The Association Between Long-term Exposure to Traffic-related Air Pollution and Cardiovascular Mortality in Ontario, Canada

Hong Chen

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University, Montreal

A dissertation submitted to McGill University in partial fulfilment of the requirements of the degree of Doctor of Philosophy in Epidemiology.

Abstract

Introduction: Cardiovascular disease is the leading cause of mortality worldwide. There is suggestive evidence that chronic exposure to traffic-related air pollution may increase the incidence and mortality from cardiovascular disease. However, in few studies has the health effects of traffic-related air pollution been examined at the relatively lower concentrations of pollution such as observed in Canadian cities. In addition, the few studies using land use regression models to assess exposures of traffic-related air pollution at fine geographic resolutions have had modest sample sizes, so statistical power was limited. The purpose of this dissertation was to further our understanding of the aetiology of traffic-related air pollution in relation to cardiovascular outcomes through improved estimates of exposure derived through the juxtaposition of accurate information on residential addresses with improved methods of estimates of spatial concentrations of traffic-related air pollution.

Methods: To achieve the research objectives of this dissertation, three studies were conducted. First, I conducted a systematic review of epidemiological evidence between 1950 and 2011 regarding the chronic health effects of ambient air pollution. Second, I developed three new methods of estimating historical exposure to traffic-related air pollution at fine geographic scales through extrapolating "current" land use regression models back in time. These three extrapolation methods entailed multiplying the predicted concentrations of NO_2, a marker of traffic-related air pollution, by the ratio of past estimates of concentrations from fixed-site monitors, such that they reflected the change in the spatial structure of NO_2 from measurements at fixed-site monitors. Third, I conducted a population-based cohort study to determine whether an association exists between traffic-related air pollution and cause-specific cardiovascular mortality among adults living in Ontario, Canada. This population-based cohort study used as a sampling frame the Canadian Federal income tax file and subjects living in Hamilton, Toronto, and Windsor were included in the analyses. I made use of estimates of exposure to traffic-related air pollution using land use regression models and the back extrapolation methods. I carried out adjusted Cox regression models included known individual risk factors and selected ecological covariables and I carried out sensitivity analyses using indirect methods to adjust for smoking that was not available on individuals in the cohort.

Results and Discussion: In the Ontario tax cohort study, I found that for each increase of 5 parts per billion (ppb) of NO_2, the rate ratios for mortality for all cardiovascular disease and for ischemic heart disease, using different models, varied between 1.04 (95%CI: 1.00-1.09) and 1.10 (95%CI: 1.00-1.21). I found no associations between traffic pollution and cerebrovascular mortality (excess rates of mortality of 0.95 (95%CI: 0.89-1.02)). These results support the hypothesis that long-term exposure to traffic-related air pollution increases the mortality of cardiovascular disease.

Résumé

Introduction: Les maladies cardiovasculaires représentent la principale cause de mortalité dans le monde. Il existe une preuve suggestive que l'exposition chronique à la pollution atmosphérique d'origine automobile puisse augmenter l'incidence et la mortalité attribuable aux maladies cardiovasculaires. Cependant, peu d'études ont été menées à ce jour pour évaluer les effets sanitaires de la pollution atmosphérique liée au trafic à des concentrations relativement faibles telles que celles mesurées dans les villes canadiennes. De plus, les quelques études qui se sont servi de modèles de régression de type *Land-use regression* pour évaluer l'exposition de la pollution atmosphérique liée au trafic à haute résolution spatiale ont été réalisées à partir d'échantillons de taille modeste, et sont conséquemment de puissance statistique limitée. Le but de cette thèse est d'approfondir notre compréhension du lien entre la pollution atmosphérique liée au trafic et les maladies cardiovasculaires, ce à l'aide de meilleures estimations de l'exposition environnementale obtenues par la juxtaposition d'information précise sur les adresses résidentielles avec de meilleures estimations des concentrations spatiales de pollution atmosphérique liée à la circulation.

Méthodes: Afin d'atteindre les objectifs de recherche de cette thèse, trois études ont été menées. Tout d'abord, un examen systématique des données épidémiologiques associées aux effets chroniques de la pollution de l'air sur la santé entre 1950 et 2011 a été effectué. Deuxièmement, trois nouvelles méthodes ont été développées pour estimer rétrospectivement l'exposition à la pollution atmosphérique liée au trafic à fine résolution spatiale par extrapolation des modèles de régression 'actuels' de type *Land use regression.* Ces trois méthodes d'extrapolation impliquaient la multiplication des concentrations prédites de NO_2, un marqueur de la pollution atmosphérique liée au trafic, par le ratio des estimations des concentrations passées aux sites d'échantillonnage fixes, de sorte qu'elles reflètent les variations spatiales des concentrations de NO_2. Dans un dernier temps, cette thèse présente une étude de cohorte menée afin de déterminer s'il existe une association entre la pollution atmosphérique liée au trafic et des causes spécifiques de mortalité cardiovasculaire chez les adultes vivant en Ontario, Canada. L'échantillon de la cohorte est tiré des bases de données fédérales d'impôt sur le revenu des particuliers. Nous avons utilisé les valeurs estimées de l'exposition à la pollution atmosphérique automobile générées par des modèles de régression

de type *Land-use regression* et des méthodes de rétro-extrapolation. Les modèles de Cox ajustés incluent des facteurs de risque individuels connus et des covariables environnementales, alors que des analyses de sensibilité ont été réalisées avec des méthodes indirectes afin d'ajuster pour le tabagisme.

Résultats et discussion: Peu d'études ont été menées pour explorer les associations entre la pollution atmosphérique liée au trafic dans les villes et la mortalité cardiovasculaire. Les résultats de cette thèse ont montré que, pour chaque augmentation de 5 parties par billiard (ppb) de NO_2, les estimations du risque accru de cardiopathies ischémiques et de mortalité pour l'ensemble des maladies cardiovasculaires variaient entre 1.04 (95%CI: 1.00-1.09) and 1.10 (95%CI: 1.00-1.21). Par ailleurs, nous n'avons trouvé aucune association statistiquement significative entre la pollution liée au trafic et la mortalité cérébrovasculaire (taux de mortalité en excès de 0.95 (95%CI: 0.89-1.02). Cette thèse soutient l'hypothèse selon laquelle l'exposition à long terme à la pollution causée par le trafic accroit la mortalité cardiovasculaire, et plus particulièrement la mortalité due à une cardiopathie ischémique.

Statement of Originality

Cardiovascular disease is the leading cause of mortality worldwide. There is a growing body of evidence suggesting that chronic exposure to traffic-related air pollution, at its current level, may increase the incidence and mortality from cardiovascular disease. Policies and programs to reduce exposure to traffic pollution are hampered by the gaps of evidence in the empirical literature. The work contained in this dissertation represents an original and important contribution to the research of associations between traffic-related air pollution and mortality from cardiovascular disease. I made two methodological contributions to the field as well as a substantive one, as follows.

One issue is that it is very difficult to estimate historical exposures to air pollution to individuals enrolled in cohort studies and in case-control, especially estimates that are at fine geographical scales. An original contribution of my thesis was the development of three new methods that can be used in these types of studies to reconstruct historical exposure to intraurban air pollution at fine spatial resolutions (2). I have applied these methods in this thesis to the Ontario tax cohort study in estimating rates ratios for cardiovascular mortality and exposure to nitrogen dioxide, a marker for traffic-related air pollution. In addition, my colleagues and I used these methods in a case-control of postmenopausal breast cancer conducted in Montreal in the mid-1990s where we found that risk increased by 30% per a 5ppb increase in nitrogen dioxide (3).

Another methodological contribution that I made in this thesis was refining methods to account in cohort studies for unmeasured confounding. The Ontario tax cohort study did not have information on smoking so we needed a method to take into account the potential confounding effects of smoking, which may occur if there are differential prevalences of smoking between the exposed and the unexposed groups. The classic method of indirect adjustment for smoking proposed by Axelson (4) and others (5) estimated the possible confounding effect as a ratio of the incidence rate of the outcome of interest between the exposed and the unexposed, assuming that there is no exposure effect. The work in this dissertation improves upon the classic indirect method by generalizing its application from categorized exposure variables to continuously measured exposure variables. In addition, it

accounts for random error or uncertainties about the prevalence of unmeasured confounding variables in the study population and the relation of unmeasured confounding variables with health outcomes of interest.

In addition, the substantive contribution of my research furthers our knowledge of the extent to which long-term exposure to traffic-related air pollution may affect cardiovascular disease at the relatively lower levels of pollution that are observed in Canadian cities. The findings of this research provides support to the hypothesis that exposure to traffic-related air pollution may increase the mortality of cardiovascular disease and in particular, from ischemic heart disease. This finding is critical to the development of an appropriate response to reduce the harmful health effects of urban air pollution.

The research objectives and findings contained in my dissertation represent original research. With the guidance and feedback provided by members of my thesis committee, I conceived of the research questions, developed the study protocols and methods, performed all data management and statistical analyses, interpreted the findings, and wrote all three manuscripts. The dissertation comprises a systematic review of the literature on the chronic effects of air pollution on health, a paper describing the historical reconstruction of exposures, and the substantive manuscript on cardiovascular mortality.

Contributions of Authors

Manuscript 1: Chen H, Goldberg MS, Villeneuve PJ. A Systematic Review of Relation between Long-term Exposure to Ambient Air Pollution and Chronic Diseases. Reviews On Environmental Health 2008;23:243-96

Manuscript 2: Chen H, Goldberg MS, Crouse DL, Burnett RT, Jerrett M, Villeneuve PJ, Wheeler AJ, Labrèche F, Ross NA. Back-extrapolation of estimates of exposure from current land-use regression models. Atmospheric Environment 2010;44:4346-4354.

Manuscript 3: Chen H, Goldberg MS, Burnett RT, Jerrett M, Wheeler AJ, Villeneuve PJ. A cohort study in Ontario, Canada, of exposure to traffic-related air pollution and cardiovascular mortality. Submitted to *Epidemiology*.

Dr. Mark Goldberg is full professor in the Department of Medicine, McGill University, is a member of the Division of Clinical Epidemiology, Royal Victoria Hospital, MUHC, and is an associate member in the Department of Epidemiology and Biostatistics, the McGill School of Environment, and the Department of Oncology. Dr. Goldberg is the main investigator of the case-control study of postmenopausal breast cancer in Montreal and the Ontario T1 Family File Tax Cohort study. As thesis supervisor, Dr. Goldberg oversaw all aspects of the development of my study protocols, provided guidance on the statistical analyses and the interpretation of the results, and provided considerable editorial inputs for all three manuscripts.

Dr. Paul Villeneuve is senior research scientist of Health Canada, and is assistant professor of Division of Occupation and Environmental Health, Dalla Lana School of Public Health, University of Toronto. Dr. Villeneuve is the primary investigator of the Ontario T1 Family File Tax Cohort study and secured the funding of the cohort study. As a member of my thesis committee, Dr. Villeneuve provided input on the development of my study protocols, statistical analyses, and the interpretation of the results, and he also edited all three manuscripts and suggested revisions.

Dr. Richard Burnett is a biostatistician, a senior research scientist of Health Canada, and is adjunct professor in University of Ottawa. As a member of my thesis committee, Dr. Burnett provided guidance on statistical analyses and advised how to interpret findings. Dr. Burnett also provided editorial inputs for two manuscripts.

Dr. Daniel Crouse is postdoctoral fellow of Health Canada. As a co-author of my second manuscript, Dr. Crouse contributed to the research related to this manuscript the collection of air pollution data, creation of geographic information system (GIS)-based variables of land use and vehicular traffic, development of the original land use regression model, and commented on the revisions of this manuscript.

Dr. Michael Jerrett is associate professor of University of California, Berkeley. As a research collaborator of the Ontario T1 Family File Tax Cohort study, Dr. Jerrett provided land use regression models to the Ontario T1 Family File Tax Cohort study, and he also edited two of the three manuscripts and suggested revisions.

Dr. Amanda Wheeler is research scientist of Health Canada. As a research collaborator of the Ontario T1 Family File Tax Cohort study, Dr. Wheeler provided land use regression models to the Ontario T1 Family File Tax Cohort study, and she also edited two of the three manuscripts and suggested revisions.

Dr. France Labrèche is research scientist of Institut de recherche en santé publique de l'Université de Montréal, Montréal, Québec, Canada. She is co-investigator of the case-control study of postmenopausal breast cancer in Montreal, and contributed to its study design. Dr. Labrèche provided substantial feedback on the second manuscript.

Dr. Nancy Ross is associate professor of Department of Geography, McGill University. As a research collaborator the case-control study of postmenopausal breast cancer in Montreal, Dr. Ross provided constructive feedback on the second manuscript.

Acknowledgements

First and foremost, I would like to thank my thesis supervisor, Mark Goldberg, for his wonderful wisdom, advice, and encouragement. This thesis has benefited from his invaluable guidance at every step. I thank him for having confidence in my work, for helping me become a better writer, and for training me into an independent researcher. Working with Mark has been a real joy for me, and I feel very lucky to have had an advisor who cares so much about his students: not only their academic development but also their well beings. Thank you for being such a great mentor to me over these many years.

Second, I would like to thank my thesis committee for their important contributions. Very special thanks to Paul Villeneuve, for providing me a great opportunity to participate in the Ontario tax cohort study for which my thesis is based on. Thank you for your crucial support to my data analysis at Statistics Canada and your ideas and suggestions to the development of my thesis. Many thanks to Richard Burnett for his brilliant insights and ideas that lent crucial guidance in my doctoral work, particularly statistical analysis, without which I could not have completed this dissertation.

Thanks also to Daniel Crouse for spending many hours preparing crucial data that allowed me to develop new exposure methods. I have enjoyed collaborating with Dan. Many thanks to Mike Jerrett, Amanda Wheeler, France Labrèche, and Nancy Ross for their invaluable advice and support to the work. Their considerable experience in environmental and occupational health and/or geography lent important perspective to my dissertation.

My research was supported through Canadian Institute of Health Research (CIHR). I am grateful to have received the funding (CIHR doctoral research award). I would also like to acknowledge the support from Hung Pham of Statistics Canada. His exceptional assistance ensured that each of my trips to Statistics Canada was fruitful.

The administrative staffs in the Department of Epidemiology and Biostatistics at McGill have been highly supportive and dependable. I thank for their assistance, in particular Katherine Hayden, André Yves Gagnon, and Suzanne Lariviere. Also I want to thank the

Department of Epidemiology and Biostatistics for offering excellent graduate courses which laid foundation for me to embark my research in epidemiology.

I have enjoyed the company of some great colleagues and friends at McGill over the years. I particularly want to thank my friends Teddy, Dean, Ian, Peter, Julie, and Lisa for our adventures to the Thomson House and for countless conversations – some even about work. I thank all of my other friends, and hope to soon have more time to spend with them than I have been able to over the course of last five years. I also thank Lea Heguy for her help with translating my abstract into French.

Finally, I extend my warmest thanks to my wonderful family: I want to thank my parents for their constant love and support from afar, for which I am profoundly grateful. I want to thank my sister, Lin, for her tolerance of numerous missed calls because of my busy schedule – even at the weekends. Lastly, I am grateful to my wife Lynn for her support and tolerance and for being always there for me in all the good and the bad times. I would especially like to thank my baby son Jeff for kissing me goodbye and keeping my door close whenever he saw me studying at home.

Table of Contents

Chapter 1

INTRODUCTION

1.1 Cardiovascular Disease

Cardiovascular disease is the leading cause of mortality worldwide (6). In Canada, cardiovascular disease accounted for 30%-35% of overall mortality, followed by cancer (~29%) and chronic obstructive pulmonary diseases and allied conditions (5%) in 2000-2007 (7). Approximately 55% of deaths from cardiovascular diseases were coded, as an underlying cause of death, to coronary heart disease and 20% were due to cerebrovascular disease (7). Men are more likely to suffer death from ischemic heart disease and heart attacks, but more women than men die of stroke (7).

Between 1950 and 2007 the age-standardized mortality rates per 100,000 (standardized to the 1991 Canadian population) from cardiovascular disease dropped from 702 to 193 in men, and from 562 to 119 in women in Canada (7, 8). The largest decrease in mortality rates occurred in the oldest age group (80 years and older), which declined from about 14,000 to 7000 deaths per 100,000 standardized population per year in both men and women (7, 8). At younger ages (35-44 years), mortality rate decreased from about 100 to 25 deaths per 100,000 standardized population per year in both men and women (7, 8). In the past 10 years, however, the rate at which cardiovascular death rates have been decreasing has begun to level off (7).

In Canada, there are important regional variations in cardiovascular mortality rates (7). Between 2000 and 2007, Newfoundland and Labrador had the highest rates of cardiovascular mortality and in particular, mortality from ischemic heart disease (7). The province of Ontario ranked in the middle among the 13 provinces and territories. Smoking and unemployment may account in part for the regional differences in mortality rates (9).

Mortality from ischemic heart disease may occur suddenly or within one hour of the onset of symptoms (10). A Canadian study showed that 42% of all ischemic heart disease deaths were observed in men dying within one hour of the onset of their symptoms or found dead in bed (11). These results underscore the need for disease prevention.

Cardiovascular disease has multiple risk factors. The major personal risk factors are age (12), sex (12), tobacco smoking (13-15), high blood pressure (16-18), obesity (19-22), and diabetes (23, 24). More recently, urban air pollution has been implicated in increasing the incidence and mortality from cardiovascular disease (25-29).

1.2 Urban Air Pollution As a Risk Factor of Cardiovascular Disease

Present-day urban air pollution comprises hundreds of substances, including sulphur dioxide, ozone, nitric oxide, nitrogen dioxide (NO_2), carbon monoxide, carbon dioxide, particulate matter (PM), rubber dust, polycyclic aromatic hydrocarbons (PAHs), and many different volatile organic compounds (VOCs). Particles are a heterogeneous mixture of solid and liquid droplets with wide distributions of size and mass. Coarse particles, between 2.5μm and 10μm in median aerodynamic diameter, derive from a variety of sources including windblown dust and grinding operations, and fine particles (particles less than 2.5μm in diameter or $PM_{2.5}$) are primarily due to the combustion of fossil fuel (30). Common constituents of particulates include elemental and organic carbon, sulphates, nitrates, pollen, microbial contaminants, and metals (30). Fine particles may also react with sulphur dioxide and oxides of nitrogen in the atmosphere to form strong acid aerosols, such as sulphuric acid, nitric acid, and hydrochloric acid (30). In addition, some VOCs (e.g., benzene and 1,3-butadiene) and some PAHs (e.g., benzo-a-pyrene) contained in urban air are also considered carcinogenic (31).

Short-term (i.e., day to day) effects of urban air pollution on cardiovascular morbidity and mortality have been investigated in a number of studies conducted in Canada and in other parts of the world (32-47). Associations have been observed consistently between daily numbers of cardiovascular deaths, number of cardiovascular-related hospitalizations, and number of emergency room visits, occurring on the day of the event or the preceding days, and a number of air pollutants, such as nitrogen dioxide and $PM_{2.5}$ (32-47). In contrast, the long-term (i.e., years) effects of air pollution on cardiovascular disease have not been studied as extensively, and the bulk of this work has been focused on health effects relating to

particulate matter, especially $PM_{2.5}$ (27, 28, 48-54). Conclusions from several literature reviews of the health effects of $PM_{2.5}$ are that present-day levels of $PM_{2.5}$ contribute to cardiovascular morbidity and mortality (55, 56). With respect to the health effects of other air pollutants, the evidence is less consistent.

1.3 Assessment of Exposure to Urban Air Pollution

A major challenge in studying associations between long-term exposure to urban air pollution and cardiovascular disease is assigning concentrations of air pollutants to individuals. In the earlier studies (26, 50, 57, 58), rates of cardiovascular mortality were compared between cities, and exposure was assigned using city-specific concentrations of pollutants estimated near the beginning of the follow-up period. Finer-scale estimates of air pollution within cities have been made in some recent studies of cardiovascular mortality (28, 49), and within-city estimates of exposure showed stronger associations with cardiovascular mortality than the between-city estimates of exposure. For example, in the Women's Health Initiative Study (28), Miller and colleagues showed that the rate ratio of mortality from cardiovascular disease for an increase of $10\mu g/m^3$ of $PM_{2.5}$ that was estimated within cities was 2.28 (95% confidence interval (CI): 1.10-4.75), compared to 1.63 (95% CI: 1.10-2.40) for an increase of $10\mu g/m^3$ of $PM_{2.5}$ that was estimated across cities. The differences in the cardiovascular effects between the inter-city and the intra-city estimates of exposure may be due to measuring different components of air pollution (2), but the differences may also be due to chance, different types and levels of errors in measuring exposure and confounding, and possibly that the inter-city estimates of exposure measure regional differences in air pollution that result from many different sources whereas the intra-city estimates of exposure measure spatial variations in local concentrations of air pollution, in particular from traffic sources, and that traffic-related air pollutants may have higher toxicity than from other sources (59).

The concentrations of ambient air pollution can be considered as containing a regional background and a local component (60). The regional background of air pollution is spatially homogeneous within cities, while the local component of air pollution varies substantially over short distances (60). As well, different pollutants have different within-city variabilities,

with fine particles being much more homogeneous spatially than nitrogen dioxide. In urban areas, the spatial variations in the concentrations of air pollution are influenced mainly by vehicular traffic (61).

Several methods have been developed to measure exposure to traffic-related air pollution at small spatial scales. Proximity to roadways is one (49, 62-64) and spatial interpolation at fine geographic resolutions of concentrations of pollutants from fixed-site air pollution monitors that measure criteria air pollutants is another (49, 65, 66). (These fixed-site monitors are often used for the purposes of measuring compliance to regulations by government agencies; in Canada, Environment Canada has this general responsibility, although the provinces are in fact responsible for environmental conditions. The monitoring network in Canada is referred to as the National Air Pollution Surveillance Network) (67).) Proximity to roadways is limited because it does not account for factors related to the dispersion of pollutants that are driven by topography, weather conditions, and that the different types of vehicles and engines (e.g., gas, diesel) may influence emissions (61, 68). The accuracy of spatial interpolation of concentrations of pollutants from fixed-site air pollution monitors requires a dense distribution of monitoring stations that cover the geographic area, and this is rarely available.

A third type of approach is to predict concentrations of selected pollutants using air dispersion models. These models make use of deterministic equations of hydrodynamics to characterize the atmospheric processes that disperse air pollutants emitted by vehicular traffic. Using estimated emission rates from traffic sources and atmospheric and meteorological inputs, air dispersion models are used to predict downwind concentrations of air pollution at selected locations (69). Limitations of these methods include a lack of emission and micro-climatic indicators to make accurate predictions at small scales as well as a dearth of validation studies to ensure the models function appropriately (61).

Over the past decade, there is considerable interest in using geographic information systems (GIS)-based land use regression models to assess exposures of traffic-related air pollution at fine geographic resolutions (64, 70, 71). This method involves measuring selected ambient pollutants, usually using a dense sampling campaign of nitrogen dioxide, and then

developing a statistical prediction model whereby the measured concentrations of air pollutants are regressed against proximate characteristics of land use and vehicular traffic (60, 70-75). Land use regression has been shown to produce reasonably good predictions of small-scale spatial variability of concentrations of NO_2, a valid marker of traffic-related air pollution (63). These models often have explained a relatively large amount of variability in concentrations of NO_2 explained by the models (often above 80%) (60, 71, 72, 76). Because land use regression models have often been developed toward the end of a study (e.g., at the end of follow-up in a cohort study (77-81), it is important to understand whether these models can characterize exposure adequately during the relevant etiological periods in the past. This is particularly important in urban areas where there may be non-homogeneous spatially-dependent temporal trends in air pollution. The validity of the methodology is thus critical for health outcomes that have a long latency (e.g., many cancers).

1.4 Epidemiologic Studies of Cardiovascular Outcomes with Urban Air Pollution

The effect of long-term exposure to urban air pollution on increasing cardiovascular mortality has been investigated in an increasing number of cohort studies (25, 26, 28, 49, 50, 54, 57, 58, 62, 65, 82-86) and case-control studies (87, 88). Among them, the inter-city comparison of exposure to air pollution was made in some studies (26, 50, 54, 57, 58, 65, 82, 86) while in others intraurban exposure to air pollution was estimated (25, 28, 49, 62, 83-85, 87, 88). Positive associations of cardiovascular mortality with long-term exposure to air pollution, in particular $PM_{2.5}$, were found in most of these studies, with a range of between 4% (25) and 76% (28) for each increase of 10 $\mu g/m^3$ of $PM_{2.5}$. In addition, the association between urban air pollution and cardiovascular mortality appeared to be linear (on a log scale) (28, 49, 50, 58, 65, 83, 85).

It is important to note, that while there has been relatively consistent evidence relating particulate air pollutants to cardiovascular events, in few studies (25, 62, 77, 80, 88) has the health effects of traffic-related air pollution been evaluated at the relatively lower levels of pollution that are observed in some areas, such as in Canadian cities (73, 80). In addition, the few studies using land use regression models have had modest sample sizes, so statistical power was limited (73). It is for these reasons that we decided that a new Canadian study of

the association between intra-urban contrasts in the concentrations of air pollution and cause-specific cardiovascular mortality was warranted, which is the subject of my dissertation.

Chapter 2

RESEARCH OBJECTIVES

2.1 Objectives of the thesis

The purpose of this dissertation was to determine the association between traffic-related air pollution and mortality from cardiovascular disease making use of new improved methods of estimates of spatial concentrations of traffic-related air pollution. The specific research objectives were:

1) ***To conduct a comprehensive, systematic review of epidemiological studies used to investigate associations between "long-term" exposure to ambient air pollution and nonaccidental mortality and the incidence and mortality from cardiovascular disease, cancer, and respiratory disease***. The purpose of this objective was to improve upon previous literature reviews by identifying and synthesizing the quantitative findings from all cohort and case-control studies in which exposure-response functions were estimated between specific air pollutants and the incidence and mortality from cardiovascular diseases, cancers, and respiratory diseases. The review included all studies published between 1950 and 2007 and it has been published (1). An update is provided in Chapter 3.

2) ***To develop new methods to extrapolate into the past estimates of concentrations of nitrogen dioxide derived from land use regression models that are based on dense environmental sampling schemes in urban areas***. There is a considerable interest for epidemiologists to assess period-specific exposures to traffic-related air pollution, especially exposures early in life (89-91). The rationale for this objective was to design new methods that could be used to estimate historical exposure to traffic-related air pollution measured at fine spatial resolutions in selected Canadian cities. I first developed and applied these methods to a case-control study of traffic-related air pollution and postmenopausal breast cancer in Montreal and I and my collaborators have published two papers on this (2, 3). I then used these methods in a cohort study in Ontario of traffic-related air pollution and cardiovascular mortality that is the subject of this thesis (objective-3).

3) ***To determine the association between traffic-related air pollution and cause-specific cardiovascular mortality among adults living in three cities in Ontario, Canada***. Very few studies have been conducted to examine the intra-urban variation of traffic-related air pollution on mortality rates of cardiovascular disease. I made use of a population-based cohort study that used as a sampling frame the Canadian federal income tax file "T1 Family File". More than 200,000 adults who lived in either Toronto, Hamilton, or Windsor, Ontario were followed-up to assess vital status between 1982 and 2004. Exposures to traffic-related air pollution were ascertained using concentrations of NO_2 estimated from land use regression models developed previously by other Canadian investigators (73-75) coupled with the above-mentioned back-extrapolation methods.

2.2 Organization of the Dissertation

This dissertation is organized around three core research manuscripts. Two of the manuscripts have been published in the peer-reviewed journals *Reviews on Environmental Health* and *Atmospheric Environment*, respectively (1, 2), and the third has been submitted for publication. The three manuscripts each present research that responds directly to one of the three objectives.

Chapter 3 includes the first manuscript that described a systematic review of all studies published between 1950 and 2007 of the associations between long-term exposure to specific components of ambient air pollution and risks in adults of nonaccidental mortality and the incidence and mortality from cancer and cardiovascular and respiratory diseases. I conducted a comprehensive search of the literature focusing on cohort and case-control studies. The quantitative findings of identified studies were pooled and tested for heterogeneity and pooled estimates were obtained using random-effect models. Because this study included articles that were published before 2008, I have since updated this review by identifying and summarizing quantitative findings from articles published between January 1, 2008 and March 30, 2011 (the most recent date when possible). My dissertation is focused on cardiovascular mortality and thus in this update I only included the recent studies of cardiovascular diseases.

Chapter 4 presents the rationale for estimating concentrations of NO_2 in the assessment of exposure to traffic-related air pollution.

Chapter 5 comprises the second manuscript that describes three new methods to extrapolate a "current" land use regression model back in time by incorporating historical trends in spatially-dependent concentrations of pollutants as well as temporal changes in land use and vehicular traffic. For the purpose of illustration, a current land use regression model that was developed using measurements of concentrations of NO_2 from a dense monitoring campaign conducted in 2005 and 6 in Montreal was back-extrapolated approximately 10 years (to 1996) and 20 years (to 1985). As a specific application of these new methods, these models were then applied to data from a case-control study of postmenopausal breast cancer that was conducted in the mid-1990s (92, 93).

Chapter 6 includes the third manuscript (to be submitted for publication), which builds upon the work developed in the first two manuscripts. It presents results from a population-based cohort study in which the association between traffic-related air pollution and cause-specific cardiovascular mortality among adults living in Toronto, Hamilton, and Windsor, Ontario was investigated. The estimates of traffic-related air pollution for the three cities were derived from dense measurement campaigns of NO_2 that were used specifically to develop land use regression models. The land use regression model for Toronto was then back-extrapolated using the methods that were described above. For Hamilton and Windsor, however, the data from fixed-site monitors were too sparse to allow the use of these back-extrapolation methods. Mortality from cardiovascular and cerebrovascular disease were ascertained for all subjects over the period of 1982-2004 using the Canadian Mortality Database. Adjusted Cox regression models included known individual risk factors and ecological covariables, and smoking was controlled for indirectly using a series of sensitivity analyses.

Chapter 7 concludes the dissertation. First, it summarizes the most important findings from the three core chapters. Second, it discusses the strengths and limitations of the dissertation.

Third, it summarizes the implications for public health of long-term exposure to present-level air pollution, and offers recommendations for further research.

2.3 Ethics and Confidentiality of the Data

The thesis made use of data from two studies: a case-control study conducted by Dr. Mark Goldberg and his collaborators (92, 93) on postmenopausal breast cancer and the tax cohort study in Ontario, for whom Dr. Paul Villeneuve (Health Canada) is principle investigator. For the former study, ethical approval was granted by the McGill University Faculty of Medicine Institutional Review Board as well as all hospitals and institutions that participated in the study. For the latter study, ethical approval was obtained by Health Canada's Research Ethics Board and was approved by the Chief Statistician. Statistics Canada believed that this was a valid use of these data and it did not infringe on any individual's rights and privileges. In addition, I received explicit approval from the McGill University Faculty of Medicine Institutional Review Board to conduct the analyses for my thesis.

One may be interested in knowing how I could have obtained use of data from federal income tax files. Statistics Canada made an agreement with Revenue Canada that it could make use of nominal data from the 1982-2004 T1 family income tax file. Statistics Canada was responsible for a number of record linkage procedures so that the cohort could be followed for mortality and cancer incidence using the Canadian Mortality Database and the Canadian Cancer Registry, respectively. All of these linkages were conducted in-house by Statistics Canada, in collaboration with Dr. Villeneuve. In addition, I signed the Research Data Centres Microdata Research Contract with Statistics Canada.

All statistical analyses of mortality that I conducted in the offices of Statistics Canada in Ottawa made use of a linked, denominalized data set. We were allowed to remove only statistical summaries of the results of the analyses from the offices of Statistics Canada.

I also made use of publicly-available air pollution monitoring data from Environment Canada and U.S. Environmental Protection Agency, digital elevation data from Natural

Resource of Canada, and land use and road network data through DMTI spatial, Inc. (Markham, Ontario, Canada).

Chapter 3

Background

This chapter addresses the first research objective of the dissertation, namely, to conduct a comprehensive, systematic review of epidemiological studies used to investigate associations between "long-term" exposure to ambient air pollution and nonaccidental mortality and the incidence and mortality from cancer and cardiovascular and respiratory disease. This study was conducted to enhance our understanding of the current state of science regarding the chronic health effects of ambient air pollution. The goal of this review was to support policy making and to identify areas where the available evidence is insufficient and where further studies are required. As well, results from this study would assist in interpreting the results of the cohort study. This manuscript has been published in the journal of Reviews on Environmental Health as follows:

Chen H, Goldberg MS, Villeneuve PJ. A Systematic Review of Relation between Long-term Exposure to Ambient Air Pollution and Chronic Diseases. Reviews On Environmental Health 2008;23:243-96

A Systematic Review of Relation between Long-term Exposure to Ambient Air Pollution and Chronic Diseases

Hong Chen, MSc[1]

Mark S. Goldberg, PhD[1,2]

Paul J. Villeneuve, PhD[3]

[1] Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec

[2] Department of Medicine, McGill University, Montreal, Quebec

[3] Population Studies Division, Health Canada, Ottawa, Ontario

Address for correspondence:
Mark Goldberg, PhD
Division of Clinical Epidemiology
McGill University Health Center – RVH
687 Pine Avenue West, R4.29
Montreal, Quebec H3A 1A1
Tel: (514) 934-1934, ext 36917
Fax: (514) 843-1493
Email: mark.goldberg@mcgill.ca

Online data: The search strategy, detailed results and additional tables of the pooled analyses are available from the authors at http://www.med.mcgill.ca/epidemiology/goldberg/Review of Outdoor Air Pollution

ABSTRACT

We conducted a systematic review of all studies published between 1950 and 2007 of associations between long-term exposure to ambient air pollution and the risks in adults of nonaccidental mortality and the incidence and mortality from cancer and cardiovascular and respiratory diseases. We searched the bibliographic databases for cohort and case-control studies. We synthesized the quantitative findings in tabular and graphical form. We assessed heterogeneity, we estimated pooled effects for specific pollutants, and we conducted sensitivity analyses according to certain characteristics of the studies. We conclude that long-term exposure to $PM_{2.5}$ may increase the risk of nonaccidental mortality by 6% per a 10μg/m^3 increase, independent of age, gender, and geographic region. Exposure to $PM_{2.5}$ is also associated with an increased risk of mortality from lung cancer (range: 15%-21% per a 10μg/m^3 increase) and total cardiovascular mortality (range: 12%-14% per a 10μg/m^3 increase). In addition, living close to busy traffic appears to be associated with elevated risks of these three outcomes. There is also suggestive evidence that exposure to $PM_{2.5}$ is positively associated with fatal coronary heart diseases and exposure to SO_2 increases lung cancer mortality. For other pollutants and health outcomes, there is a paucity of data to draw conclusions.

Keywords: air pollution; chronic disease; long-term exposure; review; adult

Published as: Chen H, Goldberg MS, Villeneuve PJ. A Systematic Review of Relation between Long-term Exposure to Ambient Air Pollution and Chronic Diseases. Reviews On Environmental Health 2008;23:243-96

INTRODUCTION

The acute episodes of air pollution in London, Europe, and the USA in the 1950s demonstrated that the confluence of adverse weather conditions and extremely high levels of pollution from ambient particles and sulphur dioxide (SO_2) can cause immediate and dramatic increases in mortality (94, 95). In subsequent decades, changes in fuels (e.g., low sulphur fuels) and improved combustion technology has, especially in economically developed countries, led to significant reductions in the levels of ambient air pollution. Unfortunately, the situation in other less developed countries is not as rosy. For example, Delhi, India is subjected frequently to high levels of total suspended particulates (TSP), with an annual mean concentration well exceeding 600-$\mu g/m^3$ reported (96). In many other parts of the world such as Mexico City and Beijing similar levels of TSP are observed frequently (97, 98). In contrast, the maximum annual level of TSP in Windsor, Ontario, one of the most polluted cities in Canada, is usually under 120-$\mu g/m^3$ (99).

Present-day urban air pollution comprises a complex mixture of substances, including SO_2, ozone (O_3), nitrogen oxide (NO), NO_2, carbon monoxide (CO), carbon dioxide (CO_2), particulate matter (PM), and volatile organic compounds. PM is a heterogeneous mixture of solid and liquid droplets with wide distributions of size and mass. It has been measured and regulated according to its physical properties. Particles greater than 2.5μm in median aerodynamic diameter derive from a variety of sources including windblown dust and grinding operations, whilst fine particles ($PM_{2.5}$) are primarily from the combustion of fossil fuels (100). Common constituents of particles include elemental and organic carbon, sulphates, nitrates, pollen, microbial contaminants, and metals (100). In addition, fine particles may react with SO_2 and NO_x in the atmosphere to form strong acids, such as sulphuric acid, nitric acid, hydrochloric acid, and acid aerosols (100). Urban air also contains benzene and 1,3-butadiene that are carcinogenic (31) and it is thus plausible that long-term exposure to outdoor air pollution, even at low concentrations, can increase the risk of developing a variety of cancers.

Over the past 25 years, considerable epidemiological research has been conducted regarding the acute and chronic effects on health of exposure to ambient air pollution. Associations between exposures occurring on the day of the event or the preceding days and daily non-accidental mortality, cardio-respiratory deaths, hospitalizations, and emergency room visits have been found. There is also evidence that certain subgroups of the population are at higher risk. As a result, one may conclude that short-term elevations of ambient air pollution are associated with a variety of acute health events and that certain subgroups of the population are at much higher risk (101). In contrast, chronic health effects of air pollution have not been studied as extensively as acute effects. Several reviews have been conducted to assess the associations between long-term exposure to ambient air pollution and chronic disease (56, 102-104) and most of these authors concluded that there were reasonable grounds that chronic exposure to air pollution, in particular fine particulates, is associated with the risk of overall mortality from cardiovascular diseases. They differed in their conclusions regarding associations for other causes of death, such as lung cancer.

Given the number of new reports each year as well as the need to capture all published reports in the literature, we conducted a systematic review of all studies published between 1950 and 2007. The focus of the paper is on associations for nonaccidental mortality and the incidence and mortality from cancer and cardiovascular and respiratory diseases.

METHODS

Strategy for searching the literature and the selection of studies

We searched the MEDLINE, EMBASE, and RISK ABSTRACT databases for epidemiological studies for studies of the associations between long-term exposure to ambient air pollution and the risk in adults (men and women who were at least 19 years of age) of developing or dying from cancer, cardiovascular disease, and respiratory disease. We included only cohort and case-control studies. Studies published in the peer-review literature from January 1, 1950 through December 31, 2007 were included regardless of the language of publication. In addition, we perused the bibliographies of all obtained articles and of

previously published reviews, and additional information was sought where necessary from the authors.

We searched these bibliographic databases using the key words "ambient air pollution", which included any component of air pollution, such as O_3, SO_2, NO and NO_2, black smoke, TSP, particles with an aerodynamic diameter of less than 10 (PM_{10}) and $PM_{2.5}$, CO, benzene, and polycyclic aromatic hydrocarbons. We combined this search with the following terms: 'urban'; 'outdoor'; 'ambient'; 'traffic exposure'; 'mobile source'; 'gaseous'; 'particulate'; 'automobile exhaust'; 'proximity'; or 'near' and with the following health outcomes: natural mortality; cancer; cardiovascular disease; and respiratory disease. We used exploded Medical Subject Headings (MeSH) terms for searching health outcomes, which included 'mortality', 'chronic disease', 'cardiovascular diseases', 'neoplasms', and 'respiratory tract diseases'. In addition, we added terms for more specific health outcomes: 'stroke'; 'ischemic heart diseases'; 'lung neoplasms'; 'lung diseases'; 'bronchial diseases'; 'COPD'; and 'death'. To distinguish studies of chronic exposure to ambient air pollution from those of short-term exposure, we only considered studies having exposure periods of at least one year. We excluded occupational studies and those that made use of prevalent outcomes.

We included all related publications from each investigation. If identical outcomes and exposures were analyzed for the same base population, we included reports of the longest follow-up period and/or the largest study population. When a study was reanalyzed, we included results from the latest article. We first screened abstracts and titles, and we obtained the full text of those papers that may have met our eligibility criteria. The second screening was performed with reviewing the paper in its entirety and the reasons for excluding the study were recorded.

Data abstraction and qualitative synthesis

We entered the following data into a relational database: study design; study location; characteristics of the study population; characteristics of subject recruitment and follow-up (if applicable); definitions of outcomes and exposures; methods of exposure assessment and outcome ascertainment; methods of statistical analysis; covariates included in the analyses;

and adjusted odds and/or rate ratios including confidence intervals and/or p-values. In addition, for the cohort studies, we abstracted start and end dates of follow up, percentage of losses to follow-up, response rates, characteristics of non-responders, and characteristics of subjects lost to follow up. For the case-control studies, we extracted descriptions of the study base, inclusion and exclusion criteria, sources of cases and controls, methods of sampling cases and controls, and response rates.

The epidemiologic design of each article was recorded. Nested case-control or case-cohort studies were considered as cohort studies. Population-based case-control studies were defined as to whether the person-time experience of the target population was defined in a well-defined fixed geographic area over a specific period of time, with cases being subjects who were selected randomly from the base population (105). Hospital-based studies were defined as those in which cases were identified from selected hospitals and controls were those who were admitted to the same hospitals for other reasons (105).

We performed a synthesis of the findings by summarizing in tabular and graphical form the quantitative results. We assessed the heterogeneity of effect across studies using Cochran's Q statistic (106) and the coefficient of inconsistency (I^2) (107). The statistic Q measures differences of each study's estimate of effect from the pooled estimate, weighting the contribution of each study by its inverse variance. Under the null hypothesis of homogeneity, Q follows a χ^2 distribution; we considered p-values under 0.1 as indicating heterogeneity (108). The I^2 statistic indicates the proportion of total heterogeneity among studies that is attributed to between-study variance; $I^2<25\%$ is often considered low (109). In an effort to assess publication bias, we produced funnel plots and examined visually whether the funnel plots were asymmetric. We also performed the Begg test (110) and the Egger test (111) to evaluate statistically the symmetry of the funnel plots. The Begg test makes fewer assumptions than the Egger method, but its power to detect bias is more sensitive to the number of studies (112). For those studies in which age, sex, and income stratum-specific effects were reported, we calculated an overall adjusted estimate of effect by weighting the rate ratios by their estimated inverse variance, and then included this summary estimate in subgroup analysis. Where possible, we translated the effect estimates into relative risks or

rate ratios (RRs) for a 10 $\mu g/m^3$ increase in the concentration of the air pollutant, and we assumed, as most authors, that the response function was linear.

Subgroup analysis

To investigate whether the variability in the estimated RRs could be explained by characteristics of the design and conduct of the studies, we conducted separate analyses according to: outcomes (based on the ICD-9 (113)); types of pollutants; indices of exposure (e.g., within-city or between-city comparisons); and adjustment of covariates. We also stratified the cohort studies with respect to participation rates, duration of follow-up, and losses to follow-up. We grouped case-control studies according to the types of study base (e.g., population-based or hospital-based), latency period, and response rate. We then assessed between-study heterogeneity within each subgroup by both visually examining graphic presentations of the estimates of effect and Cochran's Q statistic and the I^2 statistics. We decided *a priori* to use the Dersimonian and Laird random effects method to pool the effect estimates across studies (114).

To determine whether the rate ratios varied by these characteristics, we performed univariate meta-regression, in which we regressed the study-specific RR, on a natural logarithmic scale, by the above-mentioned factors and we used restricted maximum likelihood for estimating variability between studies (115). In addition, we conducted sensitivity analyses by excluding studies according to the results of the meta-regression analyses, or if a characteristic of design was different from other investigations.

RESULTS

Search of the Literature

Of the 1,154 citations identified, 983 were from MEDLINE, 159 were from EMBASE, and 12 were from perusing the bibliographies of the articles. We excluded 1,088 citations after reviewing the abstracts and titles, leaving 66 articles for full review (**Figure 1**). We further

excluded 12 articles for the reasons listed in Figure 1. In the remaining 54 articles, 17 cohort studies and 20 case-control studies were described. In view of a European cohort study whose case-cohort analysis was reported in 2002 (83) but a full-cohort analysis was published in 2008 (25), we made an exception by including the latter paper because identical exposure and outcomes were examined and the direction and magnitude of estimates of associations were different from the previous report.

The cohort studies were used mostly to estimate rate ratios for all-cause mortality (25, 48-51, 57, 58, 62, 85, 116-121), incidence and mortality from lung cancer (25, 49, 50, 57, 58, 84, 85, 116, 119, 120, 122-127) and cardiovascular diseases (25, 27, 28, 49, 50, 57, 58, 62, 64, 82, 84, 85, 116-119). The case-control studies were designed to investigate lung cancer (128-140), breast cancer (141-143), myocardial infarction (87, 144, 145), and asthma(146) and chronic obstructive pulmonary disease (COPD)(147).

Characteristics of the Design and Conduct of the Cohort Studies

Table 1 presents a summary of selected design characteristics of the cohort studies. All studies were conducted prospectively except for three studies (64, 84, 118, 124), that we classified as "database studies" because the sampling of cohorts and the collection of covariate information was based exclusively on administrative databases and no information was obtained directly from subjects.

The duration of follow-up, in most cohort studies was longer than 10 years (median≈17 years), except for the American Legion Study (123), the GenAir Study (127), the Oslo Cohort Study (84), and the US Veterans' Cohort Mortality Study (51), for which follow-up was only about five years. Such short follow-up periods will reduce statistical power and will attenuate rate ratios if induction times are longer than these follow-up periods.

Of the 17 cohort studies, 10 studies were conducted in North America (27, 28, 48-51, 58, 64, 82, 116-123, 125, 148-153) and seven in Europe (25, 57, 62, 84, 85, 124, 126, 127). Three cohort studies included men only (51, 85, 123, 126), two comprised women (28, 62), and the other 12 included both men and women (25, 27, 48-50, 57, 58, 64, 82, 84, 116-122, 124, 125,

127, 148-153). Substantial differences were found in the age distributions of the 17 cohorts, ranging from an average age of 42 years at recruitment (range: 25-59 years) in the study by Filleul et al. (57) to an average age of 65 years in the study by Enstrom (48). The number of participants also varied considerably, ranging from about 1,000 (127) to about a million and half (124).

In addition, there were differences in the levels of education between studies: the proportions of subjects not graduating from high school were: about 4% in the Women's Health Initiative (WHI) study (28), about 12% in the American Cancer Society Study (ACS study) (27, 49, 58); 25-30% in the Harvard Six Cities study (50, 119) and the California Cancer Prevention Study (CA CPS I) (48); 31-39% in four studies (57, 57, 62, 85, 126, 127); ~50% in the Adventist Health Study on the Health Effects of Smog (ASHMOG) (82, 116, 117, 120, 122, 125, 149-153); and ~67% in the Oslo Cohort study (84). In the seven other studies (25, 51, 64, 118, 121, 123, 124, 148), educational status of subjects was not presented.

Most of the studies were population-based cohort studies with the exception of the US Veterans' Cohort Mortality Study (51), for which the source population included veterans who were hypertensive (100%). This cohort also had a disproportionately large number of current or former smokers (81%) and African-Americans (35%), and subjects had a lower SES compared to the U.S. general population.

Mortality was ascertained in these studies from underlying causes of death as recorded on death certificates. The incidence of cancer, cardiovascular or respiratory diseases was ascertained using postal questionnaires (28, 116, 149, 151-153), medical records (25, 28, 127), clinical examinations or interviews (148), and data from cancer registries (116, 122, 124, 126, 127).

Participation and losses to-follow-up

The three database studies (64, 84, 118, 124) had high participation and follow-up rates (above 98%). In the remaining 14 studies for which subjects were contacted directly, participation rates were reported in five studies (25, 62, 82, 85, 116, 117, 120, 122, 123, 125,

126, 149-153) and they varied considerably: 36% in the Dutch NLCS Study (25); 50% in the American Legion Population (123); 63% in the Cohort of Norwegian Men (85, 126); 70% in the German Women's Health Study (62); and 87% in the ASHMOG study (82, 116, 117, 120, 122, 125, 149-153). In addition, losses to follow-up were reported in 16 studies: about 50% in the Study of residents of Lancaster and Glendora, California (148); 25% in the Pennsylvania Cohort Study (121); about 20% in the French Cohort Study (57); 0% to 51% in the ASHMOG study (depending on the analysis) (82, 116, 117, 120, 122, 125, 149-153); and under 10% in 12 other cohort studies (25, 27, 28, 48, 49, 51, 58, 62, 64, 84, 85, 118, 119, 123, 124, 126).

Exposure assessment

Table 2 shows details of the methods used to assess exposure. Area-wide measurements of air pollution at residences and/or workplaces during the follow-up periods were used throughout to infer long-term personal exposure to ambient air pollution. It was thus assumed that all subjects who resided in a defined geographic area shared a common level of exposure to ambient air pollution.

We refer to comparisons of exposure across metropolitan regions (e.g., at the scale of the city or larger) as "interurban" studies (27, 48, 50, 51, 58, 119, 121, 123, 124, 148) and those at a finer geographic scale (e.g., residential district or neighbourhood) as "intraurban" studies (25, 28, 49, 57, 62, 64, 82, 84, 85, 116-118, 120, 122, 125-127, 149-153).

In four out of the eight interurban studies (121, 123, 124, 148), cities were classified on an ordinal scale according to historical observations of multiple air pollutants. In the other four studies (27, 48, 50, 51, 58, 119), as new monitoring data were available, annual mean levels of specific pollutants were estimated by averaging across all monitoring sites within each city or county and the estimates were then assigned to the subjects according to their city or county of residence.

The density of fixed-site monitors varied considerably among the interurban studies (27, 48, 50, 51, 58, 119, 121, 123, 124, 148). In the original Harvard Six Cities Study (119) one

centrally located air-monitoring station per city was used. In the other American cohort studies, use was made of the network of fixed-site monitoring stations operated by the US Environmental Protection Agency (EPA). In the California Cancer Prevention Study (48), one to three monitoring sites were used per county to estimate the concentrations of $PM_{2.5}$. An expanded version of the EPA network of monitors was used in the continued follow-up of the Harvard Six Cities Study (50) (referred to as the "extended" study), the ACS Study (27, 58), and the US Veterans' Cohort Mortality Study (51).

Data from relatively dense monitoring networks were available in the nine intraurban cohort studies (25, 28, 49, 57, 62, 64, 82, 84, 85, 116-118, 120, 122, 125-127, 149-153). Gehring et al. (62) and Filleul et al. (57) estimated exposure of participants at their residences by averaging concentrations in each residential area defined by jurisdictional boundaries. In the GenAir Study (127) and in the WHI study (28) levels of pollutants were calculated from one or more monitoring stations closest to each subject's home, regardless of the boundaries. In the GenAir Study (127), the estimates were then assigned to geocoded addresses of subjects' residences and in the Women's Health Initiative Study they were attributed to centroids of Zip-code areas (28).

Statistical modeling (25, 49, 64, 82, 116-118, 120, 122, 125, 149-153)and dispersion models (84, 85, 126) using air pollution monitors and emission inventories were used in six cohort studies . In three studies (49, 64, 82, 116-118, 120, 122, 125, 149-153), concentrations of air pollutants at unsampled locations were estimated as a weighted average between adjacent monitoring sites. In these three studies, spatial interpolation models using inverse distance weighting and kriging (109) were used. In the Dutch NLCS study (25), levels of air pollution were estimated from the spatial dependency of concentrations of air pollutants, accounting for land use and vehicular traffic. This latter method consisted of inverse distance weighting in conjunction with a regression model. Dispersion models which simulated the physical and chemical processes that affect air pollutants as they disperse and react in the atmosphere were used in two studies (84, 85, 126).

Distance to nearest major roads (25, 49, 62, 127) and local traffic density (25, 51) was also estimated. Exposure was assigned using zip-code centroids in one study (49) and in the other

three studies it was assigned to geocoded addresses of subjects' residences. Various definitions were used for close proximity to roads: within 500m in the ACS-LA study (49); within 50m in the German Women's Health Study (62) and in the Dutch NLCS study (25); and living on a "major" road in the GenAir study (127). In the Veterans' Cohort Mortality Study (51), traffic density was calculated as county-average vehicular traffic density and three indices derived from the density of traffic were used in the Dutch NLCS study (25).

There were also substantial differences in exposure periods. In three studies (84, 121, 123), exposure was assessed for a period prior to the enrolment of subjects. In the remaining 14 studies (25, 27, 28, 48-51, 57, 58, 62, 64, 82, 85, 116-120, 122, 124-127, 148-153), with the exception of four studies (25, 82, 116, 117, 120, 122, 125, 127, 148-153), exposure was assessed for only a portion of the follow-up period. In four studies (57, 62, 64, 118, 126) exposure was assessed for an early portion of the follow-up, in another four studies (48, 50, 85, 119, 124) it was conducted for the mid-years of the follow-up, in three other studies (28, 49, 51) it was for the last a few years of the follow-up, and in one study (27, 58) it was an average of different periods.

Characteristics of the Design and Conduct of the Case-control Studies

Table 3 summarizes essential design characteristics of the 20 case-control studies. The health outcomes evaluated in the case-control studies were: lung cancer (13 studies: (128-140)); female breast cancer (2 studies: (141-143)); COPD (1 study: (147)); adult asthma (1 study: (146)); and myocardial infarction (3 studies: (87, 144, 145)).

