(3):466-476.

27. Wilker EH, Preis SR, Beiser AS, et al. Long-term exposure to fine particulate matter, residential proximity to major roads and measures of brain structure. *Stroke*. 2015;46(5):1161-1166. doi:10.1161/STROKEAHA.114.008348.

28. Heusinkveld HJ, Wahle T, Campbell A, et al. Neurodegenerative and neurological disorders by small inhaled particles. *Neurotoxicology*. 2016;56:94-106.

29. Gorelick PB, Scuteri A, Black SE, et al. Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2011;42(9):2672-2713.

30. Kinney PL. Interactions of Climate Change, Air Pollution, and Human Health. *Curr Environ Heal Reports*. 2018:179-186. doi:10.1007/s40572-018-0188-x.

31. Ford B, Val Martin M, Zelasky SE, et al. Future Fire Impacts on Smoke Concentrations, Visibility, and Health in the Contiguous United States. *GeoHealth*. 2018;2(8):229-247. doi:10.1029/2018gh000144.

32. Silva RA, West JJ, Lamarque JF, et al. Future global mortality from changes in air pollution attributable to climate change. *Nat Clim Chang*. 2017;7(9):647-651. doi:10.1038/nclimate3354.

33. West J, Zhang Y, Smith S, et al. Cobenefits of global and domestic greenhouse gas emissions for air quality and human health. *Lancet*. 2017;389:S23.

34. Cacciottolo M, Wang X, Driscoll I, et al. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. *Transl Psychiatry*. 2017;7(1):e1022-8. doi:10.1038/tp.2016.280.

35. Carey IM, Anderson HR, Atkinson RW, et al. Are noise and air pollution related to the incidence of dementia? A cohort study in London, England. *BMJ Open*. 2018;8(9):1-11. doi:10.1136/bmjopen-2018-022404.

36. Li CY, Li CH, Martini S, Hou WH. Association between air pollution and risk of vascular dementia: A multipollutant analysis in Taiwan. *Environ Int*. 2019;133(November):105233. doi:10.1016/j.envint.2019.105233.

37. Oudin A, Forsberg B, Adolfsson AN, et al. Traffic-related air pollution and dementia incidence in Northern Sweden: A longitudinal study. *Environ Health Perspect*. 2016;124(3):306-312. doi:10.1289/ehp.1408322.

38. Smargiassi A, Sidi EAL, Robert L-E, et al. Exposure to ambient air pollutants and the onset of dementia in Québec, Canada. *Environ Res*. 2020:109870.

39. Wu Y-C, Lin Y-C, Yu H-L, et al. Association between air pollutants and dementia risk in the elderly. *Alzheimer's Dement Diagnosis, Assess Dis Monit*. 2015;1(2):220-228. doi:10.1016/j.dadm.2014.11.015.

40. Cerza F, Renzi M, Gariazzo C, et al. Long-term exposure to air pollution and hospitalization for dementia in the Rome longitudinal study. *Environ Heal*. 2019;18(1):72.

41. Chang KH, Chang MY, Muo CH, Wu TN, Chen CY, Kao CH. Increased risk of dementia in patients exposed to nitrogen dioxide and carbon monoxide: A population-based retrospective cohort study. *PLoS One*. 2014;9(8). doi:10.1371/journal.pone.0103078.

42. Chen H, Kwong JC, Copes R, et al. Exposure to ambient air pollution and the incidence of

dementia: a population-based cohort study. *Environ Int*. 2017;108:271-277.

43. Chen H, Kwong JC, Copes R, et al. Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study. *Lancet*. 2017;389(10070):718-726.

44. Grande G, Ljungman PLS, Eneroth K, Bellander T, Rizzuto D. Association between Cardiovascular Disease and Long-term Exposure to Air Pollution with the Risk of Dementia. *JAMA Neurol*. 2020:1-9. doi:10.1001/jamaneurol.2019.4914.

45. Jung CR, Lin YT, Hwang BF. Ozone, particulate matter, and newly diagnosed Alzheimer's disease: A population-based cohort study in Taiwan. *J Alzheimer's Dis*. 2015;44(2):573-584. doi:10.3233/JAD-140855.

46. Kioumourtzoglou MA, Schwartz JD, Weisskopf MG, et al. Long-term PM2.5exposure and neurological hospital admissions in the northeastern United States. *Environ Health Perspect*. 2016;124(1):23-29. doi:10.1289/ehp.1408973.