Five of the studies (129, 132, 141-144) were conducted in North America, 12 in Europe (87, 128, 131, 133-137, 139, 145-147), and three in Asia (130, 138, 140). Six studies involved men only (128, 129, 135-137, 145), five included only women (132, 133, 140-143), and the remaining nine included both men and women (87, 130, 131, 134, 138, 139, 144, 146, 147). There were substantial differences in the number of participants, ranging from 96 (132) to 15,326 (144).

Fifteen case-control studies were population-based (87, 128, 130-132, 135-137, 139-146) and five were hospital-based (129, 133, 134, 138, 147). In the population-based studies, controls were identified from population registries (8 studies: (87, 130, 132, 136, 137, 144-146)), death registries (3 studies: (131, 139, 140)), driver's license files (2 studies: (141-143)), a registry at a local university (1 study: (135)), and selected number of hospitals (1 study: (128)). Random sampling or incidence density sampling was used to select control subjects, except in the study by Stocks et al. (128) in which a convenient sample was taken. In one study (134), both population- and hospital-based controls were recruited. In another study (137), living and deceased controls from the general population were enrolled. In both of these studies, the two control series were similar and were combined in the analyses. In the population-based study in Ontario by Holowaty et al. (132), cases were sampled from seven out of nine hospitals located in the study region (two hospitals did not participate).

In the case-control studies of lung cancer and breast cancer, cases were identified from regional/national cancer registries (141, 142), from hospital records (128-130, 132-134, 138), and from death certificates (139, 140). In five of the seven studies that relied on hospital records, cases were confirmed histologically (129, 130, 133, 134, 138). In the remaining two studies, 88% of cases were confirmed histologically (132) and Stocks et al. (128) did not report confirmation of cases.

Participation rates for control subjects varied considerably: 41% in the study by Jockel et al. (134); 58% in the study by Holowaty et al. (132); 64% in the study by Jedrychowski et al. (131); 67% in the two studies of breast cancer (141-143); 68% in the study by Pawlega et al. (136); 70% in two studies by Rosenlund et al. (87) and Grazuleviciene et al. (145); about 85% in two studies by Nyberg et al. (137) and Barbone et al. (135); and above 90% in six other studies (130, 133, 139, 140, 144, 147). In four other studies (128, 129, 138, 146), participation of controls were not described. Although high response rates (100%) for cases and controls were reported in the studies by Chiu et al. (140) and Tonne et al. (144), no information was obtained directly from subjects.

Exposure assessment

Table 4 summarizes the key characteristics of the assessment of exposure used in the 20 case-control studies. Long-term exposure to air pollution was assigned according to area-specific levels of air pollution at subjects' residences and/or workplaces. Of particular importance, and in contradistinction to the cohort studies, was that individual air pollutants were not measured routinely.

In six studies of lung cancer (128, 130, 131, 135, 136, 140), one study on adult asthma (146), and two studies of myocardial infarction (144, 145) residential addresses at time of diagnosis were used. Among the case-control studies of lung cancer, exposure was assessed at the time of diagnosis in four studies (128-130, 140) and was lagged by about five years in three other studies (131, 133, 145). In the remaining six studies of lung cancer (128, 132, 136-139), lagging exposure by over 20 years was reported. In addition, in one case-control study of adult asthma (146) and one study of myocardial infarction (144), exposure was assessed for a period that corresponded to the last few years of the study period.

In one study (134), exposure was assessed at the county level. Due to the lack of adequate monitoring data, two semi-quantitative indices of ambient exposure were derived from energy consumption and historical emission data. Exposure was assessed at the city level in five studies (128, 132, 138, 140, 147). In four of these five studies, exposure was classified on ordinal scales (128, 138) or by the number of years of exposure (132, 147) at each residence. In the study by Chiu et al. (140), direct measurements were used to derive a composite index of air quality which was assigned to subjects' residences at the time of diagnosis.

Health effects at the neighbourhood scale were also assessed. In four of 15 intraurban case-control studies (135, 136, 139, 145), exposure was evaluated by residential district. Specifically, Barbone et al. (135) classified exposure according to residence in different parts of a city (the highest level of exposure was assigned to the urban centre) and Besso et al. (139) considered exposure if subjects lived in two parishes closest to a local non-ferrous metal smelter. In Grazuleviciene et al. (145), the estimates of exposure were derived from concentrations of NO_2 measured at fixed-site stations (one site per residential district). Pawlega et al. (136) assigned exposure from annual levels of TSP and SO_2 obtained by averaging measurements across several monitoring sites in each district.

In the remaining 11 case-control studies, assessment of exposure was conducted at subjects' homes. Modig et al. (146) made direct measurement of specific pollutants at residential addresses for one week and then converted the one-week measurements to annual measurements using correlations between temperature and NO_2 derived from time series data (local traffic density was also considered as an index of exposure). In three other studies (130, 143, 144), proximity to traffic and industries was used. In addition, Jedrychowski et al. (131) and Vena et al. (129) reported the plotting of isopleth maps from fixed-monitoring sites to interpolate measurements of air pollution at home addresses. Air dispersion models (87, 137, 141) and inverse distance weighting at adjacent fixed-site stations (133, 142) were also used to estimate levels of air pollution at home addresses.

Adjustment for Potential Risk Factors

Age and sex were accounted for in all of the cohort and case-control studies. In most of the studies, cigarette smoking was associated with an elevated risk of all-cause mortality, lung cancer, and respiratory and cardiovascular diseases. Smoking was accounted for in all but two cohort studies (64, 84, 118) and four case-control studies (138, 140, 143, 144), including one study on breast cancer (143) (it is controversial as to whether smoking is a risk factor for breast cancer). In another case-control study of breast cancer, smoking was not included in the analysis in one paper (142) but was in another paper (141).

In the Danish Cohort Study (124), smoking was controlled indirectly using smoking information obtained from a population survey. In four other cohort studies (25, 51, 121, 123), smoking was classified as a binary factor (51, 123) or as a three-level categorical variable (25, 121). In the remaining 10 cohort studies, adjustment for smoking included various indices of intensity and duration.

Individual-level socioeconomic status (SES) was measured usually according to income or education. Where individual-level SES was not available, area-based measures derived from census data were used (25, 51, 64, 118, 144, 146). In seven cohort studies (25, 51, 64, 118,

121, 123, 124, 148) and six case-control studies (128, 129, 131, 132, 134, 139), SES was not accounted for.

Control for occupational exposure was reported in nine cohort studies (27, 28, 49, 57, 58, 62, 82, 84, 85, 116, 117, 120, 122, 124-127, 149-153) and eight case-control studies (129, 131, 134-137, 139, 145). In addition, body mass index was accounted for in nine cohort studies (27, 28, 49-51, 57, 58, 62, 64, 82, 116-118, 120, 122, 125, 127, 149-153) and four case-control studies (141, 142, 145, 146).

Substantive Findings

Figs 2 – 5 show the results from all studies according to selected outcomes and pollutants. Not included in these figures are the results for cohort and case-control studies for which quantitative estimates of risk by pollutant were not published. Results for the case-control studies are shown in **Table 5**. In our analyses of $PM_{2.5}$, NO_2, and SO_2 we did not find any evidence of publication bias (using funnel plots and the Begg test and Egger test). We could not test for publication bias for other pollutants because there were too few reports.

All-cause Mortality

Figure 2 shows the adjusted rate ratios or relative risks (RR) for all-cause mortality from 11 cohort studies. One cohort study was excluded because RRs were not reported (121) and, obviously, there were no case-control studies for this outcome. Proportional hazards models were used in all the studies except that of Morris et al. (121) who reported standardized mortality ratios (SMR) comparing mortality rates in the cohort to that of the general population. It was assumed in all studies that the response functions were linear.

Gaseous pollutants

Associations for CO, NO, NO_x, and O_3, were reported in very few studies (**Figure 2**). Null associations were observed in most of the studies except for the Norwegian Cohort Study of Men (85) (NO_x: RR for a 10 $\mu g/m^3$ increase (RR_{10})=1.08; 95% CI: 1.06-1.11).

Associations for NO_2 were reported in six studies (25, 51, 57, 58, 62, 117) and in two of the studies (25, 62) the estimates were greater than unity. The pooled estimate for NO_2 was unity (95%CI: 0.99-1.02), but there was substantial heterogeneity (p=0.023; 61.8% of the total variance among studies (I^2)). The meta-regression analysis suggested that one study of women (62) had a substantially different estimate of effect (RR_{10}=1.11; 95%CI: 1.01-1.23).

Estimates for SO_2 were reported in seven studies (25, 51, 57, 58, 85, 117, 118)), and four of these studies (57, 58, 117, 118) showed rate ratios greater than unity. Six studies were included in the subgroup analysis because the estimates from Finkelstein et al. (118) could not be translated to a 10 $\mu g/m^3$ increase (**Fig 1**). The pooled RR_{10} for SO_2 from the six studies was 1.01 (95%CI: 0.98-1.03) and there was significant between-study heterogeneity (p<0.001, I^2=79.2%). We did not find in the meta-regression analyses any characteristic that explained the heterogeneity, but we found slightly higher relative risks in North American studies (RR_{10}=1.03; 95%CI: 0.98-1.08) (51, 58, 117) and in interurban studies (RR_{10}=1.03; 95%CI: 0.97-1.09) (51, 58).

We conducted further sensitive analyses by excluding studies that explained the heterogeneity in the meta-regression analysis (57, 62) as well as the US Veterans' Cohort Mortality Study (51), for which there were concerns over the source population and the short period of follow-up. We found that the RRs for NO_2 and SO_2 were essentially unchanged (RR≈1.01) (see on-line data).

Particulates

Five (25, 48-50, 58, 120) of six cohort studies showed positive associations for $PM_{2.5}$. The pooled RR_{10} for $PM_{2.5}$ was 1.06 (95%CI: 1.03-1.10) in both fixed- and random-effects analyses (heterogeneity: p=0.22; I^2=28.1%) (see on-line appendix). We found some differences in relative risks according to completeness of follow-up: specifically, the RR_{10} for the extended Harvard Six Cities Study (50) (loss to follow-up was not reported) was 1.16 (95% CI: 1.07-1.26) as compared to the five studies with >90% of follow-up (RR=1.05; 95% CI: 1.03-1.07) (meta-regression p-value=0.02). In sensitivity analysis, we excluded the extended Harvard Six Cities Study (50) and the US Veterans' Cohort Study (51), but did not find any meaningful differences in the pooled estimate (RR_{10}≈1.05).

Very few studies were used to evaluate mortality in relation to other types of particulates. Null associations were found in two out of the three studies of PM_{10} (58, 117) and TSP (57, 58), and in one of the two studies on black smoke (57). A strong positive association between sulphates and all-cause mortality was reported in two U.S. cohort studies (58, 117) (pooled RR_{10}=1.06; 95%CI: 0.96-1.19).

Other measures of air pollution

Associations by indices measuring traffic density were reported in four cohort studies (25, 49, 51, 62) (see on-line appendix). Significant positive associations were found in two studies in which traffic density was estimated by number of vehicles per day passing on the nearest road (RR=1.03; 95%CI: 1.00-1.08) (25) and by county-specific annual estimates of vehicle-miles driven (RR=1.06; 95%CI: 1.02-1.09) (51), respectively. Among the three studies (25, 49, 62) in which distance to roads was used as another metric of traffic exposure, living within 50 m to a major road was positively associated with all-cause mortality (RR=1.05 (95%CI: 0.97-1.12) (25) and 1.29 (95%CI: 0.93-1.78) (62)) but not in one study (49) in which a crude measure was used (RR=0.99; 95%CI: 0.88-1.11).

Incidence and Mortality from Lung Cancer

Figure 3 shows adjusted rate ratios or odds ratios for the incidence and mortality from lung cancer (nine cohort studies (25, 49, 50, 57, 58, 84, 85, 117, 122, 125-127) and one case-control study (137)). Excluded from this summary are two cohort studies (123, 124) and 12 case-control studies (128-136, 138-140) because in these studies quantitative estimates of risk by pollutant were not reported.

Gaseous pollutants

Associations between CO, NO, NOx, and O_3, and the incidence and mortality from lung cancer were assessed in very few studies (**Figure 3**). Null associations were reported for CO and NO. NOx was positively associated with incidence (RR_{10}=1.08; 95%CI: 1.02-1.15) and mortality (RR_{10}=1.11; 95%CI: 1.03-1.19) of lung cancer in the Cohort Study of Norwegian Men (85, 126). For O_3, a 37% of increase in mortality in men (95%CI: 1.00-1.88) and a 10%

decrease in mortality in women (95%CI: 0.66-1.22) was found in the California ASHMOG study (117) and null associations were reported in the ACS study (49, 58).

Associations between NO_2 and mortality from lung cancer were reported in five studies (25, 57, 58, 84, 117). In two of these studies (84, 117) the estimates were greater than unity (**Figure 3**). The pooled RR_{10} for NO_2 was 1.01 (95%CI: 0.94-1.09), but this excluded the Oslo Cohort Study (84) because the risk estimates could not be translated to a 10 $\mu g/m^3$ increase. There was heterogeneity across these four studies ($p=0.005$; $I^2=65.3\%$), but this heterogeneity could not be explained in the meta-regression analysis, possibly due to a lack of power (see on-line appendix). In further sensitivity analysis, we excluded Filleul et al. (57) because of its relatively high losses to follow-up, but this had little effect on the pooled estimate. Positive associations between NO_2 and the incidence of lung cancer were found (122, 127, 137), and the confidence intervals of the estimates were wide and barely included the null (pooled $RR_{10}=1.11$; 95%CI: 0.99-1.24).

In five studies (25, 57, 58, 117, 126) associations between mortality from lung cancer and SO_2 were reported and two of these studies (58, 117) showed rate ratios greater than unity. The pooled estimate for SO_2 was 1.07 (95%CI: 0.96-1.19) (**Table 6)** and there was significant between-study heterogeneity ($p<0.001$; $I^2=85.3\%$). The North American studies had stronger associations ($RR_{10}=1.58$; 95%CI: 0.66-3.76) (58, 117) as compared to the European ones ($RR_{10}=1.00$; 95%CI: 0.96-1.19) (25, 57, 126), although the difference was not significant ($p>0.1$). In sensitivity analysis, we excluded the study by Filleul et al. (57) and this increased the RR_{10} from 1.07 to 1.12 (95%CI: 0.97-1.30).

In addition, associations between SO_2 and the incidence of lung cancer were estimated in four studies (85, 122, 127, 137) and positive associations were observed in all of the studies. The pooled RR_{10} was 1.12 (95%CI: 0.98-1.29), and there was substantial heterogeneity across these studies ($p<0.001$ and $I^2=84.4\%$). This heterogeneity was attributable mostly to the ASHMOG study (122) ($RR_{10}=3.43$; 95%CI: 2.12-5.55).

Particulates

The association between $PM_{2.5}$ and mortality from lung cancer was estimated in five studies (25, 49, 50, 58, 84, 120) and all studies showed positive effects (pooled RR_{10}=1.21; 95%CI: 1.10-1.32) (**Table 6**) and there was little heterogeneity (p=0.26, I^2=24.5%). However, smoking was not taken into account in the Oslo Cohort Study (84) and this study showed a stronger association (RR_{10}=1.34; 95%CI: 1.17-1.54) than in those studies for which smoking was accounted for (pooled RR_{10}=1.15; 95%CI: 1.06-1.24) (25, 50, 58, 120) (meta-regression p=0.046).

Very few studies were carried out to assess the effects of the other particulates. Among the three studies of PM_{10}, strong positive associations with mortality from lung cancer were found in two studies (84, 117) (**Figure 3**). Sulphates were examined only in the ACS study (58) and a 15% excess mortality was found (95%CI: 1.06-1.24). Null associations were reported for black smoke and TSP.

Other measures of air pollution

Exposure to air pollution was also assessed by residential areas with high levels of air pollution or by duration of exposure at "high" levels, and findings using these two types of exposure indices were reported in 12 case-control studies (128-136, 138-140) (**Table 5**) and in two cohort studies (119, 124). It is difficult to summarize succinctly the results because of the variety of indices that were used. In summary, and referring to **Table 5**, positive associations with lung cancer (including incidence and mortality) were reported in 10 case-control studies (128-135, 139, 140) and in the two cohort studies. In three case-control studies (130, 135, 140) and one cohort study (124) statistically significant associations were found. All 12 case-control studies were substantially under-powered (i.e., wide confidence intervals). There were also concerns of low response rates in four case-control studies (<70%) (131, 132, 134, 136), classification of exposure according to residence at the time of diagnosis in six case-control studies (128, 130, 131, 135, 136, 140), not accounting for smoking in two case-control studies (138, 140) and SES in six case-control studies (128, 129, 131, 132, 134, 139).

Exposure to air pollution was also estimated in three cohort studies (25, 49, 127) according to distance from the subjects' residences to a major road (see on-line appendix). The

incidence of lung cancer was elevated by 31% (95%CI: 0.82-2.09) for living on a "major" road (127), mortality was increased by 44% (95%CI: 0.94-2.21) for living within 500m of freeways (49), and a 20% excess in mortality (95%CI: 0.98–1.47) was found for living within 50m of a local road (25). A positive association with lung cancer mortality was also found with local traffic density (RR=1.07; 95%CI: 0.93–1.23) (25).

Incidence and Mortality from Other Sites of Cancer

Associations between air pollution and breast cancer were assessed in two case-control studies (141-143) (**Table 5).** Air pollution was positively associated with the incidence of breast cancer in both studies, but the 95% confidence intervals were wide and they included the null. In addition, digestive cancer and other sites of cancer were investigated in the ACS study (49) and in the California ASHMOG study (116, 125), and no associations were reported for these outcomes.

Incidence and Mortality from Cardiovascular Diseases

Figure 4 summarizes the adjusted rate ratios for associations with cardiovascular and cardiopulmonary disease from 10 cohort studies (25, 27, 28, 49, 50, 57, 58, 62, 82, 84, 85, 117, 118). **Table 5** shows risk estimates for three case-control studies (87, 144, 145) of myocardial infarction. Mortality was estimated in the following studies: total cardiovascular disease in four cohort studies (25, 27, 50, 84); cardiopulmonary diseases (i.e., disorders that affects the normal functions of the heart and lungs) in six cohort studies (25, 49, 57, 58, 62, 117, 119); coronary heart diseases in four cohort studies (27, 28, 49, 82, 85); and cerebrovascular diseases in four cohort studies (27, 28, 64, 85).

Gaseous pollutants

Associations between gaseous pollutants and any cardiovascular event (incidence and mortality) were estimated in one study (28) (**Fig 4**). Increased risks were found for SO_2 (RR_{10}=1.16; 95%CI: 0.95-1.40), NO_2 (RR_{10}=1.04; 95%CI: 0.96-1.12), and O_3 (RR_{10}=1.05; 95%CI: 0.99-1.11). Null association was reported for CO (RR_{10}=1.00; 95%CI: 0.81-1.22).

The effects of gaseous pollutants (NO_2, SO_2) on total cardiovascular mortality were reported in two studies (25, 84). Effect estimates greater than unity were found for NO_2 but not SO_2 (RR_{10}=0.97; 95%CI: 0.91–1.03).

In five studies (49, 57, 58, 62, 117, 118) associations with cardiopulmonary mortality and gaseous pollutants were estimated. The association with NO_2 was estimated in four studies (57, 58, 62, 117) of which one study, comprising only women (62), had a substantially different estimate (RR_{10}=1.41; 95%CI: 1.18-1.70) from the pooled one (RR_{10}=1.02; 95%CI: 0.98-1.07) (meta-regression p<0.001) (see on-line appendix for details). In sensitivity analysis, we excluded this study and the pooled RR_{10} decreased to 1.00 (95%CI: 0.99-1.02).

SO_2 was positively associated with cardiopulmonary mortality in three (58, 117, 118) studies but not in a fourth one (57). We excluded the study by Finkelstein et al. (118) (the RR could not be converted to a 10 $\mu g/m^3$ increase) and the pooled estimate was 1.02 (95%CI: 0.95-1.09). Considerable heterogeneity was present (p=0.02; I^2=75.6%) and was due to the null study by Filleul et al. (57), the pooled RR_{10} after excluding this study was 1.06 (95%CI: 1.02-1.09) and heterogeneity was eliminated (p=0.42, I^2=0%).

In addition, null associations with cardiopulmonary mortality were reported in one study of CO (RR_{10}=0.56; 95%CI: 0.35-0.90) (58), one study of NO (RR_{10}=1.00; 95%CI: 0.98-1.02) (57), and two studies of O_3 (pooled RR_{10}=1.00; 95%CI: 0.98-1.03) (58, 117).

The effects of gaseous pollutants on mortality from cerebrovascular diseases were estimated only in the Norwegian Cohort Study of Men (85), where a 4% excess (95%CI: 0.94-1.15) and a 2% excess (95%CI: 0.93-1.12) were reported.

The association between NO_2 and the incidence of myocardial infarction was estimated in two case-control studies (87, 145) (**Table 5)**. Increased risks were reported in one study (145) (RR_{10}=1.14; 95%CI: 1.01-1.30 for 25-64 years old and 1.29 (1.07-1.55) for 55-64 years old), but not in the other study (87) (RR_{10}=0.96; 95%CI: 0.87-1.06). In the latter study (87), the RR_{10} for NO_2 was 1.15 (95%CI: 0.99-1.33).

Particulates

$PM_{2.5}$ was positively associated with total cardiovascular mortality in all four studies (25, 27, 50, 84). The pooled RR_{10} was 1.14 (95%CI: 1.09-1.18) and there was moderate between-study heterogeneity (p=0.13; I^2=46.3%) (**Table 7**). In the meta-regression analysis, we found a higher estimate in the study that did not report losses to follow-up (RR_{10}=1.28; 95%CI: 1.13-1.45) (50) (meta-regression p=0.037). Excluding this study and the Oslo Cohort Study (84) yielded a pooled RR_{10} of 1.12 (95% CI: 1.09-1.15). Including the estimate of mortality from the combined outcome of coronary heart disease and cerebrovascular disease (RR_{10}=1.76)) in the WHI study (28) increased the pooled RR_{10} slightly (pooled RR_{10}=1.15; 95%CI: 1.09-1.23).

Mortality from coronary heart disease was investigated in three cohort studies (27, 28, 49, 82). The pooled RR_{10} for $PM_{2.5}$ was 1.16 (95%CI: 0.96-1.40) and there was considerable heterogeneity (p=0.013; I^2=77%) (**Table 7**). The meta-regression analysis indicated that the estimates of relative risk reported in the WHI Study (28) (RR_{10}=2.21) were considerably higher than the pooled RR_{10} (RR=1.10; 95%CI: 0.94-1.29) from the two other studies that included men and women (meta-regression p=0.05). In addition, the intraurban studies showed stronger associations (RR_{10}=1.24) (28, 49, 82) than the interurban study (RR_{10}=1.18; 95%CI: 1.14-1.23) (27), although the difference was not significant (meta-regression p=0.21).

Associations between $PM_{2.5}$ and fatal cerebrovascular disease were assessed in two studies (27, 28) (**Figure 4**), with one study showing a 83% increase in mortality (95%CI: 1.11-3.00) (28) and the other (27) showed a 2% (95%CI: 0.95-1.10) increase.

Other measures of air pollution

Associations between living close to traffic and cardiovascular outcomes were evaluated in four cohort studies (25, 49, 62, 64) and one case-control study (144). Significant positive RRs for cardiopulmonary mortality were reported in two studies (25, 62).In the first one (25), traffic intensity on the nearest road was used to measure traffic (RR=1.06; 95%CI: 1.00-1.12) and in the other study (62), living within 50 m from a major road was reported (RR=1.70; 95%CI: 1.02-2.81). In addition, a 4% increase (95%CI: 1.02-1.06) in nonfatal myocardial infarction was shown in one study (144), a 79% increase in mortality from cerebrovascular

disease was found in one (64), and there were positive associations with total cardiovascular mortality (RR=1.37; 95%CI: 1.05-1.79 and RR=1.05; 95%CI: 0.93-1.18) in two studies (25, 64) (see on-line appendix).

Incidence and Mortality from Respiratory Diseases

Figure 5 summarizes the estimates of association between air pollution and respiratory diseases. The effects of air pollution on respiratory diseases were reported in very few studies. A 1% (95%CI: 0.94-1.09) increase in mortality from all respiratory diseases for NO_2 (25, 117), a 16% (95%CI: 1.06-1.26) excess in mortality for NOx (126), and a 3% increase in mortality for O_3 (95%CI: 0.95-1.11) (117) were found. In three studies (25, 117, 126), null associations with respiratory mortality and SO_2 were reported (pooled RR_{10}=0.99; 95%CI: 0.91-1.07). We found positive associations with respiratory mortality and particulates (e.g., pooled $RR_{PM2.5}$=1.03; 95%CI: 0.89-1.20), although the confidence intervals of all of these estimates included unity. The long-term effects of air pollution on adult asthma were less clear because the estimates were imprecise (e.g., O_3: RR_{10}=2.16 (1.24-3.75); SO_4: RR_{10}=4.51 (1.09-18.71)) (149, 153).

DISCUSSION

Interpretation of the Findings

Studies of the effects of ambient air pollution on human health are complicated by the fact that individuals are exposed to a complex mixture of toxic and non-toxic substances that vary in their make-up in space and in time (154) and even within one city there can be extensive spatial variability (155). The pollutants derive mostly from fixed and moving combustion sources and from other non-anthropogenic sources. Few compounds are measured routinely (e.g., particles, criteria gaseous pollutants) and many other pollutants are measured infrequently (e.g., volatile organic compounds, pollen). Particles are also a heterogeneous mixture varying in space and time.

One of us (154) has argued that it is likely that the causal agents in this mixture cannot be identified through standard epidemiological designs. This problem is common in environmental epidemiology, cigarette smoke and radon gas are prototypical examples. It is likely, however, that some pollutants may be considered as markers for the etiological components. In addition, regulatory action on certain selected pollutants can benefit health, as such actions can modify the complex mixture (154).

All-cause mortality

Our analysis showed that long-term exposure to $PM_{2.5}$ may increase the risk of nonaccidental mortality by 6% per a 10μg/m^3 increase, independent of age, gender, and geographic region. Living close to busy traffic appears also to be associated with elevated nonaccidental mortality. The evidence for the other measured gaseous and particulate pollutants was not persuasive.

Lung cancer

We conclude that long-term exposure to $PM_{2.5}$ is associated with an increased risk of mortality from lung cancer, that ranges between 15% and 21% per a 10μg/m^3 increase. There is also suggestive evidence that exposure to SO_2, (range: 7%-12% per a 10μg/m^3 increase) increases mortality but not sufficient evidence to make conclusions for the other pollutants. Living near heavy volumes of traffic may also increase risk. In addition, although elevated risks according to living in residential areas with high levels of air pollution were found in the 12 case-control studies, strong conclusions cannot be made because of concerns with bias (e.g., misclassification of exposure, confounding by strong risk factors).

Other sites of cancer

There is insufficient data to make any conclusions for any pollutant or site.

Cardiovascular diseases

We also found that increased concentrations of $PM_{2.5}$ increase mortality rates from cardiovascular illnesses (range: 12%-14% per a 10μg/m^3 increase). In addition, $PM_{2.5}$ may be positively associated with mortality from coronary heart diseases; this conclusion is based on only three studies, but all studies showed increased relative risks (range: 10%-16% per a

10μg/m^3 increase). For other particulate and gaseous pollutants, there is a paucity of data to draw conclusions. Living close to busy roads may increase the mortality of cardiovascular diseases, but there are only four studies (all showed positive effects).

Respiratory diseases

There is insufficient data to make any conclusions for any pollutant or site.

Potential Sources of Heterogeneity

Theoretically, one may expect that differences between studies in certain characteristics of the design and conduct of studies may lead to heterogeneity. For example, the composition of the study population, exposure assessment, duration of follow-up, differences in age, gender, socioeconomic status, and location, residual confounding, and misclassification of health outcomes may be important. However, one of the major issues with this literature is the small number of studies, and this limits the ability to detect at high levels of confidence departures of risk estimates from the average.

In terms of the composition of the cohort, all studies were drawn from the general population except for the US Veterans' Cohort Mortality Study (51) that enrolled veterans who were hypertensive. However, when we excluded this study from the pooled analyses of pollutants, we did not find any meaningful differences from the combined estimate.

Errors in measuring exposure can lead to differences in estimates of effect, particularly when different methods of exposure assessment are used. For example, different methods were used in the intraurban and interurban studies and we found some differences in effect. Specifically, the ACS-LA study (49) showed stronger associations between $PM_{2.5}$ and mortality of lung cancer and of all causes than did the ACS study (58)). In general, however, we did not find important differences by type of exposure assessment.

We found some differences in effect that were related to losses to follow-up. For example, in all-cause mortality and exposure to $PM_{2.5}$ we found a large difference in the RRs for the study that did not report loss of follow-up as compared to the other studies that reported

follow-up of over 90% of subjects. The interpretation of this finding, however, could be clouded by other unidentified sources of heterogeneity.

In terms of residual confounding, the risk estimates for lung cancer mortality from the Olso Cohort study (84), for which smoking was not accounted for, was dramatically higher than the other studies. This is an interesting finding because we observed that adjustment for smoking in the cohort studies lowered the relative risks between exposure to gaseous and particulate pollutant (e.g., NO_2, $PM_{2.5}$) and mortality from lung cancer (and any other causes) by only 2%-10% (27, 28, 48-50, 57, 58, 119, 126). (In addition, when exposure to traffic-related air pollution was estimated using proximity to busy roads, it was weakly associated with smoking because of low correlations between smoking and traffic in the studies (25, 62).) One possible explanation, other than residual confounding by smoking, is the lower socioeconomic status of subjects in the Oslo study.

The WHI study (28) showed higher RRs of the association with mortality from coronary heart diseases and exposure to $PM_{2.5}$ than the other studies, for which men and women were recruited. Differences in the distribution of gender in coronary diseases may contribute to this heterogeneity, such as in the WHI study, in which the analysis was restricted to postmenopausal women. Another possible source of the above-mentioned heterogeneity is the measurement of outcomes. In the WHI study, cardiovascular mortality was defined by review of medical records, including those from emergency, outpatient, and inpatient departments, and emergency medical services; autopsy and coroner records; and death certificates. Such extensive records were not available in the earlier studies (27, 82) that relied only on death certificates. It has been estimated that in the U.S. coronary heart disease may be overrepresented by approximately 20% overall and by as much as two-fold in older persons ($\geq$85 years old) as a cause of death on death certificates (156, 157). Such misclassification, if present in these studies, would likely have resulted in underestimate of the risk.

The misclassification of outcome in death certificates may have also attributed to the stronger association with incidence of lung cancer (ascertained using cancer registries) and NO_2 and SO_2 than that was found between mortality of lung cancer (ascertained using death

certificate) and the two pollutants. Although accuracy of death certificates in assigning lung cancer as the underlying cause of death is usually high (158), it is not surprise to observe such magnitude of differences in RRs because there were very few lung cancer cases in the studies. For example, there were a total of 36 lung cancer cases out of 6,338 subjects in the ASHMOG study (117, 122) and 382 out of 16,209 subjects in the Cohort study of Norwaygian Men (85, 126). As a result, the estimated RRs of lung cancer were more susceptible to the specificity of death certificates.

Biological Mechanisms

The mechanisms by which air pollution influences the risk of cardiovascular disease are still under investigation. Most of the research in the past 15 years has focused on the effects of particulates despite the evidence of associations for other pollutants, especially for acute effects. In any event, several potential pathophysiological pathways for the effects of particles have been suggested (56, 101). First, repeated inhalation of ambient PM may result in low-to-moderate-grade pulmonary oxidative stress and inflammation (159-162), which subsequently triggers systemic inflammatory responses with a cascade of reactions: production and mobilization of proinflammatory leukocytes and platelets into the circulation (163, 164); increase in circulating inflammatory mediators such as interleukins (IL)-6 (165, 166); and stimulation of the production of acute phase proteins such as C-reactive protein and fibrinogen (159). This induced systemic inflammatory response may in turn lead to increasing blood coagulability, accelerating atherosclerosis progression, and ultimately precipitating or aggravating cardiovascular events (167-169). This possible biological mechanism, proposed by Seaton et al. (168) and van Eeden et al. (169), has been supported by accumulated epidemiological evidence (170, 171) which have shown that exposure to ambient fine particles is positively associated with development of atherosclerosis. Second, it has been hypothesized that sustained inflammation of lung may also exacerbate preexisting lung diseases such as COPD, which further contributes to cardiovascular risk (172). A third hypothesized pathophysiological pathways suggests that ultrafine particulates and soluble components of fine particles may cross the pulmonary epithelium into the circulation, thus confer a direct effect on cardiovascular system by adversely altering cardiac autonomic function: contribute to the instability of a vascular plaque or initiate cardiac arrhythmias (56,

101). There is evidence that decreases in resting heart rate, an indicator of cardiac autonomic function, are strongly associated with increased risk of cardiovascular morbidity and mortality in the elderly and those with significant heart disease (173, 174). This pathway may explain both the acute and the long-term cardiovascular effects of particulate pollution.

In contrast, it would appear that the gaseous pollutants should not play a great role in carcinogenesis because they are not mutagenic, although certain volatile organic compounds (e.g., benzene) and polycyclic aromatic hydrocarbons are both mutagenic and carcinogenic. The biological mechanisms underlying the association between particulate pollution and lung cancer have not been elucidated. It has been hypothesized that lung cancer develops through a series of progressive pathological changes occurring in the respiratory epithelium as a result of direct genotoxicity effects of particulate air pollution (175, 176). Many studies have shown that urban air pollution, in particular polycyclic aromatic hydrocarbons, may cause cancer mediated through the formation of DNA adducts (177-180). Polycyclic aromatic hydrocarbons are associated with fine particulates (142). It was found that more than 90% of the particulate phase of polycyclic aromatic hydrocarbons are associated physically with particulate <3.3 μm (181). In addition, there is evidence that other constituents of fine particles may cause oxidative DNA damage (182, 183).

Limitations of the data

Although we identified a total of 17 cohort studies and 20 case-controls studies between 1950 and 2007, not all the studies could be included in our analysis due to diverse types of pollutants and metrics used to characterize air pollution exposure. In addition, we did not include the results from Finkelstein et al. (64, 118) in our subgroup analysis because their RRs could not be translated to a 10 ug/m^3 increase. This hampered us to explain fully heterogeneity of effect estimates across the studies. We identified that difference in gender and losses to follow-up may be potential sources of heterogeneity. However, we could not confirm whether the heterogeneity was caused by these characteristics, or some other related design characteristics.

Recommendations

In comparison to other environmental exposures, such as cigarette smoking or radon gas, there is indeed a paucity of epidemiological data. Although the data suggest that certain outcomes are indeed associated with fine particles, there is insufficient information to make solid conclusions regarding many other pollutants or the mixture itself. We thus recommend that considerable effort be made into developing new studies with large sample sizes and high quality assessments of exposure.

Figure Legends

Figure 1: Flow diagram of search strategy and selection process in the systematic review of studies of the relationship between long-term exposure to ambient air pollution and chronic diseases, 1950-2007.

Figure 2. Associations between all-cause mortality (ICD9: <800) and long-term exposure to (A) gaseous pollutants, (B) particulates with 10 $\mu g/m^3$ difference of concentrations (M: men; W: women; L.I.: subjects with below-median income; H.I.: subjects with above-median income; NS/NR: estimates were reported as non-significant, but no quantitative results were provided; * Total deaths in this study may include accidental deaths; ‡ RRs and 95% CI were based on comparison between above- and below-median concentration of air pollutants)

Figure 3. Associations between lung cancer (ICD9: 162) and long-term exposure to (A) gaseous pollutants, (B) particulate pollutants with 10 $\mu g/m^3$ difference of concentrations (M: men; W: women; Mort.: mortality; Incid.: incidence; NS/NR: RRs were reported as non-significant, but no quantitative results were provided; NR: not reported; UCI: upper 95% CI; * Naess et al. 2007 /66/ reported estimates based on per quartile increase which were grouped by gender and age groups. For $PM_{2.5}$, the difference between quartiles was about 4 $\mu g/m^3$ and thus the estimates were translated to a 10 $\mu g/m^3$ increase in the subgroup analysis (see Results section))

Figure 4. Associations between (A) cardiovascular mortality (ICD9: 400-440), (B) cardiopulmonary mortality (ICD9: 400-440 and 460-519), (C) mortality from coronary heart disease (ICD9: 410-414) or cerebrovascular disease (ICD9: 430-438) and long-term exposure to gaseous and particulate pollutants with 10 $\mu g/m^3$ difference of concentrations (M: men; W: women; UCI: upper 95% CI; * Outcome included first occurrence of myocardial infarction, coronary revascularization, stroke, and death from coronary heart or cerebrovascular disease, § Outcome included only deaths from coronary heart disease and cerebrovascular disease; || Outcome included all cardiovascular diseases plus diabetes; ‡ Naess et al. 2007 /66/ reported estimates based on per quartile increase which were grouped by gender and age groups. For $PM_{2.5}$, the difference between quartiles was about 4 $\mu g/m^3$

and thus the estimates were translated to a 10 μg/m^3 increase in the subgroup analysis (see Results section); † Finkelstein et al. 2003 /52/ reported estimates based on comparison between above- and below-median concentration of air pollutants)

Figure 5. Associations between (A, B) mortality from all respiratory diseases (ICD9: 460-519), (C) incidence of asthma (ICD9: 493) and long-term exposure to gaseous and particulate pollutants with 10 μg/m^3 difference of concentrations (M: men; W: women; NS/NR: RRs were reported as non-significant, but no quantitative results were provided; UCI: upper 95% CI; § Outcome included respiratory diseases with ICD9: 485-49)

Figure 6: Flow diagram of search strategy and selection process in the systematic review of studies of the relationship between long-term exposure to ambient air pollution and chronic diseases, 01/2008-04/2011.

Figure 7. Associations between (A) cardiovascular mortality, (B) cardiopulmonary mortality, (C) mortality from coronary heart disease or cerebrovascular disease, (D) Incidence of coronary heart disease and in particular myocardial infarction and long-term exposure to gaseous and particulate pollutants with 10 μg/m^3 difference of concentrations (* For organic carbon, black carbon, SO_4, levels of traffic particles, and cumulative traffic count, the rate ratios were estimated per inter-quartile range increase, depending on the study. For CO, the rate ratios were estimated for a change in the concentration from the fifth to the 95th percentile.)

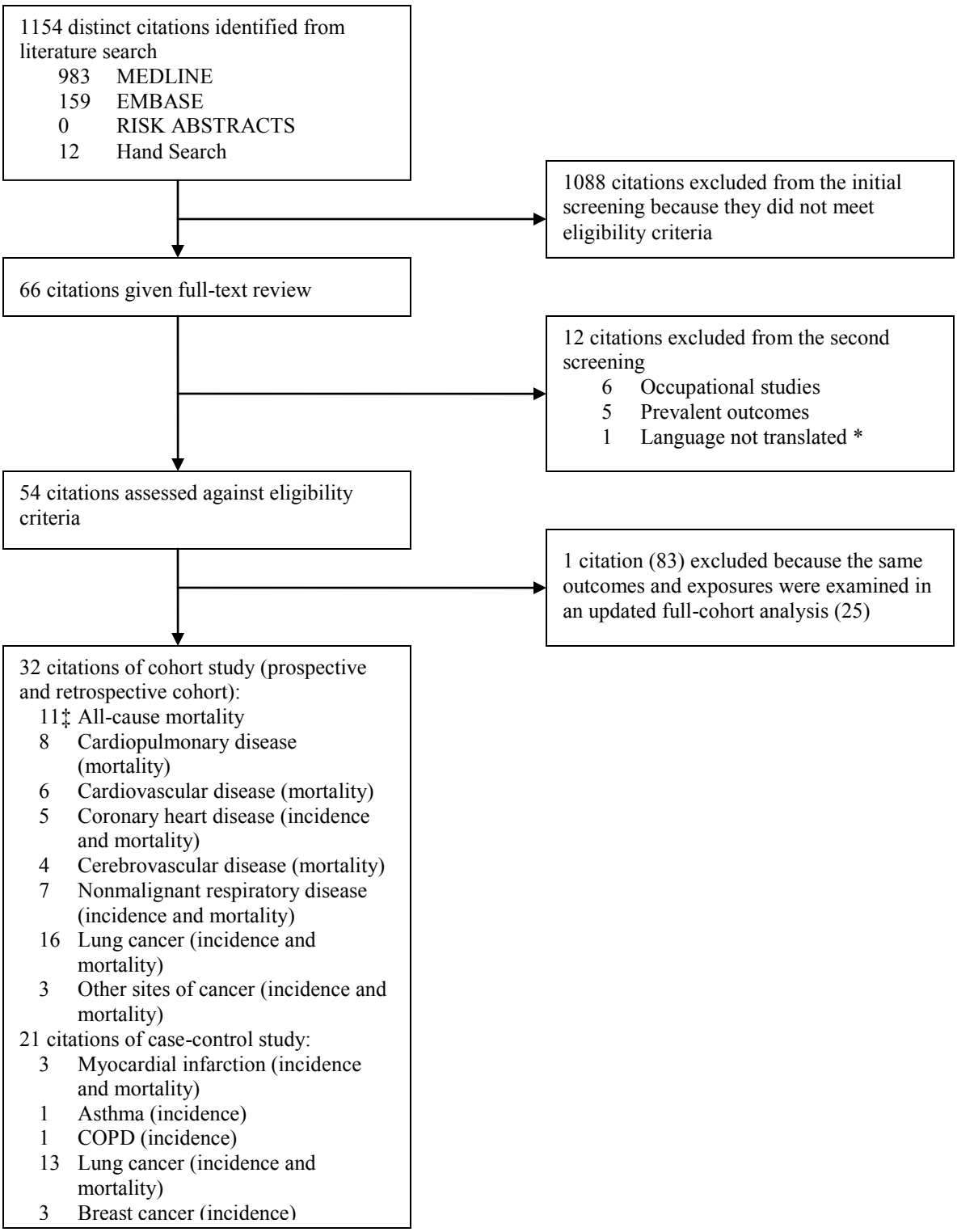
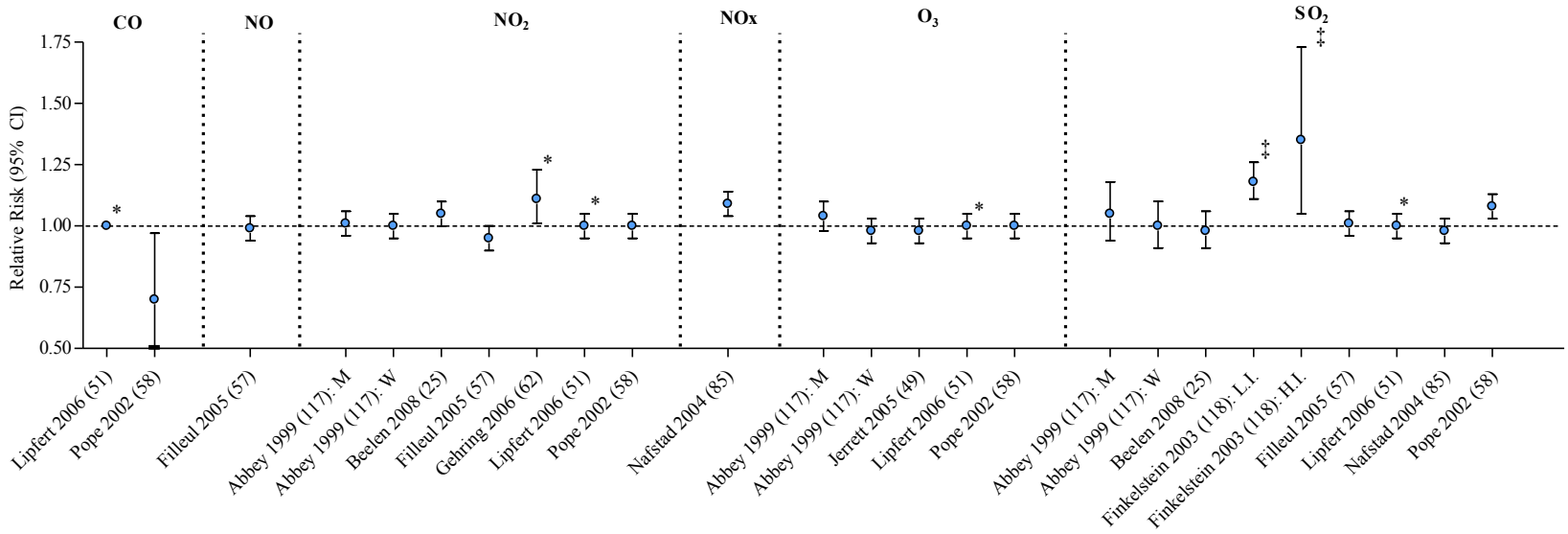


Figure 1: Flow diagram of search strategy and selection process in the systematic review of studies of the relationship between long-term exposure to ambient air pollution and chronic diseases, 1950-2007.

* There are a total of three non-English articles identified through literature search: 2 Italian (183, 184) and 1 Russian (185). We were able to translate the Italian articles but not the Russian one.

‡ Numbers refer to number of papers. The numbers under each category do not total 32 because multiple outcomes may have been presented in a publication. Each study may generate more than one publication.

A. Gaseous pollutants



B. Particulate pollutants

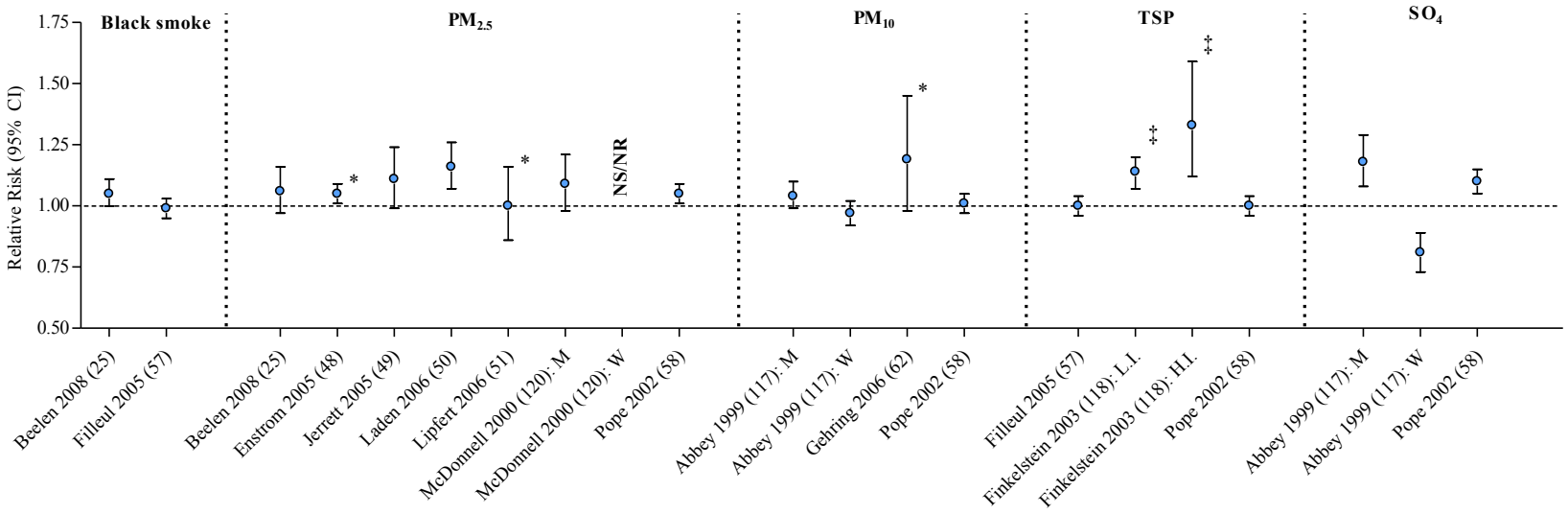
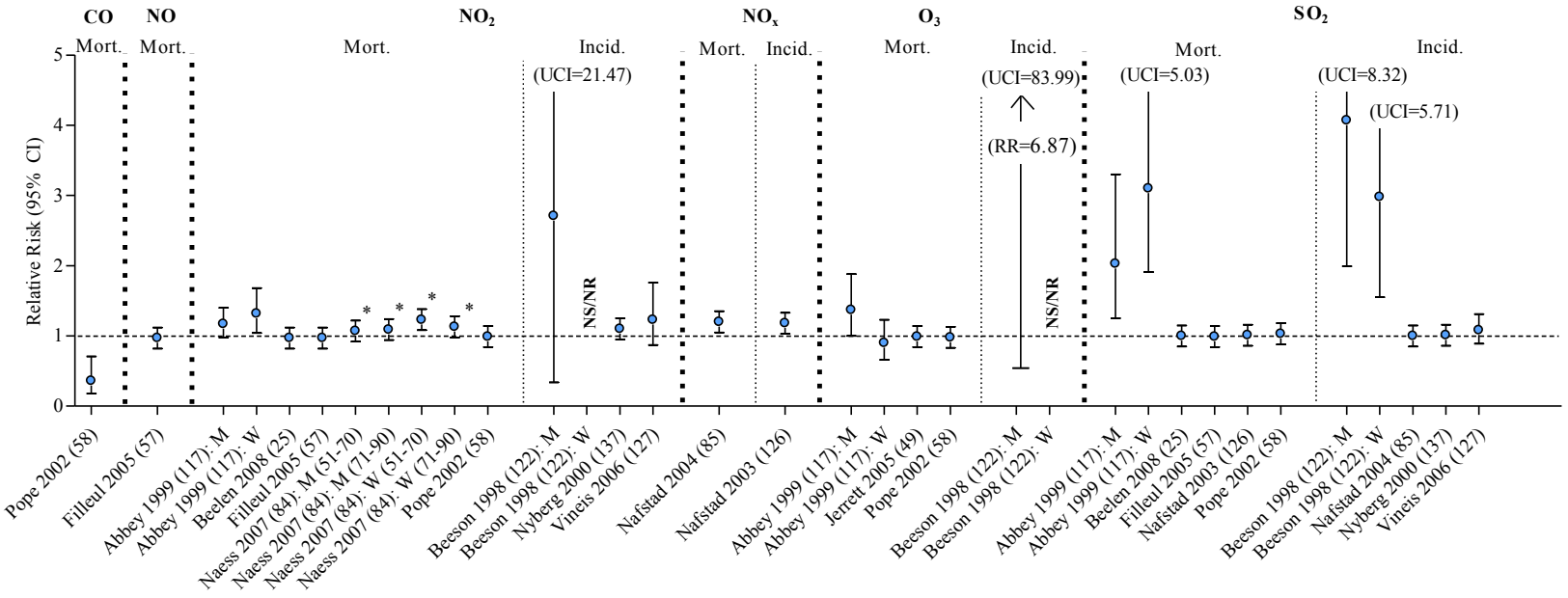


Figure 2. Associations between **all-cause mortality (ICD9: <800)** and long-term exposure to **(A) gaseous pollutants**, **(B) particulates** with 10 μg/m^3 difference of concentrations (M: men; W: women; L.I.: subjects with below-median income; H.I.: subjects with above-median income; NS/NR: estimates were reported as non-significant, but no quantitative results were provided; * Total deaths in this study may include accidental deaths; ‡ RRs and 95% CI were based on comparison between above- and below-median concentration of air pollutants)

A. Gaseous pollutants



B. Particulate pollutants

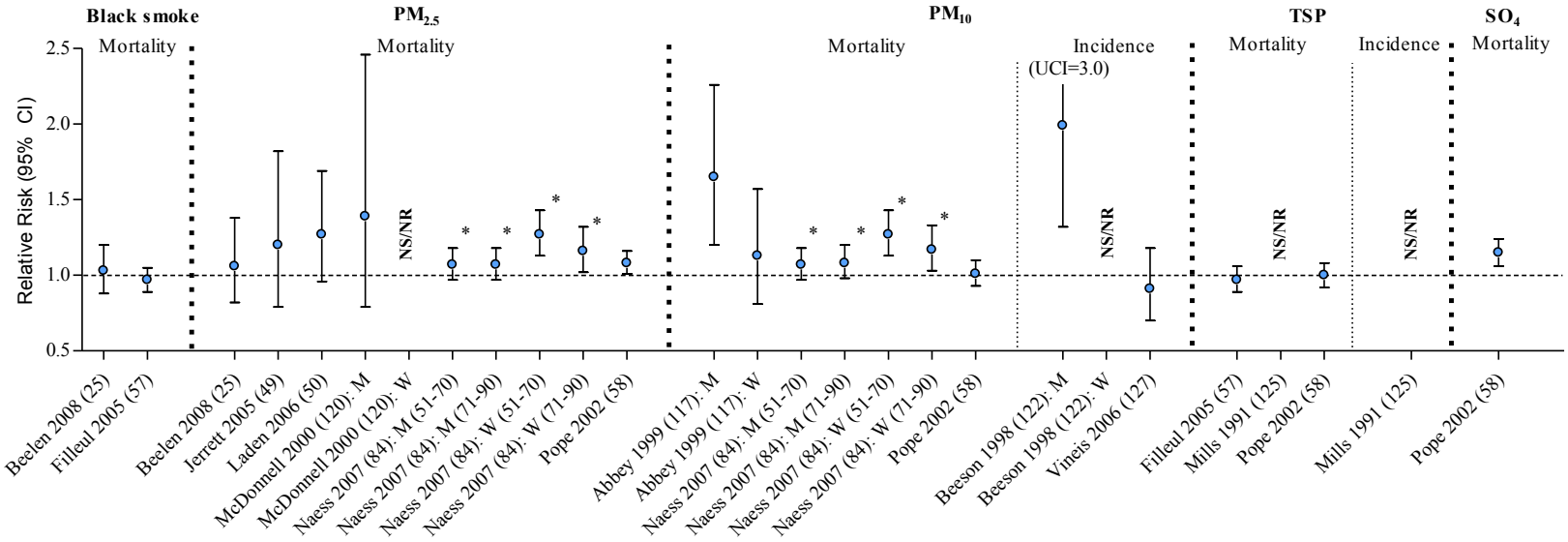
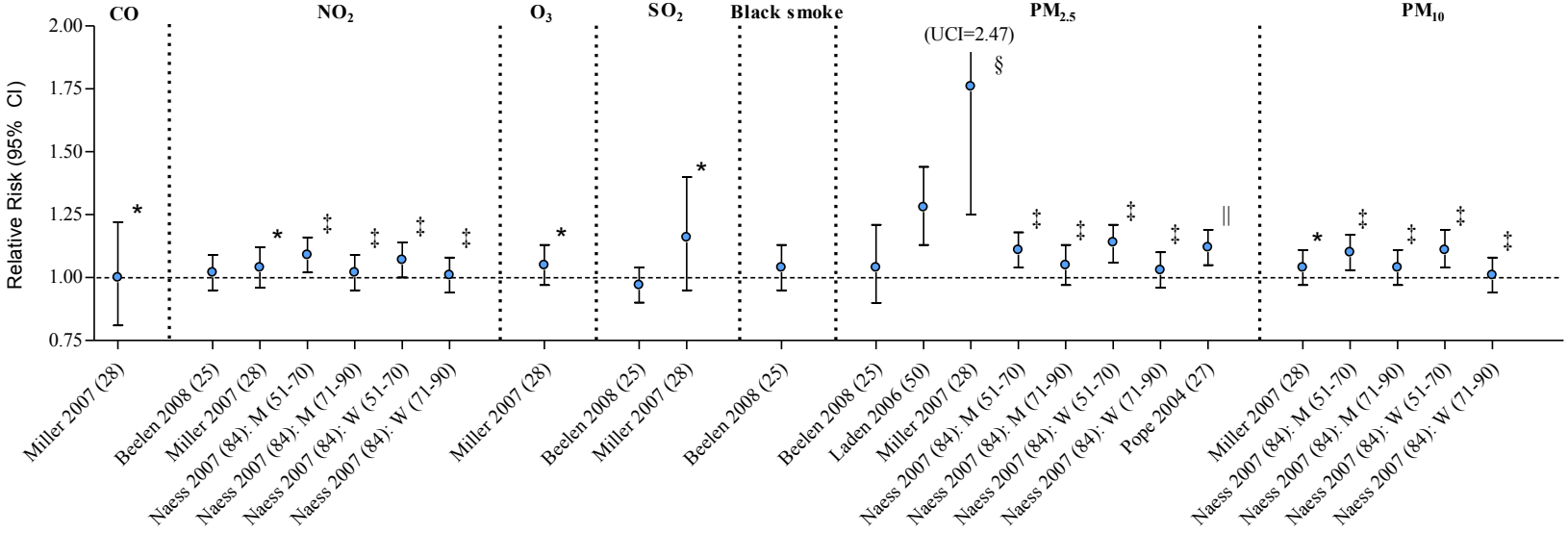
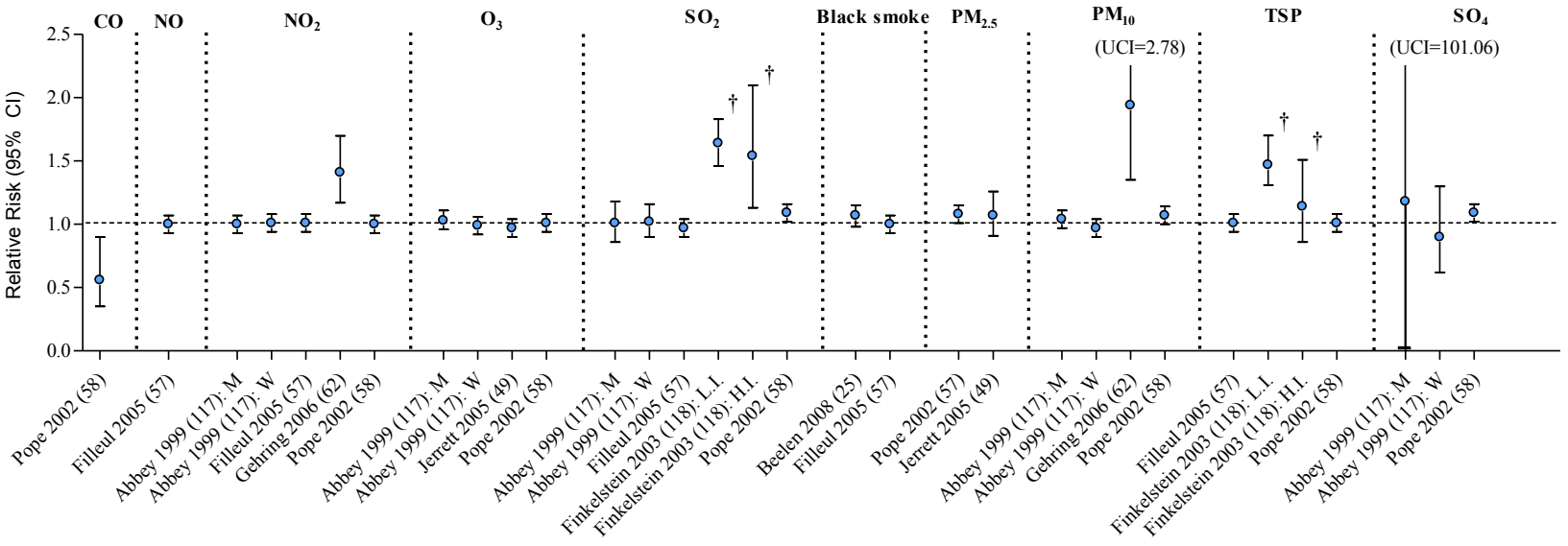


Figure 3. Associations between **lung cancer (ICD9: 162)** and long-term exposure to **(A) gaseous pollutants, (B) particulate pollutants** with 10 μg/m^3 difference of concentrations (M: men; W: women; Mort.: mortality; Incid.: incidence; NS/NR: RRs were reported as non-significant, but no quantitative results were provided; NR: not reported; UCI: upper 95% CI; * Naess et al. 2007 /66/ reported estimates based on per quartile increase which were grouped by gender and age groups. For $PM_{2.5}$, the difference between quartiles was about 4 μg/m^3 and thus the estimates were translated to a 10 μg/m^3 increase in the subgroup analysis (see Results section))

A. Cardiovascular mortality



B. Cardiopulmonary mortality



C. Mortality from coronary heart disease or cerebrovascular disease

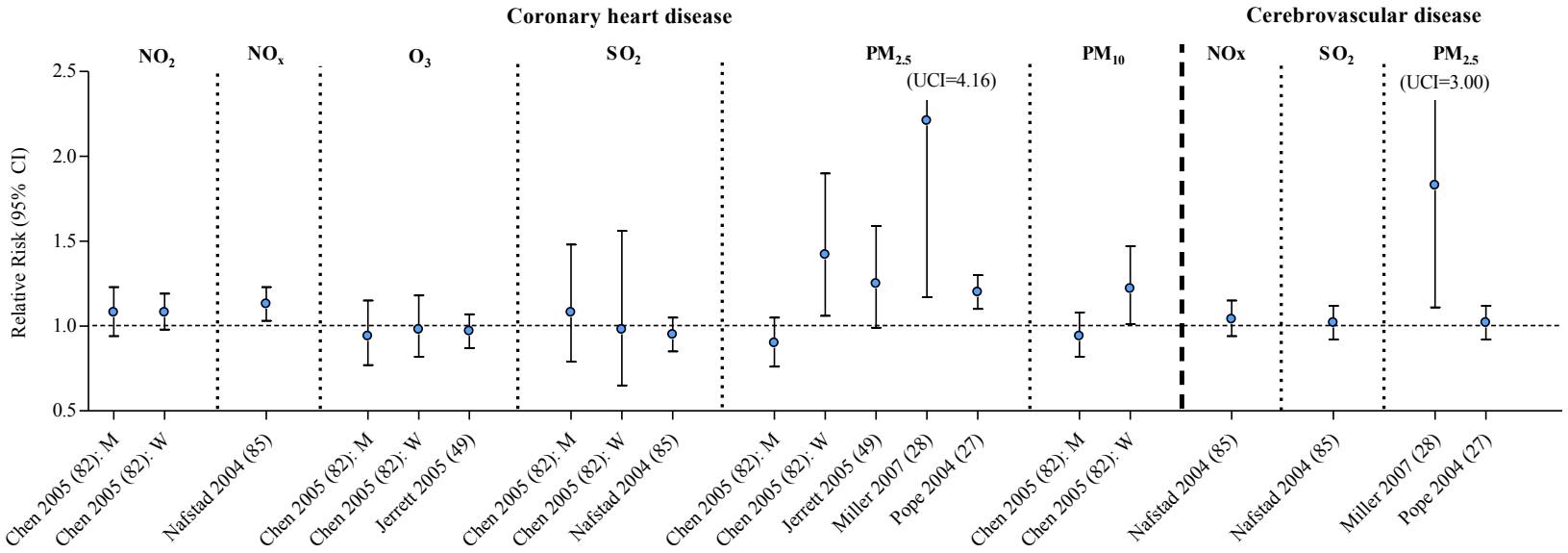
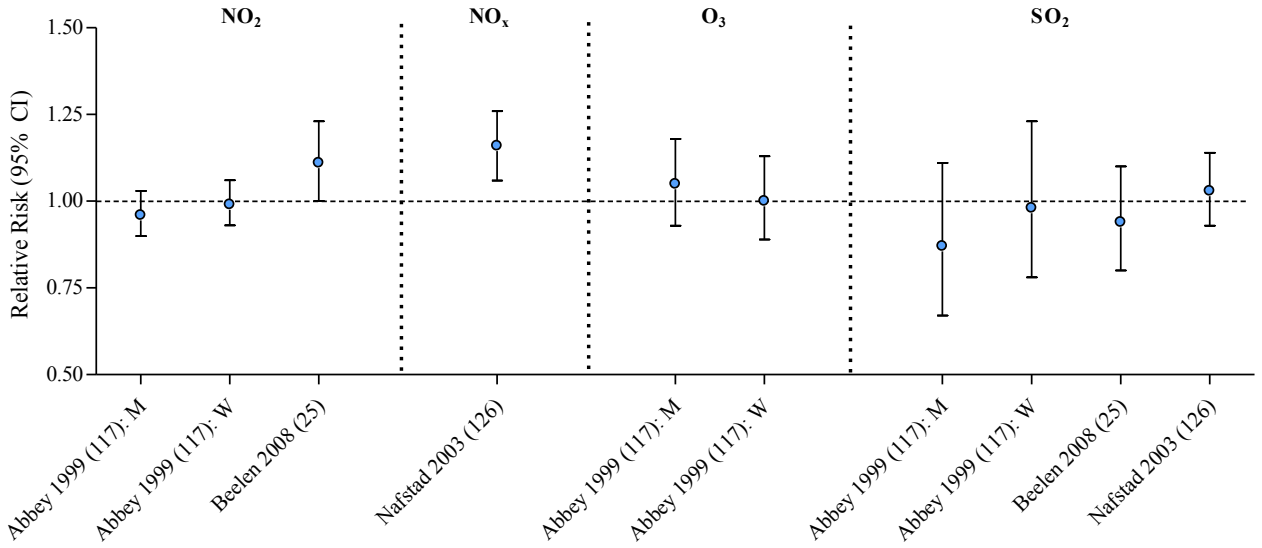
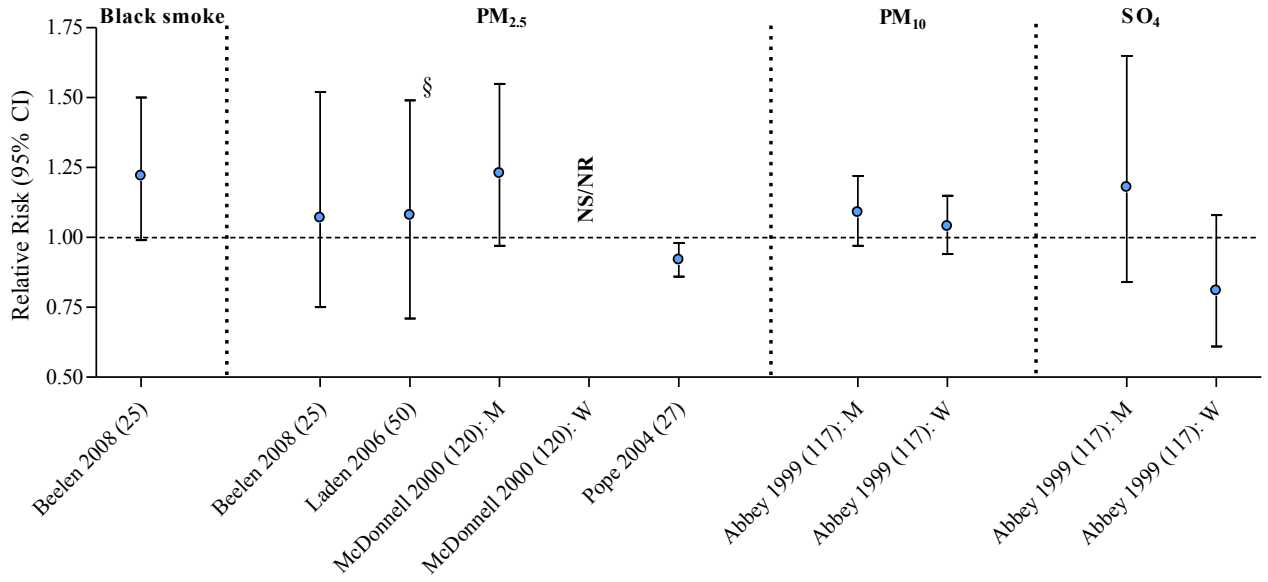


Figure 4. Associations between **(A) cardiovascular mortality (ICD9: 400-440), (B) cardiopulmonary mortality (ICD9: 400-440 and 460-519), (C) mortality from coronary heart disease (ICD9: 410-414) or cerebrovascular disease (ICD9: 430-438)** and long-term exposure to **gaseous and particulate pollutants** with 10 μg/m^3 difference of concentrations (M: men; W: women; UCI: upper 95% CI; * Outcome included first occurrence of myocardial infarction, coronary revascularization, stroke, and death from coronary heart or cerebrovascular disease, § Outcome included only deaths from coronary heart disease and cerebrovascular disease; || Outcome included all cardiovascular diseases plus diabetes; ‡ Naess et al. 2007 (84) reported estimates based on per quartile increase which were grouped by gender and age groups. For $PM_{2.5}$, the difference between quartiles was about 4 μg/m^3 and thus the estimates were translated to a 10 μg/m^3 increase in the subgroup analysis (see Results section); † Finkelstein et al. 2003 (118) reported estimates based on comparison between above- and below-median concentration of air pollutants)

A. Mortality from all respiratory diseases and gaseous pollutants



B. Mortality from all respiratory diseases and particulate pollutants



C. Incidence of asthma and gaseous and particulate pollutants

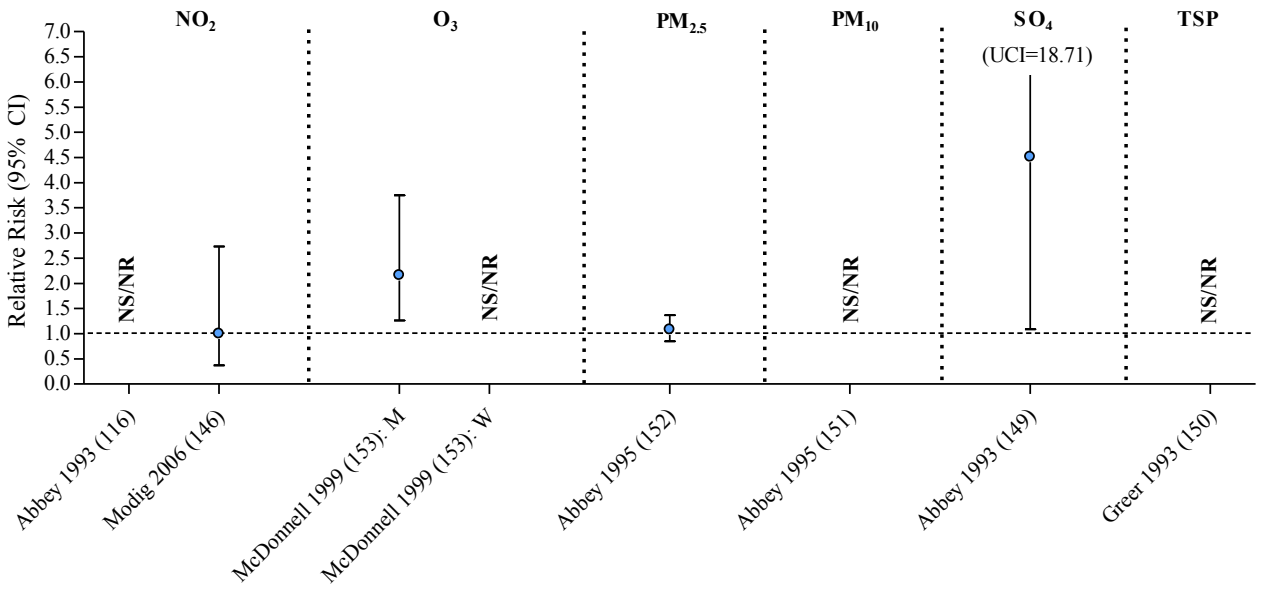


Figure 5. Associations between **(A, B) mortality from all respiratory diseases (ICD9: 460-519), (C) incidence of asthma (ICD9: 493)** and long-term exposure to **gaseous and particulate pollutants** with 10 μg/m^3 difference of concentrations (M: men; W: women; NS/NR: RRs were reported as non-significant, but no quantitative results were provided; UCI: upper 95% CI; § Outcome included respiratory diseases with ICD9: 485-49)

TABLE 1. Selected characteristics of cohort studies of long-term exposure to ambient air pollution and chronic disease, according to distinct study cohort, 1950-2007

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
Adventist Health Study on the Health Effects of Smog (ASHMOG)											
Mills 1991 (125)	US	6340	≥25	36	Incidence and mortality of cancer: all malignant neoplasm colon rectum lung leukemia/lymphoma breast cervix	Incidence: 301 39 15 17 27 65 5	Incidence: 112 17 3 12 14 0 0	*Incidence:* postal questionnaire, medical records; *Mortality:* death certificates, church records	*Incidence*: 1977-82 *Mortality*: 1977-86	87	99
Greer 1993 (150)	US	3914	≥25	36	Incidence: asthma	78	27	Postal questionnaire (MD‡ or symptoms)	1977-87	87	~91
Abbey 1993 (116)	US	3914	≥25	36	Incidence: all cancers myocardial infarction obstructive airway disease chronic bronchitis symptom asthma Mortality: all-natural causes	NR‡	NR	*Incidence*: postal questionnaire, tumor registry *Mortality*: death certificates, NDI‡, church records	1977-87	87	77
Abbey 1993 (149)	US	3914	≥25	36	Incidence: obstructive airway disease chronic bronchitis symptom asthma	NR	NR	Postal questionnaire (MD or symptoms)	1977-87	87	77
Abbey 1995 (151)	US	3914	≥25	36	Incidence: obstructive airway disease chronic bronchitis symptom asthma	289 355 84	114 153 30	Postal questionnaire (MD or symptoms)	1977-87	87	77
Abbey 1995 (152)	US	1868	≥25	NR	Incidence: obstructive airway disease chronic bronchitis symptom asthma	135 117 40	NR	Postal questionnaire (MD or symptoms)	1977-87	87	77
Beeson 1998 (122)	US	6338	29-95	36	Incidence: lung cancer	36	16	Cancer registry, medical records	1977-92	87	98
McDonnell 1999	US	2881	27-87	35	Incidence: asthma	111	32	Postal	1977-92	87	49

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
(153)								questionnaire (MD or symptoms)			
Abbey 1999 (117)	US	6338	29-95	36	Mortality: all natural causes	1575	610	Death certificates	1977-93	87	98
					cardiopulmonary	1029	398				
					lung cancer	30	18				
					any mention of nonmalignant respiratory disease	410	164				
McDonnell 2000 (120)	US	3769	27-87	36	Mortality: all natural causes	917	359	Death certificates, church records	1977-92	87	98
					lung cancer	244	98				
					any mention of nonmalignant respiratory disease	24	14				
Chen 2005 (82)	US	3239	Mean = 57.4	36	Mortality: coronary heart disease	250	95	Death certificates, NDI, church records	1977-98	87	~100
American Cancer Society (ACS) study											
Pope 2002 (58)	US	359000	≥30	NR	Mortality: all causes	80775	NR	Death certificates, NDI	1982-99	NR	93
					cardiopulmonary	40388					
					lung cancer	6462					
					other causes	33926					
Pope 2004 (27)	US	319000	≥30	NR	Mortality: all cardiovascular + diabetes	32371	NR	Death certificates, NDI	1982-98	NR	93
					all respiratory disease	5886					
					ischemic heart disease	17011					
					cerebrovascular disease	4881					
					COPD‡ and allied conditions	2943					
Jerrett 2005 (49) (Los Angeles)	US	22905	≥30	NR	Mortality: all causes	5856	NR	Death certificates, NDI	1982-00	NR	~86
					ischemic heart disease	NR					
					cardiopulmonary	NR					
					lung cancer	NR					
					digestive cancer	NR					
					other cancers	NR					
Study of American Legion Population											
Buell 1967 (123)	US	69868	≥25	100	Mortality: lung cancer	304	304	California death file	1958-62	~50	~99
California Cancer Prevention Study (CA CPS I)											
Enstrom 2005 (48)	US	38524	43-99 (mean = 65)	42	Mortality: all causes	28441	13532	California death file, nationwide Social Security Death Index	1973-02	NR	93

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
Cohort of Norwegian Men											
Nafstad 2003 (126)	Norway	16209	40-49	100	Incidence: lung cancer	418	418	Cancer registry	1972-98	63	~100
Nafstad 2004 (85)	Norway	16209	40-49	100	Mortality: all natural causes	4227	4227	National death registry	1972-98	63	~100
					respiratory diseases	200	200				
					lung cancer	382	382				
					cerebrovascular disease	258	258				
					ischemic heart disease (includ. sudden death)	1508	1508				
Danish Cohort Study on Lung Cancer, Smoking, and Environment (retrospective)											
Engholm 1996 (124)	Denmark	1413600	30-64	66	Incidence: lung cancer	NR	22534	Cancer registry	1970-87	100	~100
French PAARC study											
Filleul 2005 (57)	France	14284	25-59 (mean =42)	48	Mortality: all natural causes	2396	1590	National death registry, National Inst. of Health and Medical Research	*All-cause:* 1974-01 *Cause-specific:* 1974-98	NR	81
					cardiopulmonary	546	406				
					lung cancer	178	153				
GenAir study											
Vineis 2006 (127)	Ten European countries	1008	35-74	~33	Incidence: lung cancer	271	91	Cancer registry, death registry, personal contact, hospital discharge data	1993-98	NR	NR
Harvard Six Cities Study											
Krewski 2005 (119)	US	8111	25-74 (mean =49)	46	Mortality: all causes	1430	830	Personal contact, death certificates	1974-89	NR	~99
					cardiopulmonary	772	NR				
					lung cancer	114	NR				
					all others	543	NR				
Laden 2006 (50)	US	8096	25-74	46	Mortality: all causes	2732	NR	Next-of-kin reports, Social Security records, NDI	1974-98	NR	NR
					cardiovascular	1196					
					respiratory	195					
					lung cancer	226					
					all others	1115					
Netherlands Cohort Study on Diet and Cancer (NLCS)											
Beelen 2008 (25)	Nether.	120852	58-66 (mean =61)	48	Mortality: all causes	17286	NR	Central Bureau of Genealogy, Central Bureau	1987-96	36	~100
					cardiovascular	7153					
					cardiopulmonary	6137					

First author and year (ref. no.)	Location	Study population: Sample size	Age range	% men	Outcome(s)	No. of cases*: Total	Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
					respiratory	1016		of Statistics			
					lung cancer	1888					
German Women's Health Study in North Rhine-Westphalia											
Gehring 2006 (62)	Germany	4874	50-59	0	Mortality: all causes	399	0	Registration	1985-03	70	97
					cardiopulmonary	139	0	office, death			
					non-cardiopulmonary and non-lung cancer	229	0	certificates			
Ontario Cohort Study (retrospective)											
Finkelstein 2003 (118)	Canada	5228	≥40	44	Mortality: all natural causes	604	NR	Ontario death	1992-99	100	~98
					cardiopulmonary	341		registry			
Finkelstein 2005 (64)	Canada	5228	≥40	44	Mortality: cardiovascular	252	NR	Ontario death	1992-01	100	~98
					cerebrovascular	58		registry			
					respiratory	232					
Oslo Cohort Study (retrospective)											
Naess 2007 (84)	Norway	143842	51-90	41	Mortality: cardiovascular	14964	6538	Death certificates	1992-98	100	~100
					lung cancer	1453	873				
					COPD	1355	636				
Pennsylvania Cohort											
Morris 1976 (121)	US	NR	≥30	NR	Mortality: all causes	253	165	Personal contact, local news paper, death registry	1960-72	75	75
Residents of Lancaster and Glendora in California											
Detels 1987 (148)	US	4275	7-59	NR	Lung function:	NR	NR	Lung function	*Lancaster*	NR	*Lancaster*
					FEV_1‡			test,	1972-77		47
					$FEV_{25-75\%}$‡			Interview/	*Glendora*		*Glendora*
					V_{50}‡			questionnaire	1977-82		58
					V_{75}‡						
					$\Delta N_{2\ 750-1250}$‡						
					Incidence: symptoms of cough, wheezing, physician-diagnosed asthma, bronchitis and emphysema						
USEPRI-Washington University Veterans' Cohort Mortality Study											
Lipfert 2006 (51)	US	67108	36-63 (mean =51)	100	Mortality: all causes	NR	NR	U.S. Veterans Administration's Beneficiary Identification Records	1997-01	NR	~100

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
								Location System			
Women's Health Initiative Study											
Miller 2007 (28)	US	65893	50-79	0	Incidence: any first cardiovascular event (i.e. myocardial infarction, coronary revascularization, stroke) Mortality: any cardiovascular event (coronary heart or cerebrovascular)	1816 261	0 0	Postal questionnaire, medical records, NDI	1994-03	NR	90

* No. of cases are listed in the same order as outcomes

‡ NR, not reported; MD, refers to physician; NDI, U.S. national death index; COPD, chronic obstructive pulmonary disease; FEV_1, forced expiratory volume in one second; $FEV_{25-75\%}$, forced expiratory flow 25% to 75%; $\Delta_N 2_{750-1250}$, change in the percentage of nitrogen in the expired air at 750 and 1250 ml of expirate

TABLE 2. Exposure assessment in cohort studies of long-term exposure to ambient air pollution and chronic disease, according to distinct study cohort, 1950-2007

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Adventist Health Study on the Health Effects of Smog (ASHMOG)					
Mills 1991 (125)	*Incidence:* 1977-82 *Mortality:* 1977-86	1973-77	TSP‡, O_3	1. Interpolated using IDW‡ with $1/R^2$ (R: distance from centroid of a Zip Code area to a nearby fixed-site station) from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barriers* as the nearby stations	TSP: (1) cumulative annual mean concentration (μg/m³), (2) annual average exceedance frequency above several cutoffs (sum of hr/yr above cutoffs) O_3: (1) cumulative annual mean concentration (pphm‡), (2) annual average exceedance frequency above several cutoffs (sum of hr/yr above cutoffs)
Greer 1993 (150)	1977-87	1977-87	TSP, O_3	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations	TSP: cumulative annual mean concentration (μg/m³) O_3: cumulative annual mean concentration (pphm)
Abbey 1993 (116)	1977-87	1973-87	NO_2	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations 3. Estimated NO_2 was further adjusted with a regression model formed in an unrelated study using building characteristics and individual activity patterns	NO_2: (1) cumulative monthly average concentration (pphm), (2) cumulative exceedance frequency above several cutoffs (sum of hr/yr above cutoffs)
Abbey 1993 (149)	1977-87	1977-87	SO_4	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations	SO_4: (1) cumulative monthly average concentration (μg/m³), (2) cumulative exceedance frequency above several cutoffs (sum of d/yr above cutoffs)
Abbey 1995 (151)	1977-87	1973-87	PM_{10}‡	1. PM_{10} was imputed from TSP measurements at fixed-site stations (1973-87) using regression equations 2. PM_{10} was then interpolated using IDW approach with $1/R^2$ 3. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations	PM_{10}: (1) cumulative monthly average concentration (μg/m³), (2) cumulative exceedance frequency above several cutoffs (sum of hr/year above cutoffs)

First author, year (ref)	Follow-up period	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Abbey 1995 (152)	1977-87	1966-86	$PM_{2.5}$‡	$PM_{2.5}$ was imputed from adjacent airport visibility data (1966-86) at the centroid of Zip Code areas using season- and site-specific regression estimations	$PM_{2.5}$: (1) cumulative monthly average concentration (μg/m³), (2) cumulative exceedance frequency above several cutoffs (sum of hr/year above cutoffs)
Beeson 1998 (122)	1977-92	1973-92	PM_{10}, O_3, SO_2, NO_2	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations 3. For PM_{10} prior to 1987, it was imputed from TSP at fixed site stations using site- and season-specific regression	PM_{10}, SO_2, NO_2: cumulative monthly average concentration PM_{10}: cumulative exceedance frequency above several cutoffs (sum of d/yr above cutoffs) O_3: cumulative monthly average of the daily 8-hr average (9am-5pm) O_3: cumulative exceedance frequency above several cutoffs (sum of hr/yr above cutoffs)
McDonnell 1999 (153)	1977-92	1973-92	O_3	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations	O_3: (1) cumulative monthly average concentration (ppb‡), (2) 8-hr average concentration between the hours of 9 am to 5 pm (ppb), (3) cumulative exceedance frequency above several cutoffs (sum of hr/yr above cutoffs)
Abbey 1999 (117)	1977-93	PM_{10}, O_3, SO_2, NO_2: 1966-92 SO4: 1973-92	PM_{10}, O_3, SO_2, NO_2, SO_4	1. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 2. Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations 3. For PM_{10} prior to 1987, it was imputed from site- and season-specific regression based on TSP	PM_{10}, O_3, SO_2, NO_2, SO_4: cumulative monthly average concentration PM_{10}: cumulative exceedance frequency above several cutoffs (sum of d/yr above cutoffs) O_3: cumulative exceedance frequency above several cutoffs (sum of hr/yr above cutoffs)
McDonnell 2000 (120)	1977-92	1973-77	PM_{10}, $PM_{2.5}$, $PM_{2.5-10}$	1. For PM_{10} prior to 1987, it was imputed from site- and season-specific regression based on TSP. For PM_{10} after 1987, measurements of PM_{10} at central sites were used 2. For $PM_{2.5}$, it was imputed using season- and site-specific regression estimation from adjacent airport visibility data between 1966 and 1987 3. Interpolated using IDW with $1/R^2$ from ≤3 nearest stations 4. Interpolation was restricted to Zip Code centroid areas within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations	PM_{10}, $PM_{2.5}$, $PM_{2.5-10}$: cumulative monthly average concentration (μg/m³)

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Chen 2005 (82)	1977-98	1973-98	PM_{10}, $PM_{2.5}$, O_3, NO_2, SO_2	1. PM_{10}, NO_2, O_3 and SO_2: (1) IDW with $1/R^2$ from ≤ 3 nearest stations; (2) Interpolation was restricted to Zip Code centroids within 50 km of a monitoring station and on the same side of topographic barrier as the nearby stations 2. For $PM_{2.5}$ at the centroid of Zip Code areas, it was imputed from adjacent airport visibility data (1966-87) using season- and site-specific regression estimations	PM_{10}, $PM_{2.5}$, O_3, NO_2, SO_2: cumulative monthly average concentration
American Cancer Society (ACS) study					
Pope 2002 (58)	1982-99	$PM_{2.5}$: Avg. of 1979-83 and 1999-00 PM_{10}: 1987-96 $PM_{15\text{-}2.5}$, PM_{15}, TSP: 1979-83 SO_4: 1980-81 SO_2: 1980 NO_2: 1980 CO, O_3: 1982-98	$PM_{2.5}$, PM_{10}, $PM_{15\text{-}2.5}$, TSP, SO_4, SO_2, NO_2, CO, O_3	1. Annual mean concentrations of each pollutant at each fixed site in a metropolitan area were averaged to obtain metropolitan area-specific concentrations 2. Each subject was assigned a metropolitan area of residence based on address at time of enrollment and 3-digit Zip Code area 3. Each subject was assigned metropolitan area-specific concentrations	$PM_{2.5}$, PM_{10}, $PM_{15\text{-}2.5}$, PM_{15}, TSP, SO_4, SO_2, NO_2, CO, O_3: cumulative annual mean concentration
Pope 2004 (27)	1982-98	1979-83 1999-00	$PM_{2.5}$	1. Annual mean concentrations of each pollutant at each fixed site in a metropolitan area were averaged to obtain metropolitan area-specific concentrations 2. Each subject was assigned a metropolitan area of residence based on address at time of enrollment and 3-digit Zip Code area 3. Each subject was assigned metropolitan area-specific concentrations	$PM_{2.5}$: cumulative annual mean concentrations in several exposure periods ($\mu g/m^3$)
Jerrett 2005 (49) (Los Angeles)	1982-00	$PM_{2.5}$: 2000 O_3: 1999-01 Distance to freeways: 1982	$PM_{2.5}$, O_3, distance to freeways	1. $PM_{2.5}$ was interpolated using universal kriging with multiquadric models 2. O_3 was interpolated using universal kriging based on the average of the 4 highest 8-hr conc. (2000) or the expected peak daily conc. of O_3 (1999-01) 3. Distance to freeways was defined as nearest distance from Zip Code centroids from freeways	$PM_{2.5}$, O_3: cumulative annual mean concentration Distance to freeways: binary classification ($<$500m vs. $\geq$500m)
American Legion Study Population					
Buell 1967 (123)	1958-62	In 1950s	Oxidants	1. Based on observations of the gross presence of oxidants in the 1950s, LA County had the highest conc. of oxidants than other areas of California 2. Exposure was assigned based on county of residence (no. of years) from postal questionnaire	Contrast between regions of residence: binary classification (Los Angeles County and other counties of California excluding major metropolitan areas)
California Cancer Prevention Study (CA CPS I)					

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Enstrom 2005 (48)	1973-02	1979-83	$PM_{2.5}$	1. $PM_{2.5}$ data for each county of California were averaged during 1979-83 and across the available monitoring stations 2. The average county-level $PM_{2.5}$ value was assigned to the subjects alive as of 1973 based on their county of residence as of late 1972	$PM_{2.5}$: cumulative annual average concentration in selected exposure periods ($\mu g/m^3$)
Cohort of Norwegian Men					
Nafstad 2003 (126)	1972-98	1974-78	NOx, SO_2	1. Within Oslo: (1) SO_2: air dispersion model based on measurements of SO_2 in central Oslo and emission data for industry, heating, and traffic; (2) NOx was imputed from modeled SO2, emissions of SO_2, and observed emissions of NOx from heating and traffic, and other sources. 2. Outside Oslo: (1) for less populated areas, measurements at background stations were used directly; (2) for larger cities and industrial areas, exp was obtained by using regional concentrations from background air pollution monitoring stations multiplied by an emission index	NOx, SO_2: cumulative annual average concentration ($\mu g/m^3$)
Nafstad 2004 (85)	1972-98	1974-78 1979-83 1984-88 1989-93	NOx, SO_2	1. Within Oslo: (1) SO_2: air dispersion model based on measurements of SO_2 in central Oslo and emission data for industry, heating, and traffic; (2) NOx was imputed from modeled SO_2, emissions of SO2, and observed emissions of NOx from heating and traffic, and other sources. 2. Outside Oslo: (1) for less populated areas, measurements at background stations were used directly; (2) for larger cities and industrial areas, exp was obtained by using regional concentrations from background air pollution monitoring stations multiplied by an emission index	NOx, SO_2: cumulative annual mean concentrations in selected exposure periods ($\mu g/m^3$)
Danish Cohort Study on Lung Cancer, Smoking, and Environment					
Engholm 1996 (124)	1970-87	1980-85	Classification of areas of residence	1. Residence history obtained from baseline questionnaires and no update was made during follow-up 2. Regions of residence were described as capital, suburbs, towns, and rural	Contrast between capital, suburbs, town, and rural areas (rank order)
French PAARC study					
Filleul 2005 (57)	*All cause:* 1974-01 *Cause-specific:* 1974-98	1974-76	TSP, black smoke, NO, NO_2, SO_2	1. Selected air pollutants were measured (1974-76) at centrally located air pollution monitoring stations in each chosen area with each area varied in diameter from 0.5 to 2.3 km 2. Means of concentrations over this period were assigned to subjects based on their areas of residence	TSP, black smoke, NO, NO_2, SO_2: cumulative annual mean concentration ($\mu g/m^3$)
GenAir study					
Vineis 2006 (127)	1993-98	1980-99	PM_{10}, NO_2, SO_2, proximity of residence to major streets	1. PM_{10}, NO_2, and SO_2: exposure was assessed based on the average conc. at the nearest background monitoring stations and assigned to subjects according to the residence address at the time of enrollment 2. Traffic air pollution: home address located in a major street (yes/no)	PM_{10}, NO_2, SO_2: cumulative annual mean concentration ($\mu g/m^3$) Traffic air pollution: binary classification of exposure (yes/no)

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Harvard Six Cities Study					
Krewski 2005 (119)	1974-89	TSP, SO_2, NO_2, O_3: 1977-85 $PM_{2.5}$, PM_{10}: 1980-85 SO_4: 1979-84	TSP, PM_{10}, $PM_{2.5}$, SO_4, SO_2, NO_2	Exposure was defined by city-specific annual mean concentration of TSP, PM_{10}, $PM_{2.5}$, SO_4, SO_2, NO_2 measured at fixed site stations, one site per city	Contrast between most vs. least polluted cities based on $PM_{2.5}$ (equivalent to per 18.6 increase $\mu g/m^3$ for $PM_{2.5}$, 28.3 $\mu g/m^3$ for PM_{10}, 55.7 $\mu g/m^3$ for TSP, 7.5 $\mu g/m^3$ for SO_4, 15.8 $\mu g/m^3$ for NO_2, or 19.8 $\mu g/m^3$ for SO_2) (Note that though the least polluted cities for SO_4 and SO_2 were different from that for $PM_{2.5}$, study population in the two cities were similar)
Laden 2006 (50)	1974-98	1980-98	$PM_{2.5}$	1. Prior to 1985: city-specific annual mean concentration (1980-1985) of $PM_{2.5}$ measured at fixed site stations, one site per city 2. In 1985-1998: $PM_{2.5}$ was imputed using city-specific regression calibration equations based on humidity-corrected visibility data from adjacent air ports, indicators for season, and concentrations of PM_{10} at monitors within an 80 km radius of the city	$PM_{2.5}$: cumulative annual mean concentration ($\mu g/m^3$)
Netherlands Cohort Study on Diet and Cancer (NLCS)					
Beelen 2008 (25)	1987-96	NO_2, black smoke, SO_2: 1976-96 $PM_{2.5}$: 1992-96 Local traffic exposure: 1987-96	NO_2 Black smoke SO_2 $PM_{2.5}$ local traffic exposure	1. $PM_{2.5}$ was imputed from PM_{10} using regression equations 2. Regional backgrounds of each pollutant were interpolated based on regional background stations using IDW ($1/R^2$) 2. Additional urban backgrounds were estimated by land use regression models with residual concentrations for all regional background and urban monitoring sites as dependent variables and population density and land use variables as predictors 3. Regional backgrounds + urban backgrounds = background concentrations 4. Local traffic contributed exposure included (1) a local traffic component was estimated using regression models with traffic variables as predictors and the residuals of NO_2, black smoke, and $PM_{2.5}$ (measured concentration - estimated background concentration) as dependent variables (2) another local component was estimated based on regression models in which the association of measured concentrations and truck traffic intensity and distance to the nearest motorway was defined 5. All above were added to produce an overall exposure estimate for each pollutant at the subject's home addresses	NO_2, black smoke, SO_2, $PM_{2.5}$: cumulative annual mean concentration ($\mu g/m^3$) Local traffic exposure: (1) traffic intensity on the nearest road (no. of motor vehicle/24 hr), (2) sum of traffic intensity in a buffer of 100m (no. of motor vehicle/24 hr), (3) living near a major road (i.e. living within 100m of a motor way or 50m of a local road with traffic density > 10000 motor vehicle per 24 hr): yes/no
German Women's Health Study in North Rhine-Westphalia (retrospective)					

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Gehring 2006 (62)	1985-03	The year at baseline survey or in a 5-yr period including the year of baseline and the preceding 4 years	PM_{10}, NO_2, distance to nearest main roads	1. NO_2 was directly measured at central sites 2. PM_{10} was derived from conc. of TSP: PM_{10}=0.71*TSP 3. Distance of home to a major road: measured using GIS‡ 4. Missing air pollution data in one area was replaced by air pollution data in an adjacent area 5. Missing air pollution data in two areas prior to 1990 were imputed by a linear regression model with autoregressive covariance component 6. Exposure was assigned to subjects based on baseline home addresses	NO_2 and PM_{10}: cumulative annual mean concentrations in selected exposure periods ($\mu g/m^3$) Distance of home to a major road: binary classification (>50 vs. ≤50m)
Ontario Cohort Study (retrospective)					
Finkelstein 2003 (118)	1992-99	TSP: 1992-94 SO_2: 1993-95	TSP, SO_2	1. Three-year avg. conc. of SO_2 and TSP were interpolated using universal kriging based on 19 fixed sites for SO_2 and 29 for TSP 2. Conc. assigned to subjects based on the centroids of postal code areas	Binary classification with the cutoff at median
Finkelstein 2005 (64)	1992-01	TSP: 1992-94 SO_2: 1993-95	Pollution index, traffic indicator	1. Three-year avg. conc. of SO_2 and TSP was estimated using universal kriging based on 19 fixed sites for SO_2 and 29 for TSP 2. A pollution index was created as sum of the standardized scores, $(x - \mu)/\sigma$, for each of the pollutants at postal code addresses of the subjects 3. Traffic indicator: subjects were considered exposed if postal code addresses were within 50m of a major urban road or within 100m from a highway	Pollution index: mean value at postal code Traffic indicator: binary classification (lived near or not near a major road/ highway)
Oslo Cohort Study (retrospective)					
Naess 2007 (84)	1995-98	1992-95	$PM_{2.5}$, PM_{10}, NO_2	1. For each of 470 administrative neighborhoods, air dispersion model was used to estimate hourly conc. of pollutants in 1992-1995 based on hourly emissions and meteorological data 2. The hourly averaged conc. for each neighborhood were calculated as a weighted average according to population density within each neighborhood 3. Exposure was assigned to subjects based on baseline home addresses	$PM_{2.5}$, PM_{10}, NO_2: cumulative annual mean concentration ($\mu g/m^3$)
Pennsylvania Cohort					
Morris 1976 (121)	1960-72	1959-60	SO_2, SO_4, dust fall, TSP	Air sampling conducted in 1959-60 showed that Seward had higher concentrations of SO_2, SO_4, dust fall, TSP levels than New Florence	Contrast between two cities: Seward and New Florence
Residents of Lancaster and Glendora in California					
Detels 1987 (148)	1972-77 (Lancaster)1977-82 (Glendora)	1972-82	O_3, SO_2, NO_2, NO, SO_4, particulates	1. According to measurements from central site of O_3, SO_2, NO_2, NO, SO_4, and particulates at monitoring stations, Glendora had higher concentrations of pollutants thus it was defined as exposed area. 2. Exposure was assigned to subjects based on residence at baseline	Contrast between two cities: Lancaster and Glendora
USEPRI-Washington University Veterans' Cohort Mortality Study					

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Lipfert 2006 (51)	1997-01	∑$PM_{2.5}$ components: 2002 $PM_{2.5}$ (TEOM‡ and gravimetric measurements): 2000-03 O_3, NO_2, CO, SO_2: 1999-01 Traffic density: 1997	∑$PM_{2.5}$ components (Al, As, Ba, Ca, Cl, Cr, Cu, Fe, Pb, Mn, Ni, Se, Si, V, Zn, organic and elemental carbon (OC, EC) and sulfate and nitrate ions (SO_4, NO_3)), $PM_{2.5}$ (TEOM and 24-hr gravimetric measurements), O_3, NO_2, CO, SO_2, traffic density	1. Annual conc. of $PM_{2.5}$ (specific components of $PM_{2.5}$, TEOM measurements, and 24-hr gravimetric filter measurements, respectively), O_3, NO_2, CO, SO_2 were measured at central sites and then averaged by sites and by county 2. ∑$PM_{2.5}$ components = $1.375SO_4+1.29NO_3+1.4OC+EC$ 3. Pooled $PM_{2.5}$ include both TEOM and 24-hr gravimetric measurements 4. Traffic density was calculated as county-average vehicular traffic density, defined as, annual vehicle-miles traveled per unit of land area	∑$PM_{2.5}$ components, $PM_{2.5}$ (TEOM and gravimetric measurements), and O_3, NO_2, CO, SO_2: cumulative annual mean concentrations Traffic density: continuous variable (per annual vehicle-miles traveled per unit of land area)
Women's Health Initiative Study					
Miller 2007 (28)	1994-03	2000	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3	Exposure was assigned to five-digit Zip Code areas based on annual averaged concentration measured at the nearest monitor within 48 km of the residences	$PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, O_3: cumulative annual mean concentration ($\mu g/m^3$)

‡ TSP, total suspended particulate; IDW, inverse distance weighting; topographic barriers, such as mountains, cause effects such as an air mass is forced from a low elevation to a higher elevation as it moves over rising terrain; PM_{10}, Particulate Matter less than 10 microns in diameter; $PM_{2.5}$, Particulate Matter less than 2.5 microns in diameter; pphm, part per hundred million; GIS, geographic information systems; TEOM, tapered element oscillating microbalance; ppb, part per billion

TABLE 3. Selected characteristics of case-control studies of long-term exposure to ambient air pollution and chronic disease, according to outcome and distinct study, 1950-2007

First author and year (ref. no.)	Location	Age range	Outcome(s)	Outcome assessment	No. of cases		No. of controls		Recruitment of participants					Design
									Year of diagnosis	Response rate		Selection of Cases	Selection of controls§	
					Total	Men	Total	Men		Cases (%)	Controls (%)			
Asthma														
Modig 2006 (146)	Sweden	20-60	Incidence: asthma	Clinical examination, hospital records	199	NR‡	196	NR	1995-99	NR	NR	Primary health care, local/ county hospitals, private clinics, occupational physicians	Population registry	Population-based
Breast Cancer														
Lewis-Michl 1996 (143)	US	20-79	Incidence: breast cancer	Cancer registry	793	0	966	0	1984-86	88	67	Cancer registry	Driver's license records	Population-based
Western New York Exposures and Breast Cancer Study:														
Bonner 2005 (142)	US	35-79	Incidence: breast cancer	Hospital records	1166	0	2105	0	1996-01	71	62	Cancer registry	Driver's license records	Population-based
Nie 2007 (141)	US	≥44	Incidence: breast cancer	Hospital records	1068	0	1944	0	1996-01	71	62	Cancer registry	Driver's license records	Population-based
Chronic obstructive pulmonary disease (COPD)														
Tzonou 1992 (147)	Greece	50-60	Incidence: COPD	Hospital records	110	98	400	352	1984-84	97	98	Four selected hospitals	Same hospitals (systematic sample)	Hospital-based
Lung Cancer														
Stocks 1955 (128)	Britain	45-74	Mortality: lung cancer	Hospital records	725	725	2726	2726	1952-54	90	NR	All hospitals, GP's patient lists, clinical register	Selected no. of hospitals (convenient sample)	Population-based
Jedrychowski 1990 (131)	Poland	Any	Mortality: lung cancer	Death certificates	1099	901	1073	875	1980-85	65	64	Death registry	Death registry	Population-based
Barbone 1995 (135)	Italy	36-98	Mortality: lung cancer	Cancer registry	755	755	755	755	1979-81 1985-86	81	83	Cancer registry	Registry at a local university	Population-based
Chiu 2006 (140) *	Taiwan	50-69	Mortality: lung cancer	Death certificates	972	0	972	0	1994-03	100	100	Death registry	Death registry	Population-based

First author and year (ref. no.)	Location	Age range	Outcome(s)	Outcome assessment	No. of cases		No. of controls		Recruitment of participants					Design
									Year of diagnosis	Response rate		Selection of Cases	Selection of controls§	
					Total	Men	Total	Men		Cases (%)	Controls (%)			
Vena 1982 (129)	US	≥30	Incidence: lung cancer	Hospital records	417	417	752	752	1957-65	NR	NR	One selected hospital	Same hospital	Hospital-based
Xu 1989 (130)	China	30-69	Incidence: lung cancer	Cancer registry	1249	729	1345	788	1985-87	95	>95	All hospitals	All residents	Population-based
Holowaty 1991 (132)	Canada	≤75	Incidence: lung cancer	Hospital records	51	0	45	0	1983-85	71	58	Seven out of nine hosp. in Niagara plus major referral hosp. in Toronto	Municipal assessment lists (1983-84)	Population-based
Katsouyanni 1991 (133)	Greece	≥35	Incidence: lung cancer	Hospital records	101	0	89	0	1987-89	96	91	Seven selected hospitals	Same or nearby hospitals (density sample)	Hospital-based
Jockel 1992 (134)	Germany	38-85	Incidence: lung cancer	Hospital records	194	146	388	292	NR	NR	Hospital control: NR Population controls: 41	Seven selected hospitals	*Hosp. control:* same hospitals *Population control:* residence registry	Hospital-based
Pawlega 1997 (136)	Poland	≥30	Incidence: lung cancer	Cancer registry	176	176	341	341	1992-94	70	68	Cancer registry	Electoral roll (1992-94)	Population-based
Nyberg 2000 (137)	Sweden	40-75	Incidence: lung cancer	Cancer registry	1042	1042	2364	2364	1985-90	87	Population Control: 88 Dead control: 82	Cancer registry	Population registry and cause-of-death registry (density sample)	Population-based
Gupta 2001 (138)	India	Any	Incidence: lung cancer	Hospital records	265	235	525	435	1995-97	NR	NR	One selected hospital	Hospital visitors, attendants	Hospital-based
Besso 2003 (139)	Sweden	Any	Diagnosis of cancer of the bronchus or lung at any time	Cancer registry	316	209	727	518	Any time	94	91	Death registry (1961-90)	Death registry (1961-90)	Population-based
Myocardial Infarction (MI)														
Grazuleviciene 2004 (145)	Lithuania	25-64	Incidence: first nonfatal MI‡	Hospital records	448	448	1777	1777	1997-00	77	71	All hospitals	Population Registry (density sample)	Population-based
Rosenlund 2006 (87)	Sweden	45-70	Incidence: first MI Mortality:	Hospital records, death certificates	1397	922	1870	1216	1992-94	Woman: 72 Man:	Woman: 70 Man:	All emergency hospitals, hospital	Population registry	Population-based

First author and year (ref. no.)	Location	Age range	Outcome(s)	Outcome assessment	No. of cases		No. of controls		Recruitment of participants					Design
										Response rate				
					Total	Men	Total	Men	Year of diagnosis	Cases (%)	Controls (%)	Selection of Cases	Selection of controls§	
			first MI							81	75	discharge list, death registry		
Tonne 2007 (144)	US	Case: 25+ Control : 17+	Incidence: acute MI	Hospital records	5049	2848	10277	5837	1995, 1997, 1999, 2001, 2003	100	100	All acute care general hospitals	Resident lists	Population-based

* All collected information was based on administrative databases and no interview or questionnaire was used

‡ NR, not reported; MI, myocardial infarction

§ Selection of controls was conducted by random sampling unless specified otherwise

TABLE 4. Exposure assessment in case-control studies of long-term exposure to ambient air pollution and chronic disease, according to outcome and distinct study, 1950-2007