47. Lee M, Schwartz J, Wang Y, Dominici F, Zanobetti A. Long-term effect of fine particulate matter on hospitalization with dementia. *Environ Pollut*. 2019;254:112926. doi:10.1016/j.envpol.2019.07.094.

48. Seaton A, Tran L, Chen R, Maynard RL, Whalley LJ. Pollution, Particles, and Dementia: A Hypothetical Causative Pathway. *Int J Environ Res Public Health*. 2020;17(3):862.

49. Oudin A. Short review: Air pollution, noise and lack of greenness as risk factors for Alzheimer's disease-epidemiologic and experimental evidence. *Neurochem Int*. 2020;134:104646.

50. Thompson CA, Zhang ZF, Arah OA. Competing risk bias to explain the inverse relationship between smoking and malignant melanoma. *Eur J Epidemiol*. 2013;28(7):557-567. doi:10.1007/s10654-013-9812-0.

51. Weuve J, Tchetgen Tchetgen EJ, Glymour MM, et al. Accounting for Bias Due to Selective Attrition. *Epidemiology*. 2012;23(1):119-128. doi:10.1097/EDE.0b013e318230e861.

52. Lau B, Cole SR, Gange SJ. Competing Risk Regression Models for Epidemiologic Data. *Am J Epidemiol*. 2009;170(2):244-256. doi:10.1093/aje/kwp107.

53. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA, Tchetgen EJT, Hernán MA. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med*. 2020;39(8):1199-1236. doi:10.1002/sim.8471.

54. VanderWeele TJ. Mediation analysis: a practitioner's guide. *Annu Rev Public Health*. 2016;37:17-32. doi:10.1146/annurev-publhealth-032315-021402.

55. Ailshire JA, Clarke P. Fine particulate matter air pollution and cognitive function among U.S. older adults. *Journals Gerontol - Ser B Psychol Sci Soc Sci*. 2015;70(2):322-328. doi:10.1093/geronb/gbu064.

56. Loop MS, Kent ST, Al-Hamdan MZ, et al. Fine Particulate Matter and Incident Cognitive Impairment in the REasons for Geographic and Racial Differences in Stroke (REGARDS)

Cohort. *PLoS One*. 2013;8(9):1-9. doi:10.1371/journal.pone.0075001.

57. Chen J-C, Schwartz J. Neurobehavioral effects of ambient air pollution on cognitive performance in US adults. *Neurotoxicology*. 2009;30(2):231-239. doi:10.1016/j.neuro.2008.12.011.

58. Gatto NM, Henderson VW, Hodis HN, et al. Components of air pollution and cognitive function in middle-aged and older adults in Los Angeles. *Neurotoxicology*. 2014;40(3):1-7. doi:10.1016/j.neuro.2013.09.004.

59. Ailshire JA, Crimmins EM. Fine particulate matter air pollution and cognitive function among older US adults. *Am J Epidemiol*. 2014;180(4):359-366. doi:10.1093/aje/kwu155.

60. Ailshire J, Karraker A, Clarke P. Neighborhood social stressors, fine particulate matter air pollution, and cognitive function among older U.S. adults. *Soc Sci Med*. 2017;172(2):56-63. doi:10.1016/j.socscimed.2016.11.019.

61. Zeng Y, Gu D, Purser J, Hoenig H, Christakis N. Associations of environmental factors with elderly health and mortality in china. *Am J Public Health*. 2010;100(2):298-305. doi:10.2105/AJPH.2008.154971.

62. Schikowski T, Vossoughi M, Vierkötter A, et al. Association of air pollution with cognitive functions and its modification by APOE gene variants in elderly women. *Environ Res*. 2015;142:10-16. doi:10.1016/j.envres.2015.06.009.

63. Ranft U, Schikowski T, Sugiri D, Krutmann J, Krämer U. Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly. *Environ Res*. 2009;109(8):1004-1011. doi:10.1016/j.envres.2009.08.003.

64. Tzivian L, Dlugaj M, Winkler A, et al. Long-term air pollution and traffic noise exposures and mild cognitive impairment in older adults: a cross-sectional analysis of the Heinz Nixdorf recall study. *Environ Health Perspect*. 2016;124(9):1361.

65. Boogaard H, Walker K, Cohen AJ. Air pollution: The emergence of a major global health risk factor. *Int Health*. 2019;11(6):417-421. doi:10.1093/inthealth/ihz078.

66. Rudolph J, Lesko C, Naimi A. Causal Inference in the Face of Competing Events. *Curr Epidemiol Reports*. 2020. doi:10.1007/s40471-020-00240-7.