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Asthma					
Modig 2006 (146)	1995-99	NO_2: 1999-00 Traffic count: NR‡	NO_2, traffic flow	1. NO_2: measured for one week at home for each case-control pair and then the measurements were standardized based on annual average temperature of the study area using a linear regression equation of NO_2 and temperature in 1999-00 2. Traffic flow: at the residence of each subject, the total traffic count (vehicles/24 hr period, weekday) was enumerated within a radius of 200 m	NO_2: weekly concentration ($\mu g/m^3$) Traffic flow: vehicles per 24 hr period on a weekday
Breast Cancer					
Lewis-Michl 1996 (143)	1984-86	Proximity to industries: in 1965 and 1975 Proximity to highway: 1990-92	Proximity to industries, proximity to heavily traffic highways	1. Proximity to industries was based on the industrial directory in 1965 and 1975, for each 1-km^2 grid cell: (1) total no. of chemical-industry facilities and no. of other manufacturing facilities; (2) total no. of facilities of each type in the 8 nearby cells. 3. Proximity to traffic: traffic counts for principle roads were used to compute traffic density and assigned to each 25-km^2 grid cell which was re-sampled to produce 1-km^2 grid cells. 4. Grid cell values were assigned to cases and controls based on their residences in each year 5. The duration of residence was used to compute weighted avg. exposure over 20 years for each subject	Proximity to industries: no. of facilities in a grid cell of residence Proximity to traffic: traffic density
Western New York Exposures and Breast Cancer Study:					
Bonner 2005 (142)	1996-01	1960-97	TSP‡	1. PAH‡ was surrogated by concentrations of TSP 2. Interpolated using IDW‡ with 45-degree angle from 7 closest fixed site stations 3. 10-year averaged conc. of TSP in 1960-69 and annual averaged TSP conc. in 1970-97 were estimated 4. Lifetime cumulative exposure was assigned to subjects based on residential history	TSP: annual mean concentration at selected exposure time points ($\mu g/m^3$)

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Nie 2007 (141)	1996-01	1960-01	Benzo[a]pyrene	1. Exposure to PAH was surrogated by Benzo[a]pyrene 2. Benzo[a]pyrene was estimated using dispersion model based on region-specific meteorological and traffic data to estimate traffic emissions along road segments in the two study counties, which producing exposure estimates for each historical residence for each subject at various exposure time points 3. The estimate of Benzo[a]pyrene was a relative rather than absolute concentration	Benzo[a]pyrene: relative values (continuous variable) at selected exposure time points
Chronic obstructive pulmonary disease (COPD)					
Tzonou 1992 (147)	1984-84	Life long	Classification of areas of residence	Exposure was classified based on residential history as three categories: (1) born and having lived virtually all of their lives in a rural area, (2) born in a rural area but having lived some part of their lives in urban areas, (3) born and having lived exclusively in urban areas	Contrast between regions of residence by rank order (residence in urban exclusively vs. in rural exclusively and mixed)
Lung Cancer: Mortality					
Stocks 1955 (128)	1952-54	1954-55	Total smoke, trace elements, 3:4 benzopyrene, PAHs	1. Sampled levels of total smoke, trace elements, 3:4 benzopyrene, and other PAHS at 10 air quality monitoring stations in the study area 2. Assigned exposure to subjects based on last residence	Contrast between regions of residence by rank order (urban, mixed vs. rural area)
Jedrychowski 1990 (131)	1980-85	1973-80	TSP, SO_2	1. Classification of exposure to air pollution was based on measured TSP and SO_2: isopleths for pollutants were constructed from data from 20 sampling stations 2. Based on isopleth maps, conc. of each pollutant was assigned to the last residence of subjects 3. A combined exposure index of air pollution based on TSP and SO_2 was developed by categorizing TSP and SO_2 and then summarized into a categorical variable of three levels: high, medium, and low	Contrast between levels of exposure by rank order: binary classification
Barbone 1995 (135)	1979-81 1985-86	1972-77	Particulate deposition, classifications of areas of residence	1. Conc. of particulate deposit was assigned to last residence based on measurements from nearest fixed site stations 2. Last residence as classified as center of the city, industrial, mixed, residential, and rural	Particulate deposition: contrast by exposure of rank order (tertile classification) Contrast between regions of residence: categories of residence
Chiu 2006 (140)	1994-03	1994-03	Air quality index	1. Air quality index was derived from PM_{10}‡, O_3, CO, NO_2, and SO_2 together 2. Annual avg. measurement of each pollutant was divided by its national standard 3. Each ratio was scaled to 100 and all ratios were averaged to generate the index 4. A single air pollution index was assigned to each	Contrast between air quality index by rank order (tertile classification)

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
				municipality, and all municipalities were classified into three levels of exposure 5. Air quality index was assigned to subjects based on the municipality in which they resided	
Lung Cancer: Incidence					
Vena 1982 (129)	1957-65	1961-63	TSP	1. TSP was measured in 1961-62 at 21 sampling stations 2. An isopleth was constructed to characterize the levels of air pollution based on the two-year averaged concentration of TSP 3. Based on the concentration of TSP, towns of residence were classified into high, medium, and low air pollution	Contrast years of residence in high or medium air pollution by rank order
Xu 1989 (130)	1985-87	1985-87	Outdoor air pollution index, Proximity to industries	1. Outdoor air pollution index was defined as self-perceived smokiness within a 100m radius of residence (yes or no) 2. Proximity to industries: nearest distance from the closest residences to a major smelter	Self-perceived outdoor smokiness: binary classification (smoky and not smoky) Proximity to industries: binary classification (<1km and ≥3 km)
Holowaty 1991 (132)	1983-85	Lifelong	Classification of areas of residence	Exposure was defined as living in large town/city as compared to country/small town or living for extended period in Niagara Region	Contrast between regions of residence: (1) duration of residence in exposed areas (<40 vs. ≥40 years), (2) regions of residence in childhood or adulthood (country/small town vs. large town)
Katsouyanni 1991 (133)	1987-89	1983-85	Smoke, NO_2	1. Levels of smoke and NO_2 at each borough were estimated from annual mean measurements at 14 monitoring stations 2. Zero air pollution line was defined at the highest points on the surrounding mountains 3. At each borough the air pollution level was the average of three nearest stations, or the two nearest and the zero air pollution line, weighted by the inverse distance from the borough's center to the measurement points 4. All boroughs were classified into five levels of exposure	Contrast between boroughs by rank order of exposure
Jockel 1992 (134)	NR	Emission index: 1955-80 Semiquantitative index: 1895-84	Emission index, semiquantitative index	1. Emission index was derived for SO_2 based on energy consumption which yielded county-level SO_2 emission on a 5-yr basis 2. A semiquantitative index was formed based on BaP‡, TSP and SO_2 from energy consumption, SO_2 emission, use of coal, and degree of industrialization. All counties were classified into eight pollution categories for each 10-year interval 3. County-level emission and semiquantitative index were classified into two categories	Emission index: binary classification Semiquantitative index: binary classification

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Pawlega 1997 (136)	1992-94	1973-80	TSP, SO_2	1. Classification of exposure to air pollution was based on measured annual averaged concentrations of TSP and SO_2 2. Each district of residence was assigned annual mean concentrations of TSP and SO_2 (approx. 5 sites per district)	Contrast between districts of residence by rank order of TSP and SO_2, respectively
Nyberg 2000 (137)	1985-90	1960-90	SO_2, NO_2	1. Dispersion calculations for annual mean SO_2 and NOx were based on traffic and heating emission databases 2. Derived NO_2 from NOx using a non-linear regression based on measurements of NOx in the early 1980s 3. Annual levels of SO_2 and NOx/NO_2 were computed for each year in 1950-1990 by linear extrapolation and interpolation based on three databases, historical traffic counts for NO_2, and available trend data for SO_2	SO_2, NO_2: annual mean concentration ($\mu g/m^3$)
Gupta 2001 (138)	1995-97	lifelong	Classification of areas of residence	Exposed was defined as lifetime residence ($\geq$75% of his/her life) in urban areas (with 50,000+ population) according to baseline questionnaires	Contrast between regions of residence by rank order
Besso 2003 (139)	1961-90	1930-90	Metal depositions	Subjects resided in two selected parishes closet to a smelter was classified as exposed	Contrast residential history in two selected parishes by rank order (yr)
Myocardial Infarction (MI)					
Grazuleviciene 2004 (145)	1997-00	1993-00	NO_2	1. For each residential district, ambient level of NO_2 was obtained based on measurements at fixed site stations, one station per each residential district 2. Classified residential districts into tertiles based on ambient levels of NO_2	Contrast between residential districts by rank order
Rosenlund 2006 (87)	1992-94	NO_2, CO, SO_2: 1960-90 $PM_{2.5}$, PM_{10}: 2000	NOx, NO_2, PM_{10}, $PM_{2.5}$‡, CO, SO_2	1. Regional source-specific contributions to NO_2, CO, SO_2, PM_{10}, and $PM_{2.5}$ were estimated using air dispersion model based on traffic and heating emission databases 2. Derived NO_2 from NOx from a non-linear regression based on measurements in the early 1980s 3. Annual means of NO2, CO, SO_2, PM_{10}, and $PM_{2.5}$ were computed for each year in 1950-1990 by linear extrapolation and interpolation	NO_2, PM_{10}, $PM_{2.5}$, CO, SO_2: annual mean concentrations ($\mu g/m^3$)

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Tonne 2007 (144)	1995, 1997, 1999, 2001, and 2003	2002	Cumulative traffic within 100 m of subjects' residence; Distance from major roadway	1. Estimated cumulative traffic by creating a 100-m radius buffer around each subject's residence and summing the product of road segment length and estimated annual average daily traffic which were based on actual traffic counts for major roadways but were only estimated according to regional traffic for more local roads 2. Distance from subjects' homes to major roadway: used GIS to measure	Cumulative traffic: vehicle-km Distance: km

‡ NR, not reported; TSP, total suspended particulate; PAH, polycyclic aromatic hydrocarbon; IDW, inverse distance weighting; PM_{10}, particulate matter less than 10 microns in diameter; $PM_{2.5}$, particulate matter less than 2.5 microns in diameter; BaP, aquatic pollutant benzo[a]pyrene

TABLE 5. Estimates of relationship between long-term exposure to ambient air pollution and the risk of chronic diseases in published case-control studies, according to outcomes

First author, year (ref)	Location	Pollutants and increment of exposure	RR / OR (95% CI)	Stat. analysis	Covariates
Incidence Of Lung Cancer					
Vena 1982 (129)	US	TSP‡: ≤29 vs. >50 years of residence in high or medium level of TSP	1.26 (CI: NR‡)	M-H‡	Age, sex, occupation, and smoking
Xu 1989 (130)	China	Air pollution index: self-perceived outdoor smokiness (smoky vs. not smoky), proximity to industries: <1 vs. ≥ 3 km	Men / Women Air pollution index: 1.30 (1.70-2.90) / 2.50 (1.80-3.50) Distance to a smelter: 3.00 (1.50-6.00) / 1.00 (0.40-2.50)	Uncond. logistic	Age, education, smoking, and indoor air pollution index
Holowaty 1991 (132)	US	Contrast duration of residence: <40 vs. ≥40 years Contrast regions of residence: lived in large town vs. country/small town in childhood and adulthood, respectively	≥40 vs. <40 years: 6.10 (0.96-102.00) Large town vs. country/small town (childhood residence): 7.80 (0.62-98.00) Large town vs. country/small town (adulthood residence): 2.30 (0.49-13.70)	Cond. logistic	Age, sex, and smoking
Katsouyanni 1991 (133)	Greece	Smoke, NO_2: most vs. least polluted boroughs	Nonsmokers: 0.81 (CI: NR) Smokers (15-yr duration): 1.35 (CI: NR) Smokers (30-yr duration): 2.23 (CI: NR)	Uncond. logistic	Age, diet, smoking, education, and interviewer
Jockel 1992 (134)	Germany	Emission index, semiquantitative index: high vs. low levels	Emission index: 1.01 (0.53-1.91) Semiquantitative index: 1.16 (0.64-2.13)	Cond. logistic	Age, sex, occupation, and smoking
Pawlega 1997 (136)	Poland	TSP, SO_2: most polluted vs. least regions	TSP: 0.24 (0.12-0.48) SO2: 0.24 (0.12-0.48)	Uncond. logistic	Age, smoking, occupation exposure, diet, and education
Nyberg 2000 (137)	Sweden	NO_2 (10-year average exposure with 20 years lag): 10 μg/m^3 SO_2 (10-year average exposure with 20 years lag): 10 μg/m^3	NO_2: 1.10 (0.97-1.23) SO_2: 1.01 (0.98-1.03)	Uncond. logistic	Age, selection year, smoking, radon, socioeconomic grouping, occupational exposure to diesel exhaust, other combustion products, asbestos, and employment in risk occupations
Gupta 2001 (138)	India	Contrast between regions of residence: urban vs. rural	Men / Women Lifetime residence 0.82 (0.53-1.27) / 0.29 (0.07-1.17) Residence in first 20 yr of life 0.97 (0.61-1.53) / 0.38 (0.09-1.53)	Uncond. logistic	Age and education

First author, year (ref)	Location	Pollutants and increment of exposure	RR / OR (95% CI)	Stat. analysis	Covariates
Besso 2003 (139)	Sweden	Metal depositions: ≥20 years of vs. never First residence before 1940 vs. never	Men Women ≥20 yr vs. never: 1.28 (0.77-2.12) 1.00 (0.48-2.12) First residence before 1940 vs. never: 1.51 (0.90-2.54) 0.78 (0.35-1.74)	Cond. logistic	Age, smoking, and occupation
Mortality Of Lung Cancer					
Stocks 1955 (128)	Britain	Total smoke, trace elements, 3:4 benzopyrene, PAHs‡: contrast between types of residence (urban, mixed vs. rural)	Non-smokers: 9.35 (CI: NR) Smokers: NR	Direct Standard-ization	Age, smoking, and sex
Jedrychowski 1990 (131)	Poland	A composite exposure index from TSP, SO_2: high, medium vs. low level	Men (high vs. low): 1.48 (1.06-1.99) Women (medium, high vs. low): 1.17 (0.70-1.96)	Uncond. logistic	Age, sex, occupation, and smoking
Barbone 1995 (135)	Italy	Particulate deposition: highest tertile vs. lowest tertile, Contrast between types of residence: urban, industrial, mixed, rural vs. residential	Particulate deposition: 1.40 (1.10-1.80) Urban centre vs. residential: 1.50 (1.00-2.20) Industrial vs. residential: 1.40 (1.00-2.10) Rural vs. residential: 0.60 (0.40-1.00) Mixed vs. residential: 1.20 (0.90-1.50)	Uncond. logistic	Age, occupation, SES‡, and smoking
Chiu 2006 (140)	Taiwan	Air quality index: highest vs. lowest category	1.28 (1.02-1.61)	Cond. logistic	A socioeconomic index (composed of population density, age, mobility, income, education, environmental factors, etc)
Incidence Of Breast Cancer					
Lewis-Michl 1996 (14)	US	Proximity to industries: no. of facilities in a grid cell of residence Proximity to traffic: traffic density (95th percentile vs. all other)	Postmenopausal women: *In Nassau county* Proximity to industries: 1.11 (0.83-1.48) Proximity to heavily traffic highways: 1.29 (0.77-2.15) *In Suffolk county* Proximity to industries: 1.12 (0.72-1.74) Proximity to heavily traffic highways: 0.89 (0.40-1.99)	Uncond. logistic	Age at first live birth, family history of breast cancer, history of benign breast disease, age at diagnosis, and years of education
Western New York Exposures and Breast Cancer Study					
Bonner 2005 (142)	US	TSP (continuous): 30 ug/m^3 increase at several time points of observation TSP (quartile): 4th quartile vs. 1st quartile at several time points of observation	Premenopausal Postmenopausal *Exposure at birth address* 30 ug/m^3 0.92 (0.76-1.11) 1.20 (1.04-1.38) 4th quartile vs. 1st quartile 1.78 (0.62-5.10) 2.42 (0.97-6.09) *Exposure at menarche address* 30 ug/m^3 NR‡ 1.08 (0.96-1.21)	Uncond. logistic	Age, race, education, age at first birth, age at menarche, age at menopause, parity, previous benign breast disease, family history of breast cancer, and BMI‡

First author, year (ref)	Location	Pollutants and increment of exposure	RR / OR (95% CI)	Stat. analysis	Covariates
			4th quartile vs. 1st quartile 0.66 (0.38-1.16) 1.45 (0.74-2.87) *Lifetime cumulative exposure* 30 ug/m³ NR NR 4th quartile vs. 1st quartile NR 3.50 (1.40-8.90)		
Nie 2007 (141)	US	Benzo[a]pyrene: 4th quartile vs. 1st quartile at several time points	Premenopausal Postmenopausal At menarche address 2.07 (0.91-4.72) NR At first birth address 1.22 (0.44-3.36) 2.58 (1.15-5.83) Lifetime exposure (20 years prior) 1.29 (0.59-2.82) 0.82 (0.58-1.18) Lifetime exposure (10 years prior) 1.49 (0.65-3.43) NR	Uncond. logistic	Age, race, education, age at first birth, age at menarche, age at menopause, BMI, year at interview, smoking, family history of breast cancer, and number of births
Myocardial Infarction (MI)					
Grazuleviciene 2004 (145)	Lithuania	NO_2: 10 μg/m³	Incidence of first MI: 25-64 years old: 1.14 (1.01-1.30) 55-64 years old: 1.29 (1.07-1.55)	Uncond. logistic	Age, education, marital status, smoking, blood pressure, BMI, and stress
Rosenlund 2006 (87)	Sweden	NO_2, PM_{10}‡, CO, SO_2: 10 μg/m³ $PM_{2.5}$‡: NR	Incidence of first MI NO_2: 0.96 (0.87-1.06) PM_{10}: 0.85 (0.50-1.42) CO: 1.00 (0.99-1.01) SO_2: 0.99 (0.92-1.07) $PM_{2.5}$: NR Mortality of first MI NO_2: 1.15 (0.99-1.33) PM_{10}: 1.93 (0.88-4.25) CO: 1.01 (1.00-1.01) SO_2: 1.06 (0.94-1.19) $PM_{2.5}$: NR	Uncond. logistic	Age, sex, hospital catchment area, smoking, diabetes, physical inactivity, and SES
Tonne 2007 (144)	US	Cumulative traffic: vehicle-km Distance: km	Incidence of first MI: Cumulative traffic (IQR increase): 1.05 (1.02–1.08) Living near major roadway: 1.04 (1.02–1.06)	Uncond. logistic	Age, sex, section of study area, % of open space, neighborhood-level socioeconomic characteristics (% persons below the federally defined poverty line), and $PM_{2.5}$ point-source emission density
Incidence Of Adult Asthma					

First author, year (ref)	Location	Pollutants and increment of exposure	RR / OR (95% CI)		Stat. analysis	Covariates
Modig 2006 (146)	Sweden	NO_2: 10 µg/m^3 increase Traffic flow: above 75th percentile vs. all other	NO_2: Traffic flow:	1.0 (0.37-2.73) 1.5 (0.90-1.10)	Cond. logistic	Age, sex, urbanization, skin-pick test, family history of asthma, BMI and smoking
Incidence Of Chronic Obstructive Pulmonary Disease (COPD)						
Tzonou 1992 (147)	Greece	Contrast between regions of residence: urban exclusively vs. rural exclusively and mixed	2.00 (1.20-3.30)		Uncond. Logistic	Age, sex, education, no. of children, and smoking

‡ M-H, Mantel-Haenszel estimation; NR, not reported; PM_{10}, particulate matter less than 10 microns in diameter; $PM_{2.5}$, particulate matter less than 2.5 microns in diameter; BMI, body mass index; TSP, total suspended particulate; SES, socioeconomic status; PAH, polycyclic aromatic hydrocarbon

TABLE 6. Subgroup analysis of estimates for **mortality of lung cancer** associated with long-term exposure to $PM_{2.5}$ and NO_2

Subgroup analysis	Number of studies	Summary RR (95% CI)	Measure of heterogeneity			*p*-value Meta-regression¶
			Q-value	*p*-value	I^2	
		$PM_{2.5}$				
All studies	5	1.21 (1.10-1.32)	5.30	0.258	24.5%	N/A
Location						
North America	3	1.15 (1.07-1.25)	0.93	0.627	0%	0.171
Europe	2	1.23 (0.98-1.54)	2.49	0.115	60%	Ref.*
Gender						
Both men and women	4	1.20 (1.08-1.33)	5.01	0.171	40.1%	0.635
Men only	1	1.39 (0.79-2.46)	N/A*	N/A	N/A	Ref.
Exposure assessment						
Intraurban comparison	3	1.26 (1.08-1.47)	2.57	0.276	22.3%	0.137
Interurban comparison	2	1.15 (1.06-1.25)	0.51	0.473	0%	Ref.
Adjustment for smoking						
Yes	4	1.15 (1.06-1.24)	1.31	0.726	0%	0.046
No	1	1.34 (1.17-1.54)	N/A	N/A	N/A	Ref.
Completeness of follow up						
≥ 90%	4	1.20 (1.07-1.35)	5.08	0.166	41.0%	0.747
Not reported	1	1.27 (0.96-1.69)	N/A	N/A	N/A	Ref.
Exclusion of studies						
- Naess et al. 2007 (84)	4	1.15 (1.06-1.24)	1.31	0.726	0%	N/A
		SO_2				
All studies	5	1.07 (0.96-1.19)	27.30	<0.001	85.3%	N/A
Location						
North America	2	1.58 (0.66-3.76)	25.10	<0.001	96.0%	0.142
Europe	3	1.00 (0.96-1.19)	0.04	0.980	0%	Ref.
Gender						
Both men and women	4	1.12 (0.97-1.30)	26.89	<0.001	88.8%	0.526
Men only	1	1.00 (0.92-1.08)	N/A*	N/A	N/A	Ref.
Exposure assessment						
Intraurban comparison	4	1.13 (0.96-1.33)	27.09	<0.001	88.9%	0.650
Interurban comparison	1	1.03 (0.98-1.08)	N/A	N/A	N/A	Ref.
Completeness of follow up						
≥ 90%	4	1.12 (0.97-1.30)	26.45	<0.001	88.7%	0.358
< 90%	1	0.99 (0.92-1.07)	N/A	N/A	N/A	Ref.
Exclusion of studies						
- Filleul et al. 2005 (57)	4	1.12 (0.97-1.30)	26.45	<0.001	88.7%	N/A

* N/A, not available because only 1 study; Ref., reference group

¶ This meta-regression analysis used the method of reml (restricted maximum likelihood using an iterative procedure) to estimate the additive between studies variance tau^2.

TABLE 7. Subgroup analysis of estimates for the associations between $\mathbf{PM_{2.5}}$ and mortality of all cardiovascular diseases and coronary heart disease

Subgroup analysis	Number of studies	Summary RR (95% CI)	Measure of heterogeneity			*p*-value Meta-regression¶
			Q-value	*p*-value	I^2	
		Mortality of cardiovascular diseases				
All studies	4	1.14 (1.09-1.18)	5.59	0.134	46.3%	N/A
Location						
North America	2	1.18 (1.04-1.34)	4.37	0.037	77.1%	0.906
Europe	2	1.12 (1.05-1.19)	1.2	0.272	17.00%	Ref.*
Exposure assessment						
Intraurban comparison	2	1.12 (1.05-1.19)	1.2	0.272	17.00%	0.906
Interurban comparison	2	1.18 (1.04-1.34)	4.37	0.037	77.1%	Ref.
Adjustment for smoking						
Yes	3	1.14 (1.04-1.26)	5.51	0.063	63.7%	0.787
No	1	1.13 (1.09-1.19)	N/A	N/A	N/A	Ref.
Completeness of follow up						
≥ 90%	3	1.12 (1.09-1.15)	1.24	0.537	0%	0.037
Not reported	1	1.28 (1.13-1.45)	N/A	N/A	N/A	Ref.
Inclusion/Exclusion of studies						
+ Miller et al. 2007 (28)	5	1.15 (1.09-1.23)	12.10	0.017	66.9%	N/A
- Laden et al. 2006 (50)	3	1.12 (1.09-1.15)	1.24	0.537	0%	N/A
- Naess et al. 2007 (84)	2	1.12 (1.09-1.15)	0.92	0.337	0%	N/A
		Mortality of coronary heart diseases				
All studies	3	1.16 (0.96-1.40)	8.68	0.013	77.0%	N/A
Gender						
Both men and women	2	1.10 (0.94-1.29)	4.8	0.028	79.2%	0.049
Women only	1	2.21 (1.17-4.17)	N/A	N/A	N/A	Ref.
Exposure assessment						
Intraurban comparison	3 **	1.24 (0.91-1.67)	7.42	0.024	73.0%	0.212
Interurban comparison	1	1.18 (1.14-1.23)	N/A	N/A	N/A	Ref.
Exclusion of studies						
- Miller et al. 2007 (28)	2	1.10 (0.94-1.29)	4.8	0.028	79.2%	N/A

* N/A, not available because only 1 study; Ref., reference group

** Included Jerrett 2005 (49) in this subgroup

§ Pooled estimate was not calculated due to significant heterogeneity

¶ This meta-regression analysis used the method of reml (restricted maximum likelihood using an iterative procedure) to estimate the additive between studies variance tau^2.

UPDATED REVIEW

I made use of the same strategy as in the original systematic review (1) to search the literature and select studies in this updated review, except that the search was limited to cardiovascular outcomes and included studies between January 1, 2008 and March 30, 2011. From the identified papers, I abstracted the data that characterized the conduct of the studies and I performed a synthesis of the findings by summarizing in tabular and graphical form the quantitative results.

Of the 251 citations identified, 249 were from MEDLINE and two were from perusing the bibliographies of the articles. I excluded 239 for the reasons listed in **Figure 6**, leaving 12 articles that consisted of eight cohort studies (53, 77, 80, 86, 187-191) and three case-control studies (81, 88, 192).

Table 8 presents a summary of selected design characteristics of the cohort studies. Of the nine cohort studies, six were conducted in North America (53, 77, 80, 86, 187-189), one in Europe (190) and one in Asia (191). Most of these studies were focused on cardiovascular mortality and four of the cohort studies were conducted prospectively (53, 86, 188, 189, 191). **Table 9** presents a summary of selected design characteristics of the case-control studies. All three case-control studies were conducted to examine the effects on the incidence of nonfatal cardiovascular events and all were based on existing administrative databases (81, 88, 192) (Table 2). **Tables 10-11** shows details of the methods used to assess exposure in the selected cohort and case-control studies. Long-term exposure to nitrogen dioxide was estimated in seven studies (77, 80, 81, 88, 187, 190, 191) and land use regression models were used in five of these studies (77, 80, 187, 190, 191).

Figure 7 summarizes the adjusted rate ratios for associations with cardiovascular and cardiopulmonary disease from the cohort and case-control studies that were published between January 2008 and March 2011. Increased risks for NO_2 were found in three studies of overall cardiovascular mortality (77, 187, 191) and in five studies of mortality from ischemic heart disease (80, 88, 187, 190, 191). Pooling the quantitative findings from two of recently published intraurban studies (187, 191) and the two other intraurban studies that

were published prior to 2008 (25, 84) resulted in a summary estimates of RR_{10}=1.04 (95%CI: 1.0-1.1) (test of heterogeneity: p=0.23) for all cardiovascular mortality for an increase of 10 $\mu g/m^3$ of NO_2. Jerrett 2009 (77) was not included for the meta analysis because the study population consisted of patients who attended a respiratory disease clinic who are likely more susceptible the chronic effects of air pollution than the general population. Similarly, pooling the findings from these recently published intraurban studies (80, 88, 187, 190, 191) and the four other intraurban studies that were published prior to 2008 (82) on mortality from ischemic heart disease resulted in a pooled RR_{10}= 1.06 (95%CI: 1.03-1.08) (test of heterogeneity: p=0.20). For the effects of NO_2 on nonfatal incidence of ischemic heart disease, nonfatal incidence of myocardial infarction, and cerebrovascular mortality, they were examined in few studies between 1950 and 2001 (80, 87, 88, 190, 191), and there was little evidence of increased risk on these cardiovascular outcomes: nonfatal incidence of ischemic heart disease, RR_{10} is 0.96 (95%CI: 0.94-0.99) in one study (80) and 1.01 (95%CI: 0.97-1.05) in the other (190); nonfatal incidence of myocardial infarction, RR_{10}=0.98 (95%CI: 0.96-1.00) (88); and cerebrovascular mortality, RR_{10}=1.09 (95%CI: 0.94-1.27) (191).

The effects of particulate pollutants on cardiovascular outcomes were reported in 12 studies (25, 27, 28, 49, 50, 53, 80-82, 84, 86, 88, 187-189), nine of which are general population-based studies (25, 27, 49, 50, 53, 80, 81, 84, 86, 88, 188, 189). The pooled results from the population-based studies of all cardiovascular mortality, for each increase of 10 $\mu g/m^3$ of $PM_{2.5}$, was pooled RR_{10}=1.13 (95%CI: 1.08-1.18) (test of heterogeneity: p=0.14) (25, 27, 50, 84) excluding Jerrett 2009 (188) because its analysis was based on the same cohort as an earlier study (27). For the effects on mortality from ischemic heart disease, most of the general population-based studies (27, 49, 53, 80, 86, 188, 189) showed positive associations, however, there were substantial heterogeneity among these studies. This may be due to considerable improvement in exposure assessment in two of the studies (53, 86, 189), because exposure was assigned to monthly updated addresses. For PM_{10}, organic carbon, and elemental carbon, there appeared to be increased risks for mortality from ischemic heart disease, but the findings were based on few studies: PM_{10}, RR_{10} varied between 1.02 to 1.43 (53, 88, 187); organic carbon, RR=2.23 (95%CI: 1.79-2.29) for each increment of 0.83 $\mu g/m^3$ (86); and elemental carbon, RR=1.06 (95%CI: 1.03-1.09) for each increment of 0.75 $\mu g/m^3$ (80).

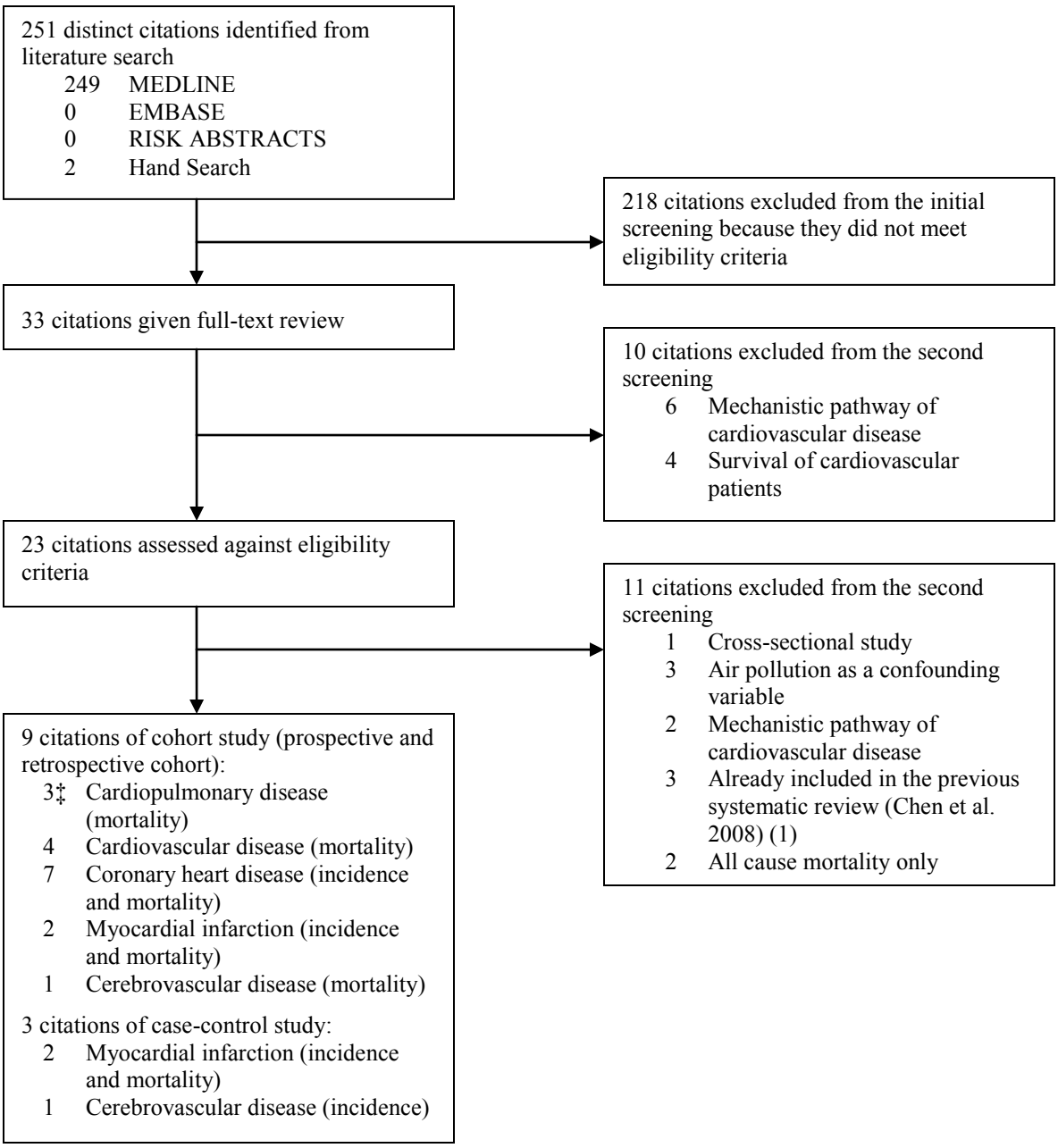
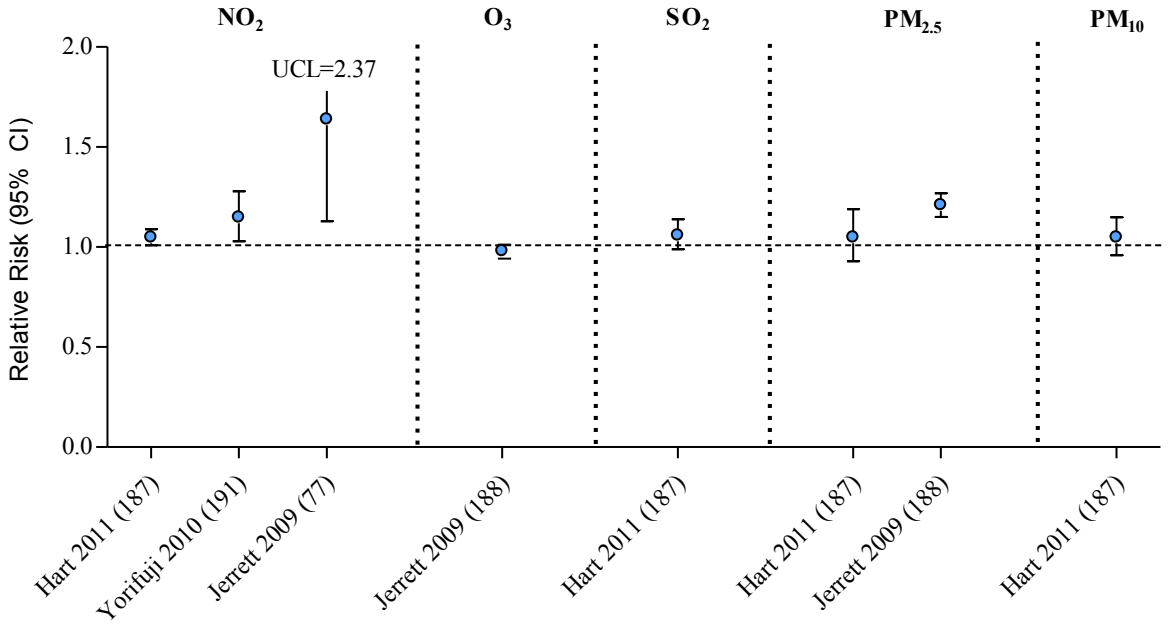
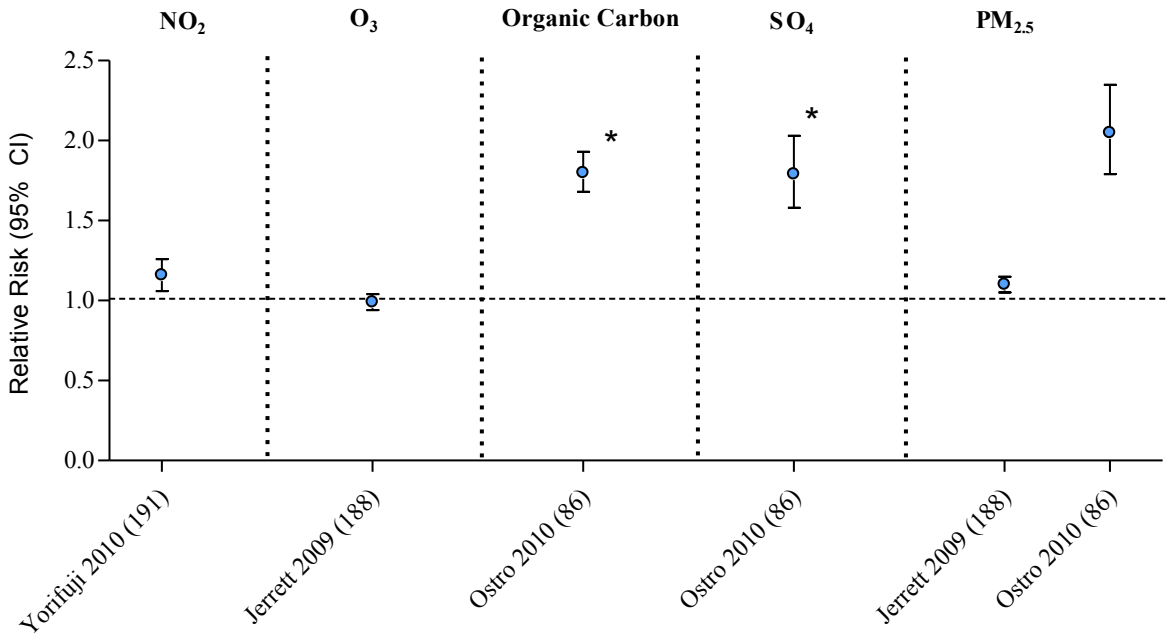


Figure 6: Flow diagram of search strategy and selection process in the systematic review of studies of the relationship between long-term exposure to ambient air pollution and chronic diseases, 01/2008-04/2011.
‡ Numbers refer to number of papers. The numbers under each category do not total 9 because multiple outcomes may have been presented in a publication.

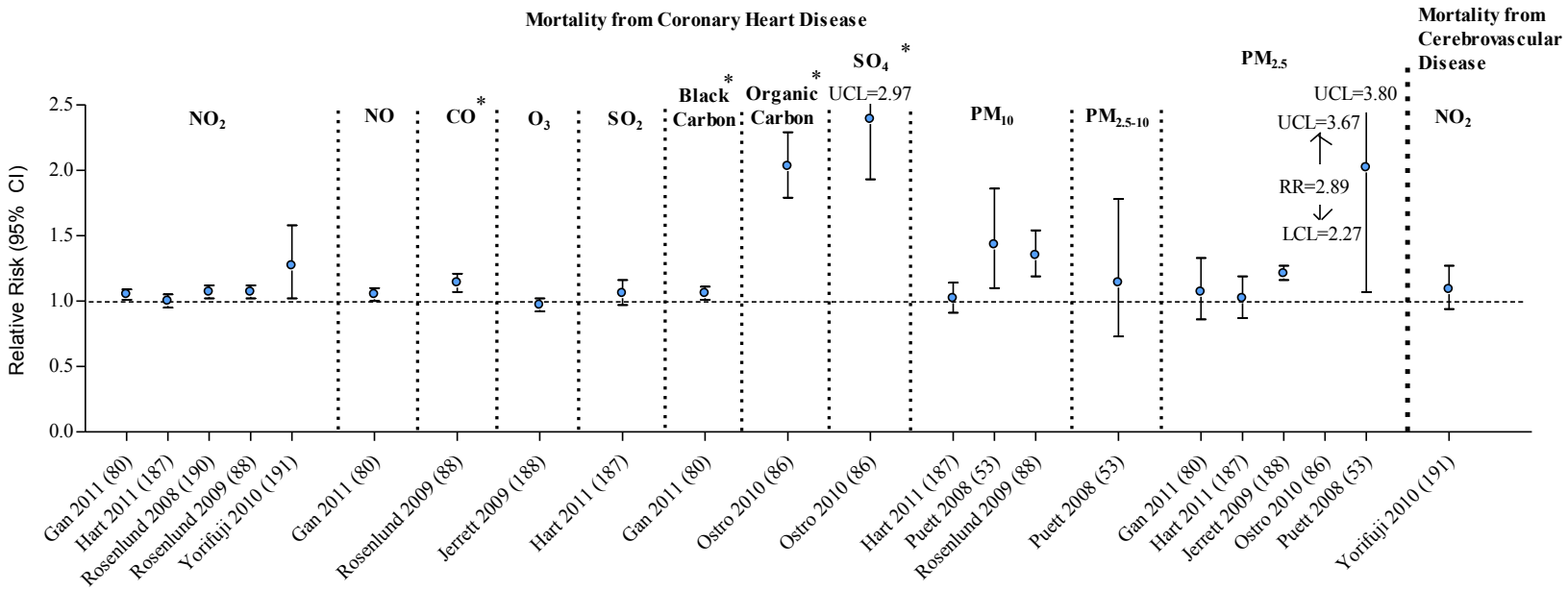
A. Cardiovascular mortality



B. Cardiopulmonary mortality



C. Mortality from coronary heart disease or cerebrovascular disease



D. Incidence of coronary heart disease and myocardial infarction

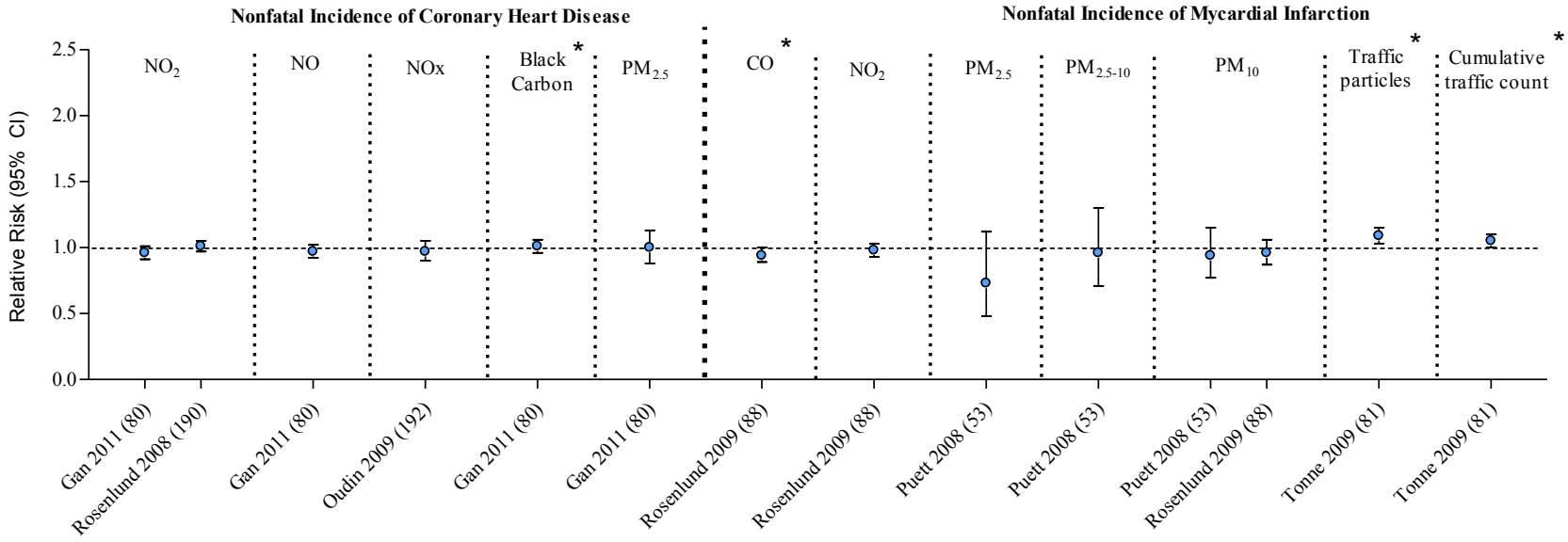


Figure 7. Associations between **(A) cardiovascular mortality, (B) cardiopulmonary mortality, (C) mortality from coronary heart disease or cerebrovascular disease, (D) Incidence of coronary heart disease and in particular myocardial infarction** and long-term exposure to **gaseous and particulate pollutants** with 10 $\mu g/m^3$ difference of concentrations (* For organic carbon, black carbon, SO_4, levels of traffic particles, and cumulative traffic count, the rate ratios were estimated per inter-quartile range increase, depending on the study. For CO, the rate ratios were estimated for a change in the concentration from the fifth to the 95th percentile.)

TABLE 8. Selected characteristics of cohort studies of long-term exposure to ambient air pollution and cardiovascular disease, according to distinct study cohort, 2008-2011

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
American Cancer Society (ACS) study											
Jerrett 2009 (188)	US	448850	≥30	43.4	Mortality: cardiopulmonary cardiovascular ischemic heart disease	58775 48884 27642	NR ‡	Death certificates, NDI ‡	1982-99	NR	93
California Teachers Study (CTS)											
Ostro 2010 (86)	US	44847	31-104 (median = 54)	0	Mortality: cardiopulmonary ischemic heart disease	1397 474	0	California death file, US Social Security Death master file, NDI	08/2002-07/2007	NR	93
Nurses' Health Study (NHS)											
Puett 2008 (53)	US	66250	Mean= 62	0	Incidence: myocardial infarction Mortality: ischemic heart disease (excluded those with history of MI at the time of entry and censored incident MI)	854 494	0	Incidence: self-report and medical record Mortality: NDI	07/1992-07/2002	94	NR
Puett 2009 (189)	US	66250	Mean= 62	0	Incidence: myocardial infarction Mortality: ischemic heart disease (excluded those with history of MI at the time of entry)	854 379	0	Incidence: self-report and medical record Mortality: NDI	07/1992-07/2002	94	NR
Rome Cohort Study (retrospective)											
Rosenlund 2008 (190)	Italy	4461669	35-84	45.7	Mortality: first-time coronary heart disease (incl. out-of-hospital deaths and fatal hospitalizations) Incidence: first-time coronary heart disease	4654 (=3598 + 1056) 6513	2918 (=2289 + 629) 4711	The regional cause-of-death registry, the regional hospital-discharge registry, the municipal population registry	1998-2000	100	~100
Ontario Cohort Study (retrospective)											

First author and year (ref. no.)	Location	Study population: Sample size	Study population: Age range	Study population: % men	Outcome(s)	No. of cases*: Total	No. of cases*: Men	Outcome assessment	Follow-up period	Response rate (%)	Completeness follow-up (%)
Jerrett 2009 (77)	Canada	2360	Median= 60	48	Mortality: cardiovascular disease	80	NR	Ontario mortality registry	1992-2002	NR	~100
Shizuoka Elderly Cohort Study											
Yorifuji 2010 (191)	Japan	13444	65-84 (mean= 74)	51	Mortality: cardiopulmonary cardiovascular disease ischemic heart disease cerebrovascular disease	593 394 91 191	NR	Mortality database of the Ministry of Health, Labor, and Welfare of Japan	12/1999-03/2006	63	88
Trucking Industry Particle Study (retrospective)											
Hart 2010 (187)	US	54319	Mean= 42	100	Mortality: cardiovascular disease ischemic heart disease	1682 1109	1682 1109	NDI	01/1985-12/2000	NR	NR
Vancouver Cohort Study (retrospective)											
Gan 2011 (80)	Canada	406232	45-85 (Mean= 59)	45	Incidence: hospitalization of coronary heart disease Mortality: coronary heart disease	10312 3104	6847 1909	Provincial hospital records and death registration records	01/1999-12/2002	NR	NR

* No. of cases are listed in the same order as outcomes

‡ NR, not reported; NDI, U.S. national death index

TABLE 9. Selected characteristics of case-control studies of long-term exposure to ambient air pollution and chronic disease, according to outcome and distinct study, 2008-2011

First author and year (ref. no.)	Location	Age range	Outcome(s)	Outcome assessment	No. of cases		No. of controls		Recruitment of participants					Design
									Year of diagnosis	Response rate		Selection of Cases	Selection of controls [§]	
					Total	Men	Total	Men		Cases (%)	Controls (%)			
Scania Case-control Study of Ischemic Stroke (administrative database)														
Oudin 2009 (192)	Scania, Sweden	Case: median =72 control: NR	First-time ischemic stroke hospital admissions	National stroke register	Phase I: 4904 Phase II: 4375	2778 NR	Phase I: 556912 Phase II: 4716	Phase I: 273489 Phase II: NR	2001-2005	NR‡	NR	Limited to those who were born between 1923 and 1965 (Phase I: all cases; Phase II: those with completed personal confounding information)	Phase I: Population registry Phase II: A health survey (Self-selection was present)	Phase I: Population-based Phase II: controls may not be selected from the same study base as the cases of phase II
Stockholm County Case-control Study (administrative database)														
Rosenlund 2009 (88)	Stockholm, Sweden	Case and control: 15-79	Incidence: myocardial infarction Mortality: myocardial infarction (inhospital and out-of-hospital fatal cases)	Registries of hospital discharges and deaths	43275	29860	511065	250422	1985-1996	NR	NR	Included all first-time cases from the registry (excluded prevalent cases)	Registry of the total population of Stockholm County: matched on age (5 yrs), sex, and calendar year, and randomly selected from the registry	Population-based
Worcester Heart Attack Study														
Tonne 2009 (81)	Worcester, US	Case: 25+ Control : 17+	Incidence: acute myocardial infarction (first time or recurrent	Records from all hospitals in the study area	5049	2848	10277	5837	1995, 1997, 1999, 2001, 2003	100	100	All acute care general hospitals	Resident lists	Population-based

First author and year (ref. no.)	Location	Age range	Outcome(s)	Outcome assessment	No. of cases		No. of controls		Recruitment of participants					Design
										Response rate				
					Total	Men	Total	Men	Year of diagnosis	Cases (%)	Controls (%)	Selection of Cases	Selection of controls [§]	
			cases)											

‡ NR, not reported

§ Selection of controls was conducted by random sampling unless specified otherwise

TABLE 10. Exposure assessment in cohort studies of long-term exposure to ambient air pollution and chronic disease, according to distinct study cohort, 2008-2011

		Exposure Assessment			
First author, year (ref)	Follow-up period	Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
American Cancer Society (ACS) study					
Jerrett 2009 (188)	1982-99	$PM_{2.5}$: 1999-00 O_3: 1977-00	$PM_{2.5}$, O_3	1. Annual mean concentrations of each pollutant at each fixed site in a metropolitan area were averaged to obtain metropolitan area-specific concentrations 2. Each subject was assigned a metropolitan area of residence based on address at time of enrolment and 3-digit Zip Code area 3. Each subject was assigned metropolitan area-specific concentrations	$PM_{2.5}$: two-year average of annual mean concentration in 1999-2000 O_3: long-term average of two quarterly average of daily hourly maximum concentration between the second (April-June) and third (July-September) from 1977 to 2000 ($\mu g/m^3$)
California Teachers Study (CTS)					
Ostro 2010 (86)	08/2002-07/2007	08/2002-07/2007	$PM_{2.5}$, EC, OC, SO_4, NO_3, Fe, K, Si, Zn	1. For each individual and each pollutant, long-term average of monthly mean concentration during 08/2002-07/2008 in the nearest available monitoring station was assigned as a measure of long-term exposure 2. Residences within 8km and 30km buffer of a monitor	$PM_{2.5}$, EC, OC, SO_4, NO_3, Fe, K, Si, Zn: long-term average of monthly mean concentration within two selected buffers of a nearest monitor ($\mu g/m^3$)
Nurses' Health Study (NHS)					
Puett 2008 (53)	07/1992-07/2002	1988-2002	PM_{10}	1. Spatiotemporal multilevel land use regression model 2. Various exposure time windows: last month; previous 3 month moving average; previous 12 month moving average; and previous 48 month moving average	PM_{10}: monthly mean concentration ($\mu g/m^3$)
Puett 2009 (189)	07/1992-07/2002	1988-2002	$PM_{2.5}$, $PM_{2.5-10}$	1. Spatiotemporal multilevel land use regression model 2. Various exposure time windows: last month; previous 3 month moving average; previous 12 month moving average; and previous 48 month moving average	$PM_{2.5}$, $PM_{2.5-10}$: monthly mean concentration ($\mu g/m^3$)
Rome Cohort Study (retrospective)					
Rosenlund 2008 (190)	1998-2000	1995-6	NO_2	1. Land use regression model of NO_2 was developed using data from an intensive monitoring campaign over three 7-day periods in 06/1995, 11/1995, and 03/1996 (N=68 sites; R^2=0.69) 2. Exposure to NO_2 was assigned to the subjects' census block of residence in 1995-1996	NO_2: annual mean concentration ($\mu g/m^3$)
Ontario Cohort Study					

First author, year (ref)	Follow-up period	Exposure Assessment: Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Jerrett 2009 (77)	1992-2002	2002-2004	NO_2	1. Land use regression model of NO_2 was developed using data from an intensive monitoring campaign over two 2-week periods in 09/2002 and 05/2004 (N=100 sites; R^2=0.70) 2. Exposure to NO_2 was assigned to the subjects' six-character postal code addresses in 1992-2002	NO_2: annual mean concentration ($\mu g/m^3$)
Shizuoka Elderly Cohort Study					
Yorifuji 2010 (191)	12/1999-03/2006	04/2000-03/2006	NO_2	1. Land use regression model was developed using long-term average of NO_2 from a network of fixed-site monitors in 04/2000-03/2006 (N=67 sites; R^2=0.54) 2. Exposure to NO_2 was assigned to the subjects' census block of residence in 1995-1996	NO_2: long-term average concentration ($\mu g/m^3$)
Trucking Industry Particle Study (retrospective)					
Hart 2010 (187)	1985-2000	PM_{10}, SO_2, NO_2: 1985-2000 $PM_{2.5}$: 2000	PM_{10}, $PM_{2.5}$, SO_2, NO_2	1. PM_{10}, SO_2, NO_2: Spatiotemporal multilevel land use regression model 2. $PM_{2.5}$: annual mean concentration at the nearest fixed-site monitor, 2000 3. Two exposure time windows: current calendar year; long-term average (1985-2000)	PM_{10}, $PM_{2.5}$, SO_2, NO_2: long-term average concentration ($\mu g/m^3$)
Vancouver Cohort Study (retrospective)					
Gan 2011 (80)	1999-2002	Black carbon: 07/14/2005-08/16/2005 $PM_{2.5}$, NO_2, NO: 02/24/2003-03/14/2003; 09/08/2003-09/26/2003	Black carbon $PM_{2.5}$ NO_2 NO	1. NO and NO_2: land use regression model for annual average concentrations was developed using data from two 2-week intensive monitoring campaigns in 2003 (N=116, R^2=0.62 for NO, R^2=0.56 for NO_2) 2. $PM_{2.5}$: land use regression model for annual average concentrations was developed using data from two 2-week intensive monitoring campaigns in 2003 (N=25, R^2=0.52) 3. Black carbon: land use regression model for annual average concentrations was developed using data from a mobile monitoring campaign in 2005 (N=39, R^2=0.56) 4. Twelve monthly spatial surfaces were derived using month-year adjustment factor using monitoring data from fixed-site monitors 4. Monthly average concentrations of each of the four pollutants was assigned to the subjects' historical addresses (six-character postal code) on the monthly basis during 1994 to 1998 5. For each subject and for each of the four pollutants, exposure to the long-term average of monthly concentrations was estimated 6. Subjects who had missing exposure data in more than a total of any 15 months or in more than 3 consecutive months in 1994-1998 were excluded	NO_2, NO, $PM_{2.5}$: average concentrations during the 5-year exposure period ($\mu g/m^3$) Black carbon: average concentrations during the 5-year exposure period (10^{-5}/m)

‡ PM_{10}, Particulate Matter less than 10 microns in diameter; $PM_{2.5}$, Particulate Matter less than 2.5 microns in diameter; GIS, geographic information systems; OC, organic carbon; EC, elemental carbon; Fe, iron; K, potassium; Si, silicon; Zn, zinc

TABLE 11. Exposure assessment in case-control studies of long-term exposure to ambient air pollution and chronic disease, according to outcome and distinct study, 2008-2011

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
Scania Case-control Study of Ischemic Stroke					
Oudin 2009 (192)	2001-2005	2003	NOx	1. An air dispersion model for 2003 was developed, but details of the model were not provided 2. Spatial resolution: 500 meters 3. For cases, exposure was assigned according to their addresses in the year prior to diagnosis 4. For phase I controls, exposure was assigned to their addresses in 2002, and for phase II controls, exposure was assigned to their addresses in 2004	NOx: annual mean concentration ($\mu g/m^3$)
Stockholm County Case-control Study (administrative database)					
Rosenlund 2009 (88)	1985-1996	NO_2, CO_2: 1960-90 PM_{10}: 2000	NO_2, CO, PM_{10}	1. Regional source-specific contributions to NO_2, CO, and PM_{10}, were estimated using air dispersion model based on traffic and heating emission databases 2. Derived NO_2 from NOx from a non-linear regression based on measurements in the early 1980s 3. Annual means of NO_2, CO, and PM_{10} were computed for each year in 1950-1990 by linear extrapolation and interpolation	NO_2, PM_{10}, CO: annual mean concentrations ($\mu g/m^3$)
Worcester Heart Attack Study					
Tonne 2009 (81)	1995, 1997, 1999, 2001, and 2003	09/2003-04/2005	1. $PM_{2.5}$ mass, $PM_{2.5}$ filter absorbance (proxy for elemental carbon), and NO_2 2. Proximity to traffic	1. 1-week integrated concentrations of $PM_{2.5}$ mass, $PM_{2.5}$ filter absorbance, and NO_2 were measured at a total of 36 locations in the study area in 09/2003-04/2005: during any given week, pollutant concentration data were collected at six to 10 sites for a total of 36 different locations. 2. Latent variable modeling (similar to factor analysis) was used with dependent variable being the spatial variation common to measured NO_2 and $PM_{2.5}$ filter absorbance to represent sources of both pollutants such as traffic, and independent variables comprising of spatial coordinate, urbanization density. 3. The spatial correlation of the latent variable was specified using a penalized spline formulation. 4. Proximity to traffic: an indicator of local scale traffic (the	Traffic particles: latent variable of NO_2 and $PM_{2.5}$ filter absorbance (10^{-5}/m) Cumulative traffic: vehicle-km

First author, year (ref)	Diagnosis period(s)	Exposure Assessment			
		Exposure period(s)	Exposure(s)	Method(s) of assessment	Exposure metrics
				sum of the product of road segment length and AADT within 100 m radius buffer the residence)	

‡ NR, not reported; PM_{10}, particulate matter less than 10 microns in diameter; $PM_{2.5}$, particulate matter less than 2.5 microns in diameter; NOx: nitrogen oxides; NO_2: nitrogen dioxide; CO: carbon monoxide

Chapter 4

Nitrogen Dioxide as a Marker for Traffic-related Air Pollution

In the previous chapter, the epidemiological evidence regarding the effects of ambient air pollution on chronic health effects was reviewed. The results support the hypothesis that exposure to ambient air pollution, in particular $PM_{2.5}$, may increase cardiovascular morbidity and mortality. In addition, there is suggestive evidence that exposure to traffic-related air pollution was associated with increased rates of cardiovascular mortality. However, the possible cardiovascular effects of traffic-related air pollution have been examined in relatively few studies. This is a research area where further studies are required and it is addressed specifically in this dissertation.