67. Brook RD, Rajagopalan S, Pope CA, et al. Particulate Matter Air Pollution and Cardiovascular Disease. *Circulation*. 2010;121(21):2331-2378. doi:10.1161/CIR.0b013e3181dbece1.

68. Harrison SL, Ding J, Tang EYH, et al. Cardiovascular Disease Risk Models and Longitudinal Changes in Cognition: A Systematic Review. Norris CM, ed. *PLoS One*. 2014;9(12):e114431. doi:10.1371/journal.pone.0114431.

69. Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012-1014. doi:10.1093/ije/dys223.

70. Díaz-venegas C, Downer B, Langa KM, Wong R. *Contextualizing Health and Aging in the Americas*. Springer International Publishing; 2018. doi:10.1007/978-3-030-00584-9.

71. Holland PW. Statistics and causal inference. *J Am Stat Assoc*. 1986;81(396):945-960.

72. Hernán MA, Robins JM. Causal Inference: What If (Harvard book). 2019. https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2019/11/ci_hernanrobins_10nov19.pdf.

73. Rehkopf DH, Glymour MM. Epidemiology : When a Rose is Not a Rose. *Curr Epidemiol Rep*. 2016;3(1):63-71. doi:doi:10.1007/s40471-016-0069-5.

74. Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, Hernán MA. A causal framework for classical statistical estimands in failure time settings with competing events. 2018:1-55. http://arxiv.org/abs/1806.06136.

75. Edwards JK, Hester LL, Gokhale M, Lesko CR. Methodologic Issues when Estimating Risks in Pharmacoepidemiology. *Curr Epidemiol Reports*. 2016;3(4):285-296. doi:10.1007/s40471-016-0089-1.

76. Schuster NA, Hoogendijk EO, Kok AAL, Twisk JWR, Heymans MW. Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. *J Clin Epidemiol*. 2020;122:42-48. doi:10.1016/j.jclinepi.2020.03.004.

77. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med*. 2017;36(27):4391-4400. doi:10.1002/sim.7501.

78. Weuve J. Are we ready to call exposure to air pollution a risk factor for dementia? *Neurology*. 2020;94(17):727-728. doi:10.1212/WNL.0000000000009318.

79. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. *Circulation*. 2016;133(6):601-609. doi:10.1161/CIRCULATIONAHA.115.017719.

80. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94(446):496-509. doi:10.1080/01621459.1999.10474144.

81. Lesko CR, Lau B. Bias Due to Confounders for the Exposure–Competing Risk Relationship. *Epidemiology*. 2017;28(1):20-27. doi:10.1097/EDE.0000000000000565.

82. Hernán MA, Robins JM. Causal inference: what if. *Boca Rat Chapman Hill/CRC*. 2020;2020.

83. Alpérovitch A, Amouyel P, Dartigues JF, et al. Vascular factors and risk of dementia: Design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology*. 2003;22(6):316-325. doi:10.1159/000072920.

84. Vienneau D, De Hoogh K, Bechle MJ, et al. Western european land use regression incorporating satellite- and ground-based measurements of NO2 and PM10. *Environ Sci Technol*. 2013;47(23):13555-13564. doi:10.1021/es403089q.

85. Menut L, Goussebaile A, Bessagnet B, Khvorostiyanov D, Ung A. Impact of realistic hourly emissions profiles on air pollutants concentrations modelled with CHIMERE. *Atmos Environ*. 2012;49(x):233-244. doi:10.1016/j.atmosenv.2011.11.057.

86. Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger T V. Performance of logistic regression modeling: Beyond the number of events per variable, the role of data structure. *J Clin Epidemiol*. 2011;64(12):1463-1464. doi:10.1016/j.jclinepi.2011.06.013.

87. Latouche A, Porcher R, Chevret S. Sample size formula for proportional hazards modelling of competing risks. *Stat Med*. 2004;23(21):3263-3274. doi:10.1002/sim.1915.

88. Austin PC, Allignol A, Fine JP. The number of primary events per variable affects estimation of the subdistribution hazard competing risks model. *J Clin Epidemiol*. 2017;83:75-84. doi:10.1016/j.jclinepi.2016.11.017.

89. Chaix B, Evans D, Merlo J, Suzuki E. Commentary: Weighing up the dead and missing: Reflections on inverse-probability weighting and principal stratification to address truncation by death. *Epidemiology*. 2012;23(1):129-131. doi:10.1097/EDE.0b013e3182319159.

90. Tchetgen EJT, Glymour MM, Shpitser I, Weuve J. Rejoinder: To weight or not to weight?: On the relation between inverse-probability weighting and principal stratification for truncation by death. *Epidemiology*. 2012;23(1):132-137. doi:10.1097/EDE.0b013e31823b5081.

91. Cortese G, Andersen PK. Competing risks and time-dependent covariates. *Biometrical J*. 2010;52(1):138-158.

92. Lau B, Cole SR, Gange SJ. Parametric mixture models to evaluate and summarize hazard ratios in the presence of competing risks with time-dependent hazards and delayed entry. *Stat Med*. 2011;30(6):654-665. doi:10.1002/sim.4123.

93. Ishwaran H, Gerds TA, Kogalur UB, Moore RD, Gange SJ, Lau BM. Random survival forests for competing risks. *Biostatistics*. 2014;15(4):757-773.

94. Prince M, Guerchet M, Prina M. *Policy Brief: The Global Impact of Dementia 2013–2050*.; 2013. https://www.alz.co.uk/research/GlobalImpactDementia2013.pdf.

95. Carey IM, Anderson R, Atkinson RW, et al. Are noise and air pollution related to the incidence of dementia? A cohort study in London, England. *BMJ Open*. 2018;8:22404. doi:10.1136/bmjopen-2018-022404.

96. Weuve J, Puett RC, Schwartz J, Yanosky JD, Laden F, Grodstein F. Exposure to particulate air pollution and cognitive decline in older women. *Arch Intern Med*. 2012;172(3):219-227. doi:10.1001/archinternmed.2011.683.

97. Cullen B, Newby D, Lee D, et al. Cross-sectional and longitudinal analyses of outdoor air pollution exposure and cognitive function in UK Biobank. *Sci Rep*. 2018;8(1):1-14. doi:10.1038/s41598-018-30568-6.

98. Weuve J. Invited commentary: how exposure to air pollution may shape dementia risk, and what epidemiology can say about it. *Am J Epidemiol*. 2014;180(4):367-371.

99. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ Heal*. 2013;12(1):43. doi:10.1186/1476-069X-12-43.

100. Martinelli N, Olivieri O, Girelli D. Air particulate matter and cardiovascular disease: A narrative review. *Eur J Intern Med*. 2013;24(4):295-302. doi:10.1016/j.ejim.2013.04.001.

101. Pope CA, Burnett RT, Thurston GD, et al. Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution. *Circulation*. 2004;109(1):71-77. doi:10.1161/01.CIR.0000108927.80044.7F.

102. Béland Y. Canadian Community. *Heal reports*. 2002;13(3).

103. Liisa Jaakkimainen R, Bronskill SE, Tierney MC, et al. Identification of physician-diagnosed Alzheimer's disease and related dementias in population-based administrative data: A validation study using family physicians' electronic medical records. *J Alzheimer's Dis*. 2016;54(1):337-349. doi:10.3233/JAD-160105.

104. Yeung DF, Boom NK, Guo H, Lee DS, Schultz SE, Tu J V. Trends in the incidence and outcomes of heart failure in Ontario, Canada: 1997 to 2007. *Can Med Assoc J*. 2012;184(14):E765-E773. doi:10.1503/cmaj.111958.

105. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; 2015.

106. Lange T, Hansen J V. Direct and indirect effects in a survival context. *Epidemiology*. 2011;22(4):575-581. doi:10.1097/EDE.0b013e31821c680c.

107. Aalen O. A Model for Nonparametric Regression Analysis of Counting Processes. In: Springer, New York, NY; 1980:1-25. doi:10.1007/978-1-4615-7397-5_1.

108. Thomas Scheike M. *Title Flexible Regression Models for Survival Data*.; 2019. https://github.com/scheike/timereg.git. Accessed May 24, 2019.

109. Therneau T. *Mixed Effects Cox Models*.; 2018. https://cran.r-project.org/web/packages/coxme/vignettes/coxme.pdf. Accessed May 23, 2019.

110. Weuve J, Proust-Lima C, Power MC, et al. Guidelines for reporting methodological challenges and evaluating potential bias in dementia research. *Alzheimers Dement*. 2015;11(9):1098-1109. doi:10.1016/j.jalz.2015.06.1885.

111. van Donkelaar A, Martin R V, Spurr RJD, et al. Optimal estimation for global ground-level fine particulate matter concentrations. *J Geophys Res Atmos*. 2013;118(11):5621-5636.