In metropolitan areas, vehicular traffic is the main contributor of ambient air pollution (61). The emissions from motor vehicles comprise large quantities of gases, particles, volatile organic compounds (VOCs), and polycyclic aromatic hydrocarbons (PAHs) (61). The majority of emissions, by mass, are gaseous pollutants, with NO_2 being an important component (193). NO_2 is itself toxic and as demonstrated by toxicological studies it may exert adverse effects on lung metabolism (194, 195), structure (196), function (197), and inflammation (198). As important as its toxicity, NO_2 is widely used in epidemiological studies as a marker for the complex mixture of traffic-related pollutant (77, 88, 101, 199-201). This is because NO_2 is generally found in the atmosphere in close association with other primary pollutants (202). For example, a study using data from fixed-site monitors in 10 Canadian cities from 1981 to 2000 (193) has shown that in comparison to $PM_{2.5}$, concentrations of NO_2 had stronger co-locational associations with a range of toxic pollutants such as VOCs, PAHs, and particle-bound organic carbon generated from local mobile sources. This study (193) also showed that concentrations of NO_2 exhibited a distinct diurnal pattern that correlated well with the volume of traffic: a sharp increase in the morning associated with rush hour, followed by a decline due to less traffic, and then increases again in the late afternoon. In another study (203), it was found that variation in concentrations of NO_2 around an expressway in Toronto characterized well that of CO, ultrafine particles (particles size less than 0.1 micron in diameter), black carbon, and VOCs.

As a result, NO_2 is a valid marker for traffic-related air pollution and it is the air pollutant that I used in this dissertation.

In the next chapter, three new methods of back-extrapolating land use regression models of NO_2 will be described. The purpose of developing these new methods was to estimate historical exposure to traffic-related air pollution within cities.

Chapter 5

Back-extrapolation of Estimates of Exposure from Current Land-use Regression Models

In this chapter I address the second objective of this dissertation, namely, to develop new methods to extrapolate into the past estimates of exposure from "current" land use regression models of NO_2. The work presented in this chapter draws from and contributes to the literature related to the measurement and modelling of traffic-related air pollution *within* cities. I made use of estimated concentrations of NO_2 in Montreal obtained from a dense monitoring campaign and the resultant land use regression map that was developed by Daniel Crouse and collaborators (71). In the manuscript I make use of an example where we show that postmenopausal breast cancer was associated with these estimated concentrations (2). This manuscript has been published in the journal *Atmospheric Environment* as:

Chen H, Goldberg MS, Crouse DL, Burnett RT, Jerrett M, Villeneuve PJ, Wheeler AJ, Labrèche F, Ross NA. Back-extrapolation of estimates of exposure from current land-use regression models. Atmospheric Environment 2010;44:4346-4354.

Research Paper

Back-extrapolation of estimates of exposure from current land-use regression models

Hong Chen,[1] Mark S Goldberg,[2,3] Dan L Crouse,[4] Richard T Burnett,[5] Michael Jerrett,[6] Paul J Villeneuve,[7, 8] Amanda J Wheeler,[9] France Labrèche,[10] and Nancy A Ross[4]

[1] Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec

[2] Department of Medicine, McGill University, Montreal, Quebec

[3] Division of Clinical Epidemiology, McGill University Health Center, Montreal, Quebec

[4] Department of Geography, McGill University, Montreal, Quebec

[5] Environmental Health Directorate, Health Canada, Ottawa, Ontario

[6] Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California

[7] Population Studies Division, Health Canada, Ottawa, Ontario

[8] Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario

[9] Air Health Science Division, Health Canada, Ottawa, Ontario

[10] Institut de recherche en santé publique de l'Université de Montréal, Montréal, Québec

ABSTRACT

Land use regression has been used in epidemiologic studies to estimate long-term exposure to air pollution within cities. The models are often developed toward the end of the study using recent air pollution data. Given that there may be spatially-dependent temporal trends in urban air pollution and that there is interest for epidemiologists in assessing period-specific exposures, especially early-life exposure, methods are required to extrapolate these models back in time. We present herein three new methods to back-extrapolate land use regression models. During three two-week periods in 2005-2006, we monitored nitrogen dioxide (NO_2) at about 130 locations in Montreal, Quebec, and then developed a land-use regression (LUR) model. Our three extrapolation methods entailed multiplying the predicted concentrations of NO_2 by the ratio of past estimates of concentrations from fixed-site monitors, such that they reflected the change in the spatial structure of NO_2 from measurements at fixed-site monitors. The specific methods depended on the availability of land use and traffic-related data, and we back-extrapolated the LUR model to 10 and 20 years into the past. We then applied these estimates to residential information from subjects enrolled in a case-control study of postmenopausal breast cancer that was conducted in 1996.

Observed and predicted concentrations of NO_2 in Montreal decreased and were correlated in time. The estimated concentrations using the three extrapolation methods had similar distributions, except that one method yielded slightly lower values. The spatial distributions varied slightly between methods. In the analysis of the breast cancer study, the odds ratios were insensitive to the method but varied with time: for a 5ppb increase in NO_2 using the 2006 LUR the odds ratio (OR) was about 1.4 and the ORs in predicted past concentrations of NO_2 varied (OR~1.2 for 1985 and OR~1.3-1.5 for 1996). Thus, the ORs per unit exposure increased with time as the range and variance of the spatial distributions

decreased, and this is due partly to the regression coefficient being approximately inversely proportional to the variance of exposure. Changing spatial variability complicates interpretation and this may have important implications for the management of risk. Further studies are needed to estimate the accuracy of the different methods.

Keywords: Land use regression model; back extrapolation; traffic-related air pollution; historical exposure; Montreal

Published as: Chen H, Goldberg MS, Crouse DL, Burnett RT, Jerrett M, Villeneuve PJ, Wheeler AJ, Labrèche F, Ross NA. Back-extrapolation of estimates of exposure from current land-use regression models. Atmospheric Environment 2010;44:4346-4354.

Introduction

A particular challenge for epidemiologic studies is to accurately estimate historical exposure to traffic-related air pollution within cities. Land use regression (LUR) is a method to predict concentrations of pollutants at locations within cities for which measurements were not taken (60, 70-75, 204-208). The method involves measuring ambient pollutants, usually using a dense environmental sampling campaign, and then developing a prediction model whereby the measured concentrations of air pollutants are regressed against proximate characteristics of land use and vehicular traffic. This method has been used in cohort studies to estimate the association between long-term exposure to air pollution and chronic health outcomes (25, 77-79, 191, 209-213), but can also be relevant to case-control studies. Often, the LUR is developed toward the end of the study (e.g., at end of follow-up) so it is important to understand whether these models can characterize exposure adequately during the relevant etiological periods. The validity of the methodology is especially critical for outcomes that have long latency (e.g., most cancers). Given that there may be non-homogeneous, spatially-dependent temporal trends in urban air pollution and that there is interest for epidemiologists in assessing period-specific exposures, especially early life exposure, methods are required to back-extrapolate these models into the past.

We present here three new methods to extrapolate a "current" LUR back in time by incorporating historical trends in spatially-dependent concentrations of pollutants as well as temporal changes in land use and vehicular traffic. For the purpose of illustration, we sought to back-extrapolate approximately 10 years (to 1996) and 20 years (to 1985) a current LUR that we developed using measurements of nitrogen dioxide (NO_2) in Montreal, Quebec, Canada, from a dense monitoring campaign that we conducted in 2005 and 2006 (71). As a

specific application of these new methods to an epidemiologic study, we applied the models to data from a case-control study of breast cancer that we conducted in the mid-1990s (92, 93, 214).

Methods

Fixed-site monitoring data

Environment Canada, in collaboration with the City of Montreal, administers the Canadian National Air Pollution Surveillance (NAPS) network, a network of fixed-site monitors in the Montreal region (Figure 1). The number of fixed-site monitors used to measure ambient concentrations of NO_2 varied from 8 to 13 during the period of 1985 and 2006. Measurements of NO_2 were made every hour and were analyzed using chemiluminescence (Thermo Environmental Instruments (*TEI*) Model *42C*).

We obtained mean daily concentrations of NO_2 measured at the fixed-site monitors in 1985, 1996, and 2006 (from 10-12 monitors, depending on the year). For each year, we included all fixed-site monitors in the Montreal region. This allowed us to incorporate all sources of spatial variability that were represented in the monitoring network. The social geography of Montreal is quite complex, with some low and high income areas located close to expressways running through the city.

Development of the LUR model for 2005 - 2006

We conducted three separate monitoring campaigns measuring NO_2 between November 2005 and August 2006 (71). Two-week integrated samples were measured using Ogawa passive samplers (Ogawa and Co., USA) that were deployed at 129 locations within the island of Montreal.

In a previous analysis of these data (71), we made use of data from 2006 for land use, population density, and vehicular traffic to develop a LUR model to predict annual mean concentrations of NO_2 on the island of Montreal. We developed the LUR model (hereafter referred to as the "original 2006 LUR model") through an approach that maximized the predictions ($R^2 = 0.80$). We used this LUR model in two of the extrapolation methods presented below.

We also developed an alternative LUR model for 2006 (hereafter called the "alternative 2006 LUR model"), using a supervised forward stepwise model selection procedure (60, 73, 205) that led to a reduced number of covariables as well as a lower R^2. The rationale for developing this latter model was that one of our extrapolation methods involved replacing the spatial predictors in the 2006 LUR model with those from 1985 and 1996. Because of limited land use data in the past, not all spatial variables used in 2006 were available for 1985 and 1996. Another reason for using this approach was that we also developed two additional models using measurements of NO_2 in 1985 and 1996 from the relatively few number of fixed-site monitors, and only a handful of spatial covariables could be included in these models.

Back-extrapolation of the LUR model

Motivated by the analytical strategy that was advocated by Hoek et al. (2002) and Brauer et al. (2003) to normalize temporal misalignment in measured concentrations, we propose that back extrapolation of a "current" land use regression model should have the following general form:

$$\hat{Z}(s)_a = LUR(s)_b \times f\{Z(i)_a\} / f\{Z(i)_b\} \qquad (1)$$

where $\hat{Z}(s)_a$ denotes estimated annual mean concentrations of NO_2 for year *a* at spatial locations *s*, which vary continuously throughout study area, $LUR(s)_b$ represents estimates from a land use regression for year *b* at the spatial locations *s*, and *a* and *b* refer to two separate years with $a < b$. $f\{Z(i)_a\}$ and $f\{Z(i)_b\}$ are two models used to predict NO_2 throughout the study area from the observed concentrations, *Z(i)*, from the same fixed-site monitors, i=1, 2,..., N, in year *a* and *b*, respectively.

A principle rationale for our back extrapolation models was that the surface of NO_2 derived from a LUR model would change over time in proportion to what we observe from fixed-site monitoring stations in the study area. As a result, we sought to model the change in the spatial structure of NO_2 using measured concentrations at the fixed-site monitors, and then we applied these changes to the surface derived from the LUR. Depending on the availability of land use and traffic-related data, we developed three back-extrapolation methods. In each method, we employed a different spatial exposure technique to estimate $f\{Z(i)_a\}$ and $f\{Z(i)_b\}$.

To demonstrate the three methods, we extrapolated estimates of exposure from the original 2006 LUR model and from the alternative 2006 LUR model to 1996 (time of interview in the breast cancer study) and to 1985 (10 years before diagnosis) (Table 1). We obtained spatial data from 1996 and 1985 for land use, population density, and vehicular traffic. Land use and road network data were acquired through DMTI spatial, Inc. (Markham, Ontario, Canada). Population data were obtained from the Canadian censuses of 1986 and 1996 (Statistics Canada, Canada).

For the first method (see Figure 2), we multiplied the surface of NO_2 produced by the original 2006 LUR model at each grid cell (5× 5 m) by the ratio of two surfaces derived using

first-order, inverse-distance weighted interpolation (IDW) of annual concentrations of NO_2 measured at the same nine fixed-site monitors in 1985 and 2006 and at the same 10 monitors in 1996 and 2006. This method is the simplest and requires the least amount of data.

For the second method, we took into consideration recent land use and vehicular traffic (Table 1). In this method, we multiplied the surface of NO_2 produced by the original 2006 LUR model by the ratio of estimates from the two following models. The model for the numerator, $f\{Z(i)_a\}$, was developed by regressing the observed annual concentrations of NO_2 at 10 fixed-site monitors in 1985 against a reduced set of the spatial variables that were used to develop the original 2006 LUR model. We selected the variables using a supervised forward stepwise model selection procedure. For the denominator, $f\{Z(i)_b\}$, we used as dependent variable predicted concentrations from the original 2006 LUR model at the 10 fixed-site monitors for which there were observations in 1985, and then we regressed these predicted values against the same covariables that we selected for the numerator. For 1996, we used 12 fixed-site monitors.

For the third method, we accounted in more detail for historical land use and vehicular traffic (Table 1). In this method, we applied the regression coefficients of the alternative LUR model for 2006 to land use and traffic variables measured in 1985 (referred to as the "refitted alternative LUR"). (This step is analogous to the local calibration method described by Briggs et al. (2000) (215).) As in the second method, we then multiplied this refitted alternative LUR by the ratio of two models in which the numerator, $f\{Z(i)_a\}$, was computed by regressing the observed annual concentrations of NO_2 at 10 fixed-site monitors in 1985 against the reduced set of spatial variables that were used for the refitted alternative LUR model. For the denominator, $f\{Z(i)_b\}$, we used as dependent variable predicted

concentrations at the 10 fixed-site monitors from the refitted alternative LUR model, and then we regressed these predicted values against the same covariables that we selected for the numerator. For 1996, we used 12 fixed-site monitors.

Illustration of the methods: A case-control study of postmenopausal breast cancer

We illustrate the extrapolation methods using data from a hospital-based, case-control study of postmenopausal breast cancer among women who in 1996 and 1997 lived in the greater metropolitan Montreal area (92, 93, 214). Briefly, case subjects were postmenopausal women diagnosed with a primary malignant breast cancer from all of the 18 hospitals in the Montreal region that treated breast cancer. Eligible control subjects had one of 32 other types of cancers, diagnosed in the same hospitals as the cases, and they were frequency-matched to the cases by age and date of diagnosis. The response rates for cases and controls were 81.1% and 75.7%, respectively. Participants completed a structured interview which elicited information on age, weight, height, smoking history, alcohol use, education, reproductive history, occupational history, and other individual characteristics, and home address at the time of interview. We also made use of census data to estimate characteristics of neighbourhoods that subjects lived in at time of interview. For the present analysis, we included 799 subjects whose homes at time of interview were on the island of Montreal.

The research protocol was approved by the McGill University Faculty of Medicine Institutional Review Board and the ethics committees of all participating hospitals and institutions, and informed consent was obtained from participating subjects.

Statistical methods

Using these three extrapolation methods, we predicted concentrations of NO_2 for each grid cell (5× 5 m) on the island of Montreal for 1985 and 1996 and we also computed the mean estimates of NO_2 during the ten-year period between 1985 and 1996 for the island.

Estimates of concentrations of NO_2 were then linked to each subjects' residence at time of interview. Logistic regression was used to estimate adjusted odds ratios (OR) and associated 95% confidence intervals (CI). The covariables included accepted and suspected risk factors for breast cancer as well as occupational exposures that we have found to be associated with risk including occupational exposure to organic solvents with reactive metabolites and polycyclic aromatic hydrocarbons from petroleum sources (214), and neighbourhood-level socioeconomic characteristics.

Predicted concentrations of NO_2 were included as continuous variables and we used natural cubic spline functions having 1 to 3 degrees of freedom to verify that the response function for the association with the risk of breast cancer was linear (216). To account for differences in the absolute value of the distributions (measured concentrations at the fixed-site monitors have decreased by up to 50% over the past 20 years), we computed the ORs for an increase equal to the interquartile range (IQR) of each predicted distribution.

We conducted sensitivity analyses with method-1 (IDW-based extrapolation) that excluded one fixed-site monitor very close to a highway. We predicted concentrations of NO_2 for 1985 and 1996 using the remaining fixed-site monitors, and we repeated the logistic regression analysis with these new estimates of exposure.

Results

Alternative LUR model for 2006

Table 2 shows the results of the alternative 2006 LUR model developed using the supervised forward stepwise model selection approach. Population density (partial R^2=0.29), road network (partial R^2=0.23), point source pollution (partial R^2=0.06), and land use/cover measures (i.e., area of open water, distance to shoreline) (partial R^2=0.07) included in the model explained 65% of the intraurban variability in the natural logarithm of NO_2. In contrast, the original 2006 LUR model included 47 spatial variables and its R^2 was 0.80 (71).

Observed historical trends of NO_2 at the fixed-site monitors

Over the 22 year period between 1985 and 2006, the two fixed-site monitors located in the central areas of Montreal showed a 50% reduction in concentrations of NO_2 (Figure 3). Concentrations decreased less (between 5% to 40%) at fixed-site monitors located in other parts of the city.

Predicted spatial variation of concentrations of NO_2 over time

Estimates of concentrations and the variability of concentrations of NO_2 also decreased in time (Table 3). With each year, the estimated concentrations using the three extrapolation methods had similar distributions, except that method-3 (refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$) produced a lower maximum concentration than the other methods. Bland-Altman plots (Bland and Altman 1986) (Figure 4) show explicitly that the differences in the models are at the higher end of exposure.

Predicted concentrations across different periods were correlated with each other (Table 4). We found for the most part that the spatial patterns were similar across different periods, except that the surface maps produced using method-3 showed that concentrations of NO_2 decreased more in the central areas than in other parts of the city (Figure 5).

Example: Association between estimated past exposures to NO_2 and postmenopausal breast cancer

Table 5 shows the distribution of predicted exposures among the 799 cases and controls estimated at their residences at time of interview (1996). In contrast with the exposure distribution for the entire area, the predicted maximum concentration using method-3 was higher than the other methods. Table 6 shows associations between the incidence of postmenopausal breast cancer and the various metrics of exposure to ambient NO_2. For a 5 ppb increase in NO_2, the fully-adjusted ORs varied between 1.17 and 1.55, with the exposure metric based on the alternative 2006 LUR model showing the highest OR. In contrast, when the ORs were calculated using an increment in exposure equal to the calendar year-specific inter-quartile range (IQR; OR_{IQR}) in the study population, the ORs were more similar, ranging between 1.20 and 1.39. The OR_{IQR} using the exposure metrics developed for the earlier years were higher ORs and had narrower confidence intervals than that in 2006. When the one monitor very close to an inner-city expressway was excluded, the results for method-1 (IDW-based extrapolation) were similar to those including all monitors (data not shown).

Discussion

We presented three new methods to extrapolate into the past estimates of exposure from current land use regression models. By incorporating non-homogeneous spatially-dependent temporal trends in ambient concentrations and changes in land use and vehicular traffic, these extrapolation methods allowed the spatial gradient of estimated concentrations from present models to vary over time and space.

The issue of accurately reconstructing historical exposure to traffic-related air pollution is challenging. In our three new methods, we assumed that we could estimate past exposures by rescaling the current LUR model based on the change in spatial distributions of concentrations of NO_2 derived from fixed-site monitoring stations. These methods for estimating historical spatial distributions (IDW or LUR) are limited by the number of fixed-site monitors. The rescaling entailed multiplying the current LUR model by a ratio which was inherently less accurate, as the final error would be a function of the sum of the errors of each surface. (Indeed, if the actual errors were known, the final error could be estimated using the delta method for the propagation of errors (217). Nevertheless, the expectation is that these extrapolation methods should be more accurate than simply using estimates derived from data collected well after the relevant periods.

The accuracy of extrapolating LURs in other locations will depend on the accuracy of the LUR, the number and spatial distribution of available historical monitors, as well as the method of spatial extrapolation. Among the three extrapolation methods, the IDW-based approach may introduce the largest errors because IDW is an exact interpolator (218) (i.e., predicted concentrations will be constrained by the range of the observed values) and it may produce over-smoothed estimates of spatial variations of NO_2 as compared to land use regression. The land use regression models using sparse monitoring data may be constrained, as in our study, by a small number of spatial covariables that could be included in the model. In addition, relatively few monitors may not provide the full range of urban background concentrations and land use. But when more monitoring data are available, the LUR-based extrapolation approach should produce more accurate estimates. One direction in the

further development of the LUR-based extrapolation may be to explore different ways of incorporating historical land use and vehicular traffic.

Although the primary objective of the current paper was to introduce the new extrapolation methods, validations are required. Empirical comparisons of land use regression models prospectively in time is one method. One can also conduct numerical simulation studies to simulate underlying spatial processes of NO_2 in the past (H Chen et al., in preparation, 2010), by which errors in the back-extrapolated estimates can be evaluated.

To our knowledge, there is only one other published study of back-extrapolation of land use regression models (219) that compared, in the Netherlands, predicted concentrations of NO_2 between two LURs, one of which was developed in 1999-2000 and the other in 2007. It was found that the predicted concentrations of NO_2 for the measurement sites using the two LURs were strongly correlated (R^2=0.81). In addition to back-extrapolating LURs, one other method for reconstructing historical exposure is to model annual mean concentrations of NO_2 for each long-term monitoring site as a function of time and location (220-223). In the study by Clougherty et al. (220), secular trends were modelled by regressing annual mean concentrations at each monitor against an indicator variable for each sampling year and they modeled the location effects by regressing annual mean concentrations against land use, population density, and vehicular traffic. The two models were then integrated into a final prediction model using a multilevel modelling approach (Goldstein 2003) (224). Hart et al. (2009) (221) developed a model including indicator variables for calendar year, monitoring locations, and proximate land use and traffic for the prediction of NO_2 in the continental United States, 1985-2000, and they fitted the model using generalized additive models. For all the methods that were used to reconstruct historical exposures, the number of monitors

is crucial for estimating fine-scale spatial variability: in our method, we made use of 129 monitors whereas Clougherty et al. (2007) relied on 13 fixed-site monitoring stations (220).

One limitation of the breast cancer study is that historical addresses of subjects were not available, so that estimates of past exposures were limited to the 10 years before diagnosis. However, an important observation from our analyses of the breast cancer study was that the ORs per unit exposure increased with time as the range and variance of the spatial distributions decreased. This increase in the per-unit risk occurred because the regression coefficient is approximately inversely proportional to the variance of exposure (conditional on the distribution of cases and controls being constant). The approximate invariance of the ORs according to a measure of the spread of each distribution (i.e., calendar year-specific interquartile range) was likely a consequence of this. Often in air pollution epidemiologic studies relative risks are reported as per-unit (or per 5 or 10 unit) increment in exposure. Our observation about per-unit risk increasing with time implies that without further information on actual personal exposures arising from ambient pollution, it is difficult to know which slope should apply, and this may have important implications for the management of risk, such as in setting standards. In addition, for chronic diseases, one needs also to know the appropriate etiological period. Thus, the use of spatially-derived exposure metrics is more complicated to interpret than in situations for which personal exposures are used.

Conclusions

In summary, we developed three methods to extrapolate current land use regression models back in time, and we showed in a concrete example that plausible estimates of risk were obtained. Using these methods in other jurisdictions will depend on the availability in the

past of the number and spatial distribution of fixed-site monitoring stations. Further studies are needed to estimate the accuracy of the different methods.

Acknowledgements

Support for this project was provided by grants from the Canadian Breast Cancer Initiative and the Canadian Institutes of Health Research (CIHR). Hong Chen gratefully acknowledges receipt of a CIHR Doctoral Research Award.

Figure Legends

Figure 1. Fixed-site monitors by the National Air Pollution Surveillance (NAPS) network in the Montreal region, 1985-2006. The number of fixed-site monitors varied between 8 and 13 depending on the year.

Figure 2. Illustration of ratio-based extrapolation methods. The map on the right is the LUR map generated for 2006. The maps on the left were generated either through inverse distance weighted interpolation or land-use regression, depending on the specific method (see **Table 1**). The ratio is taken by dividing, for each grid cell (5m × 5m), the predicted value of concentrations of NO_2 in 1985 or 1996 on the top map (numerator) by the predicted value for 2006 on the bottom map (denominator), and then multiplying each grid cell by the corresponding predicted value of the 2006 LUR prediction map.

Figure 3. Trends in observed annual average concentrations of NO_2 (ppb) across 9 fixed-site monitors in the Montreal region, 1985-2006. Among them, two fixed-site monitors, "De Maisonneuve" and "Duncan-Décarie", are located in the central areas of Montreal. Seven fixed-site monitors that operated for less than half of the period (<11 years) are not shown.

Figure 4. Bland-Altman Plots for comparing predicted spatially-dependent NO_2 for 1985 between (A) extrapolation method 1 and method 2, (B) extrapolation method 1 and method 3, and (C) extrapolation method 2 and method 3. (Results for 1996 are similar to those for 1985.) Method 1: Original LUR_{2006} × ratio of IDW_{fixed}; Method 2: Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$; Method 3: Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$

Figure 5. Spatial surface concentrations of NO_2 (ppb) produced in (A) 2006 using the alternative 2006 LUR model, and in (B) 1996, (C) 1985, and (D) the average of 1985 and 1996, using the extrapolation method based on refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$. The colours in the maps are annual mean concentrations (ppb) of NO_2 classified in deciles.

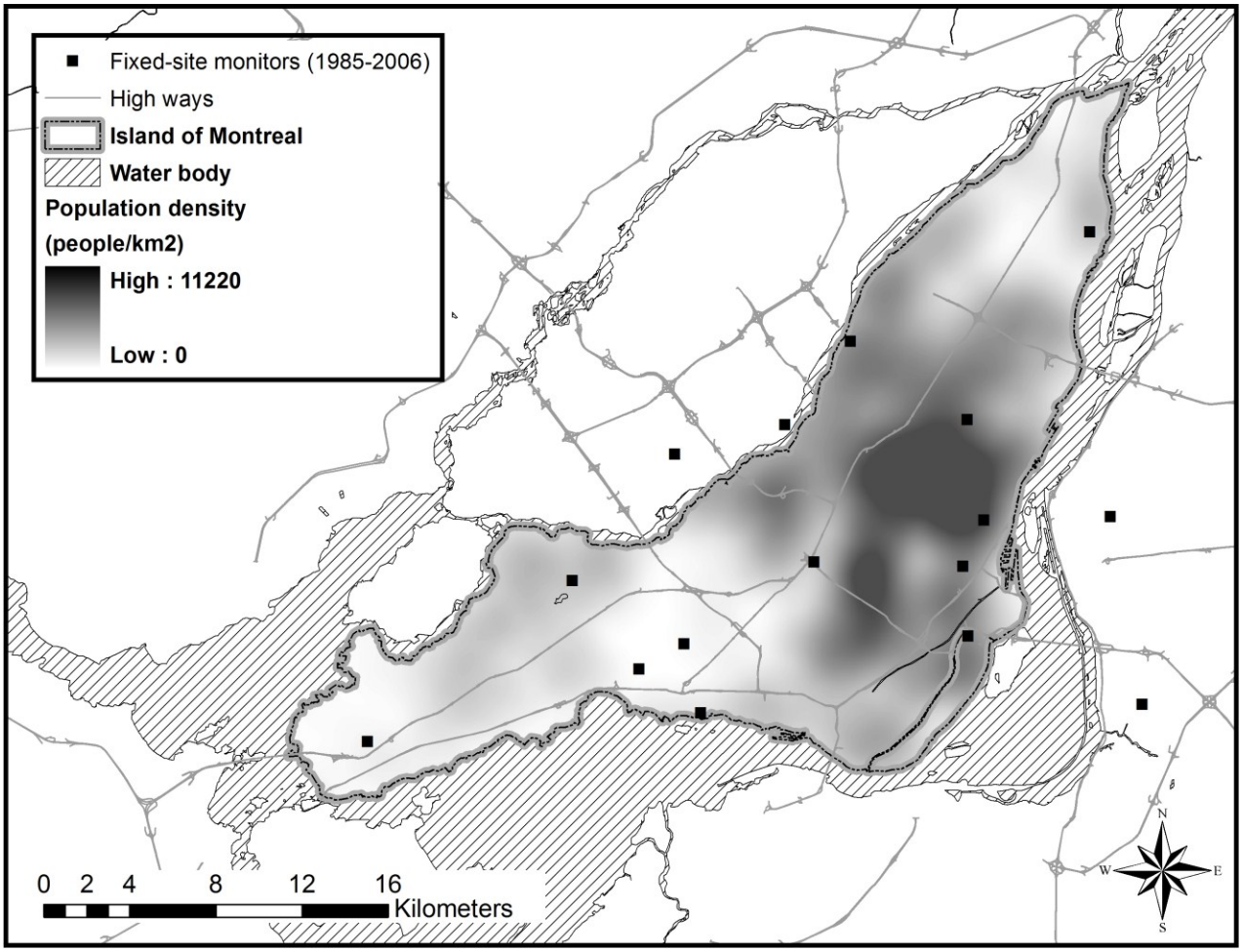


Figure 1. Fixed-site monitors by the National Air Pollution Surveillance (NAPS) network in the Montreal region, 1985-2006.

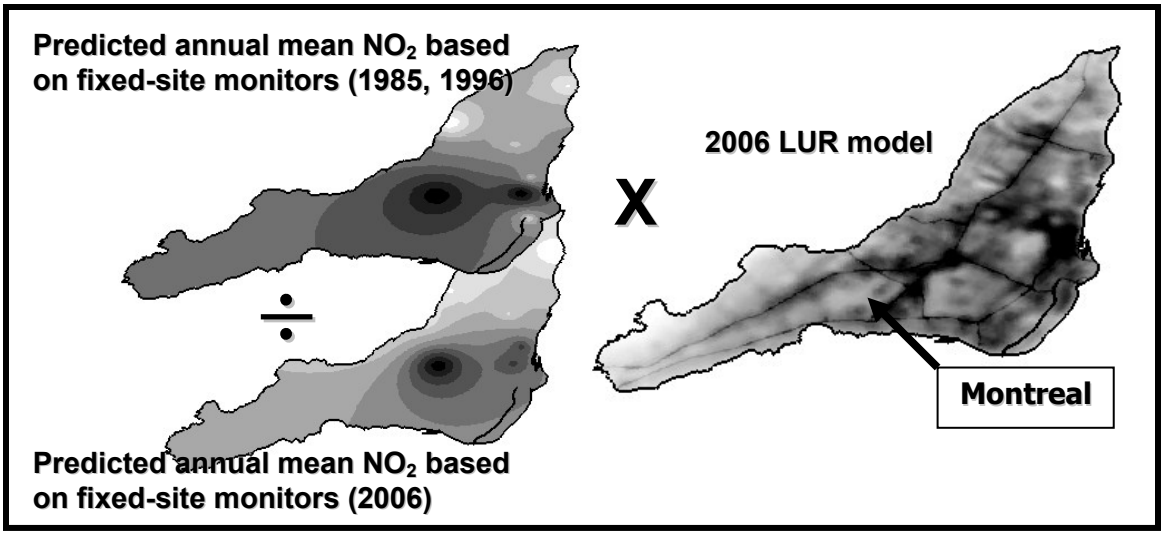


Figure 2. Illustration of ratio-based extrapolation methods.

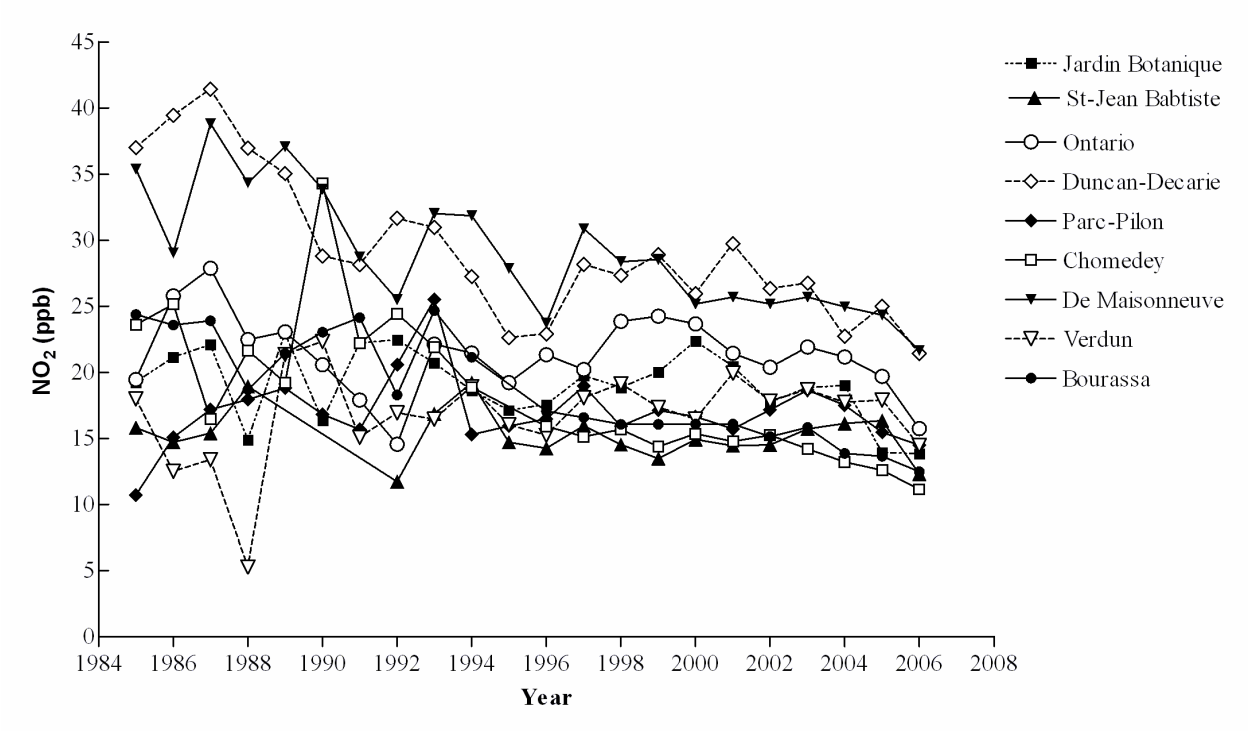
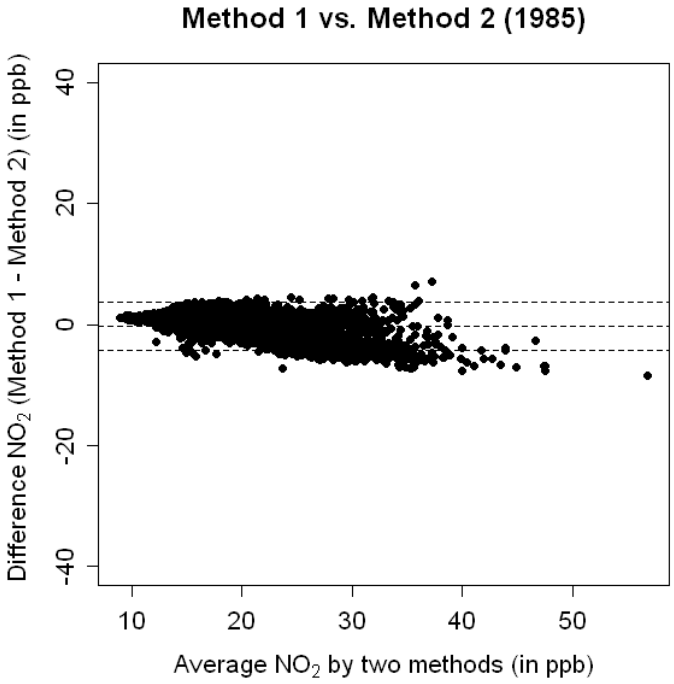
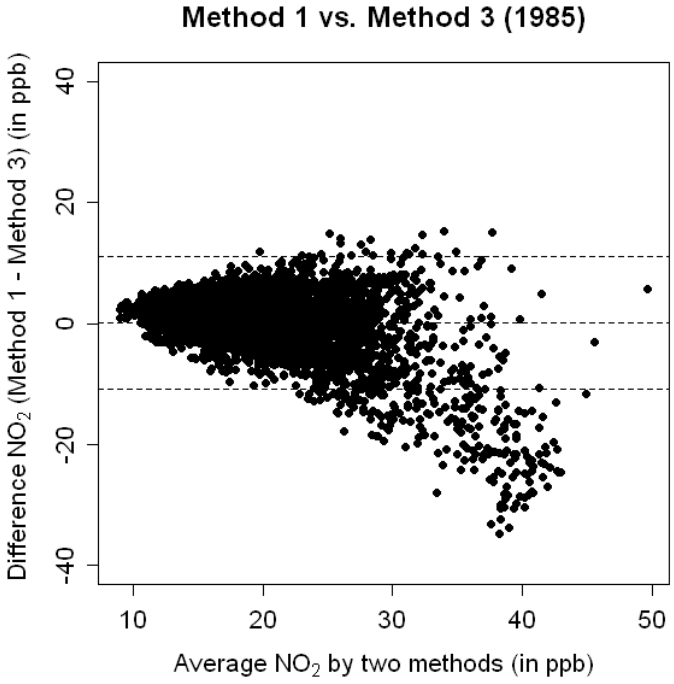


Figure 3. Trends in observed annual average concentrations of NO_2 (ppb) across 9 fixed-site monitors in the Montreal region, 1985-2006.


Method 1 vs. Method 2 (1985)
Difference NO_2 (Method 1 - Method 2) (in ppb)
40
20
0
-20
-40
10
20
30
40
50
Average NO_2 by two methods (in ppb)


Method 1 vs. Method 3 (1985)
Difference NO_2 (Method 1 - Method 3) (in ppb)
40
20
0
-20
-40
10
20
30
40
50
Average NO_2 by two methods (in ppb)

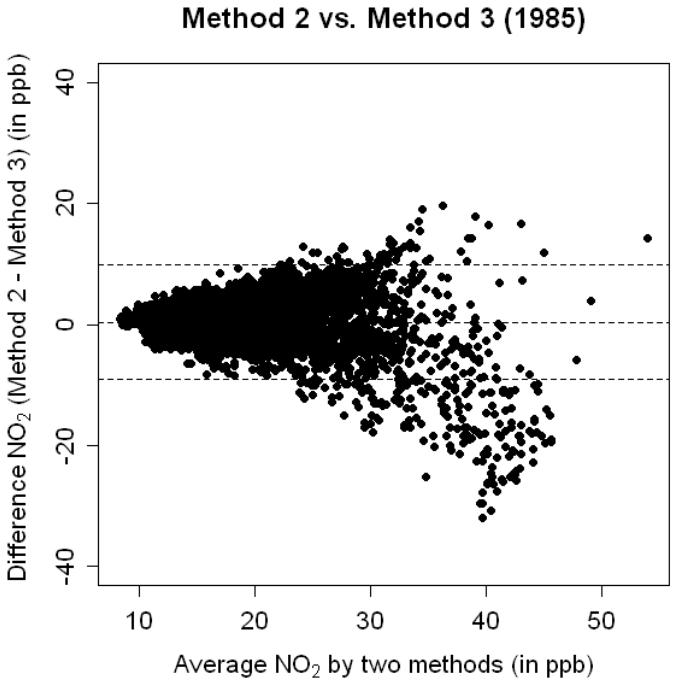
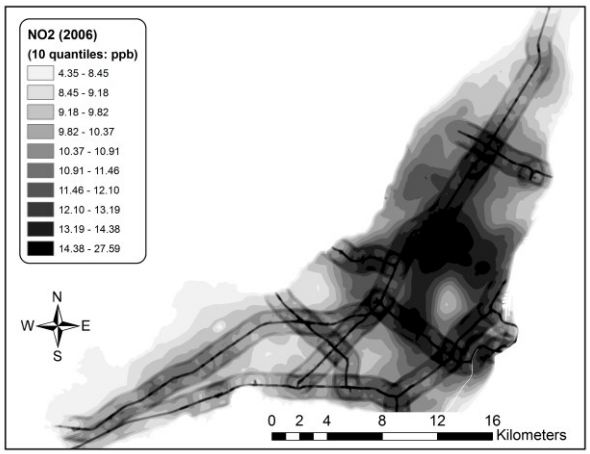
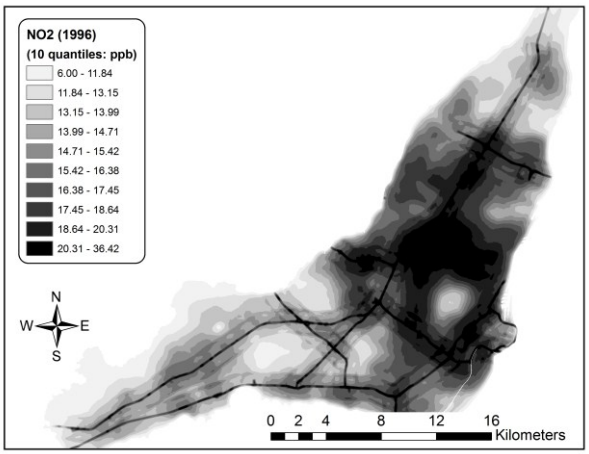


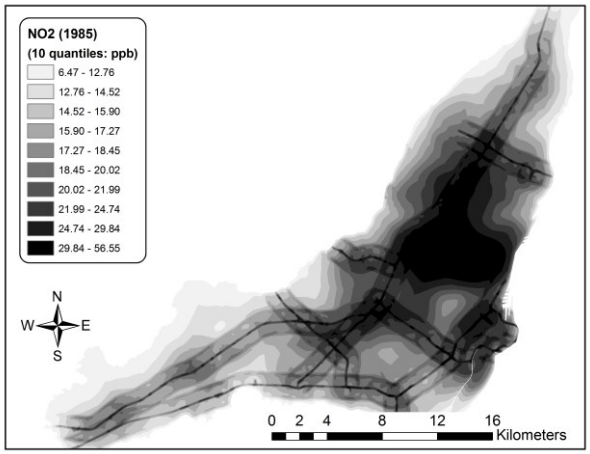
Figure 4. Bland-Altman Plots for comparing predicted spatially-dependent NO_2 for 1985 among three extrapolation methods. Results for 1996 are similar to those for 1985.



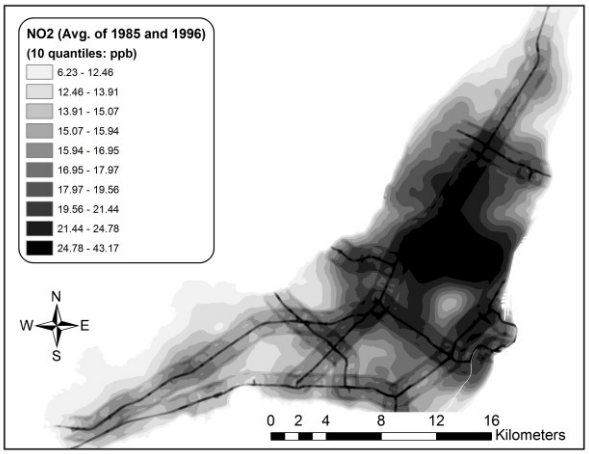
(A) NO_2 (2006): Alternative 2006 LUR model (R^2=0.65)



(B) NO_2 (1996): Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$



(C) NO_2 (1985): Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$



(D) NO_2 (1985-96): mean of 1985 and 1996 using the methods in (B) and (C)

Figure 5. Spatial surface concentrations of NO_2 (ppb) produced in (A) 2006 using the alternative 2006 LUR model, and in (B) 1996, (C) 1985, and (D) the average of 1985 and 1996, using the extrapolation method based on refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$.

Table 1. Description of different methods for extrapolating back in time (1985, 1996) current (2006) estimates of exposure from land use regression models

	Short title for method	Description of the method of extrapolation
1	Original LUR_{2006} × ratio of IDW_{fixed} [a, b]	$\textit{Original } LUR_{2006} \times \frac{\text{IDW observed } [NO_2] \text{ at fixed site monitors, 1985 or 1996}}{\text{IDW observed } [NO_2] \text{ at fixed site monitors, 2006}}$ [c]
2	Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$	$\textit{Original } LUR_{2006} \times \frac{\text{LUR observed } [NO_2] \text{ at fixed site monitors, 1985 or 1996}}{\text{LUR predicted } [NO_2] \text{ at fixed site monitors from original 2006 LUR model}}$ [d, e]
3	Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$	*Step 1: Refitted alternative LUR* = applied regression coefficients from alternative LUR_{2006} to land use and traffic in 1985 or 1996 [f] *Step 2:* $\textit{Refitted alternative LUR} \times \frac{\text{LUR observed } [NO_2] \text{ at fixed site monitors, 1985 or 1996}}{\text{LUR predicted } [NO_2] \text{ at fixed site monitors from refitted alternative LUR}}$ [g, h]

a. IDW: inverse-distance weighted interpolation; LUR: land-use regression model; $[NO_2]$: NO_2 concentration

b. Original 2006 LUR model: $R^2 = 0.80$

c. Extrapolation to 1985: same 9 fixed-site monitors in 1985 and 2006; extrapolation to 1996: same 10 fixed-site monitors in 1996 and 2006

d. For extrapolation to 1985, the variables included in the two LURs of the ratio are lengths of expressway (2006) and distance to shoreline (R^2=0.57). For extrapolation to 1996, the variables included in the two LURs of the ratio are lengths of all roads (2006), lengths of expressway (2006), and distance to shoreline (R^2=0.89)

e. Extrapolation to 1985: same 10 fixed-site monitors in 1985 and 2006; extrapolation to 1996: same 12 fixed-site monitors in 1996 and 2006

f. This step is similar to the local calibration method described by Briggs et al. (2000). Alternative 2006 LUR model: $R^2 = 0.65$.

g. For extrapolation to 1985, the variables included in the two LURs of the ratio are lengths of expressway (1985) and distance to shoreline (R^2=0.57). For extrapolation to 1996, the variables included in the two LURs of the ratio are lengths of all roads (1996), lengths of expressway (1996), and distance to shoreline (R^2=0.89).

[h] Extrapolation to 1985: same 10 fixed-site monitors in 1985 and 2006; extrapolation to 1996: same 12 fixed-site monitors in 1996 and 2006

Table 2. Summary of regression results for the LUR model in Montreal, 2006, developed using a supervised forward stepwise model selection approach (NO_2 in ppb) [a]

	Coefficient	Standard Error	t	R^2 (partial correlation) [b]
ln(NO_2) (Constant)	2.224	0.055	40.85	
Length of expressway within 100m	0.343	0.077	4.45	0.14
Length of expressway between 750 and 1000m	0.031	0.010	2.95	0.06
Distance to expressway	-0.0000421	0.0000192	-2.20	0.03
Density of population within 2500m (Kernel estimate)	0.0000489	0.0000068	7.21	0.29
Area of open water within 300m	-0.0129	0.0076	-1.70	0.03
Distance to shoreline	0.0000250	0.0000117	2.15	0.04
Distance to nearest NPRI sites [c]	-0.0000620	0.0000220	-2.81	0.06

[a.] Number of observations=129; final model R^2=0.65
[b.] The R^2 contribution of each individual variable for the model while controlling for all other variables
[c.] NPRI: National Pollutant Release Inventory (see *http://www.ec.gc.ca/inrp-npri/* for more information)

Table 3. Distributions of estimated annual mean concentrations of NO_2 (ppb) at all grid cells (5× 5 m) in Montreal across three time periods, according to the two 2006 LUR models and the three different extrapolation methods

Exposure metrics	Year	Minimum	25th percentile	Median	75th percentile	Maximum	Mean	Coefficient of variation (%)
1. Original LUR model (R^2=0.80)	2006	4.23	8.84	10.33	12.19	35.94	10.78	24.9
2. Alternative LUR model (R^2=0.65)	2006	4.35	9.55	10.82	12.37	27.59	11.06	20.4
3. Original LUR_{2006} × ratio of IDW_{fixed}	1985	7.82	16.37	19.38	23.07	66.76	20.12	25.6
	1996	5.95	12.90	15.16	17.88	44.47	15.65	25.2
4. Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	6.96	15.76	19.17	24.00	79.39	20.40	31.0
	1996	5.75	12.87	15.39	18.04	41.36	15.69	25.5
5. Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	6.47	15.11	18.26	22.97	56.55	20.05	38.2
	1996	6.00	13.51	15.30	17.93	36.42	15.76	20.7

Table 4. Correlations (in R^2) between estimated annual mean concentrations of NO_2 across three time periods in Montreal, according to the surface maps of NO_2 produced using the LUR model for 2006 and three extrapolation methods (based on 5000 random locations on the island of Montreal) [a]

	R^2 between periods		
	1985	1996	2006
Original LUR_{2006} × ratio of IDW_{fixed} [b]			
1985	1	0.92	0.94
1996		1	0.96
Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$ [b]			
1985	1	0.87	0.89
1996		1	0.90
Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$ [c]			
1985	1	0.71	0.70
1996		1	0.90

[a.] The linear relationship between concentrations of NO_2 in two separate years was confirmed from visual inspection of the scatter plots. Similar correlations were observed for the 799 residential locations included in the analysis of the breast cancer case-control study.

[b.] Original 2006 LUR model: R^2=0.80

[c.] Alternative 2006 LUR model: R^2=0.65

Table 5. Distributions of estimated annual mean concentrations of NO_2 (ppb) at the 1996 home addresses of cases and controls (n=799) in Montreal across three time periods, according to the two 2006 LUR models and the three different extrapolation methods

Exposure metrics	Year	Minimum	25th percentile	Median	75th percentile	Maximum	Mean	Coefficient of variation (%)
1. Original LUR model (R^2=0.80)	2006	5.71	10.34	11.80	13.96	25.16	12.27	22.4
2. Alternative LUR model (R^2=0.65)	2006	6.17	10.45	11.66	13.54	19.86	11.94	29.4
3. Original LUR_{2006} × ratio of IDW_{fixed}	1985	10.48	18.25	20.86	24.28	45.89	21.48	22.0
	1996	7.97	14.38	16.46	19.30	34.32	17.00	22.1
4. Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	9.33	18.00	21.75	26.53	47.20	22.55	27.5
	1996	8.14	14.97	17.16	19.87	34.92	17.58	21.6
5. Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	9.59	17.15	21.30	26.97	55.39	23.54	40.4
	1996	9.14	15.61	17.47	19.63	29.52	17.53	17.1

Table 6. Odds ratios (OR) and associated 95% confidence intervals (CI) between the incidence of post-menopausal breast cancer and estimates of exposure to NO_2 (ppb) using different exposure metrics, per an increase of 5 ppb as well as per an increase of the interquartile range, Montreal

Exposure metrics	Year	IQR (ppb) [a]	Fully-Adjusted [b]			
			OR (5ppb)	95% CI	OR (IQR)	95% CI
1. Original LUR model (R^2=0.80)	2006	3.61	1.37	0.95 - 1.96	1.25	0.96 - 1.63
2. Alternative LUR model (R^2=0.65)	2006	3.09	1.55	0.96 - 2.50	1.31	0.98 - 1.76
3. Original LUR_{2006} × ratio of IDW_{fixed}	1985	6.03	1.17	0.95 - 1.44	1.20	0.94 - 1.55
	1996	4.92	1.32	1.01 - 1.72	1.31	1.01 - 1.71
	Mean of 1985-1996	5.50	1.23	0.97 - 1.56	1.26	0.97 - 1.63
4. Original LUR_{2006} × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	8.53	1.17	1.00 - 1.38	1.31	0.99 - 1.73
	1996	4.90	1.36	1.04 - 1.79	1.36	1.04 - 1.79
	Mean of 1985-1996	6.37	1.24	1.01 - 1.53	1.32	1.01 - 1.72
5. Refitted alternative LUR × ratio of LUR_{fixed} to $LUR_{predicted}$	1985	9.82	1.18	1.05 - 1.32	1.37	1.10 - 1.72
	1996	4.02	1.47	1.04 - 2.09	1.37	1.03 - 1.81
	Mean of 1985-1996	6.62	1.28	1.07 - 1.53	1.39	1.10 - 1.75

[a] The IQR is computed from the exposures that were assigned at the 1996 home addresses of cases and controls (n=799) and it varies between years and back-extrapolation methods (see Table 5)

[b.] Adjusted for age at diagnosis, hospital, benign breast disease, mother or sister with breast cancer, oophorectomy, ethnicity, education, age at menarche, age at full-term pregnancy, breast feeding (number of weeks), oral contraception use (year), hormone replacement therapy (number of months), body mass index, tobacco smoking, alcohol use, proxy respondent, percent without a high school diploma, median household income (in 1996), and past occupational exposures to organic solvents with reactive metabolites, extremely low frequency magnetic fields, polycyclic aromatic hydrocarbons from petroleum sources, and carbon monoxide

Chapter 6

A Cohort Study in Ontario, Canada, of Exposure to Traffic-Related Air Pollution and Cardiovascular Mortality

In the previous chapter the methods of back-extrapolation of estimates of "current" land use regression models were described. These new methods allow the spatial gradient of estimated concentrations from a land use regression model to vary over time and space, such that period-specific exposures can be estimated. They are especially useful to studies used to determine the health effects of long-term exposure to traffic-related air pollution specifically where a land use regression model is used in a cohort study to estimate exposure at the end of the follow-up period or in a case-control study at the time of interview.

In this chapter I address the third and final objective of this dissertation, namely, to determine whether an association exists between traffic-related air pollution and cause-specific cardiovascular mortality among adults living in Ontario, Canada. This chapter builds upon the study findings and methods that were acquired and developed through the research presented in the previous two manuscripts. I describe in this paper how intra-urban contrasts in traffic-related air pollution were related to cardiovascular outcomes among subjects in a cohort study conducted in Ontario. The study population was sampled randomly from the Canadian income tax database and comprised individuals in 10 cities in Ontario. The present analysis comprised a sub-population who lived in Toronto, Hamilton, or Windsor, the only three cities in which land use regression models had been developed (in 2002-2004). However, because a sufficient number of fixed-site monitoring stations were available only in Toronto, we only applied these back-extrapolation methods to subjects living in Toronto.

The data for the cohort comprised a number of personal risk factors for cardiovascular disease but information on subjects' smoking behaviour was not available. Thus, the paper also describes my extension of a method invented by Axelson (4) and extended by Steenland and Greenland (225) to adjust indirectly for the potential confounding effects of smoking.

This manuscript will be submitted in the Summer 2011:

Chen H, Goldberg MS, Burnett RT, Jerrett M, Wheeler AJ, Villeneuve PJ. A cohort study in Ontario, Canada, of exposure to traffic-related air pollution and cardiovascular mortality. To be submitted.

Original Research Article

A Cohort study in Ontario, Canada, of exposure to traffic-related air pollution and cardiovascular mortality

Hong Chen,[1] Mark S Goldberg,[2,3] Richard T Burnett,[4] Michael Jerrett,[5] Amanda J Wheeler,[6] and Paul J Villeneuve[7, 8]

[1] Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec

[2] Department of Medicine, McGill University, Montreal, Quebec

[3] Division of Clinical Epidemiology, McGill University Health Centre, Montreal, Quebec

[4] Environmental Health Directorate, Health Canada, Ottawa, Ontario

[5] Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, California

[6] Air Health Science Division, Health Canada, Ottawa, Ontario

[7] Population Studies Division, Health Canada, Ottawa, Ontario

[8] Division of Occupational and Environmental Health, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario

Corresponding Author:

Hong Chen, M.S.

Division of Clinical Epidemiology

McGill University Health Centre – RVH

687 Pine Avenue West, R4.29

Montreal, Quebec H3A 1A1

Phone: (514) 934-1934, ext 36917

Fax: (514) 843-1493

Email: hong.chen2@mail.mcgill.ca

Competing interest declaration: We declare that we have no competing financial interests.

Running head: Traffic air pollution and cardiovascular mortality in a cohort study in Ontario, Canada

Keywords: Cardiovascular mortality, Traffic air pollution, Land use regression model, Cohort study, Ontario, Canada

Acknowledgements: Support for this project was provided by grants from the Canadian Institutes of Health Research (CIHR). Hong Chen gratefully acknowledges receipt of a CIHR Doctoral Research Award.

ABSTRACT

Background: In several longitudinal studies associations have been found between air pollution and cardiovascular mortality, but in few studies has intraurban variations in traffic-related pollution been assessed, especially at the relatively lower concentrations of pollution such as observed in Canadian cities. The purpose of this study was to determine the association between traffic-related air pollution, as measured by concentrations of nitrogen dioxide (NO_2), and cause-specific cardiovascular mortality among adults living three cities in Ontario, Canada.

Methods: This population-based cohort study used as a sampling frame the income tax database which captures 95% of Canadians. More than 660,000 adults who were 35 years of age and older and lived in 10 cities in Ontario between 1982 and 1986 were randomly selected from this database. Mortality was ascertained by linking all subjects to the Canadian Mortality Database. Follow-up extended until the earliest of death, or December 31, 2004. We report herein an analysis of a sub-cohort who, at the time of entry, lived in Toronto, Hamilton, or Windsor, Ontario (205,430 subjects). Residential concentrations of NO_2 were estimated using land use regression models. Cox proportional hazard model were used to estimate associations between concentrations of NO_2 and mortality from cardiovascular and cerebrovascular disease. Associations were adjusted for individual risk factors and selected ecological covariables, and we conducted sensitivity analyses using indirect methods to adjust for smoking.

Results: For each increase of 5 parts per billion (ppb) of NO_2, the rate ratios for mortality for all cardiovascular disease and for ischemic heart disease, using different methods, varied between 1.04 (95%CI: 1.00-1.09) and 1.10 (95%CI: 1.00-1.21). We found no associations between traffic pollution and cerebrovascular mortality (excess rates of mortality of 0.95 (95%CI: 0.89-1.02).

Conclusion: Our findings support the hypothesis that long-term exposure to traffic-related air pollution increases mortality rates of cardiovascular disease and in particular, from ischemic heart disease.

INTRODUCTION

Long-term exposure to outdoor air pollution, particularly fine particles ($PM_{2.5}$), has been implicated in the aetiology of cardiovascular disease (26, 28, 49, 50, 57, 58, 62, 65, 82-85). From a meta-analysis of published studies, we estimated that long-term exposure to $PM_{2.5}$ increases the risk of mortality from cardiovascular disease by 12%-14% per a 10μg/m^3 increase (1). Several recent studies have added to this evidence (77, 80, 88, 190, 191).

In most of the cohort studies, rates of cardiovascular mortality were compared between cities, and exposure was assigned using average city-specific concentrations of pollutants estimated near the beginning of the follow-up period (26, 50, 57, 58). In some recent studies of cardiovascular mortality, finer-scale estimates of air pollution within cities have been made and these have shown stronger associations than the between-city estimates of exposure (28, 49). The inter-city studies measure regional differences in air pollution from many different sources and the intra-city estimates of exposure measure spatial variations in local concentrations of air pollution. Indeed, in the land-use regression models that are based on dense campaigns of measurements of nitrogen dioxide (NO_2) (70, 71, 204), the main component of air pollution that is being measured is related to emissions from traffic. Thus, the differences of the effects on health between the inter-city and intra-city studies may be due to measuring different components of air pollution (2), but the differences may also be due to chance, different types and levels of errors in measuring exposure and confounding, and possibly that traffic-related air pollutants may have higher toxicity than from other sources (59).

To measure exposure to traffic-related air pollution at local scales, several methods have been developed. Proximity to roadways is often considered a surrogate for traffic-related air pollution and positive associations have been found for mortality from cardiovascular disease (49, 62-64). This metric suffers from a number of sources of errors, including not accounting for factors related to dispersion of pollutants, such as topography, meteorology (61), and the different types of vehicles and engines (e.g., gas, diesel) that influence emissions (68).

Spatial interpolation, such as inverse distance weighting and kriging (226), at fine geographic scales of fixed-site monitors that measure criteria air pollutants has also been used (49, 65), but its accuracy requires a dense distribution of monitoring stations, which is rarely available.

A third approach is air dispersion models that make use of deterministic equations to characterize the atmospheric processes that disperse air pollutants emitted by vehicular traffic. Using estimated emission rates from traffic sources and atmospheric and meteorological inputs, air dispersion models predict concentrations of air pollution at selected locations (69). Limitations of these methods include a lack of emission and micro-climatic indicators to make accurate predictions at small scales as well as lack of validation studies to ensure the models function appropriately (61).

Another method that has been developed to estimate long-term averages of exposure to traffic-related pollution at fine scales is land use regression models (64, 70, 71). This method involves measuring ambient pollutants, usually using a dense sampling campaign, and then developing a statistical prediction model whereby the measured concentrations of air pollutants are regressed against proximate characteristics of land use and vehicular traffic (60, 70-75). Land use regression has been shown to produce reasonably good predictions of small-scale spatial variability of NO_2, a marker of traffic-related air pollution (63). These models function well with a relatively large amount of variability in NO_2 explained by the models (often above 80%) (60, 71, 72, 76). A few validation studies of comparing predicted concentrations of NO_2 with independent data from routine monitoring sites have been carried out and showed excellent predictions of NO_2 at the validation sites (coefficient of determination, R^2>80%) (215, 227). A minor limitation of this method is that the models are not generalizable to other locals (68).

While several cohort studies have shown associations between air pollution and cardiovascular mortality (25, 26, 28, 49, 50, 57, 58, 62, 64, 65, 77, 80, 84, 85, 190, 191), few studies (25, 62, 77, 80, 190) have been used to assess intra-urban variation in traffic-related air pollution or evaluated health effects at the relatively lower levels of pollution that are observed in some areas, such as in Canadian cities (73, 80). In addition, the few studies using land use regression models have had modest sample sizes, so statistical power was limited

(62, 77). To address these gaps, we designed the present study to determine the association between traffic-related air pollution and cause-specific cardiovascular mortality among adults living in 10 cities in Ontario, Canada. This study is a population-based cohort that used as a sampling frame the Canadian federal income tax "T1 Family File". We report herein an analysis of a subcohort of 205,430 adults who lived in Toronto, Hamilton, or Windsor, Ontario, and for whom we were able to make use of concentrations of NO_2 estimated from land use regression models developed previously (73-75).

METHODS

The Study Areas

The cities of Toronto, Hamilton, and Windsor are the largest, the ninth largest, and the fifteenth largest cities in Canada, with a population of approximately 2.5 million, 505,000, and 220,000 people, respectively (Canadian Census 2006) (**Figure 1**). In the three cities, 54% to 56% of the population are 35 years of age and above (228). Foreign-born individuals (landed immigrants) account for 50%, 25%, and 23% of the population of the three cities, respectively (228). Among them, 22%, 13%, and 18% are immigrants who landed in Canada within the past five years, respectively (228).

Similar to other large cities in North America, many expressways traverse the three cities. Of them, Toronto has the highest density of cars of any Canadian city. In addition, Windsor has the busiest international crossings between Canada and the United States, and Hamilton houses one of the largest industrial areas in Canada, which has two large steel manufacturing complexes. Indeed, the three cities have levels of pollution higher than most other Canadian cities (229), yet lower than the study areas in most other published studies and ambient pollution and cardiovascular disease (1).

Source and Study Populations

The Ontario Tax Cohort study is a retrospective cohort study, with approximately 660,000 participants who were followed up during the period of 1982 to the earliest of death, or

December 31, 2004. The source population of the study comprised all persons who filed tax returns between 1982 and 1986 and all family members, aged 35 years and above, identified on the tax returns. The study population was constructed by sampling randomly 660,000 individuals from the federal income tax database. Canadian citizens or landed immigrants, 35-85 years of age, were eligible to be included in the cohort if they had filed at least one T1 tax return when they resided in any of 10 urban centres in Ontario (Hamilton, Kingston, London, St. Catherine, Sarnia, Sudbury, Toronto, Thunder Bay, Windsor) during the period 1982 to 1986. The date of entry into the cohort was January 1 of the year of the first tax filing during 1982 to 1986. The latter years were used in the smaller cities to comply with privacy guideline for sampling of populations. Indeed, for the larger cities such as Toronto, Hamilton, and Windsor, almost every subject may have entered in the cohort in 1982.

This database was established by Statistics Canada in 1982 for the development and dissemination of social, economic and demographic statistics and indicators for sub-provincial geographic areas such as postal areas and selected census areas (230). It is thus an excellent sampling frame of the Canadian adult population (231). A comparison of the 1989 T1 tax data with census data showed that the T1 tax files provided fairly accurate estimates of the population: 98% for the overall population; 94% for central cities; and 98% for suburbs (232).

The study population for the present analysis included 205,430 participants who lived in one of three urban centres (Toronto, Hamilton, Windsor) between 1982 and 1986. Our analyses were restricted to the three cities for which land use regression models were available. The follow-up period thus extended from enrolment until December 31, 2004, the last date in which mortality data were available. Because Statistics Canada put a constraint on the maximum number of subjects to be sampled from each city, the percentage of subjects living in the largest city, Toronto, is lower than in the other two cities. Ninety-five percent of study subjects were entered into follow-up in 1982. All data are held and analyzed at Statistics Canada and individuals were not identifiable. The study was approved by the Research Ethics Board of Health Canada and by Statistics Canada.

Assessment of Exposure to Traffic-related Air Pollution

Our estimates of traffic-related air pollution for the three cities derived from dense measurement campaigns of NO_2 to develop land use regression models (73-75). Briefly, in two of the three cities (Hamilton and Toronto) a two-week integrated sampling campaign of ambient NO_2 was conducted using two-sided Ogawa passive diffusion samplers (Ogawa and Co., Pompano Beach, FL, USA). Two-sided samplers were deployed in pairs at a height of 2.5 m above roadways. In Hamilton, the Ogawa samplers were positioned at 107 locations between October 21 and November 6, 2002 (74). This period was chosen because it was the closest to the average annual weather conditions for Hamilton, rather than for the extremes of winter or summer conditions (74). In Toronto, the samplers were set up at 100 locations and the sampling campaign was repeated in two seasons, one in the fall of 2002 (September 9 to 24) and the other in the spring of 2004 (May 11 to 26) (73, 77). Fifty monitors were deployed at one set of locations during both sampling periods and the locations for 50 samplers were changed in the second round of monitoring. In Windsor, Maxxam all-season passive air samplers (Maxxam Analytics, Calgary, Edmonton) was used to measure ambient concentrations of NO_2 (75). Measurements were made in 2004 at 54 locations over four, two-week periods in February, May, August, and October as for each of the four seasons in 2004.

Manual forward-selection regression procedures were used, and included an array of land use, traffic, physical geography, and population variables. The R^2 of the land use regression models were 70%, 76%, and 77% for Toronto, Hamilton, and Windsor, respectively.

Data from the integrated sampling campaigns in Toronto and in Windsor showed that the measurements of NO_2 at the same locations were highly correlated between different seasons: Toronto (Pearson correlation coefficient (r) = 0.8) and Windsor (r = 0.7-0.9) (75, 77). As a result, in Toronto, we took the average of the two seasons and in Windsor we used the average of the four seasons. Exposure surfaces of concentrations of NO_2 were estimated based on the land use information and the measurements of NO_2.

We linked the exposure surfaces to the centroids of the areas represented by the six-character residential addresses of participants at the time of enrolment. In urban areas, these six-character addresses represent the face of a city block or a large apartment complex.

Historical Estimates of Exposure for Toronto

The above estimates of exposure were derived at the end of the period of follow-up (2004). A key assumption inherent in using these exposure data in an analysis of the cohort is that these estimates applied throughout the period of follow-up. Rather, we attempted to model exposures in the past by estimating changes in the spatial structure of concentrations of NO_2. We have developed methods to extrapolate spatial exposures derived from a land-use regression model into the past (2) that make use of concentrations measured at fixed-site monitors. These methods require fairly dense fixed-site monitoring sites and, unfortunately, for Hamilton and Windsor there were few stations, so we could only apply our methods to Toronto.

We used two related back-extrapolation methods that made use of data from Environment Canada's National Air Pollution Surveillance Network for the years 1982 and 1992 (229). In the first method, we multiplied the surface of concentrations of NO_2 produced from the land use regression model from the sampling campaign in the fall of 2002 in each grid cell (5m × 5 m) by the ratio of two surfaces derived using first-order, inverse-distance weighted interpolation (218) of the two-week mean concentrations of NO_2 measured in the fall of 1982 and 2002 at the same six fixed-site monitors. This process was repeated with data for the two-week period during the spring of 1982 and 2004. We then combined the two new exposure surfaces by taking an average at each grid cell, which resulted in an estimated surface of concentrations of NO_2 for 1982. For 1992, we used nine fixed-site monitors.

For the second extrapolation method, we took into consideration recent land use and vehicular traffic. In this method, we multiplied the surface concentrations of NO_2 produced by the land use regression model for 2002 by a ratio of estimates from the following models: 1) we developed a model for the numerator by regressing the observed two-week mean concentrations of NO_2 at six fixed-site monitors in the fall of 1982 against a reduced set of

spatial variables that were used to develop the 2002 model. We selected the variables using a supervised forward stepwise model selection procedure. 2) We developed a model for the denominator using as the dependent variable predicted concentrations at the six fixed-site monitors from the 2002 model, and then we regressed these predicted values against the same covariables that we selected for the numerator. This process was repeated with the data for the two-week period in spring 1982 and 2004. Similar to the first method, we then combined the two new exposure surfaces by taking an average at each grid cell, which resulted in a surface of NO_2 for 1982. For 1992, we used nine fixed-site monitors.

Mortality Data

Vital status was ascertained for all subjects over the follow-up period of 1982-2004 through a probabilistic record linkage to the Canadian Mortality Database that was conducted by Statistics Canada. This database provides data on all deaths of Canadians that occurred in Canada as well as most of those that occurred in approximately 20 U.S. states.(233) The cohort was linked according to first, middle, and family names, sex, date of birth, place of residence, and in some cases social insurance number.(234) Previous work suggests that under-coverage of the deaths is minimal and the accuracy of identifying deaths is around 98%.(233)

We assessed mortality from coronary heart disease (International Classification of Diseases, revision-9 (ICD-9): 410-414; ICD-10: I20-I25), cerebrovascular disease (ICD-9: 430-438; ICD-10: I60-I69), and all cardiovascular diseases combined (ICD-9: 400-440; ICD-10: I00-I99). The accuracy of coding cardiovascular diseases as an underlying cause of death on the Canadian death certificates is, however, not known. Previous studies showed that average false positive rates and false negative rates in coding acute myocardial infarction on death certificates in three Canadian provinces using 1984 data were about 5% and 1%, respectively (235, 236). In studies conducted in the U.S., death certificates showed overestimates of 7%-10% for cardiovascular diseases and 7%-20% for coronary heart disease (157, 237-239). As a result, some misclassification in outcomes is likely but we expect that it should be independent of exposure to air pollution and therefore no differential bias should be introduced into our risk estimates.

Primary Covariables

We had from the tax file the individual-level variables of age, sex, marital status, family composition, and annual personal and household income (in Canadian dollars) as recorded at the year of the start of follow-up. Family composition is a concept that combines marital status and family size and we created a categorical variable with three levels: husband-wife or common-law family; lone-parent family; and lone person (228).

Neighbourhood deprivation may confound the association between traffic-related air pollution and cardiovascular mortality if it is associated with cardiovascular mortality (240-248) and with exposure to traffic-related air pollution. Conceptually, a neighbourhood refers to a person's immediate residential environment, which is hypothesized to have both material and social characteristics potentially related to health (242, 249). Because it is a difficult task to delineate the boundaries of neighbourhoods that are 'ecologically meaningful', administratively defined areas have often been used as rough proxies for neighbourhoods. For example, in two recent studies (77, 80) of traffic-related air pollution and cardiovascular disease that were conducted in Toronto and Vancouver, Canada, neighbourhoods were defined by the census enumeration area (228). The enumeration area is a small, relatively homogeneous geographic unit composed of one or more blocks (400 to 700 persons), for which all census data are disseminated (228). In another study that was conducted in Montreal, Canada, to examine neighbourhood effects on health status, census tracts appeared to be good proxies for "natural" neighbourhood boundaries (250). Given the uncertainty about the choice of geographic scale in defining neighbourhoods, we derived ecological variables at two geographic scales: the enumeration area and the census tract, and these were included as potential confounding covariables in separate analyses.

As a result, we derived four ecological variables using data from the 1981 Canadian Census: percentage of immigrants; percentage of adults with less than a high school education; unemployment rate; and average household income. As in most studies of the effects of neighbourhood socioeconomic characteristics and cardiovascular outcomes (240-248), we assumed that the measurement of neighbourhood deprivation made in a specific time period

may be a reasonable proxy for the accumulation of neighbourhood effects over the follow-up period. As a result, we assigned values of the ecological variables to study subjects according to their residences at the year of entry (1982-1986).

Statistical Analysis

We used the Cox proportional hazards model (251) to estimate the association between traffic-related air pollution and cardiovascular mortality. We censored subjects for whom a death record was not found at the end of the follow-up period (i.e., December 31, 2004). The record linkage process we used is highly accurate (>98%) so this is a reasonable assumption (252). The time variable used in the models was survival time from date of entry into the cohort until the earliest of date of death or to the end of follow-up. We used the stratified version of the Cox model, with strata defined as one-year age categories at date of entry. We conducted separate analyses for each city. In addition, we tested the statistical heterogeneity of the estimates across cities, and we pooled the effect estimates across the cities using a fixed effect model weighted by the inverse of the city-specific variance (253).

We calculated adjusted rate ratios (RR) and associated 95% confidence intervals (CI) for an increase of 5 parts per billion (ppb) in concentrations of NO_2. The exposure variable was concentrations of NO_2 derived from the land use regression models, developed at the end of the follow-up period that were assigned to the subjects' six-character postal code addresses at time of enrolment. For Toronto, we also conducted analyses using the back-extrapolated estimates of concentrations of NO_2 for the years 1982, 1992, 1982-1992, and 1982-2004. In the latter analysis, to account for differences in the absolute value of the distributions, we computed the RRs for an increase equal to the interquartile range (IQR) of each predicted distribution as well as for an increase of 5 ppb.

We included as covariables sex, marital status, individual-level annual household income (quintiles of the distribution that are city-specific), percentage of immigrants in the neighbourhood (quintiles), percentage of adults with less than high school education in the neighbourhood (continuous), unemployment rate in the neighbourhood (continuous), and average household income of the neighbourhood (quintiles). We verified the assumption of

linearity for continuous exposure and confounding variables by fitting a quadratic term of these variables. Household income and percentage of immigrants were coded as a series of indicator variables to allow for nonlinear relationships with mortality. For each model, we tested deviations from the proportional hazards assumption by assessing whether the cross-product of each variable with the natural logarithm of the time variable was statistically important (254).

We also estimated whether age (<45 years, 45-54 years, 55-65 years, ≥65 years), sex, marital status, and individual-level household income (quintiles) modified the associations between cardiovascular mortality and traffic-related exposures.

Indirect Adjustment for Unmeasured Confounding Variables

A limitation of the above analyses was that we were unable to account for tobacco smoking, a major risk factor for cardiovascular disease. We thus made use of indirect methods to estimate the bias that may have occurred because smoking was not included directly in the models. These "indirect" methods were developed by Axelson (4) and others (5). Specifically, we made use of a Monte Carlo sensitivity analysis proposed by Steenland and Greenland (2004) (255), and we extended the method to handle continuously measured exposure variables.

The method allows for the estimation of a "bias factor" which under the null hypothesis of no association between air pollution and cardiovascular mortality implies that the rate ratio would be estimated solely due to the confounding effects of smoking (RR_{bias}) (4). In computing this, we classified the distribution of exposure to NO_2 into quintiles and a RR_{bias} was estimated for each quintile of NO_2, and the rate ratio for the lowest quintile was set to unity.

$$RR_{bias} = I_{E+}/I_{E-} = \frac{I_o \times \left(1 + \sum_{i=1}^{k} P_{c,i} \times (RR_{c,i} - 1)\right)}{I_o \times \left(1 + \sum_{i=1}^{k} P_{g,i} \times (RR_{c,i} - 1)\right)} \qquad (1)$$

where I_{E-} refers to the incidence rate of cardiovascular mortality among unexposed subjects (i.e., the lowest quintile of NO_2), I_{E+} denotes the incidence rate of cardiovascular mortality among the subjects in the cohort who are classified as being exposed to air pollution (i.e., higher quintiles of NO_2), I_o is the incidence rate among those who do not smoke, $P_{c,i}$ and $P_{g,i}$ represent the prevalence of smokers, at level i (e.g., current smoker, i=1; ex-smoker, i=2), in the exposed cohorts and the unexposed cohort, respectively. $RR_{c,i}$ is the relative risk for cardiovascular mortality for smoking at level i.

To estimate area-specific prevalences of smoking, we made use of the Canadian Community Health Survey conducted in 2001 which is a national probability sample of all households in Canada (256). We obtained the prevalence of current and former smokers in each six-character postal code area in the three cities, thereby providing estimates of $P_{g,i}$. In addition, we made use of rate ratios for the cardiovascular diseases of interest for current and former smokers using estimates form the American Cancer Society Cancer Prevention Study II, for the follow-up period 1982 until 1988 (257). These estimates were rate ratios that were adjusted for age, sex, marital status, and other risk factors in the original study. As these rate ratios for smoking and cardiovascular mortality varied by age and sex, we re-weighted the age- and sex-specific rate ratios by the underlying age and sex structure of participants of the Canadian Community Health Survey.

The classic formula of indirect adjustment for unmeasured smoking (Equation 1) allowed us to estimate a RR_{bias} for each quintile of exposure to NO_2. Because concentrations of NO_2 are on a continuous scale, we required a bias factor per each increase of 5ppb of NO_2 ($RR_{bias\text{-}5ppb}$). We thus derived a simple linear regression model with the dependent variable equal to the estimated RR_{bias} for each quintile of concentrations of NO_2 through using Equation 1. The independent variable was the concentration of NO_2 that was sampled randomly from a uniform distribution of NO_2 for each of the quintile groups. The slope ($RR_{bias\text{-}5ppb}$) obtained from fitting the linear regression model represented the estimate of the amount of confounding by smoking for each 5 ppb increase of NO_2.

Rate ratios for an increase of 5 ppb of exposure to NO_2 that were indirectly adjusted for smoking ($RR_{Indirect\ adj\text{-}5ppb}$) are computed as:

$$RR_{Indirect\text{-}adj\text{-}5ppb} = RR_{Cox\ model\text{-}5ppb} / RR_{bias\text{-}5ppb} \quad (2)$$

where $RR_{Cox\text{-}model\text{-}5ppb}$ is the rate ratio for an increase of NO_2 of 5 ppb adjusted for all variables included in the main analysis.

To compute the statistical uncertainty of $RR_{Indirect\text{-}adj\text{-}5ppb}$, we used Monte Carlo sampling (100,000 replications) to repeatedly sample from the priors of the prevalence of current and former smokers in each exposure group as well as the rate ratio for the effect of smoking on cardiovascular mortality. Our prior distribution for the prevalence of smoking was computed from a bivariate normal distribution with means equal to the logit of the proportions of current and former smokers in each exposure group. For the rate ratios relating current and former smoking to cardiovascular mortality, we specified a normal distribution with a mean value equal to the natural logarithm of the rate ratio and standard deviations equal to the standard errors from the American Cancer Society Cancer Prevention Study II (257). For each replicate, we also re-sampled the observed rate ratio for the association with concentrations of NO_2 ($RR_{Cox\ model\text{-}5ppb}$) from its estimated normal distribution with the mean and variance estimated using data from the tax cohort. We repeated the Monte Carlo sensitivity analysis for each of the three causes of death and for the study population in each of the three cities.

Similarly, obesity may also be a possible confounding variable. We also obtained the distribution of the body mass index (weight(kg)/height(m)^2) from the 2001 Canadian Community Health Survey. We assessed whether the distributions of the body mass index differed between the exposure groups. The analyses were repeated for each of the three cities.

RESULTS

Baseline characteristics. A total of 77,890 subjects in Hamilton, 58,700 in Toronto, and 68,850 in Windsor were enrolled in the study (Table 1). Because Statistics Canada put a constraint on the maximum number of subjects to be sampled from each city, the number of

subjects that were sampled in each city was not proportional to the actual size of the population of the city. The mean age of the cohort in the three cities was 52 years at date of entry, 50% were men, and more than 64% were married. The mean annual household incomes of the participants were $35,600, $43,500, and $34,200 in Hamilton, Toronto, and Windsor, respectively. The 1981 Census showed that the average percentage of immigrants of the neighbourhoods of the participants was about 40% in Toronto, 28% in Hamilton, and 24% in Windsor. In addition, the percentage of adults with less than a high school education in the neighbourhoods of the participants was 51% in Hamilton, 49% in Windsor and 45% in Toronto.

Traffic-related air pollution. Figure 2 shows the secular trends of annual mean concentrations of NO_2 at fixed-site monitors in the three cities between 1982 and 2004. The annual mean concentrations of NO_2 and the rank ordering of fixed-site monitors suggested that the spatial distribution did not change substantially over the follow-up period. The distributions in the estimated concentrations of NO_2 at the subjects' home addresses varied substantially between the three cities (Table 2). On average, subjects living in Toronto were exposed to higher concentrations of NO_2 (mean: 21.7 ppb) than those in the other two cities (mean: Hamilton=15.5 ppb and Windsor=12.1 ppb). Table 3 shows the various estimates of concentrations of NO_2 in Toronto, including the original land use regression models from 2002 and 2004 as well as the back-extrapolated ones. The estimates of concentrations of NO_2 and their variability across the follow-up period were similar: the mean concentration of NO_2 in the first 10-year follow up (1982-92) was approximately 24 ppb (inter-quartile range (IQR): 4.7 ppb) and the mean concentration of NO_2 in the entire study period was about 23 ppb (IQR: 4.4 ppb). Table 4 shows that in Toronto the land use regression model (2002 and 2004) and the back-extrapolated estimates of concentrations of NO_2 were highly correlated, as expected from the secular trends in concentrations shown in Figure 2.

Associations between exposure to NO_2 and cardiovascular mortality. During the follow-up period of 1982 to 2004, we found a total of 18,360, 12,410, and 17,360 non-accidental deaths among the study participants in Hamilton, Toronto, and Windsor, respectively. Among these, 7,370, 4,620, and 7,390 were coded on the underlying cause of death as dying from cardiovascular diseases (Table 5). The percentage of deaths from

cardiovascular diseases is according to expected figures (varied between 32% and 37% in Ontario, 2001-2007) (258). More than 50% of cardiovascular deaths were due to ischemic heart disease and substantially more of these deaths occurred among men than women.

Table 6 shows the associations between estimated concentrations of NO_2 and cardiovascular mortality. For each cardiovascular outcome, the first row shows models adjusted only for age and sex, the second row shows the effects of adding all available personal-level confounding variables, and the third row shows the rate ratios that were further adjusted for all of the selected ecological covariables. Rates ratios for all cardiovascular mortality per increase of 5ppb of NO_2 adjusted for all potential personal confounding variables were: Hamilton, 1.16 (95% CI: 1.10-1.22); Toronto, 1.01 (95% CI: 0.98-1.05); and Windsor, 1.13 (95% CI: 1.06-1.22). The addition of ecological variables reduced slightly the estimates of cardiovascular effect in Hamilton and Windsor, whereas the adjustment of these ecological variables increased the estimates for Toronto. The pooled fully-adjusted RR_{5ppb} was 1.08 (95% CI: 1.05-1.11) (test of heterogeneity, $p \simeq 0.50$).

For ischemic heart disease, the estimated effects were almost identical to those observed for all cardiovascular mortality. The pooled estimate for mortality from ischemic heart disease for an increase of 5ppb of NO_2, adjusted for all personal confounding variables, was 1.10 (95% CI: 1.01-1.21). For cerebrovascular disease, there was no association between estimated NO_2 and mortality (pooled RR_{5ppb}=0.99; 95% CI: 0.90-1.10).

Associations between cardiovascular mortality and back-extrapolated estimates of NO_2 in Toronto. As expected from the spatial stability of NO_2 in Toronto, the analyses using historical extrapolation produced in Toronto similar associations with those using the estimates of NO_2 in 2002-2004 (**Table 7**). Using the long-term average concentrations of NO_2 in 1982 and 2004, we estimated that the RR_{5ppb} was 1.07 (95% CI: 1.01-1.14) for mortality from ischemic heart disease. When the rate ratios were computed for an increased of the interquartile range (IQR) of NO_2 to account for differences in the absolute value of the distributions, the estimates were similar (**Appendix A**).

Sensitivity analyses. **Table 8** shows the city-specific estimates of prevalence of current and former smokers according to quintiles of the concentrations of NO_2. In Hamilton, the prevalence of current smoking was associated with higher concentrations of NO_2. On the other hand, in Toronto and Windsor the prevalence of current smoking was relatively constant across the different quintiles of NO_2. Indirect adjustment for smoking attenuated the estimates of risk (**Table 9**) only in Hamilton: the RR_{5ppb} for all cardiovascular mortality was reduced from 1.12 to 1.02 (95% CI: 0.91-1.11), and a similar reduction was observed for ischemic heart disease. Indirectly adjusting for smoking and including all of the other measured confounding covariates, the pooled RR_{5ppb} was 1.04 (95% CI: 1.00-1.09) and 1.05 (95% CI: 1.00-1.11) for all cardiovascular mortality and mortality from ischemic heart disease, respectively. We did not adjust for body mass index because the distributions of measured body mass index across different levels of concentrations of NO_2 were similar (**Appendix B**).

Analyses of interactions. We also examined potential effect modification by age, sex, marital status, and individual-level household income on the association of traffic-related air pollution and cardiovascular outcomes, and we found that the interaction terms with exposure were not statistically important ($p>0.05$; data not shown).

DISCUSSION

In this population-based cohort study of cardiovascular mortality with a 23-year follow up of 205,430 adults in three cities in Ontario, we found that NO_2, a marker for traffic-related air pollution, was associated with increased mortality from all cardiovascular causes. The plausible estimates of effect ranged between the fully adjusted without ecological covariables and that indirectly adjusted for smoking with all covariables: 1.04 (95% CI: 1.00-1.09) to 1.10 (95% CI: 1.00-1.21). For ischemic heart disease, the estimated effects were similar to those observed for all cardiovascular mortality. We did not find any effects for cerebrovascular disease.

Our findings are similar to the results of a recent cohort study that was carried out in Vancouver, Canada (259). In this study, Gan and colleagues (2011) examined rates of

hospitalization and mortality from ischemic heart disease in a cohort comprised of all residents aged 45-85 years in the city (n=452,735) over a four-year period (80). They reported that an increase of 8.4μg/m^3 (approximately 4.5 ppb) of NO_2 was associated with a 4% (95% CI: 1.01-1.08) increased risk of mortality from ischemic heart disease. A smaller cohort study was conducted in Toronto, Canada: it comprised 2,360 patients who attended a respiratory disease clinic in 1992 (nearly half of the subjects had been diagnosed with ischemic heart disease and 30% had chronic obstructive pulmonary disease (COPD) at the date of enrolment) (77). A 4 ppb increase in NO_2 increased the mortality rates of cardiovascular disease by 40% (95% CI: 1.05-1.86). Clearly, the constitution of the last cohort does not represent the general population and this may explain the differences between this and other studies. Many cohort and case-control studies that were conducted in other countries also showed that traffic-related exposure metrics such as concentrations of NO_2 and nitrogen oxides (NOx), traffic density, and proximity to roadways were associated with increased cardiovascular mortality (25, 57, 62, 82-85, 88).

Our study has three important strengths. First, it was population-based and it had large statistical power to detect small effects. Second, we assessed exposure to traffic-related air pollution using ambient concentrations of NO_2 derived from land use regression models. It has been shown that land use regression provides more accurate estimates of small area variations in traffic-related air pollution as compared to other methods, such as proximity to roadways and spatial interpolation (68, 73). Therefore, using this approach likely reduced misclassification of exposure. Third, some important confounding variables, especially household income and marital status, were measured accurately because they were obtained directly from the tax file. This reduced possible residual confounding due to misclassification of these risk factors for cardiovascular disease.

Adjustment for ecological risk factors modified slightly the estimates of excess risk. Interestingly, the estimates for Toronto were increased after control for the ecological variables, in particular adjusting for the percentage of immigrants in the neighbourhood. As shown in Table 1, Toronto has a very large foreign-born population (~40% in 1981). Immigrants to Canada are typically healthier than the Canadian-born population, partly due to the fact that potential immigrants are screened on medical and other health-related criteria

before they are admitted to the country (260-262). As well, immigrants tend to take up residence in larger metropolitan areas and closer to public transit or roadways where pollution levels are higher. As a result, immigrants in our study population may exert a negative confounding effect, that is, they may bias the estimates towards the null. Unfortunately, we were not able to ascertain immigrant status and therefore could not assess effects separately.

In addition to the possible "healthy immigrant effect", this study has some other limitations. First, we relied on death certificates to ascertain cardiovascular outcomes. As shown in previous studies, reporting on death certificates of cardiovascular diseases and coronary heart disease in particular (157, 237-239) appears to be over-estimated, although it is unclear whether this applies to the present study. As a result, some potential misclassification in the ascertainment of outcome may have occurred in our study. If this was not related to level of exposure, we would expect that we have underestimated the true effects.

Second, the exposure assessment used in this study relied on land use regression models that were developed towards the end of our study, so it is important to understand whether these models can characterize exposure adequately during the relevant etiological periods. Our sensitivity analysis using two back-extrapolation methods showed that the predicted concentrations in Toronto across different periods were correlated strongly with each other. This finding was supported by the fact that the annual mean concentrations of NO_2 and their rank ordering at fixed-site monitors in Toronto and in Hamilton remained largely constant during the period of 1982 to 2004. In a previous study that was conducted in Toronto, Jerrett et al. (2009) also showed that the spatial pattern of concentrations of NO_2 was relatively stable over time (77). The relative stability over time in the spatial patterns of ambient concentrations of NO_2 within cities has also been reported in studies in Montreal, Canada (71) and in the Netherlands (219). As a result, we expect that the spatial contrast in NO_2 estimated using the land use regression models over a short period would provide reasonable estimates of longer-term spatial exposures.

We did not have information on subjects' residences before they entered the study. We assumed that the subjects' previous exposures were similar to that during the follow-up

period. Indeed, although available epidemiologic data suggested larger risks of cardiovascular disease posed by long-term exposure (in years) than observed only a few days (56), it remains unclear whether exposures in the past or more recent exposures to air pollution are etiologically important. As well, it was not possible to assess personal exposures to ambient air pollution. Other factors such as individual mobility and daily activity patterns may have an important influence on the total personal exposure to traffic-related air pollution. We did not assess mobility in this study and that is a limitation. Two population surveys reported that Canadian adults who resided in major cities during 1992 and 1997 spent on average over 65% of their time each year at home (both indoors and outdoors) (263) Indoor air pollution has also been shown to correlate with outdoor air pollution (264, 265) As a result, it is reasonable to use spatially derived exposures at the residences at the beginning of follow-up as a surrogate for personal exposures. Nonetheless, given the inherent imprecision of the spatially derived exposures, our assessment of exposure by the cohort was likely subject to nondifferential misclassification. This would likely lead to further attenuation of our estimates of effects.

Third, as with many cohort studies based on administrative databases, we did not have access to information on some important individual risk factors of cardiovascular diseases such as smoking (266, 267), body weight(19-22), high blood pressure (16-18), diabetes (23, 24), and family history (268). If these risk factors varied with the levels of exposure, failure to control for them could lead to residual confounding. Smoking is one of the strongest risk factors for cardiovascular disease, because of the strength of its effect and its high prevalence, and thus has the potential to confound. To control for smoking, we made use of a classic bias formula that was proposed by Axelson and others (4, 5), and we drew on the 2001 Canadian Community Health Survey to estimate prevalence of smokers in the three cities in Ontario. As we noted above, the Canadian Community Health Survey was used to obtain information from a large probability sample of all households in Canada; thus it provided reliable estimates of the prevalence of smokers in our study area. To improve accuracy, we further restricted data from the survey to those 35-85 years of age. To account for uncertainty in the estimation of confounding due to smoking, we then applied a method that we developed based on the Monte Carlo sensitivity analysis that was proposed initially by Steenland and Greenland (2004). As expected, indirect adjustment for smoking had

almost no impact on the rate ratios of cardiovascular mortality in Toronto and in Windsor, because the prevalences of current and former smoking were weakly correlated with ambient concentrations of NO_2 in these two cities. On the other hand, substantial reduction in rate ratios was observed for Hamilton due to strong associations between smoking and the concentrations of NO_2 in the city. Combining the indirectly adjusted rate ratios from the three cites, we still found positive associations with NO_2.

The accuracy of our method of indirect adjustment will depend on the distributions of the prevalence of smoking and on the relative risks for smoking. In addition, our assumption that the bias factor is approximately constant across the strata of covariates such as age and sex will also affect accuracy. At the cost of added complexity, such assumptions could be relaxed by sampling age- and sex-specific prevalence of smoking and relative risks for smoking, and then using those to indirectly adjust for age- and sex-specific rate ratios for the association of NO_2 and cardiovascular mortality.(225) The original method of indirect adjustment by Steenland and Greenland (2004) has been used in a number of studies to assess the robustness of their findings regarding potential biases by unmeasured confounders (269-273).

With respect to unmeasured risk factors, such as high blood pressure, family history, and diabetes, we were unable to examine their potential confounding effects. Although we could not rule out potential residual confounding by these risk factors, their confounding effects would likely be small as compared to smoking. Adjustment for other risk factors such as age, sex, and smoking might reduce the potential confounding effects of these unmeasured risk factors, in particular high blood pressure and diabetes, because they are often related (274-276). In addition, there is evidence suggesting that exposure to traffic-related air pollution may increase the risk of higher blood pressure (277, 278) and diabetes (279, 280). As well, these unmeasured risk factors were unlikely to have differed in prevalence as much as smoking across the different levels of air pollution. Indeed, in an earlier study that was conducted in Toronto, Jerrett et al. (2009) showed that adjustment for smoking, diabetes, pre-existing COPD, and chronic ischemic heart disease resulted in a decrease of only 5% in the estimate of effects of NO_2 on cardiovascular mortality. More recently, there is some evidence that noise may be a risk factor of cardiovascular disease (281-284). In this study, we

were unable to measure noise, and thus a possibility of residual confounding by traffic noise cannot be ruled out, although it remains to be shown that there are substantial exposures to ambient noise because people spend a considerable amount of time indoors where street noise is highly attenuated.

The mechanisms by which air pollution influences the risk of cardiovascular diseases are still under investigation. It is hypothesized that pulmonary and or systemic oxidative stress and inflammation, the chronic promotion of atherosclerosis, and the alterations of cardiac autonomic function may be induced by exposure to air pollution, in particular $PM_{2.5}$ (55, 56). In addition, there is evidence that exposure to $PM_{2.5}$ may be capable of impairing vascular function through damaging endothelium-dependent vasodilation, particularly among higher-risk individuals (e.g., diabetes) (55). Although our study showed a positive association between NO_2 and cardiovascular outcomes, our findings should not be interpreted as meaning that NO_2 is a causal agent, rather NO_2 may be acting as a surrogate for the mixture of traffic-related air pollutants. Indeed, Brook et al. (2007) showed that NO_2 has strong co-locational associations with several air pollutants generated from local mobile sources (193).

CONCLUSIONS

Our findings support the hypothesis that exposure to traffic-related air pollution increases the mortality of cardiovascular disease and in particular, from ischemic heart disease.

Table 1. Description of selected characteristics of study participants, by city, The Ontario Tax File Cohort Study, 1982-2004

Characteristics of the cohort	Hamilton (N=77,890) [a]	Toronto (N=58,700)	Windsor (N=68,850)
Individual risk factors at time of entry (standard deviation in parentheses)			
Age at time of entry into the cohort, 1982-1986 (years)	52.0 (12.0)	52.4 (11.9)	53.3 (12.6)
Men (%)	50.6	50.4	50.4
Marital status (%)			
Married	68.5	66.7	64.7
Single	6.5	10.3	7.0
Separated, widowed, or divorced	17.7	17.8	20.4
Missing	7.3	5.2	7.9
Family composition (%)			
Husband-wife family or common-law family	77.7	73.8	74.6
Lone-parent family	5.0	5.7	5.8
Lone person	17.3	20.5	19.6
Individual mean annual income of the cohort, aged 35 years and above (in $1,000 CAN)	20.9 (34.4)	25.2 (54.7)	20.7 (27.7)
Annual mean household income of the cohort, aged 35 years and above (in $1,000 CAN) [b]	35.6 (44.7)	43.5 (91.0)	34.2 (41.3)
Neighbourhood-level risk factors from the 1981 Canadian census, at the census enumeration level [c]			
Percentage of immigrants in neighbourhood [d]	27.6	39.8	24.0
Percentage of the population 15 years and over, with less than high school education in neighbourhood	51.0	44.5	48.6
Percentage of the population 15 years and over, without employment	6.2	4.0	11.8
Average household income in neighbourhood, with all ages (in $1,000 CAN; standard deviation in parenthesis) [b]	12.9 (2.8)	14.5 (5.3)	13.2 (3.4)

[a.] Rounded to the nearest 10 individuals to protect confidentiality.

[b.] In contrast to the census, the tax cohort was restricted to those aged 35 years and above. With an older age range expecting to have higher incomes, the annual household income of the cohort may be higher than the average household income derived at the neighbourhood level from the census. In addition, a small percentage of people with low income may not have filed taxes and thus not be captured in the federal income tax file, but these households should have been captured in the census.

[c.] Two definitions of a neighbourhood were used in this study: the census enumeration area and the census tract. The dissemination area is a small, relatively homogenous geographic unit comprised of one or more blocks, for which all census data are disseminated. The census tract is the next larger geographic unit. We presented here the variables that were derived at the census enumeration area level.

[d.] Immigrant population (also known as foreign-born population) is defined in the Canadian Census as persons who are, or who have been, landed immigrants in Canada.[39]

Table 2. Distributions of estimated concentrations of NO_2 (ppb) at the addresses of subjects' homes at time of entry, derived from land use regression models, by city, The Ontario Tax File Cohort Study, 1982-2004

City	Estimates of concentrations of NO_2 derived from land use regression models (in ppb)							
	Year of sampling	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Interquartile Range
Hamilton	2002 [a]	15.49	8.31	14.05	15.22	16.93	26.18	2.88
Toronto	Mean of 2002 and 2004 [b]	21.68	10.61	19.15	21.05	23.24	42.03	4.09
Windsor	2004 [c]	12.10	8.01	10.88	12.08	13.35	27.26	2.47

[a] Concentrations of NO_2 derived from a land use regression model using measurement from a 2002 fall monitoring campaign with 107 monitors (74).

[b] Concentrations of NO_2 derived from averaged estimates from two land use regression models using measurements from 2002 fall and 2004 spring monitoring campaigns with 100 monitors (73, 77).

[c] Concentrations of NO_2 derived from a land use regression model using measurements from 2004 spring, summer, fall, and winter monitoring campaigns with 54 monitors (75).

Table 3. Distributions of estimated annual mean concentrations of NO_2 (ppb) at the addresses of subjects' homes at time of entry in Toronto across three time periods, according to the two different back-extrapolation methods, The Ontario Tax File Cohort Study, 1982-2004

Exposure metrics	Year	Mean	Minimum	25th percentile	Median	75th percentile	Maximum	Interquartile Range
1. Original LUR model [a]	Mean of 2002-2004	21.68	10.61	19.15	21.05	23.24	42.03	4.09
2. IDW-based extrapolation [b]	1982	25.25	12.37	22.65	25.16	27.47	49.78	4.82
	1992	25.24	12.07	22.53	25.17	27.53	47.54	5.00
	Mean of 1982-1992	25.25	12.22	22.67	25.13	27.44	48.12	4.77
3. LUR-based extrapolation [b]	1982	24.44	11.96	21.13	23.70	26.64	55.37	5.51
	1992	23.67	12.52	21.01	23.19	25.69	41.11	4.68
	Mean of 1982-1992	24.06	12.92	21.33	23.55	26.00	47.24	4.67
	Mean of 1982-2004	23.26	12.49	20.62	22.71	25.02	45.50	4.40

[a.] Concentrations of NO_2 derived from the average of two land use regression models using measurements from 2002 fall and 2004 spring monitoring campaigns with 100 monitors (73, 77).

[b.] IDW, inverse distance weighted interpolation; LUR, land use regression model

Table 4 Correlations between estimated annual mean concentrations of NO_2 (ppb) across three time periods in Toronto, according to the surface maps of NO_2 produced using the land use regression model for 2002-2004 and the two extrapolation methods (based on 5,000 random locations in Toronto), The Ontario Tax File Cohort Study, 1982-2004

	Pearson correlation coefficients between periods [a]		
	1982	1992	2002-2004
$LUR_{2002\text{-}2004} \times$ ratio of IDW_{fixed} [b]			
1982	1	0.93	0.89
1992		1	0.93
$LUR_{2002\text{-}2004} \times$ ratio of LUR_{fixed} to $LUR_{predicted}$ [b]			
1982	1	0.78	0.84
1992		1	0.87

[a] The linear relationship between concentrations of NO_2 in two separate years was confirmed from visual inspection of the scatter plots

[b] $LUR_{2002\text{-}2004}$ is the average of LUR 2002 (R^2=0.69) and LUR 2004 (R^2=0.71); IDW, inverse distance weighted interpolation; LUR, land use regression model

Table 5. Distribution of the number [a] of cardiovascular deaths, by cause, city, and gender, The Ontario Tax File Cohort Study, 1982-2004

Cause of death	Hamilton (N=77,890)		Toronto (N=58,700)		Windsor (N=68,850)	
	Men (741,610 person-years)	**Women** (783,975 person-years)	**Men** (574,961 person-years)	**Women** (596,595 person-years)	**Men** (642,630 person-years)	**Women** (685,468 person-years)
All non-accidental causes	11490	6880	7380	5030	10630	6730
All cardiovascular disease [b]	4700	2670	2840	1780	4520	2870
Ischemic heart disease [c]	2630	1250	1490	730	2800	1560
Cerebrovascular disease [d]	780	710	560	550	800	730

[a.] Rounded to the nearest 10 cases to protect confidentiality.