112. van Donkelaar A, Martin R V., Spurr RJD, Burnett RT. High-Resolution Satellite-Derived PM2.5 from Optimal Estimation and Geographically Weighted Regression over North America. *Environ Sci Technol*. 2015;49(17):10482-10491. doi:10.1021/acs.est.5b02076.

113. Chen H, Burnett RT, Copes R, et al. Ambient Fine Particulate Matter and Mortality among Survivors of Myocardial Infarction : Population-Based Cohort Study. 2016;124(9):1421-1428.

114. Chen H, Burnett RT, Kwong JC, et al. Risk of Incident Diabetes in Relation to Long-term Exposure to Fine Particulate Matter in Ontario, Canada. *Environ Health Perspect*. 2013;121(7):804-810. doi:10.1289/ehp.1205958.

115. Crouse DL, Peters PA, Hystad P, et al. Ambient PM2.5, O3, and NO2 Exposures and

Associations with Mortality over 16 Years of Follow-Up in the Canadian Census Health and Environment Cohort (CanCHEC). *Environ Health Perspect*. 2015;123(11):1180-1186. doi:10.1289/ehp.1409276.

116. Crouse DL, Peters PA, van Donkelaar A, et al. Risk of Nonaccidental and Cardiovascular Mortality in Relation to Long-term Exposure to Low Concentrations of Fine Particulate Matter: A Canadian National-Level Cohort Study. *Environ Health Perspect*. 2012;120(5):708-714. doi:10.1289/ehp.1104049.

117. Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(9859):2224-2260. doi:10.1016/S0140-6736(12)61766-8.

118. Environment Canada. National Air Pollution Surveillance Program (NAPS). https://www.canada.ca/en/environment-climate-change/services/air-pollution/monitoring-networks-data/national-air-pollution-program.html. Accessed May 30, 2019.

119. Hystad P, Setton E, Cervantes A, et al. Creating National Air Pollution Models for Population Exposure Assessment in Canada. *Environ Health Perspect*. 2011;119(8):1123-1129. doi:10.1289/ehp.1002976.

120. Lavigne E, Yasseen AS, Stieb DM, et al. Ambient air pollution and adverse birth outcomes: Differences by maternal comorbidities. *Environ Res*. 2016;148:457-466. doi:10.1016/j.envres.2016.04.026.

121. Cesaroni G, Porta D, Badaloni C, et al. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environ Heal*. 2012;11(1):48. doi:10.1186/1476-069X-11-48.

122. Eeftens M, Beelen R, Fischer P, Brunekreef B, Meliefste K, Hoek G. Stability of measured and modelled spatial contrasts in NO2 over time. *Occup Environ Med*. 2011;68(10):765-770. doi:10.1136/oem.2010.061135.

123. Wang R, Henderson SB, Sbihi H, Allen RW, Brauer M. Temporal stability of land use regression models for traffic-related air pollution. *Atmos Environ*. 2013;64:312-319. doi:10.1016/J.ATMOSENV.2012.09.056.

124. Hux JE, Ivis F, Flintoft V, Bica A. Diabetes in Ontario: determination of prevalence and incidence using a validated administrative data algorithm. *Diabetes Care*. 2002;25(3):512-516. doi:10.2337/diacare.25.3.512.

125. Schultz SE, Rothwell DM, Chen Z, Tu K. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. *Chronic Dis Inj Can*. 2013;33(3):160-166. http://www.ncbi.nlm.nih.gov/pubmed/23735455. Accessed May 30, 2019.

126. Tu K, Campbell NR, Chen Z-L, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med*. 2007;1(1):e18-26. http://www.ncbi.nlm.nih.gov/pubmed/20101286. Accessed May 30, 2019.

127. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; 2015.

128. Matthews KA, Xu W, Gaglioti AH, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥65 years. *Alzheimer's Dement*. 2019;15(1):17-24. doi:10.1016/j.jalz.2018.06.3063.

129. Smith KR, Woodward A, Campbell-Lendrum D, et al. *Human Health: Impacts, Adaptation, and Co-Benefits. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability*. Cambridge University Press; 2014.

130. Bell ML, Davis DL, Cifuentes LA, Krupnick AJ, Morgenstern RD, Thurston GD. Ancillary human health benefits of improved air quality resulting from climate change mitigation. *Environ Heal*. 2008;7(1):41.