[b.] All cardiovascular disease: ICD-9: 400-440; ICD-10: I00-I99.

[c.] Ischemic heart disease: ICD-9: 410-414; ICD-10: I20-I25.

[d.] Cerebrovascular disease: ICD-9: 430-438; ICD-10: I60-I69.

Table 6. Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and estimates of concentrations of nitrogen dioxide from land use regression models (2002 and 2004), The Ontario Tax File Cohort Study, 1982-2004. The rate ratios are expressed for an increase of 5ppb of NO_2.[a]

	City			
	Hamilton	Toronto	Windsor	Pooled estimate [e]
Cause of Death	RR (95% CI)	RR (95% CI)	RR (95% CI)	RR (95% CI)
All Cardiovascular Disease				
Age and sex adjusted [b]	1.28 (1.22 - 1.35)	1.05 (1.02 - 1.08)	1.25 (1.17 - 1.34)	-
+ All personal covariables [c]	1.16 (1.10 - 1.22)	1.01 (0.98 - 1.05)	1.13 (1.06 - 1.22)	-
+ All ecological covariables [d]	1.12 (1.06 - 1.19)	1.05 (1.00 - 1.09)	1.10 (1.02 - 1.19)	1.08 (1.05 - 1.11)
Ischemic Heart Disease				
Age and sex adjusted	1.32 (1.22 - 1.42)	1.07 (1.02 - 1.12)	1.25 (1.14 - 1.36)	-
+ All personal covariables	1.18 (1.09 - 1.27)	1.03 (0.98 - 1.08)	1.12 (1.02 - 1.23)	-
+ All ecological covariables	1.12 (1.02 - 1.21)	1.06 (1.00 - 1.13)	1.11 (1.00 - 1.23)	1.09 (1.04 - 1.14)
Cerebrovascular Disease				
Age and sex adjusted	1.17 (1.03 - 1.31)	0.96 (0.90 - 1.03)	1.12 (0.96 - 1.30)	-
+ All personal covariables	1.06 (0.93 - 1.19)	0.92 (0.86 - 1.00)	1.04 (0.89 - 1.18)	-
+ All ecological covariables	1.06 (0.92 - 1.22)	0.91 (0.83 - 1.00)	0.96 (0.82 - 1.18)	0.95 (0.89 - 1.02)

[a.] Rate ratios expressed for an increase in the inter-quartile range of NO_2 are shown in **Appendix C**.

[b.] The baseline hazard function in the Cox regression models stratified by 1-year age categories.

[c.] Same model as in a), but adjusted for marital status (four categories), and annual household income (quintiles).

[d.] Same model as in a), but adjusted for marital status (four categories), annual household income (quintiles), % of immigrants in neighbourhood (quintiles), % of population with less than high school education in neighbourhood (continuous), unemployment rate in neighbourhood (continuous), and average household income in neighbourhood (quintiles).

[e.] A test of heterogeneity of city-specific rate ratios that were adjusted for only age and sex yielded p values < 0.05. As a result, we did not pool the age- and sex-adjusted ratio ratios across cities. Similarly, we did not pool the rate ratios that adjusted for all individual-level covariables across cities, because the test of heterogeneity of rate ratios yielded p values < 0.05. On the other hand, the test of heterogeneity of rate ratios that adjusted for all individual-level covariables and ecological covariables yielded p values > 0.05. Therefore, the pooled estimate of fully-adjusted RRs was obtained using a fixed effect model based on the inverse variance method.

Table 7. Rate ratios (RR) for an increase of 5 ppb in NO_2 and associated 95% confidence intervals (95% CI) in Toronto for the association between cause-specific cardiovascular mortality and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, The Ontario Tax File Cohort Study, 1982-2004.

		All cardiovascular disease	Ischemic heart disease	Cerebrovascular disease
Exposure metrics	**Year**	**Fully adjusted RR (95% CI) [a]**	**Fully adjusted RR (95% CI)**	**Fully adjusted RR (95% CI)**
LURs (2002, 2004)	Mean of 2002-2004	1.05 (1.00 - 1.09)	1.06 (1.00 - 1.13)	0.91 (0.83 - 1.00)
IDW-based extrapolation	1982	1.02 (0.98 - 1.06)	1.04 (0.98 - 1.10)	0.93 (0.85 - 1.01)
	1992	1.02 (0.98 - 1.06)	1.05 (0.99 - 1.11)	0.91 (0.84 - 0.99)
	Mean of 1982-1992	1.02 (0.98 - 1.06)	1.04 (0.99 - 1.10)	0.92 (0.84 - 1.00)
LUR-based extrapolation	1982	1.03 (1.00 - 1.07)	1.05 (1.00 - 1.10)	0.95 (0.89 - 1.03)
	1992	1.05 (1.00 - 1.09)	1.06 (1.00 - 1.14)	0.95 (0.86 - 1.04)
	Mean of 1982-1992	1.05 (1.01 - 1.09)	1.07 (1.01 - 1.14)	0.94 (0.85 - 1.03)
	Mean of 1982-2004	1.05 (1.01 - 1.09)	1.07 (1.01 - 1.14)	0.93 (0.84 - 1.02)

[a] The baseline hazard function in the Cox regression models was stratified by 1-year age categories. The model was adjusted for age, sex, marriage status (four categories), annual household income (quintiles), % of immigrants in neighbourhood (quintiles), % of population with less than high school education in neighbourhood (continuous), unemployment rate in neighbourhood (continuous), and average household income in neighbourhood (quintiles).

Table 8. Prevalence of smoking status, according to five different levels of concentrations of NO_2 among the participants, 35-85 years of age, from the Canadian Community Health Survey in 2001

Exposure metrics	Hamilton			Toronto			Windsor		
	% Never smoker (95% CI)	% Current smoker (95% CI)	% Former smoker (95% CI)	% Never smoker (95% CI)	% Current smoker (95% CI)	% Former smoker (95% CI)	% Never smoker (95% CI)	% Current smoker (95% CI)	% Former smoker (95% CI)
NO_2 (ppb) [a]									
1st quintile [b]	32.4 (26.9 - 37.9)	17.6 (13.1 - 22.1)	50.0 (44.1 - 55.9)	42.9 (36.5 - 49.4)	23.0 (17.5 - 28.5)	34.1 (27.8 - 40.3)	35.5 (27.1 - 43.9)	26.6 (18.8 - 34.5)	37.9 (29.3 - 46.5)
2nd quintile	32.8 (24.8 - 40.9)	26.7 (19.1 - 34.4)	40.5 (32 - 48.9)	45.1 (38.8 - 51.4)	19.9 (15.0 - 24.8)	35.0 (29.1 - 40.8)	34.0 (24.8 - 43.2)	26.2 (17.8 - 34.6)	39.8 (30.4 - 49.2)
3rd quintile	35.6 (26.4 - 44.8)	23.1 (15.0 - 31.1)	41.4 (31.9 - 50.8)	42.2 (36.5 - 47.9)	19.4 (14.9 - 23.9)	38.4 (33.0 - 43.9)	40.9 (29.0 - 52.9)	15.2 (6.5 - 23.8)	43.9 (32.0 - 55.9)
4th quintile	30.8 (22.6 - 39.1)	38.3 (29.7 - 47)	30.8 (22.6 - 39.1)	49.2 (43.7 - 54.7)	20.1 (15.6 - 24.6)	30.7 (25.6 - 35.8)	37.4 (27.0 - 47.7)	33.7 (23.5 - 43.9)	28.9 (19.1 - 38.7)
5th quintile	28.0 (21.1 - 34.8)	39.1 (31.7 - 46.6)	32.9 (25.7 - 40.2)	38.4 (33.3 - 43.5)	24.3 (19.8 - 28.8)	37.3 (32.2 - 42.4)	39.8 (29.8 - 49.8)	28.0 (18.7 - 37.2)	32.3 (22.9 - 41.7)

[a] For Hamilton, concentrations of NO_2 were derived at the six-character postal code addresses of the participants of the 2001 Canadian Community Health Survey from a land use regression model using measurement from a 2002 fall monitoring campaign with 107 monitors. For Toronto, the levels of NO_2 were derived from the average estimate from two land use regression models using measurements from 2002 fall and 2004 spring monitoring campaigns with 100 monitors. For Windsor, the concentrations of NO_2 were derived from a land use regression model using measurements from 2004 spring, summer, fall, and winter monitoring campaigns with 54 monitors.

[b] For Hamilton, the quintiles of NO_2 (ppb) are: $\leq$13.8; 13.9-14.7; 14.8-15.8; 15.9-17.4; and $\geq$17.5. For Toronto, the quintiles of NO_2 (ppb) are: $\leq$18.7; 18.8-20.3; 20.4-21.9; 22-24; and $\geq$24.1. For Windsor, the quintiles of NO_2 (ppb) are: $\leq$10.6; 10.7-11.6; 11.7-12.5; 12.6-13.7; and $\geq$13.8.

Table 9. Indirect adjustment for possible confounding by unmeasured smoking for the association between cause-specific mortality and estimates of traffic-related air pollution from land use regression models. The rate ratios (RR) and associated 95% confidence interval (95% CI) are expressed for an increase of 5ppb of NO_2, The Ontario Tax File Cohort Study, 1982-2004.

	All cardiovascular disease		Ischemic heart disease		Cerebrovascular disease	
City	Fully-adjusted [a] RR (95% CI)	Indirect adjustment for smoking [b] RR (95% CI)	Fully-adjusted [a] RR (95% CI)	Indirect adjustment for smoking [b] RR (95% CI)	Fully-adjusted [a] RR (95% CI)	Indirect adjustment for smoking [b] RR (95% CI)
Hamilton	1.12 (1.06 - 1.19)	1.02 (0.91 - 1.11)	1.12 (1.02 - 1.21)	1.02 (0.90 - 1.13)	1.06 (0.92 - 1.22)	0.93 (0.77 - 1.10)
Toronto	1.05 (1.00 - 1.09)	1.04 (0.99 - 1.09)	1.06 (1.00 - 1.13)	1.05 (0.98 - 1.13)	0.91 (0.83 - 1.00)	0.91 (0.82 - 1.00)
Windsor	1.10 (1.02 - 1.19)	1.09 (0.98 - 1.21)	1.11 (1.00 - 1.23)	1.10 (0.98 - 1.25)	0.99 (0.82 - 1.18)	0.97 (0.79 - 1.18)
Pooled estimate [c]	1.08 (1.05 - 1.11)	1.04 (1.00 - 1.09)	1.09 (1.04 - 1.14)	1.05 (1.00 - 1.11)	0.95 (0.89 - 1.02)	0.92 (0.85 - 1.00)

[a.] The baseline hazard function in the Cox regression models was stratified by 1-year age categories. The model was adjusted for age, sex, marital status (four categories), annual household income (quintiles), % of immigrants in neighbourhood (quintiles), % of population with less than high school education in neighbourhood (continuous), unemployment rate in neighbourhood (continuous), and average household income in neighbourhood (quintiles).

[b.] Similar to (a), except that smoking was further adjusted using a Monte Carlo sensitivity analysis approach proposed by Steenland and Greenland (2004) and we extended the method so that it can handle continuously measured exposure variables. For each smoking-adjusted RR, the associated 95% CIs were computed from 100,000 replications.

[c.] A test of heterogeneity of rate ratios yielded p values > 0.05. The pooled estimate was obtained using a fixed effect model using the inverse variance method.

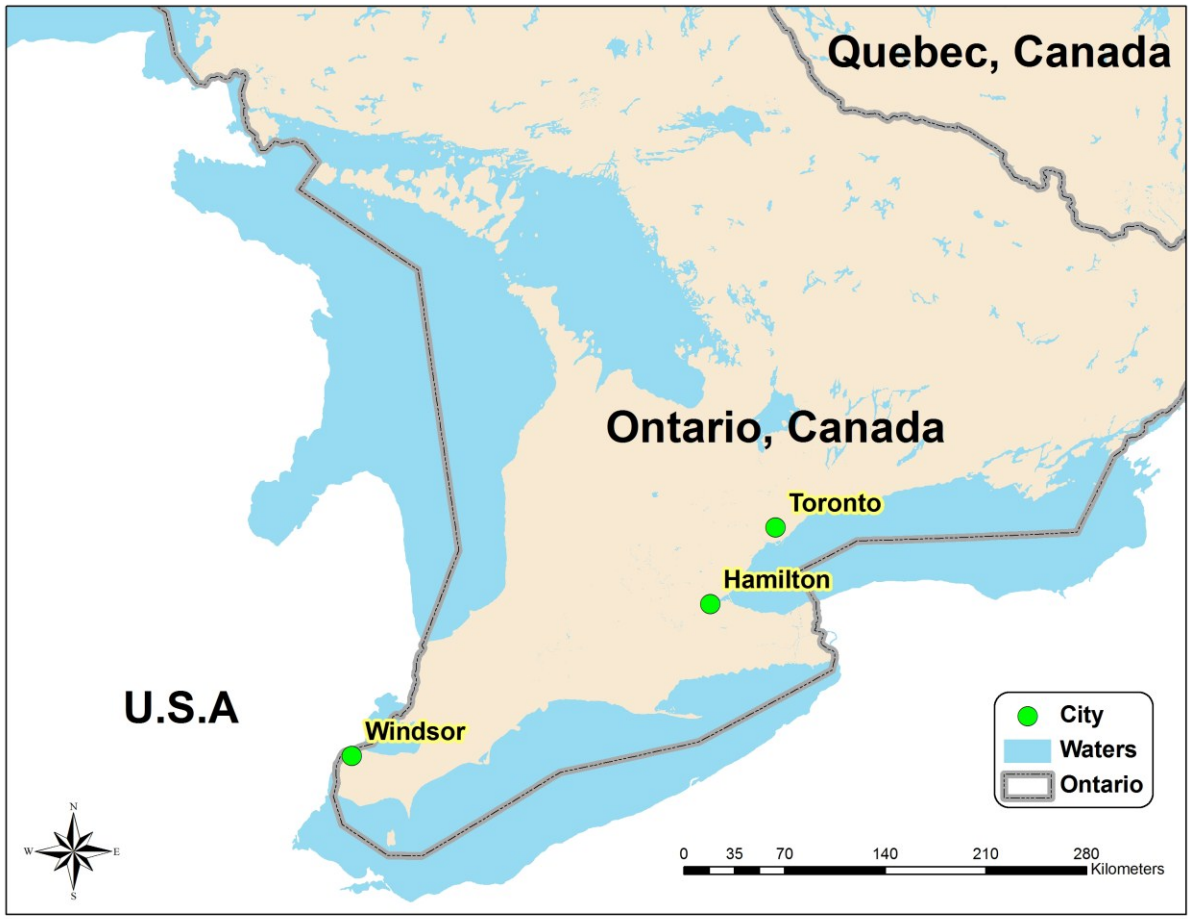
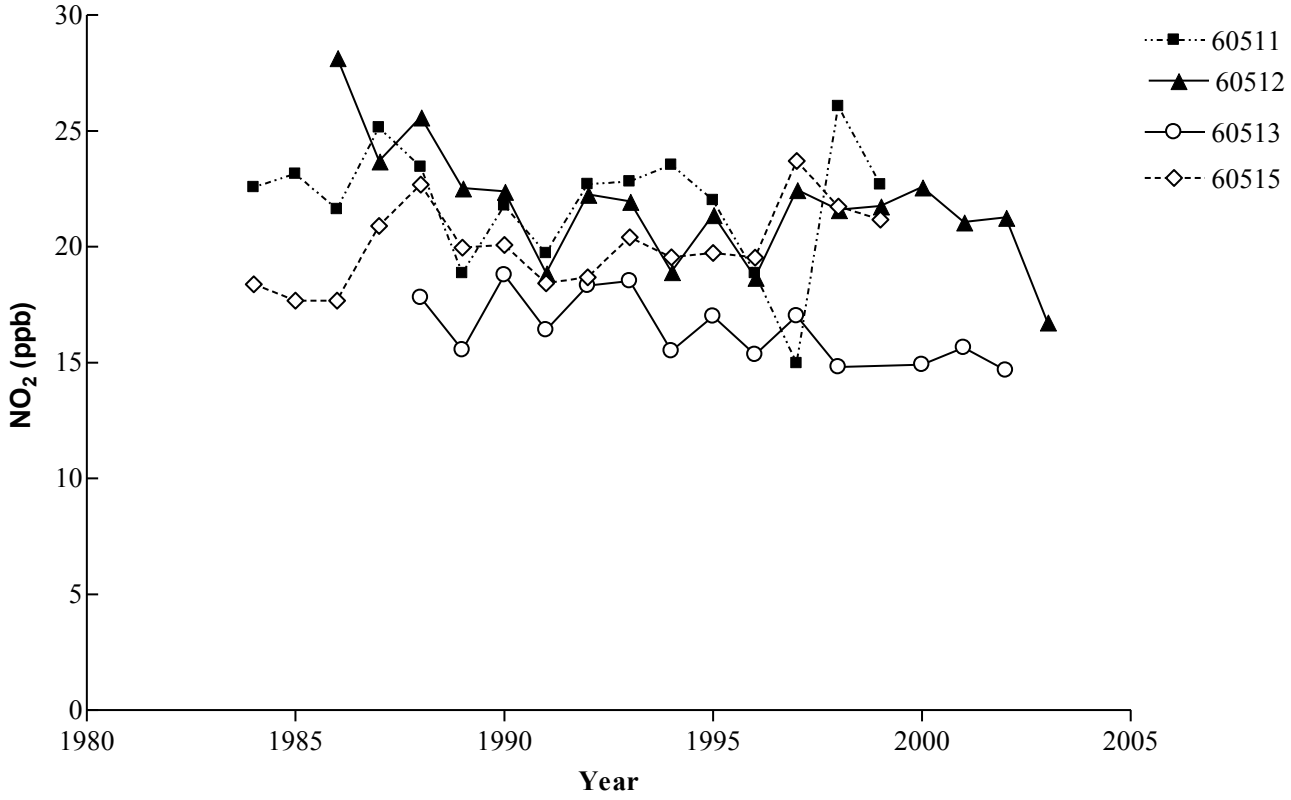
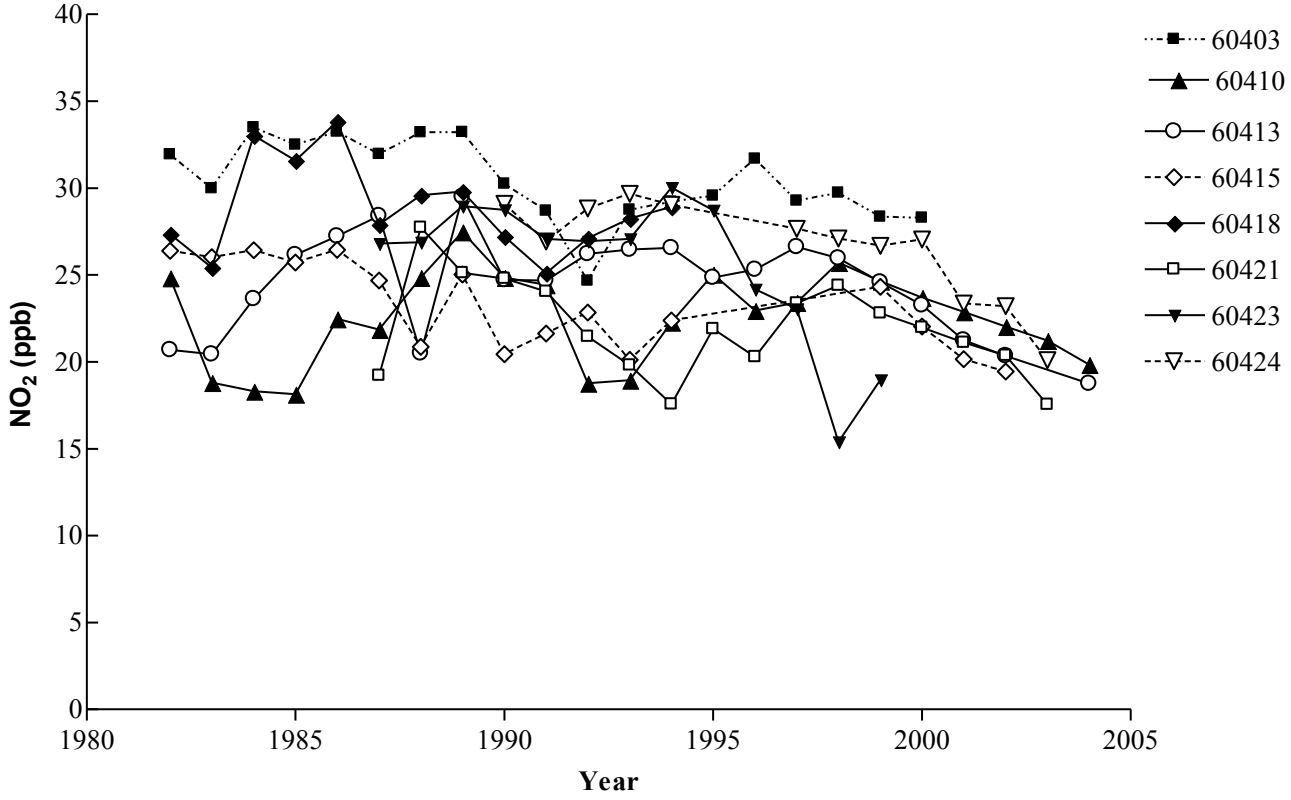


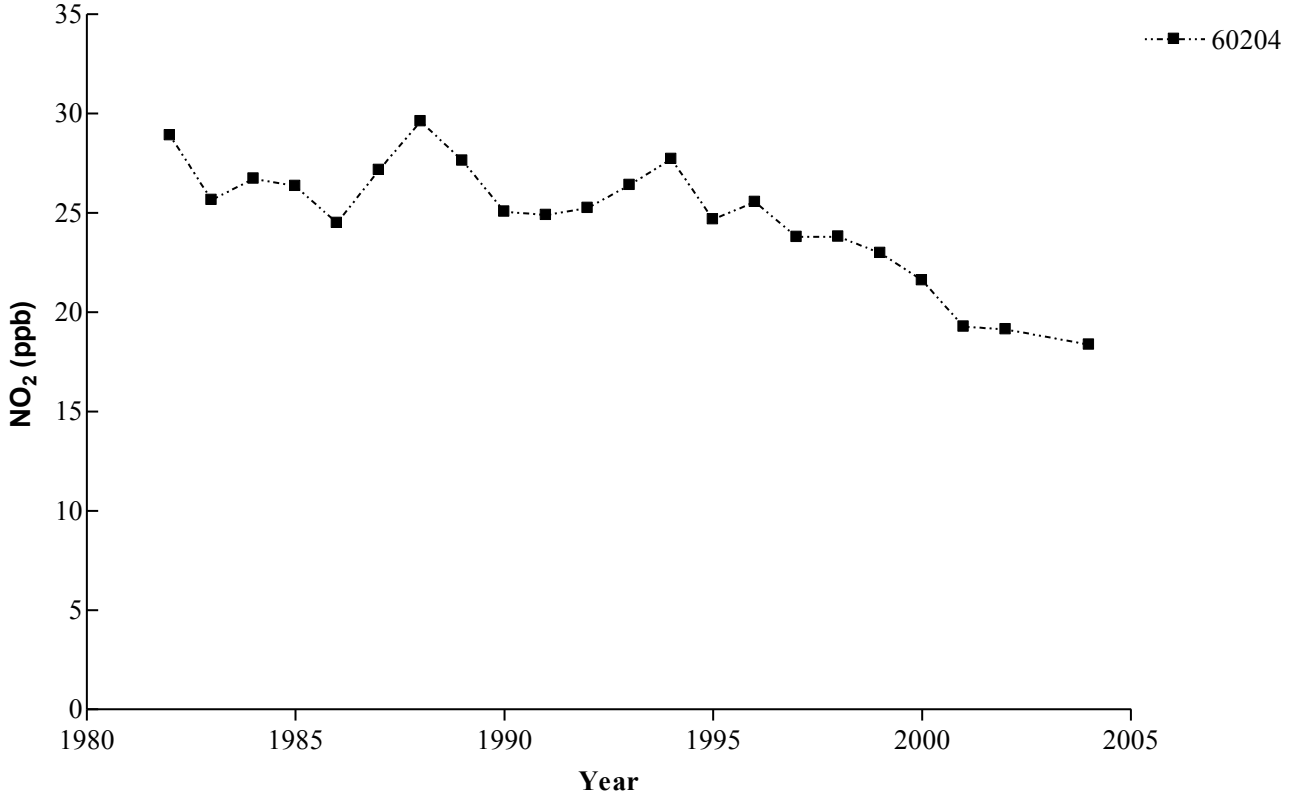
Figure 1. The cities of Toronto, Hamilton, and Windsor, Ontario, Canada


30
25
20
15
10
5
0
1980
1985
1990
1995
2000
2005
NO2 (ppb)
Year
60511
60512
60513
60515

(A) Hamilton


40
35
30
25
20
15
10
5
0
1980
1985
1990
1995
2000
2005
Year
NO2 (ppb)
60403
60410
60413
60415
60418
60421
60423
60424

(B) Toronto



(C) Windsor ‡

‡ Two fixed-site monitors were operated during the period of 1982 to 2004. However, besides the fixed-site monitor shown in the figure above, the other fixed-site monitor (60211) was operated in 2000-2004 only.

Figure 2. Trends in observed annual average concentrations of nitrogen dioxides (in ppb) across fixed-site monitors in (A) Hamilton, (B) Toronto, and (C) Windsor, 1982-2004, respectively. The fixed-site monitors are administered by the National Air Pollution Surveillance (NAPS) network in the Ontario region (site number is provided in the legend). For each city, the fixed-site monitors that operated for less than half of the period (≤12 years) are not shown.

Appendix A. Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, among study participants in Toronto, The Ontario Tax File Cohort Study, 1982-2004. The rate ratios are expressed for an increased of the interquartile range (IQR) of NO_2. [a]

			All cardiovascular disease	Ischemic heart disease	Cerebrovascular disease
Exposure metrics	**Year**	**IQR (ppb)**	**Adjusted RR (95% CI)** [b]	**Adjusted RR (95% CI)**	**Adjusted RR (95% CI)**
LURs (2002, 2004)	Mean of 2002-2004	4.09	1.04 (1.00 - 1.08)	1.05 (1.00 - 1.11)	0.93 (0.86 - 1.00)
IDW-based extrapolation	1982	4.82	1.02 (0.98 - 1.06)	1.03 (0.98 - 1.10)	0.93 (0.86 - 1.01)
	1992	5.00	1.02 (0.98 - 1.06)	1.05 (0.99 - 1.11)	0.91 (0.84 - 0.99)
	Mean of 1982-1992	4.77	1.02 (0.98 - 1.06)	1.04 (0.99 - 1.10)	0.92 (0.85 - 1.00)
LUR-based extrapolation	1982	5.51	1.03 (1.00 - 1.07)	1.06 (1.00 - 1.12)	0.95 (0.88 - 1.03)
	1992	4.68	1.04 (1.00 - 1.09)	1.06 (1.00 - 1.13)	0.95 (0.87 - 1.04)
	Mean of 1982-1992	4.67	1.04 (1.01 - 1.09)	1.07 (1.01 - 1.13)	0.94 (0.86 - 1.02)
	Mean of 1982-2004	4.40	1.04 (1.00 - 1.08)	1.06 (1.00 - 1.12)	0.94 (0.86 - 1.01)

[a] The RRs for an increase equal to the interquartile range of each predicted distribution of the different exposure metrics to account for differences in the absolute value of the distributions.

[b] The baseline hazard function in the Cox regression models was stratified by 1-year age categories. The model was adjusted for age, sex, marital status (four categories), annual household income (quintiles), % of immigrants in neighbourhood (quintiles), % of population with less than high school education in neighbourhood (continuous), unemployment rate in neighbourhood (continuous), and average household income in neighbourhood (quintiles).

Appendix B. Distribution of measured body mass index (in kg/m^2) [a] according to the quintiles of concentrations of NO_2 among the participants aged 35 years and above, from the Canadian Community Health Survey in 2001

Exposure metrics	Hamilton	Toronto	Windsor
	Mean (standard deviation)	Mean (standard deviation)	Mean (standard deviation)
NO_2 (in quintile,) [b]			
1st quintile [b]	26.3 (5.0)	24.7 (4.3)	27.3 (5.8)
2nd quintile	26.7 (4.6)	24.9 (3.9)	27.2 (6.1)
3rd quintile	27.9 (5.6)	25.2 (4.4)	26.2 (4.1)
4th quintile	26.4 (5.5)	25.5 (4.7)	26.2 (5.0)
5th quintile	26.5 (5.5)	25.4 (4.6)	26.6 (5.4)

[a] The body-mass index is the weight in kilograms divided by the square of the height in meters.

[b] For Hamilton, the quintiles of NO_2 (ppb) are: $\leq$13.8; 13.9-14.7; 14.8-15.8; 15.9-17.4; and $\geq$17.5. For Toronto, the quintiles of NO_2 (ppb) are: $\leq$18.7; 18.8-20.3; 20.4-21.9; 22-24; and $\geq$24.1. For Windsor, the quintiles of NO_2 (ppb) are: $\leq$10.6; 10.7-11.6; 11.7-12.5; 12.6-13.7; and $\geq$13.8.

Appendix C. Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and estimates of concentrations of nitrogen dioxide from land use regression models (2002 and 2004), The Ontario Tax File Cohort Study, 1982-2004. The rate ratios are expressed for an increase of the inter-quartile range (IQR) of NO_2.

	City			
	Hamilton (IQR: 2.88 ppb)	Toronto (IQR: 4.09 ppb)	Windsor (IQR: 2.47 ppb)	Pooled estimate [d]
Cause of Death	RR_{IQR} (95% CI)	RR_{IQR} (95% CI)	RR_{IQR} (95% CI)	RR_{IQR} (95% CI)
All Cardiovascular Disease				
Age and sex adjusted [a]	1.15 (1.12 - 1.18)	1.04 (1.02 - 1.06)	1.13 (1.08 - 1.15)	-
+ All personal covariables [b]	1.09 (1.06 - 1.12)	1.01 (0.98 - 1.04)	1.05 (1.02 - 1.10)	-
+ All ecological covariables [c]	1.06 (1.03 - 1.12)	1.04 (1.00 - 1.07)	1.05 (1.00 - 1.10)	1.05 (1.02 - 1.08)
Ischemic Heart Disease				
Age and sex adjusted	1.18 (1.12 - 1.22)	1.06 (1.02 - 1.10)	1.13 (1.08 - 1.15)	-
+ All personal covariables	1.09 (1.06 - 1.15)	1.02 (0.98 - 1.06)	1.05 (1.00 - 1.10)	-
+ All ecological covariables	1.06 (1.00 - 1.12)	1.05 (1.00 - 1.11)	1.05 (1.00 - 1.10)	1.05 (1.02 - 1.08)
Cerebrovascular Disease				
Age and sex adjusted	1.09 (1.03 - 1.18)	0.97 (0.92 - 1.04)	1.05 (0.98 - 1.13)	-
+ All personal covariables	1.03 (0.97 - 1.12)	0.93 (0.88 - 1.00)	1.02 (0.95 - 1.08)	-
+ All ecological covariables	1.03 (0.95 - 1.12)	0.93 (0.85 - 1.00)	0.98 (0.90 - 1.08)	0.97 (0.93 - 1.02)

[a.] The baseline hazard function in the Cox regression models stratified by 1-year age categories, because the proportional hazards assumption was not satisfied for age. In addition, the model was adjusted for sex.

[b.] Same model as in a), but adjusted for sex, marital status (four categories), and annual household income (quintiles).

[c.] Same model as in a), but adjusted for sex, marital status (four categories), annual household income (quintiles), % of immigrants in neighbourhood (quintiles), % of population with less than high school education in neighbourhood (continuous), unemployment rate in neighbourhood (continuous), and average household income in neighbourhood (quintiles).

[d.] A test of heterogeneity of city-specific rate ratios that were adjusted for only age and sex yielded p values < 0.05. As a result, we did not pool the age- and sex-adjusted ratio ratios across cities. Similarly, we did not the rate ratios that adjusted for all individual-level covariables across cities, because the test of heterogeneity of rate ratios yielded p values < 0.05. On the other hand, the test of heterogeneity of rate ratios that adjusted for all individual-level covariables and ecological covariables yielded p values > 0.05. Therefore, the pooled estimate of fully-adjusted RRs was obtained using a fixed effect model based on the inverse variance method.

Chapter 7

Discussion and Conclusions

In this dissertation, I addressed three research questions. First, I conducted a systematic review of the epidemiological literature of reported associations between ambient air pollution and incidence and mortality of cancer, and cardiovascular and respiratory diseases. Second, I developed three new methods to extrapolate "current" land use regression models back in time, which allows reconstruction of historical exposure to traffic-related air pollution. Third, I conducted a cohort study in three Ontario cities to investigate whether cause-specific cardiovascular mortality is associated with long-term exposure to traffic-related air pollution. This chapter summarizes the key findings of the dissertation, discusses the strengths and the limitations, reviews the potential implication for public health of this dissertation, and recommends directions for future research.

7.1 Summary of Key Findings

In the first published research paper I conducted a systematic review of the literature of the current state of science regarding the chronic health effects of ambient air pollution. The review showed that long-term exposure to $PM_{2.5}$ may increase the risk of nonaccidental mortality, cardiovascular mortality, and lung cancer by 6%, 12%-14%, and 7%-12% per a $10\mu g/m^3$ increase, respectively. Living close to highways or major urban roads appears also to be associated with elevated nonaccidental mortality and cardiovascular mortality. For other particulate and gaseous pollutants, the paucity of data precludes drawing strong conclusions. The updated review showed that the recently published studies added to the evidence associating elevated risk of cardiovascular mortality with $PM_{2.5}$. In addition, there appeared to be accumulating evidence suggesting that long-term exposure to NO_2 may increase the risk of all cardiovascular mortality and mortality from ischemic heart disease by 3% and 6 % for each increase of a $10\mu g/m^3$ of NO_2. For the association between NO_2 and cerebrovascular mortality or nonfatal incidence of cardiovascular disease, the data are insufficient to make any conclusions.

In the second research paper, I described three new back-extrapolation methods to reconstruct historical exposure to traffic-related air pollution at the intraurban scale, and as

an illustration, these new methods were applied to a case-control study of postmenopausal breast cancer and traffic-related air pollution in Montreal, Quebec. The paper illustrated step-by-step the procedures of these new methods. Using data from the case-control study of postmenopausal breast cancer, we also showed that the estimated concentrations using the three extrapolation methods had similar distributions, except that one method yielded slightly lower values.

In the third research paper, I carried out a cohort study of the associations between traffic-related air pollution (NO_2) and cause-specific cardiovascular mortality in Toronto, Hamilton, and Windsor in Ontario. Residential concentrations of NO_2 were estimated using land use regression models and the back-extrapolation methods were applied to Toronto where sufficient data from fixed-site monitors were available. The Cox proportional hazards model was used to estimate associations between air pollution and cardiovascular and cerebrovascular mortality. Associations were adjusted for individual risk factors and ecological covariables, and smoking was adjusted indirectly using a series of sensitivity analyses. For each increase of 5ppb of NO_2, the plausible range of the estimates of increased risk for all cardiovascular mortality varied between 4% (95% CI: 0-9%) and 10% (95% CI: 0-21%). For ischemic heart disease, the estimated risks were similar to those observed for all cardiovascular mortality. I found no statistically significant association between traffic-related air pollution and cerebrovascular mortality (excess rates of mortality of -5% (95%CI: -11% - 2%).

7.2 Strengths and Limitations

There are several important strengths of this dissertation. The systematic review of the literature is a reproducible and transparent appraisal of the current evidence of adverse health effects of long-term exposure to ambient air pollution. This process of including all available studies minimized potential selection biases and there is no reason to believe that there is a publication bias. In addition, the use of standard meta-analytic methods provided an unbiased estimate of the health effects of individual air pollutants. Additionally, the use of subgroup analysis allowed examination of the heterogeneity across individual studies and

thus provided more insights into the aetiology of chronic diseases from exposures to ambient air pollution.

The second study is one of the very few (219) to make use of back-extrapolation of land use regression models. One strength of this study is the development and introduction of three new methods for assessing historical exposure to traffic-related air pollution. By using real data from an epidemiological study, this study demonstrated the utility of these new methods.

The third study is one of the few to assess intra-urban variation in traffic-related pollution and its effects on cardiovascular mortality. This study has three important strengths. First, it was population-based and it had large statistical power to detect small effects. Second, estimates of exposure to traffic-related air pollution were obtained using land use regression models and, in Toronto, the back-extrapolation methods. Therefore, using this approach likely reduced misclassification of exposure. Third, some important confounding variables, especially household income and marital status, were measured accurately because they were obtained directly from the tax file. This reduced possible residual confounding due to misclassification of these risk factors for cardiovascular disease.

There are several limitations of this dissertation. In the first study, although a total of 25 cohort studies and 23 case-control studies were conducted between 1950 and 2011, not all studies could be included in the meta-analysis because of the diverse types of pollutants and metrics used to characterize exposure to air pollution. In addition, it appeared that the differences in gender and losses to follow-up may be potential sources of heterogeneity among the studies, but I could not be certain whether other characteristics could also explain the observed differences.

In the second study, the three new methods were developed based on an assumption that we could estimate past exposures by rescaling the current land use regression model using measured changes in spatial distributions of NO_2 as derived from fixed-site monitoring stations. These methods of estimating these historical spatial distributions (IDW or land use regression) are limited by the number of fixed-site monitors. The rescaling entailed multiplying the current land use regression model by a ratio which was inherently less

accurate, as the final error would be a function of the sum of the errors of each surface. (Indeed, if the actual errors were known, the final error could be estimated using the delta method for the propagation of errors (217).) Nevertheless, the expectation is that these extrapolation methods should be more accurate than simply using estimates derived from data collected well after the relevant periods.

The accuracy of extrapolating land use regression models in other locations will depend on the number of historical monitors available. In addition, the method of spatial extrapolation will also affect accuracy. Among the three extrapolation methods, the IDW-based approach may introduce the largest errors because IDW is an exact interpolator (218) (*i.e.*, predicted concentrations will be constrained by the range of the observed values) and it may produce over-smoothed estimates of spatial variations of NO_2 as compared to land use regression. The land use regression models using sparse monitoring data were constrained because only a small number of spatial covariables could be fitted in the model. In addition, relatively few monitors may not provide the full range of urban background concentrations and land use. But when more monitoring data are available, the land use regression-based extrapolation approach should produce more accurate estimates. One direction in the further development of the land use regression-based extrapolation method may be to explore different ways of incorporating historical land use and vehicular traffic.

In the third study, limitations are associated with the development of these methods of exposure assessment. The first one relates to the issue of accuracy of the estimates of exposure. Ideally, one would like to assess personal exposures to traffic-related air pollution, but it is clearly impossible to assess personal exposures in any large longitudinal study. Using address of residence as a proxy for personal exposures assumes that other factors, such as air infiltration, individual mobility, and daily activity patterns, do not have an important influence on total personal exposures. We did not assess mobility in this study. Although numerous studies have shown that it is reasonable to use spatially derived exposures at the residences as a surrogate for personal exposures, especially given that Canadian adults spend on average over 65% at home (indoor and outdoor) and closer to 85% indoors (263), in the end, one would expect that use of residential exposures will surly lead to misclassification and if independent of outcome will likely attenuate associations.

I used existing land use regression models to estimate residential concentrations of traffic-related air pollution. However, these models were developed towards the end of follow-up of the study, and I made use of three new methods to back-extrapolate "current" land use regression models back in time. The underlying assumption of these new methods is that the surface of NO_2 derived from a land use regression model would change over time in proportion to what we observe from fixed-site monitoring stations in the study area. As a result, the accuracy of the back-extrapolation methods depends on the accuracy of the land use regression model, and the number and spatial distribution of available historical monitors. My analysis using two back-extrapolation methods showed that the predicted concentrations in Toronto across different periods were correlated strongly with each other. This finding was supported by the fact that the annual mean concentrations of NO_2 and their rank ordering at fixed-site monitors in Toronto and in Hamilton remained largely constant during the period of study, 1982 to 2004. In a previous study that was conducted in Toronto, Jerrett et al. (2009) also showed that the spatial pattern of concentrations of NO_2 was relatively stable over time. The relative stability over time in the spatial patterns of ambient concentrations of NO_2 within cities has also been reported in studies in Montreal, Canada (3) and in the Netherlands (219). As a result, it is expected that the spatial contrast in NO_2 estimated using the land use regression models over a short period would provide reasonable estimates of longer-term spatial exposures.

The second limitation relates to unmeasured confounding variables. As with many cohort studies based on administrative databases, I did not have access to information on some important individual risk factors of cardiovascular diseases, such as smoking (266, 267), body weight (19-22), high blood pressure (16-18), diabetes (23, 24), and family history (268). If these risk factors varied with the levels of exposure, failure to control for them could lead to bias. Smoking is one of the strongest risk factors for cardiovascular disease, because of the strength of its effect and its high prevalence, and it thus had the potential to strongly confound the associations. To control for smoking, I applied a method proposed by Axelson and extended by Steenland and Greenland (2004) to use Monte Carlo sensitivity analysis of an indirect method of adjustment. As expected, indirect adjustment for smoking had almost no impact on the rate ratios of cardiovascular mortality in Toronto and in Windsor, because

the prevalences of current and former smoking were weakly correlated with ambient concentrations of NO_2. On the other hand, substantial reduction in rate ratios was observed for Hamilton due to strong associations between smoking and the concentrations of NO_2 in the city. Indeed, the accuracy of this modified method of indirect adjustment depends on the distributions of the prevalence of smoking and on the relative risks for smoking. In addition, the assumption that the bias factor is approximately constant across the strata of covariates such as age and sex will also affect accuracy.

I was unable to examine the potential confounding effects of other unmeasured risk factors, such as high blood pressure, family history, and diabetes. Although I could not rule out potential residual confounding by these risk factors, their confounding effects would likely be small as compared to smoking. In addition, as I adjusted for other risk factors such as age, sex, and smoking, and these are also correlated with other risk factors that were not measured (274-276), these adjustments probably controlled for these other factors. In addition, there is suggestive evidence that exposure to traffic-related air pollution may increase the risk of higher blood pressure (277, 278) and diabetes (279, 280), and adjustment for these potential mediators of the relationship between traffic air pollution and cardiovascular mortality may not be appropriate. As well, these unmeasured risk factors were unlikely to have differed in prevalence as much as smoking across the different levels of air pollution. Indeed, in an earlier study that was conducted in Toronto, Jerrett et al. (2009) showed that adjustment for smoking, diabetes, pre-existing COPD, and chronic ischemic heart disease resulted in a decrease of only 5% in the estimate of effects of NO_2 on cardiovascular mortality. Similarly, in the Nurses' Health Study (53, 189), Puett and colleagues showed that adjusting for personal risk factors such as smoking, body mass index, diabetes, family history, and hypertension did not have any effect on the estimates of association between particulate pollutants and mortality from coronary heart disease.

7.3 Policy Implications

Air pollution is a global problem. The increased risk for cardiovascular mortality uncovered in this thesis and observed in the other studies focused on traffic suggests strongly that traffic-related air pollution is a risk to the public's health, even at the relatively low levels

found in Canada. The development, monitoring, and enforcement of both emissions standards and air quality standards is necessary in order to improve urban air quality and to protect the health of humans and the environment. In the wake of large impact on chronic illness due to long-term exposure to ambient pollution, the weight of the evidence from the epidemiological studies strongly supports tighter standards for air pollution and emission control in Canada and in other countries. To lower emissions, one possible direction is to reduce emissions in new vehicles, increase fuel economy, use clean fuels, improve dramatically public transit, improve the design of built environment such as increasing bicycle and walking paths, and reduce the number of cars on the roadways by introducing measures that will induce people to take public transit (285, 286).

It is not a realistic option to simply increase taxes on gasoline in the hope that people will take public transit, as one needs a system that is both rapid and convenient. In fact, it is expected that gasoline prices will increase as supply of gasoline dwindles and is replaced by more expensive and carbon rich alternatives, such as from liquidification of coal. With an exponentially growing population, these types of changes may lead to much higher levels of pollution than we have today, even with improved technologies of combustion and exhaust recovery, and the effects on population health may become profound.

7.4 Future Research Priorities

Although this research showed a positive association between NO_2 and cardiovascular mortality, the findings should not be interpreted as meaning that NO_2 is a causal agent, rather NO_2 may be acting as a surrogate for the complex mixture of traffic-related air pollutants. Indeed, Brook et al. (2007) showed that NO_2 has strong co-locational associations with several air pollutants generated from local mobile sources.(193) Whether NO_2 represents the harmful effects of ultrafine particles or diesel exhaust particulates, or other components of air pollution such as VOCs or PAHs is unclear. Additional research is necessary to expand our knowledge related to the "responsible" constituents of the traffic-related mixture. Nevertheless, from a public health point of view, it does not matter what the "active" agents may be; this is similar to the case of smoking where the "causal" agents will never be identified (154).

Although available epidemiologic data suggested larger risks of cardiovascular disease posed by long-term exposure (in years) than observed only a few days (56), it remains unclear whether it is exposures in the past or more recent exposures to air pollution that are etiologically important. For a public health intervention to be effective, it is critical to take into account the relative importance of various time windows of exposure to traffic-related air pollution on causing cardiovascular events. Indeed, the key issue is to better understand what the real causal metrics are: whether it be cumulative exposure or acute exposure; this merits further investigation.

Conducting a large-scale dense monitoring campaign of ambient air pollutants can be time-consuming and costly. The back-extrapolation method developed in this research holds great value because of its possibility of extrapolating estimates from a current land use regression model into the future. In addition, the methods developed in this research can be used to address the problems and unanswered questions such as assessing relative importance of different exposure time windows. As noted earlier, the accuracy of these extrapolation methods can be improved when more monitoring data are available. One direction in the further development of the extrapolation methods may be to explore different ways of incorporating historical land use and vehicular traffic.

7.5 Conclusions

The overall impact of the research presented in this dissertation is the contribution to exposure assessment, indirect assessment of unmeasured confounding factors, as well as augmenting the body of evidence regarding associations between exposure to air pollution and adverse cardiovascular outcomes. Along with other previous studies, this dissertation indicates that exposure to traffic-related pollution is a public health concern. There is a need for pollution control strategies and continued research into the associations between human health and urban air pollution.