131. Clark LP, Millet DB, Marshall JD. Changes in transportation-related air pollution exposures by race-ethnicity and socioeconomic status: Outdoor nitrogen dioxide in the United States in 2000 and 2010. *Environ Health Perspect*. 2017;125(9):1-10. doi:10.1289/EHP959.

132. Cushing L, Faust J, August LM, Cendak R, Wieland W, Alexeeff G. Racial/ethnic disparities in cumulative environmental health impacts in California: evidence from a statewide environmental justice screening tool (CalEnviroScreen 1.1). *Am J Public Health*. 2015;105(11):2341-2348.

133. Elfassy T, Aiello AE, Schneiderman N, et al. Relation of Diabetes to Cognitive Function in Hispanics/Latinos of Diverse Backgrounds in the United States. *J Aging Health*. 2018. doi:10.1177/0898264318759379.

134. González HM, Tarraf W, Gouskova N, et al. Neurocognitive function among middle-aged and older hispanic/latinos: Results from the hispanic community health study/study of latinos. *Arch Clin Neuropsychol*. 2015;30(1):68-77. doi:10.1093/arclin/acu066.

135. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-item screener to identify cognitive impairment among potential subjects for clinical research. *Med Care*. 2002;40(9):771-781. doi:10.1097/01.MLR.0000024610.33213.C8.

136. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.

137. González HM, Mungas DAN, Reed BR, Marshall S, Haan MN. A new verbal learning and memory test for English-and Spanish-speaking older people. *J Int Neuropsychol Soc*. 2001;7(5):544-555.

138. Wechsler D. *WAIS-R Manual: Wechsler Adult Intelligence Scale-Revised*. Psychological Corporation; 1981.

139. US Environmental Protection Agency. Air Quality System (AQS). https://www.epa.gov/aqs. Accessed June 12, 2019.

140. Loizeau M, Buteau S, Chaix B, McElroy S, Counil E, Benmarhnia T. Does the air pollution model influence the evidence of socio-economic disparities in exposure and susceptibility?

Environ Res. 2018;167:650-661.

141. Kulick ER, Elkind MS V, Boehme AK, et al. Long-term exposure to ambient air pollution , APOE- ε 4 status , and cognitive decline in a cohort of older adults in northern Manhattan. *Environ Int*. 2020;136(December 2019):105440. doi:10.1016/j.envint.2019.105440.

142. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 2020;396(10248):413-446. doi:10.1016/S0140-6736(20)30367-6.

143. Shaw C, Hales S, Howden-Chapman P, Edwards R. Health co-benefits of climate change mitigation policies in the transport sector. *Nat Clim Chang*. 2014;4(6):427-433. doi:10.1038/nclimate2247.

144. Åström DO, Bertil F, Joacim R. Heat wave impact on morbidity and mortality in the elderly population: a review of recent studies. *Maturitas*. 2011;69(2):99-105.

145. Cho SH, Huang YT. Mediation analysis with causally ordered mediators using Cox proportional hazards model. *Stat Med*. 2019;38(9):1566-1581. doi:10.1002/sim.8058.

146. Huang YT, Yang HI. Causal mediation analysis of survival outcome with multiple mediators. *Epidemiology*. 2017;28(3):370-378. doi:10.1097/EDE.0000000000000651.

147. Banack HR, Kaufman JS, Wactawski-Wende J, Troen BR, Stovitz SD. Investigating and Remediating Selection Bias in Geriatrics Research: The Selection Bias Toolkit. *J Am Geriatr Soc*. 2019;67(9):1970-1976. doi:10.1111/jgs.16022.

148. Meng J, Li C, Martin R V., Van Donkelaar A, Hystad P, Brauer M. Estimated Long-Term (1981-2016) Concentrations of Ambient Fine Particulate Matter across North America from Chemical Transport Modeling, Satellite Remote Sensing, and Ground-Based Measurements. *Environ Sci Technol*. 2019;53(9):5071-5079. doi:10.1021/acs.est.8b06875.

149. Li C, Martin R V., Van Donkelaar A, et al. Trends in Chemical Composition of Global and Regional Population-Weighted Fine Particulate Matter Estimated for 25 Years. *Environ Sci Technol*. 2017;51(19):11185-11195. doi:10.1021/acs.est.7b02530.

150. Ilango SD, Mcelroy S. Recommendations for epidemiologic studies of aging populations in a changing climate. *Int J Public Health*. 2020;9:9-10. doi:10.1007/s00038-020-01453-9.