Appendix D. A statistical program of indirect adjustment for unmeasured confounders

The following program was written in R:

```
#########################################################
### Ref: Thun M 2000 JAMA - smoking and ACSII
###
### Weighted by age and sex distributions of CCHS
###
#########################################################

###################
## TORONTO
###################
#
### All cardiovascular: pooling rr for males and females and for chd and other cvds
###
# current smoker: (chd) males, females, (other) males, females
#all.logrr <- c(0.904548685376769, 1.04814670821595, 0.587786665, 0.530628251)
#all.logse <- c(0.0440772328784456, 0.0530564620641477, 0.056924375, 0.06030326)
#all.cur <- meta.summaries(all.logrr, all.logse,  method="fixed", logscale=TRUE)
# ex smoker
#all.logrr <- c(0.375699008810345, 0.336472236621213, 0.262364264467491, 0.18232155679395)
#all.logse <- c(0.0340792197555917, 0.0631901936443283, 0.0393241530171577,
0.0426158379242771)
#all.ex <- meta.summaries(all.logrr, all.logse,  method="fixed", logscale=TRUE)

#
### All cardiovascular: pooling rr for males and females and for chd and stroke
###
# current smoker: (chd) males, females, (other) males, females
all.logrr <- c(0.904548685376769, 1.04814670821595, 0.831922075219311, 1.17772880537114)
all.logse <- c(0.0440772328784456, 0.0530564620641477, 0.118626701777018, 0.0880082313841352)
all.cur <- meta.summaries(all.logrr, all.logse,  method="fixed", logscale=TRUE)
# ex smoker
all.logrr <- c(0.375699008810345, 0.336472236621213, 0, 0.338620343837828)
all.logse <- c(0.0340792197555917, 0.0631901936443283, 0.0989310815979201, 0.109874751635257)
all.ex <- meta.summaries(all.logrr, all.logse,  method="fixed", logscale=TRUE)

### CHD: pooling rr for males and females
# current smoker: males, females
chd.logrr <- c(0.904548685376769, 1.04814670821595)
chd.logse <- c(0.0440772328784456, 0.0530564620641477)
chd.cur <- meta.summaries(chd.logrr, chd.logse,  method="fixed", logscale=TRUE)
# ex smoker: males, females
chd.logrr <- c(0.375699008810345, 0.336472236621213)
chd.logse <- c(0.0340792197555917, 0.0631901936443283)
chd.ex <- meta.summaries(chd.logrr, chd.logse,  method="fixed", logscale=TRUE)
```

```
### stroke: pooling rr for males and females

# current smoker: males, females
strk.logrr <- c(0.831922075219311, 1.17772880537114)
strk.logse <- c(0.118626701777018, 0.0880082313841352)
strk.cur <- meta.summaries(strk.logrr, strk.logse,  method="fixed", logscale=TRUE)
# ex smoker: males, females
strk.logrr <- c(0, 0.338620343837828)
strk.logse <- c(0.0989310815979201, 0.109874751635257)
strk.ex <- meta.summaries(strk.logrr, strk.logse,  method="fixed", logscale=TRUE)

### output

rr.cardio.smk <- data.frame(logrr=rep(NA, 6), logse=rep(NA, 6), row.names=c("all.cur", "all.ex",
"chd.cur", "chd.ex", "strk.cur", "strk.ex"))
rr.cardio.smk[1,] <- c(all.cur$summary, all.cur$se.summary)
rr.cardio.smk[2,] <- c(all.ex$summary, all.ex$se.summary)
rr.cardio.smk[3,] <- c(chd.cur$summary, chd.cur$se.summary)
rr.cardio.smk[4,] <- c(chd.ex$summary, chd.ex$se.summary)
rr.cardio.smk[5,] <- c(strk.cur$summary, strk.cur$se.summary)
rr.cardio.smk[6,] <- c(strk.ex$summary, strk.ex$se.summary)

print(rr.cardio.smk)

#########################
# auxillary functions #
#########################
expit <- function(x) {1/(1+exp(-x))}
logit <- function(p) {log(p)-log(1-p)}

library(MASS)

####################################
#         TORONTO cohort
####################################
##### A. q2-q5

## frequency of never and current smokers in each exposure category
pop <- data.frame(pop=rep(NA, 5), row.names=c("q1", "q2", "q3", "q4", "q5"))
pop[1,] <- Here enter prevalence of never and current smokers in the reference population
pop[2,] <- same as above
pop[3,] <- same as above
pop[4,] <- same as above
pop[5,] <- same as above

## prevalence of never and current smokers in each exposure category
smk.prev <- data.frame(never=rep(NA, 5), cur=rep(NA, 5), row.names=c("q1", "q2", "q3", "q4",
"q5"))
smk.prev[1,] <- logit(c(42.92, 23.01)/100)
smk.prev[2,] <- logit(c(45.12, 19.92)/100)
smk.prev[3,] <- logit(c(42.18, 19.39)/100)
smk.prev[4,] <- logit(c(49.19, 20.06)/100)
```

```
smk.prev[5,] <- logit(c(38.42, 24.29)/100)

## min and max conc. of no2 in each quintile
no2 <- data.frame(min=rep(NA, 5),max=rep(NA, 5),row.names=c("q1", "q2", "q3", "q4", "q5"))
no2[1,] <- c(10.606,18.6919)
no2[2,] <- c(18.692,20.3201)
no2[3,] <- c(20.3202,21.8531)
no2[4,] <- c(21.8532,23.983)
no2[5,] <- c(23.9842,42.0294)


####################################
#         variance and covariance, same formula for TOR HAM and WIN
####################################

## variance of prevalence of never and current smokers (same for Toronto, Hamilton and Windsor)
var.prev <- data.frame(never=rep(NA, 5), cur=rep(NA, 5), row.names=c("q1", "q2", "q3", "q4",
"q5"))
var.prev[1,] <- (1/(sqrt(pop[1,]*expit(smk.prev[1,])*(1-expit(smk.prev[1,])))))^2
var.prev[2,] <- (1/(sqrt(pop[2,]*expit(smk.prev[2,])*(1-expit(smk.prev[2,])))))^2
var.prev[3,] <- (1/(sqrt(pop[3,]*expit(smk.prev[3,])*(1-expit(smk.prev[3,])))))^2
var.prev[4,] <- (1/(sqrt(pop[4,]*expit(smk.prev[4,])*(1-expit(smk.prev[4,])))))^2
var.prev[5,] <- (1/(sqrt(pop[5,]*expit(smk.prev[5,])*(1-expit(smk.prev[5,])))))^2

## correlation and covariance of prevalence of never and current smokers (same for Toronto,
Hamilton and Windsor)
corr.prev <- data.frame(corr=rep(NA, 5), row.names=c("q1", "q2", "q3", "q4", "q5"))
corr.prev[1,] <- -sqrt(expit(smk.prev[1,1])*expit(smk.prev[1,2])/((1-expit(smk.prev[1,1]))*(1-
expit(smk.prev[1,2]))))
corr.prev[2,] <- -sqrt(expit(smk.prev[2,1])*expit(smk.prev[2,2])/((1-expit(smk.prev[2,1]))*(1-
expit(smk.prev[2,2]))))
corr.prev[3,] <- -sqrt(expit(smk.prev[3,1])*expit(smk.prev[3,2])/((1-expit(smk.prev[3,1]))*(1-
expit(smk.prev[3,2]))))
corr.prev[4,] <- -sqrt(expit(smk.prev[4,1])*expit(smk.prev[4,2])/((1-expit(smk.prev[4,1]))*(1-
expit(smk.prev[4,2]))))
corr.prev[5,] <- -sqrt(expit(smk.prev[5,1])*expit(smk.prev[5,2])/((1-expit(smk.prev[5,1]))*(1-
expit(smk.prev[5,2]))))

cov.prev <- data.frame(cov=rep(NA, 5), row.names=c("q1", "q2", "q3", "q4", "q5"))
cov.prev[1,] <- corr.prev[1,]*sqrt(var.prev[1,1])*sqrt(var.prev[1,2])
cov.prev[2,] <- corr.prev[2,]*sqrt(var.prev[2,1])*sqrt(var.prev[2,2])
cov.prev[3,] <- corr.prev[3,]*sqrt(var.prev[3,1])*sqrt(var.prev[3,2])
cov.prev[4,] <- corr.prev[4,]*sqrt(var.prev[4,1])*sqrt(var.prev[4,2])
cov.prev[5,] <- corr.prev[5,]*sqrt(var.prev[5,1])*sqrt(var.prev[5,2])

####################################
#### MCSA
####################################

## define two vectors and two matrix
k <-1
count <- 0
```

```
bnew <- matrix(rep(NA, 1), nrow=300000, ncol=3)
bias.sum <- matrix(rep(NA, 1), nrow=300000, ncol=1)

## store values of 100,000 bias factors
bias.sum[1: 100000,1] <- "all"
bias.sum[100001: 200000,1] <- "cvd"
bias.sum[200001: 300000,1] <- "strk"

## store values of 100,000 smoke-adjusted RR
bnew[1: 100000,1] <- "all"
bnew[100001: 200000,1] <- "cvd"
bnew[200001: 300000,1] <- "strk"

## for three outcomes (all-cardio, chd, and stroke, respectively
##
for (j in 1:3) {
if (j==1) {
        whichrr <- 1
} else if (j==2) {
        whichrr <- 3
} else if (j==3) {
        whichrr <- 5
}

count <- count+1
start.indx <- (count-1)*100000 + 1

## repeat sampling prevalence, RR_smoke, no2, and calculate bias.factor(i) by 100,000 times
#
for (i in 1:100000) {

bias <- matrix(rep(NA, 1), nrow=5, ncol=1)          # for each quintile
#pc <- matrix(rep(NA, 1), nrow=5, ncol=1)           # for each quintile
no2.new <- matrix(rep(NA, 1), nrow=5, ncol=1)       # for each quintile
bias[1,] <- 0                                       # log(bias factor)
#pc[1,] <- 0

## prior normal distribution for log rate ratio (RR) for CURRENT smokers versus never smokers,
from ACS II study
bcur <- rnorm(1, rr.cardio.smk[whichrr,1], sd=rr.cardio.smk[whichrr,2])

## prior normal distribution for log RR for FORMER smokers versus never smokers, from ACS II
study
bform <- rnorm(1, rr.cardio.smk[whichrr+1,1], sd=rr.cardio.smk[whichrr+1,2])

## sample no2 from exposure distribution of each quintile
no2.new[1,] <- runif(n=1, min=no2[1,1], max=no2[1,2])

## observed exposure-disease log RR (per 1ppb increase) and standard error, adjusted for all
available variables
## except for smoking.
bexposure <- rnorm(1, rr.cardio.no2[count,1], rr.cardio.no2[count,2])
```

```
## from q2 to q5
#
for (k in 2:5) {

### Nonexposed
##  means for multivariate normal distribution of logit of proportions of never and current smokers
##  among NONexposed
munexp <- smk.prev[1,]
##  variance-covariance matrix for logits of proportions of never and current smokers, among
##  NONexposed
covnexp <- matrix(c(var.prev[1,1], cov.prev[1,], cov.prev[1,], var.prev[1,2]), ncol=2)
##  draw from a multivariate normal distribution for logits of proportions of never and current
##  smokers among NONexposed
xx <- c(munexp[1,1], munexp[1,2])
xnexp <- mvrnorm(n=1, xx, covnexp)
##  proportion of never smokers in NONexposed
pnev0 <- exp(xnexp[1])/(1+exp(xnexp[1]))
##  proportion of current smokers in NONexposed
pcur0 <- exp(xnexp[2])/(1+exp(xnexp[2]))
##  estimated proportion of former smokers in NONexposed
pform0 <- 1-pnev0-pcur0

### Exposed
##  means for multivariate normal distribution of logit of proportions of never and current smokers
##  among exposed
muexp <- smk.prev[k,]
##  variance-covariance matrix for logits of proportions of never and current smokers, among
##  exposed
covexp <- matrix(c(var.prev[k,1], cov.prev[k,], cov.prev[k,], var.prev[k,2]), ncol=2)
##  draw from a multivariate normal distribution for logits of proportions of never and current
##  smokers among exposed
xx <- c(muexp[1,1], muexp[1,2])
xexp <- mvrnorm(n=1, xx, covexp)
##  proportion of never smokers in exposed
pnev1 <- exp(xexp[1])/(1+exp(xexp[1]))
##  proportion of current smokers in exposed
pcur1 <- exp(xexp[2])/(1+exp(xexp[2]))
##  estimated proportion of former smokers in exposed
pform1 <- 1-pnev1-pcur1

### Calculated log(bias factor)
## log-transformed bias factor and percent change, pc(i)
bias[k,] <-
log((pnev1+exp(bcur)*(pcur1)+exp(bform)*(pform1))/(pnev0+exp(bcur)*(pcur0)+exp(bform)*(pfo
rm0)))
#pc[k-1,] <- (exp(bias[k,])-1)*100

### sample no2 in each quintile
no2.new[k,] <- runif(n=1, min=no2[k,1], max=no2[k,2])

} # end for k (q2-q5)
```

```
## slope of regression line for the association between no2(i) and pc(i) where i is level of quintile
data.reg <- data.frame(bias=bias, no2=no2.new)
slope <- lm(bias~no2, data=data.reg)$coefficients[2]

## indirectly adjust for smoking
bnew[start.indx,3] <- bexposure-slope
start.indx <- start.indx+1
bias.sum[start.indx] <- slope

} # end for i (repeat by 100000 iterations)
} # end for j (all-cardio, chd, stroke)

### calculate mean and 95% CI based on distributions of smoke-adjusted RR
##
## percentiles of smoking-adjusted rate ratio
bnew.out <- data.frame(Mean=rep(NA, 3), LCL=rep(NA, 3), Median=rep(NA, 3), UCL=rep(NA, 3))
bias.out <- data.frame(Mean=rep(NA, 3))

i <- 1

for (i in 1:3) {
start.indx <- (i-1)*100000+1
bnew.out[i,1] <- exp(mean(as.numeric(bnew[start.indx:(start.indx+99999), 3])))^5
bnew.out[i,2] <- exp(quantile(as.numeric(bnew[start.indx:(start.indx+99999), 3]),probs=0.025))^5
bnew.out[i,3] <- exp(quantile(as.numeric(bnew[start.indx:(start.indx+99999), 3]),probs=0.50))^5
bnew.out[i,4] <- exp(quantile(as.numeric(bnew[start.indx:(start.indx+99999), 3]),probs=0.975))^5

## mean bias parameter
#bias.out[i] <- mean(as.numeric(bias[start.indx:start.indx+99999]))

## median bias parameter
#quantile(as.numeric(bias[start.indx:start.indx+99999, 3]),probs=0.50)
}

## smoke-adjusted RR and associated 95% CI
output <- data.frame(rr=paste(round(bnew.out[,1],3), "(" , round(bnew.out[,2],3), "-
",round(bnew.out[,4],3), ")"))
output

## Data storage for output
bias.sum.tor <- bias.sum
hist(exp(as.numeric(bias.sum)), breaks=100, main = "Toronto: distribution of bias factors",
xlab="Bias factor")
bias.sum.ham <- bias.sum
hist(exp(as.numeric(bias.sum)), breaks=100, main = "Hamilton: distribution of bias factors",
xlab="Bias factor")
bias.sum.win <- bias.sum
hist(exp(as.numeric(bias.sum)), breaks=100, main = "Windsor: distribution of bias factors",
xlab="Bias factor")
## Output bias factors for later analysis
bias.out <- data.frame(bias_tor=exp(as.numeric(bias.sum.tor)),
```

```
        bias_ham=exp(as.numeric(bias.sum.ham)),
        bias_win=exp(as.numeric(bias.sum.win)))
library(foreign)
directory <- paste("E:/confounding", ".dbf", sep = "")
write.dbf(bias.out, directory)
```

References

1. Chen H, Goldberg MS, Villeneuve PJ. A Systematic Review of Relation between Long-term Exposure to Ambient Air Pollution and Chronic Diseases. Reviews On Environmental Health 2008;23:243-96.

2. Chen H, Goldberg MS, Crouse DL, et al. Back-extrapolation of estimates of exposure from current land-use regression models. Atmospheric Environment 2010;44:4346-54.

3. Crouse DL, Goldberg MS, Ross NA, et al. Postmenopausal Breast Cancer Is Associated with Exposure to Traffic-Related Air Pollution in Montreal, Canada: A CaseûControl Study. Environmental health perspectives 2010;118:1578.

4. Axelson O. Aspects on confounding in occupational health epidemiology. Scand J Work Environ Health 1978;4:98-102.

5. CORNFIELD J, Haenszel W, HAMMOND EC, et al. Smoking and lung cancer: recent evidence and a discussion of some questions. J Natl Cancer Inst 1959;22:173-203.

6. World Health Organization The global burden of disease: 2004 update. In: World Health Organization, ed. Geneva, Switzerland: WHO Press, 2008:1-160.

7. Statistics Canada. Mortality, Summary List of Causes 2007. Statistics Canada.

8. Manuel DG, Leung M, Nguyen K, et al. Burden of cardiovascular disease in Canada. Canadian Journal of Cardiology 2003;19:997-1004.

9. Filate WA, Johansen HL, Kennedy CC, et al. Regional variations in cardiovascular mortality in Canada. Canadian Journal of Cardiology 2003;19:1241-8.

10. Public Health Agency of Canada. Heart Disease and Stroke in Canada 1997. Public Health Agency of Canada.

11. Dagenais GR, Cantin B, Dagenais F, et al. Importance of outside hospital mortality as a first acute ischemic heart event: the Quebec Cardiovascular Study. The Canadian journal of cardiology 1996;12:914.

12. Wilson PWF, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837.

13. Lakier JB. Smoking and cardiovascular disease. Am J Med 1992;93:8S-12S.

14. Ambrose JA, Barua RS. The pathophysiology of cigarette smoking and cardiovascular disease An update. Journal of the American College of Cardiology 2004;43:1731-7.

15. Blair SN, Kampert JB, Kohl HW, et al. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. JAMA 1996;276:205-10.

16. Franklin SS, Larson MG, Khan SA, et al. Does the Relation of Blood Pressure to Coronary Heart Disease Risk Change With Aging? The Framingham Heart Study. Am Heart Assoc, 2001:1245-9.

17. Kannel WB, Schwartz MJ, McNamara PM. Blood Pressure and Risk of Coronary Heart Disease: The Framingham Study. Chest 1969;56:43.

18. Kannel WB. Blood pressure as a cardiovascular risk factor. JAMA: The Journal of The American Medical Association 1996;275:1571.

19. Larsson SC, Wolk A. Body mass index and risk of multiple myeloma: a meta-analysis. Int J Cancer 2007;121:2512-6.

20. Flegal KM, Graubard BI, Williamson DF, et al. Cause-specific excess deaths associated with underweight, overweight, and obesity. JAMA 2007;298:2028-37.

21. Bjorge T, Engeland A, Tverdal A, et al. Body mass index in adolescence in relation to cause-specific mortality: a follow-up of 230,000 Norwegian adolescents. Am J Epidemiol 2008;168:30-7.

22. Eckel RH, Krauss RM. American Heart Association call to action: obesity as a major risk factor for coronary heart disease. Circulation 1998;97:2099.

23. Grundy SM, Benjamin IJ, Burke GL, et al. Diabetes and cardiovascular disease: a statement for healthcare professionals from the American Heart Association. Circulation 1999;100:1134.

24. Manson JAE, Colditz GA, Stampfer MJ, et al. A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women. Archives of internal medicine 1991;151:1141.

25. Beelen R, Hoek G, Van den Brandt PA, et al. Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study). Environ Health Perspect 2008;116:196-202.

26. Dockery DW, Pope CA, III, Xu X, et al. An association between air pollution and mortality in six U.S. cities. N Engl J Med 1993;329:1753-9.

27. Pope CA3, Burnett RT, Thurston GD, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 2004;109:71-7.

28. Miller KA, Siscovick DS, Sheppard L, et al. Long-term exposure to air pollution and incidence of cardiovascular events in women. N Engl J Med 2007;356:447-58.

29. Lenters V, Uiterwaal CS, Beelen R, et al. Long-term exposure to air pollution and vascular damage in young adults. Epidemiology 2010;21:512-20.

30. Christiani D, Woodin M Urban and Transboundary Air Pollution. In: McCally M, ed. Life Support: the environment and human health. Cambridge: The MIT Press, 2002:15-37.

31. International Agency for Research on Cancer (IARC). International Agency for Research on Cancer (IARC) - Summaries & Evaluations. International Agency for Research on Cancer (IARC).

32. Analitis A, Katsouyanni K, Dimakopoulou K, et al. Short-term effects of ambient particles on cardiovascular and respiratory mortality. Epidemiology 2006;17:230.

33. Burnett RT, Dales R, Krewski D, et al. Associations between ambient particulate sulfate and admissions to Ontario hospitals for cardiac and respiratory diseases. Am J Epidemiol 1995;142:15-22.

34. Burnett RT, Dales RE, Brook JR, et al. Association between ambient carbon monoxide levels and hospitalizations for congestive heart failure in the elderly in 10 Canadian cities. Epidemiology 1997;8:162-7.

35. Burnett RT, Cakmak S, Brook JR, et al. The role of particulate size and chemistry in the association between summertime ambient air pollution and hospitalization for cardiorespiratory diseases. Environ Health Perspect 1997;105:614-20.

36. Burnett RT, Cakmak S, Raizenne ME, et al. The association between ambient carbon monoxide levels and daily mortality in Toronto, Canada. J Air Waste Manag Assoc 1998;48:689-700.

37. Burnett RT, Smith-Doiron M, Stieb D, et al. Effects of particulate and gaseous air pollution on cardiorespiratory hospitalizations. Arch Environ Health 1999;54:130-9.

38. Dominici F, Daniels M, McDermott A, et al. Shape of the Exposure-Response Relation and Mortality Displacement in the NMMAPS Database. Revised Analyses of Time-Series of Air Pollutionand Health.Special Report. Boston, MA: Health Effects Institute, 2003:91-6.

39. Goldberg MS, Burnett RT, Bailar JC, III, et al. The association between daily mortality and ambient air particle pollution in Montreal, Quebec. 2. Cause-specific mortality. Environ Res 2001;86:26-36.

40. Goldberg MS, Burnett RT, Yale JF, et al. Associations between ambient air pollution and daily mortality among persons with diabetes and cardiovascular disease. Environ Res 2006;100:255-67.

41. Katsouyanni K, Touloumi G, Samolu E, et al. Sensitivity analysis of various models of short-term effects of ambient particles on total mortality in 29 cities in APHEA2.

Revised Analyses of Time-Series of Air Pollution and Health.Special Report. Boston, MA: Health Effects Institute, 2003.

42. Le Tertre A, Medina S, Samoli E, et al. Short-term effects of particulate air pollution on cardiovascular diseases in eight European cities. Journal of epidemiology and community health 2002;56:773.

43. Omori T, Fujimoto G, Yoshimura I, et al. Effects of particulate matter on daily mortality in 13 Japanese cities. Journal of epidemiology/Japan Epidemiological Association 2003;13:314.

44. Ostro B, Broadwin R, Green S, et al. Fine particulate air pollution and mortality in nine California counties: results from CALFINE. Environmental health perspectives 2006;114:29.

45. Pope III CA, Muhlestein JB, May HT, et al. Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution. Circulation 2006;114:2443.

46. Wong CM, Vichit-Vadakan N, Kan H, et al. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. Environmental health perspectives 2008;116:1195.

47. Zanobetti A, Schwartz J. The effect of fine and coarse particulate air pollution on mortality: a national analysis. Environmental health perspectives 2009;117:898.

48. Enstrom JE. Fine particulate air pollution and total mortality among elderly Californians, 1973-2002. Inhal Toxicol 2005;17:803-16.

49. Jerrett M, Burnett RT, Ma R, et al. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 2005;16:727-36.

50. Laden F, Schwartz J, Speizer FE, et al. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. Am J Respir Crit Care Med 2006;173:667-72.

51. Lipfert FW, Baty JD, Miller JP, et al. PM2.5 constituents and related air quality variables as predictors of survival in a cohort of U.S. military veterans. Inhal Toxicol 2006;18:645-57.

52. Pope III CA, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of US adults. American journal of respiratory and critical care medicine 1995;151:669.

53. Puett RC, Schwartz J, Hart JE, et al. Chronic particulate exposure, mortality, and coronary heart disease in the nurses' health study. Am J Epidemiol 2008;168:1161-8.

54. Zeger SL, Dominici F, McDermott A, et al. Mortality in the Medicare population and chronic exposure to fine particulate air pollution in urban centers (2000-2005). Environ Health Perspect 2008;116:1614-9.

55. Brook RD, Rajagopalan S, Pope III CA, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 2010;121:2331.

56. Pope CA3, Dockery DW. Health effects of fine particulate air pollution: lines that connect. J Air Waste Manag Assoc 2006;56:709-42.

57. Filleul L, Rondeau V, Vandentorren S, et al. Twenty five year mortality and air pollution: results from the French PAARC survey. Occup Environ Med 2005;62:453-60.

58. Pope CA3, Burnett RT, Thun MJ, et al. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 2002;287:1132-41.

59. Schlesinger RB, Kunzli N, Hidy GM, et al. The health relevance of ambient particulate matter characteristics: coherence of toxicological and epidemiological inferences. Inhalation Toxicology 2006;18:95-125.

60. Beelen R, Hoek G, Fischer P, et al. Estimated long-term outdoor air pollution concentrations in a cohort study. Atmospheric Environment 2007;41:1343-58.

61. HEI Panel on the Health Effects of Traffic-related Air Pollution. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. HEI Special Report 17 2010.

62. Gehring U, Heinrich J, Kramer U, et al. Long-term exposure to ambient air pollution and cardiopulmonary mortality in women. Epidemiology 2006;17:545-51.

63. Hoek G, Beelen R, de Hoogh K, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. Atmospheric Environment 2008;42:7561-78.

64. Finkelstein MM, Jerrett M, Sears MR. Environmental inequality and circulatory disease mortality gradients. J Epidemiol Community Health 2005;59:481-7.

65. Abbey DE, Nishino N, McDonnell WF, et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med 1999;159:373-82.

66. Kunzli N, Jerrett M, Mack WJ, et al. Ambient air pollution and atherosclerosis in Los Angeles. Environ Health Perspect 2005;113:201-6.

67. Environment Canada. National Air Pollution Surveillance Network (NAPS). http://www.ec.gc.ca/natchem/default asp?lang=en&n=EE0E2169-1.

68. Jerrett M, Arain A, Kanaroglou P, et al. A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Science and Environmental Epidemiology 2004;15:185-204.

69. Esmen NA, Marsh GM. Applications and limitations of air dispersion modeling in environmental epidemiology. J Expo Anal Environ Epidemiol 1996;6:339.

70. Briggs D, Collins S, Elliott P, et al. Mapping urban air pollution GIS: a regression-based approach. International Journal of Geographic Information Science 1997;11:699-718.

71. Crouse DL, Goldberg MS, Ross NA. A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in Montreal, Canada. Atmospheric Environment 2009;43:5075-84.

72. Hochadel M, Heinrich J, Gehring U, et al. Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information. Atmospheric Environment 2006;40:542-53.

73. Jerrett M, Arain MA, Kanaroglou P, et al. Modeling the intraurban variability of ambient traffic pollution in Toronto, Canada. J Toxicol Environ Health A 2007;70:200-12.

74. Sahsuvaroglu T, Arain A, Kanaroglou P, et al. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. J Air Waste Manag Assoc 2006;56:1059-69.

75. Wheeler AJ, Smith-Doiron M, Xu X, et al. Intra-urban variability of air pollution in Windsor, Ontario--measurement and modeling for human exposure assessment. Environ Res 2008;106:7-16.

76. Gonzales M, Qualls C, Hudgens E, et al. Characterization of a spatial gradient of nitrogen dioxide across a United States-Mexico border city during winter. Science of the total environment 2005;337:163-73.

77. Jerrett M, Finkelstein MM, Brook JR, et al. A cohort study of traffic-related air pollution and mortality in Toronto, Ontario, Canada. Environmental health perspectives 2009;117:772.

78. Von Klot S, Gryparis A, Tonne C, et al. Elemental carbon exposure at residence and survival after acute myocardial infarction. Epidemiology 2009;20:547.

79. Brauer M, Hoek G, Smit HA, et al. Air pollution and development of asthma, allergy and infections in a birth cohort. European Respiratory Journal 2007;29:879.

80. Gan WQ, Koehoorn M, Davies H, et al. Long-term Exposure to Traffic-related Air Pollution and the Risk of Coronary Heart Disease Hospitalization and Mortality. Environmental health perspectives 2011;119:501-7.

81. Tonne C, Yanosky J, Gryparis A, et al. Traffic particles and occurrence of acute myocardial infarction: a case-control analysis. Occup Environ Med 2009;66:797-804.

82. Chen LH, Knutsen SF, Shavlik D, et al. The association between fatal coronary heart disease and ambient particulate air pollution: Are females at greater risk? Environmental health perspectives 2005;113:1723.

83. Hoek G, Brunekreef B, Goldbohm S, et al. Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. Lancet Vol 360(9341)()(pp 1203-1209), 2002 Date of Publication: 19 OCT 2002 2002;1203-9.

84. Naess O, Nafstad P, Aamodt G, et al. Relation between concentration of air pollution and cause-specific mortality: four-year exposures to nitrogen dioxide and particulate matter pollutants in 470 neighborhoods in Oslo, Norway. Am J Epidemiol 2007;165:435-43.

85. Nafstad P, Haheim LL, Wisloff T, et al. Urban air pollution and mortality in a cohort of Norwegian men. Environ Health Perspect 2004;112:610-5.

86. Ostro B, Broadwin R, Green S, et al. Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study. Environmental health perspectives 2010;118:363-9.

87. Rosenlund M, Berglind N, Pershagen G, et al. Long-term exposure to urban air pollution and myocardial infarction. Epidemiology 2006;17:383.

88. Rosenlund M, Bellander T, Nordquist T, et al. Traffic-generated air pollution and myocardial infarction. Epidemiology 2009;20:265.

89. Brunekreef B, Holgate ST. Air pollution and health. The Lancet 2002;360:1233-42.

90. Brauer M, Hoek G, van Vliet P, et al. Air pollution from traffic and the development of respiratory infections and asthmatic and allergic symptoms in children. American journal of respiratory and critical care medicine 2002;166:1092.

91. McConnell R, Berhane K, Yao L, et al. Traffic, susceptibility, and childhood asthma. Environmental health perspectives 2006;114:766.

92. Labreche F, Goldberg MS, Valois MF, et al. Occupational exposures to extremely low frequency magnetic fields and postmenopausal breast cancer. Am J Ind Med 2003;44:643-52.

93. Lenz SK, Goldberg MS, Labreche F, et al. Association between alcohol consumption and postmenopausal breast cancer: results of a case-control study in Montreal, Quebec, Canada. Cancer Causes Control 2002;13:701-10.

94. LOGAN WP. Mortality in the London fog incident, 1952. Lancet 1953;1:336-8.

95. BOYD JT. Climate, air pollution, and mortality. Br J Prev Soc Med 1960;14:123-35.

96. Shandilya KK, Khare M, Gupta AB. Suspended particulate matter distribution in rural-industrial Satna and in urban-industrial South Delhi. Environ Monit Assess 2007;128:431-45.

97. Romieu I, Meneses F, Sienra-Monge JJ, et al. Effects of urban air pollutants on emergency visits for childhood asthma in Mexico City. Am J Epidemiol 1995;141:546-53.

98. Wang X, Ding H, Ryan L, et al. Association between air pollution and low birth weight: a community-based study. Environ Health Perspect 1997;105:514-20.

99. Fleischmann M, McKitrick R Ambient Air Level Trends for Pollutants in Large Canadian Cities . University of Guelph, 2003:1-32.

100. Christiani D, Woodin M Urban and Transboundary Air Pollution. In: McCally M, ed. Life Support: the environment and human health. Cambridge: The MIT Press, 2002:15-37.

101. Brook RD, Franklin B, Cascio W, et al. Air pollution and cardiovascular disease: a statement for healthcare professionals from the Expert Panel on Population and Prevention Science of the American Heart Association. Circulation 2004;109:2655-71.

102. Cohen AJ. Outdoor air pollution and lung cancer. Environ Health Perspect 2000;108 Suppl 4:743-50.

103. Vineis P, Forastiere F, Hoek G, et al. Outdoor air pollution and lung cancer: recent epidemiologic evidence. Int J Cancer 2004;111:647-52.

104. Brunekreef B. Health effects of air pollution observed in cohort studies in Europe. J Expo Sci Environ Epidemiol 2007;17 Suppl 2:S61-S65.

105. Wacholder S, McLaughlin JK, Silverman DT, et al. Selection of controls in case-control studies. I. Principles. Am J Epidemiol 1992;135:1019-28.

106. Pettiti D. Meta Analysis, Decision Analysis and Cost Effectiveness Analysis. New York: Oxford University Press, 1999,

107. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.

108. Fletcher J. What is heterogeneity and is it important? Bmj 2007;334:94-6.

109. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. Bmj 2003;327:557-60.

110. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088-101.

111. Egger M, Davey SG, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. Bmj 1997;315:629-34.

112. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000;53:1119-29.

113. World Health Organization *Manual of the international statistical classification of diseases, injuries, and causes of death*. Geneva: World Health Organization, 1977.

114. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.

115. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Stat Med 1999;18:2693-708.

116. Abbey DE, Colome SD, Mills PK, et al. Chronic disease associated with long-term concentrations of nitrogen dioxide. J Expo Anal Environ Epidemiol 1993;3:181-202.

117. Abbey DE, Nishino N, McDonnell WF, et al. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. Am J Respir Crit Care Med 1999;159:373-82.

118. Finkelstein MM, Jerrett M, DeLuca P, et al. Relation between income, air pollution and mortality: a cohort study. CMAJ 2003;169:397-402.

119. Krewski D, Burnett RT, Goldberg M, et al. Reanalysis of the Harvard Six Cities Study, part I: validation and replication. Inhal Toxicol 2005;17:335-42.

120. McDonnell WF, Nishino-Ishikawa N, Petersen FF, et al. Relationships of mortality with the fine and coarse fractions of long-term ambient PM10 concentrations in nonsmokers. J Expo Anal Environ Epidemiol 2000;10:427-36.

121. Morris SC, Shapiro MA, Waller JH. Adult mortality in two communities with widely different air pollution levels. Arch Environ Health 1976;31:248-54.

122. Beeson WL, Abbey DE, Knutsen SF. Long-term concentrations of ambient air pollutants and incident lung cancer in California adults: results from the AHSMOG study.Adventist Health Study on Smog. Environ Health Perspect 1998;106:813-23.

123. Buell P, Dunn JE, Jr., Breslow L. Cancer of the lung and Los-Angeles-type air pollution. Prospective study. Cancer 1967;20:2139-47.

124. Engholm G, Palmgren F, Lynge E. Lung cancer, smoking, and environment: a cohort study of the Danish population. Bmj 1996;312:1259-63.

125. Mills PK, Abbey D, Beeson WL, et al. Ambient air pollution and cancer in California Seventh-day Adventists. Arch Environ Health 1991;46:271-80.

126. Nafstad P, Haheim LL, Oftedal B, et al. Lung cancer and air pollution: a 27 year follow up of 16 209 Norwegian men. Thorax 2003;58:1071-6.

127. Vineis P, Hoek G, Krzyzanowski M, et al. Air pollution and risk of lung cancer in a prospective study in Europe. Int J Cancer 2006;119:169-74.

128. STOCKS P, CAMPBELL JM. Lung cancer death rates among non-smokers and pipe and cigarette smokers; an evaluation in relation to air pollution by benzpyrene and other substances. Br Med J 1955;2:923-9.

129. Vena JE. Air pollution as a risk factor in lung cancer. Am J Epidemiol 1982;116:42-56.

130. Xu ZY, Blot WJ, Xiao HP, et al. Smoking, air pollution, and the high rates of lung cancer in Shenyang, China. J Natl Cancer Inst 1989;81:1800-6.

131. Jedrychowski W, Becher H, Wahrendorf J, et al. A case-control study of lung cancer with special reference to the effect of air pollution in Poland. J Epidemiol Community Health 1990;44:114-20.

132. Holowaty EJ, Risch HA, Miller AB, et al. Lung cancer in women in the Niagara Region, Ontario: a case-control study. Can J Public Health 1991;82:304-9.

133. Katsouyanni K, Trichopoulos D, Kalandidi A, et al. A case-control study of air pollution and tobacco smoking in lung cancer among women in Athens. Prev Med 1991;20:271-8.

134. Jockel KH, Ahrens W, Wichmann HE, et al. Occupational and environmental hazards associated with lung cancer. Int J Epidemiol 1992;21:202-13.

135. Barbone F, Bovenzi M, Cavallieri F, et al. Air pollution and lung cancer in Trieste, Italy. Am J Epidemiol 1995;141:1161-9.

136. Pawlega J, Rachtan J, Dyba T. Evaluation of certain risk factors for lung cancer in Cracow (Poland): A case-control study. Acta Oncologica Vol 36(5)()(pp 471-476), 1997 1997;471-6.

137. Nyberg F, Gustavsson P, Jarup L, et al. Urban air pollution and lung cancer in Stockholm. Epidemiology 2000;11:487-95.

138. Gupta D, Boffetta P, Gaborieau V, et al. Risk factors of lung cancer in Chandigarh, India. Indian Journal of Medical Research Vol 113(APR)()(pp 142-150), 2001 2001;142-50.

139. Besso A, Nyberg F, Pershagen G. Air pollution and lung cancer mortality in the vicinity of a nonferrous metal smelter in Sweden. International Journal of Cancer Vol 107(3)()(pp 448-452), 2003 Date of Publication: 10 NOV 2003 2003;448-52.

140. Chiu HF, Cheng MH, Tsai SS, et al. Outdoor air pollution and female lung cancer in Taiwan. Inhal Toxicol 2006;18:1025-31.

141. Nie J, Beyea J, Bonner MR, et al. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. Cancer Causes Control 2007;18:947-55.

142. Bonner MR, Han D, Nie J, et al. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. Cancer Epidemiol Biomarkers Prev 2005;14:53-60.

143. Lewis-Michl EL, Melius JM, Kallenbach LR, et al. Breast cancer risk and residence near industry or traffic in Nassau and Suffolk Counties, Long Island, New York. Arch Environ Health 1996;51:255-65.

144. Tonne C, Melly S, Mittleman M, et al. A case-control analysis of exposure to traffic and acute myocardial infarction. Environ Health Perspect 2007;115:53-7.

145. Grazuleviciene R, Maroziene L, Dulskiene V, et al. Exposure to urban nitrogen dioxide pollution and the risk of myocardial infarction. Scand J Work Environ Health 2004;30:293-8.

146. Modig L, Jarvholm B, Ronnmark E, et al. Vehicle exhaust exposure in an incident case-control study of adult asthma. Eur Respir J 2006;28:75-81.

147. Tzonou A, Maragoudakis G, Trichopoulos D, et al. Urban living, tobacco smoking, and chronic obstructive pulmonary disease: a study in Athens. Epidemiology 1992;3:57-60.

148. Detels R, Tashkin DP, Sayre JW, et al. The UCLA population studies of chronic obstructive respiratory disease. 9. Lung function changes associated with chronic exposure to photochemical oxidants; a cohort study among never-smokers. Chest 1987;92:594-603.

149. Abbey DE, Petersen FF, Mills PK, et al. Chronic respiratory disease associated with long-term ambient concentrations of sulfates and other air pollutants. J Expo Anal Environ Epidemiol 1993;3 Suppl 1:99-115.

150. Greer JR, Abbey DE, Burchette RJ. Asthma related to occupational and ambient air pollutants in nonsmokers. J Occup Med 1993;35:909-15.

151. Abbey DE, Hwang BL, Burchette RJ, et al. Estimated long-term ambient concentrations of PM10 and development of respiratory symptoms in a nonsmoking population. Arch Environ Health 1995;50:139-52.

152. Abbey DE, Ostro BE, Petersen F, et al. Chronic respiratory symptoms associated with estimated long-term ambient concentrations of fine particulates less than 2.5 microns in aerodynamic diameter (PM2.5) and other air pollutants. J Expo Anal Environ Epidemiol 1995;5:137-59.

153. McDonnell WF, Abbey DE, Nishino N, et al. Long-term ambient ozone concentration and the incidence of asthma in nonsmoking adults: the AHSMOG Study. Environ Res 1999;80:110-21.

154. Goldberg MS. On the interpretation of epidemiological studies of ambient air pollution. Journal of Exposure Science and Environmental Epidemiology 2007;17:S66-S70.

155. Gilbert NL, Goldberg MS, Beckerman B, et al. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. J Air Waste Manag Assoc 2005;55:1059-63.

156. Lloyd-Jones DM, Martin DO, Larson MG, et al. Accuracy of death certificates for coding coronary heart disease as the cause of death. Ann Intern Med 1998;129:1020-6.

157. Coady SA, Sorlie PD, Cooper LS, et al. Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. J Clin Epidemiol 2001;54:40-50.

158. Doria-Rose VP, Marcus PM. Death certificates provide an adequate source of cause of death information when evaluating lung cancer mortality: An example from the Mayo Lung Project. Lung Cancer 2008.

159. Ghio AJ, Kim C, Devlin RB. Concentrated ambient air particles induce mild pulmonary inflammation in healthy human volunteers. Am J Respir Crit Care Med 2000;162:981-8.

160. Salvi S, Blomberg A, Rudell B, et al. Acute inflammatory responses in the airways and peripheral blood after short-term exposure to diesel exhaust in healthy human volunteers. Am J Respir Crit Care Med 1999;159:702-9.

161. Gurgueira SA, Lawrence J, Coull B, et al. Rapid increases in the steady-state concentration of reactive oxygen species in the lungs and heart after particulate air pollution inhalation. Environ Health Perspect 2002;110:749-55.

162. Sorensen M, Daneshvar B, Hansen M, et al. Personal PM2.5 exposure and markers of oxidative stress in blood. Environ Health Perspect 2003;111:161-6.

163. Tan WC, Qiu D, Liam BL, et al. The human bone marrow response to acute air pollution caused by forest fires. Am J Respir Crit Care Med 2000;161:1213-7.

164. Weiss ST, Segal MR, Sparrow D, et al. Relation of FEV1 and peripheral blood leukocyte count to total mortality. The Normative Aging Study. Am J Epidemiol 1995;142:493-8.

165. van Eeden SF, Tan WC, Suwa T, et al. Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM(10)). Am J Respir Crit Care Med 2001;164:826-30.

166. Schols AM, Buurman WA, Staal van den Brekel AJ, et al. Evidence for a relation between metabolic derangements and increased levels of inflammatory mediators in a subgroup of patients with chronic obstructive pulmonary disease. Thorax 1996;51:819-24.

167. Ridker PM. Evaluating novel cardiovascular risk factors: can we better predict heart attacks? Ann Intern Med 1999;130:933-7.

168. Seaton A, MacNee W, Donaldson K, et al. Particulate air pollution and acute health effects. Lancet 1995;345:176-8.

169. van Eeden SF, Yeung A, Quinlam K, et al. Systemic response to ambient particulate matter: relevance to chronic obstructive pulmonary disease. Proc Am Thorac Soc 2005;2:61-7.

170. Sun Q, Wang A, Jin X, et al. Long-term air pollution exposure and acceleration of atherosclerosis and vascular inflammation in an animal model. JAMA 2005;294:3003-10.

171. Kunzli N, Kaiser R, Medina S, et al. Public-health impact of outdoor and traffic-related air pollution: a European assessment. Lancet 2000;356:795-801.

172. Sin DD, Man SF. Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. Proc Am Thorac Soc 2005;2:8-11.

173. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation 1996;93:1043-65.

174. Nolan J, Batin PD, Andrews R, et al. Prospective study of heart rate variability and mortality in chronic heart failure: results of the United Kingdom heart failure evaluation and assessment of risk trial (UK-heart). Circulation 1998;98:1510-6.

175. Vineis P, Husgafvel-Pursiainen K. Air pollution and cancer: biomarker studies in human populations. Carcinogenesis 2005;26:1846-55.

176. Brauer M, Hruba F, Mihalikova E, et al. Personal exposure to particles in Banska Bystrica, Slovakia. J Expo Anal Environ Epidemiol 2000;10:478-87.

177. Peluso M, Munnia A, Hoek G, et al. DNA adducts and lung cancer risk: a prospective study. Cancer Res 2005;65:8042-8.

178. Palli D, Russo A, Masala G, et al. DNA adduct levels and DNA repair polymorphisms in traffic-exposed workers and a general population sample. Int J Cancer 2001;94:121-7.

179. Mauderly JL, Snipes MB, Barr EB, et al. Pulmonary toxicity of inhaled diesel exhaust and carbon black in chronically exposed rats. Part I: Neoplastic and nonneoplastic lung lesions. Res Rep Health Eff Inst 1994;1-75.

180. Castano-Vinyals G, D'Errico A, Malats N, et al. Biomarkers of exposure to polycyclic aromatic hydrocarbons from environmental air pollution. Occup Environ Med 2004;61:e12.

181. Ravindra, Mittal AK, Van GR. Health risk assessment of urban suspended particulate matter with special reference to polycyclic aromatic hydrocarbons: a review. Rev Environ Health 2001;16:169-89.

182. Nagashima M, Kasai H, Yokota J, et al. Formation of an oxidative DNA damage, 8-hydroxydeoxyguanosine, in mouse lung DNA after intratracheal instillation of diesel exhaust particles and effects of high dietary fat and beta-carotene on this process. Carcinogenesis 1995;16:1441-5.

183. Risom L, Moller P, Loft S. Oxidative stress-induced DNA damage by particulate air pollution. Mutat Res 2005;592:119-37.

184. Ciccone G, Faggiano F, Falasca P. [SO2 air pollution and hospital admissions in Ravenna: a case-control study]. Epidemiol Prev 1995;19:99-104.

185. Ferrante M, Ferlito V, Scarlata S, et al. [Benzene concentration in outdoor air: estimation of leukemia risk for the population of the city of Catania]. Ann Ig 2002;14:73-9.

186. Zaridze DG, Zemlianaia GM, Aitakov ZN. [Role of outdoor and indoor air pollution in the etiology of lung cancer]. Vestn Ross Akad Med Nauk 1995;6-10.

187. Hart JE, Garshick E, Dockery DW, et al. Long-term ambient multipollutant exposures and mortality. Am J Respir Crit Care Med 2011;183:73-8.

188. Jerrett M, Burnett RT, Pope III CA, et al. Long-term ozone exposure and mortality. New England Journal of Medicine 2009;360:1085-95.

189. Puett RC, Hart JE, Yanosky JD, et al. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. Environ Health Perspect 2009;117:1697-701.

190. Rosenlund M, Picciotto S, Forastiere F, et al. Traffic-related air pollution in relation to incidence and prognosis of coronary heart disease. Epidemiology 2008;19:121-8.

191. Yorifuji T, Kashima S, Tsuda T, et al. Long-term exposure to traffic-related air pollution and mortality in Shizuoka, Japan. Occup Environ Med 2010;67:111-7.

192. Oudin A, Stroh E, Stromberg U, et al. Long-term exposure to air pollution and hospital admissions for ischemic stroke. A register-based case-control study using modelled NO(x) as exposure proxy. BMC Public Health 2009;9:301.

193. Brook JR, Burnett RT, Dann TF, et al. Further interpretation of the acute effects of nitrogen dioxide observated in Canadian time series studies. Journal of Exposure Science and Environmental Epidemiology 2007;17(suppl):36-44.

194. Ichinose T, Sagai M. Studies on biochemical effects of nitrogen dioxide:: III. Changes of the antioxidative protective systems in rat lungs and of lipid peroxidation by chronic exposure. Toxicology and applied pharmacology 1982;66:1-8.

195. Sagai M, Ichinose T, Kubota K. Studies on the biochemical effects of nitrogen dioxide:: IV. Relation between the change of lipid peroxidation and the antioxidative protective system in rat lungs upon life span exposure to low levels of NO2. Toxicology and applied pharmacology 1984;73:444-56.

196. Rombout PJA, Dormans J, Marra M, et al. Influence of exposure regimen on nitrogen dioxide-induced morphological changes in the rat lung. Environmental research 1986;41:466-80.

197. Rasmussen RE, Mannix RC, Oldham MJ, et al. Effects of nitrogen dioxide on respiratory tract clearance in the ferret. Journal of Toxicology and Environmental Health;(United States) 1994;41.

198. Hussain I, Jain VV, OÆShaughnessy P, et al. Effect of Nitrogen Dioxide Exposure on Allergic Asthma in a Murine Model. Chest 2004;126:198.

199. Emenius G, Pershagen G, Berglind N, et al. NO2, as a marker of air pollution, and recurrent wheezing in children: a nested case-control study within the BAMSE birth cohort. Occupational and environmental medicine 2003;60:876.

200. Jacquemin B, Sunyer J, Forsberg B, et al. Association between modelled traffic-related air pollution and asthma score in the ECRHS. European Respiratory Journal 2009;34:834.

201. Pattenden S, Hoek G, Braun-FahrlΣnder C, et al. NO2 and childrenÆs respiratory symptoms in the PATY study. Occupational and environmental medicine 2006;63:828.

202. Forastiere F, Peters A, Kelly FJ, et al. Nitrogen Dioxide. Air Quality Guidelines: Global Updates 2005. http://www.who.int/phe/health_topics/outdoorair_aqg/en/.

203. Beckerman B, Jerrett M, Brook JR, et al. Correlation of nitrogen dioxide with other traffic pollutants near a major expressway. Atmospheric Environment 2008;42:275-90.

204. Brauer M, Hoek G, van Vliet P, et al. Estimating Long-Term Average Particulate Air Pollution Concentrations: Application of Traffic Indicators and Geographic Information Systems. Epidemiology 2003;14:228.

205. Henderson SB, Beckerman B, Jerrett M, et al. Application of land use regression to estimate long-term concentrations of traffic-related nitrogen oxides and fine particulate matter. Environ Sci Technol 2007;41:2422-8.

206. Madsen C, Carlsen KCL, Hoek G, et al. Modeling the intra-urban variability of outdoor traffic pollution in Oslo, NorwayùA GA2LEN project. Atmospheric Environment 2007;41:7500-11.

207. Moore DK, Jerrett M, Mack WJ, et al. A land use regression model for predicting ambient fine particulate matter across Los Angeles, CA. J Environ Monit 2007;9:246-52.

208. Ross Z, Jerrett M, Ito K, et al. A land use regression for predicting fine particulate matter concentrations in the New York City region. Atmospheric Environment 2007;41:2255-69.

209. Suglia SF, Gryparis A, Wright RO, et al. Association of black carbon with cognition among children in a prospective birth cohort study. Am J Epidemiol 2008;167:280-6.

210. Gehring U, Wijga AH, Brauer M, et al. Traffic-Related Air Pollution and the Development of Asthma and Allergies During The First 8 Years of Life. Am J Respir Crit Care Med 2009.

211. Brunekreef B, Beelen R, Hoek G, et al. Effects of long-term exposure to traffic-related air pollution on respiratory and cardiovascular mortality in the Netherlands: the NLCS-AIR study. Res Rep Health Eff Inst 2009;5-71.

212. Brauer M, Lencar C, Tamburic L, et al. A cohort study of traffic-related air pollution impacts on birth outcomes. Environ Health Perspect 2008;116:680-6.

213. Ballester F, Estarlich M, I±iguez C, et al. Air pollution exposure during pregnancy and reduced birth size: a prospective birth cohort study in Valencia, Spain. Environmental Health 2010;9:6.

214. Labreche F, Goldberg MS, Valois MF, et al. Postmenopausal breast cancer and occupational exposures. Occupational and environmental medicine 2010;67:263.

215. Briggs DJ, de HC, Gulliver J, et al. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. Sci Total Environ 2000;253:151-67.

216. Cao J, Valois MF, Goldberg MS. An S-Plus function to calculate relative risks and adjusted means for regression models using natural splines. Comput Methods Programs Biomed 2006;84:58-62.

217. Casella G, Berger RL, Berger RL. Statistical inference. Duxbury Pacific Grove, CA, 2002,

218. Waller LA, Gotway CA. Applied Spatial Statistics for Public Health Data. Hoboken, New Jersey: John Wiley & Sons, Inc., 2007,

219. Eeftens M, Beelen R, Fischer P, et al. Stability of measured and modelled spatial contrasts in NO2 over time. Occupational and environmental medicine 2011.

220. Clougherty JE, Levy JI, Kubzansky LD, et al. Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology. Environ Health Perspect 2007;115:1140-6.

221. Hart JE, Yanosky JD, Puett RC, et al. Spatial modeling of PM10 and NO2 in the continental United States, 1985-2000. Environ Health Perspect 2009;117:1690-6.

222. Yanosky JD, Paciorek CJ, Schwartz J, et al. Spatio-temporal modeling of chronic PM10 exposure for the Nurses' Health Study. Atmos Environ 2008;42:4047-62.

223. Yanosky JD, Paciorek CJ, Suh HH. Predicting chronic fine and coarse particulate exposures using spatiotemporal models for the Northeastern and Midwestern United States. Environ Health Perspect 2009;117:522-9.

224. Goldstein H. Multilevel statistical models. Wiley, 2003,

225. Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. American Journal of Epidemiology 2004;160:384.

226. Waller LA, Gotway CA. Applied Spatial Statistics for Public Health Data. Hoboken, New Jersey: John Wiley & Sons, Inc., 2007,

227. Stedman JR, Vincent KJ, Campbell GW, et al. New high resolution maps of estimated background ambient NOx and NO2 concentrations in the UK. Atmospheric Environment 1997;31:3591-602.

228. Statistics Canada. 2006 Census of Canada. Statistics Canada.

229. Environment Canada. National Air Pollution Surveillance Network Data. Environment Canada.

230. Statistics Canada. Annual estimates for census families and individuals (T1 famile file). Statistics Canada.

231. Frenette M, Green D, Picot G Rising income inequality in the 1990s: An exploration of three data sources. In: Analytical Studies Branch, ed. Ottawa: Statistics Canada, 2004.

232. Spasoff RA, Gilkes DT. Up-to-date denominators: evaluation of taxation family file for public health planning. Can J Public Health 1994;85:413-7.

233. Statistics Canada. Vital Statistics - Death Database. Statistics Canada.

234. Goldberg MS, Carpenter M, Theriault G, et al. The accuracy of ascertaining vital status in a historical cohort study of synthetic textiles workers using computerized record linkage to the Canadian Mortality Data Base. Can J Public Health 1993;84:201-4.

235. Guibert R. Could the coronary heart disease mortality rates decline be artefactual? Can J Public Health 1991;82:43-5.

236. Guibert RL, Wigle DT, Williams JI. Decline of acute myocardial infarction death rates not due to cause of death coding. Can J Public Health 1989;80:418-22.

237. Engel LW, Strauchen JA, Chiazze L, Jr., et al. Accuracy of death certification in an autopsied population with specific attention to malignant neoplasms and vascular diseases. Am J Epidemiol 1980;111:99-112.

238. Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. N Engl J Med 1985;313:1263-9.

239. Lloyd-Jones DM, Martin DO, Larson MG, et al. Accuracy of death certificates for coding coronary heart disease as the cause of death. Ann Intern Med 1998;129:1020-6.

240. Smith GD, Hart C, Watt G, et al. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. British Medical Journal 1998;52:399.

241. Smith GD, Whitley E, Dorling D, et al. Area based measures of social and economic circumstances: cause specific mortality patterns depend on the choice of index. British Medical Journal 2001;55:149.

242. Diez-Roux AV. Investigating neighborhood and area effects on health. American journal of public health 2001;91:1783.

243. Vescio MF, Brookes ST, Sterne J, et al. Mortality at ages 50û59 and deprivation at early and late stages of the life course in Wales. British Medical Journal 2009;63:56.

244. Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. American Journal of Epidemiology 1997;146:48.

245. Diez-Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. New England Journal of Medicine 2001;345:99.

246. Sundquist K, Malmstrom M, Johansson SE. Neighbourhood deprivation and incidence of coronary heart disease: a multilevel study of 2.6 million women and men in Sweden. British Medical Journal 2004;58:71.

247. Cubbin C, Winkleby MA. Protective and harmful effects of neighborhood-level deprivation on individual-level health knowledge, behavior changes, and risk of coronary heart disease. American Journal of Epidemiology 2005;162:559.

248. Stjarne MK, Fritzell J, De Leon AP, et al. Neighborhood socioeconomic context, individual income and myocardial infarction. Epidemiology 2006;17:14.

249. Roux D, Ana V. Investigating neighborhood and area effects on health. American journal of public health 2001;91:1783.

250. Ross NA, Tremblay S, Graham K. Neighbourhood influences on health in Montreal, Canada. Social Science & Medicine 2004;59:1485-94.

251. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society Series B (Methodological) 1972;34:187-220.

252. Goldberg MS, Carpenter M, Theriault G, et al. The accuracy of ascertaining vital status in a historical cohort study of synthetic textiles workers using computerized record linkage to the Canadian Mortality Data Base. Canadian journal of public health Revue canadienne de santθ publique84:201.

253. Hedges LV, Olkin I. Statistical methods for meta-analysis. Academic Press New York, 1985,

254. Kleinbaum DG. Survival Analysis: A self-learning text. New York: Springer-Verlag, 1996,

255. Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Am J Epidemiol 2004;160:384-92.

256. Statistics Canada. Canadian Community Health Survey (CCHS). Statistics Canada.

257. Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. JAMA 2000;284:706.

258. Statistics Canada. Statistics Canada: CANSIM table. Statistics Canada.

259. Gan WQ, Tamburic L, Davies HW, et al. Changes in Residential Proximity to Road Traffic and the Risk of Death From Coronary Heart Disease. Epidemiology 2010;21:642.

260. DesMeules M, Gold J, McDermott S, et al. Disparities in mortality patterns among Canadian immigrants and refugees, 1980û1998: results of a national cohort study. Journal of Immigrant Health 2005;7:221-32.

261. Newbold B. Self-rated health within the Canadian immigrant population: risk and the healthy immigrant effect. Social Science & Medicine 2005;60:1359-70.

262. Chen J, Ng E, Wilkins R. The health of Canada's immigrants in 1994-95. Health Reports 1996;7:33-45.

263. Leech JA, Nelson WC, Burnett RT, et al. It's about time: a comparison of Canadian and American time-activity patterns. J Expo Anal Environ Epidemiol 2002;12:427-32.

264. Blondeau P, Iordache V, Poupard O, et al. Relationship between outdoor and indoor air quality in eight French schools. Indoor Air 2005;15:2-12.

265. Crist KC, Liu B, Kim M, et al. Characterization of fine particulate matter in Ohio: Indoor, outdoor, and personal exposures. Environmental research 2008;106:62-71.

266. Blair SN, Kampert JB, Kohl HW, et al. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. JAMA: The Journal of The American Medical Association 1996;276:205.

267. Lakier JB. Smoking and cardiovascular disease. The American journal of medicine 1992;93:S8-S12.

268. Barrett-Connor E, Khaw K. Family history of heart attack as an independent predictor of death due to cardiovascular disease. Circulation 1984;69:1065.

269. Schubauer-Berigan MK, Couch JR, Petersen MR, et al. Cohort mortality study of workers at seven beryllium processing plants: update and associations with cumulative and maximum exposure. Occupational and environmental medicine 2010.

270. Perreault S, Dragomir A, Blais L, et al. Impact of adherence to statins on chronic heart failure in primary prevention. British journal of clinical pharmacology 2008;66:706-16.

271. Goedert JJ, Schairer C, McNeel TS, et al. Risk of breast, ovary, and uterine corpus cancers among 85 268 women with AIDS. British journal of cancer 2006;95:642-8.

272. Kettani FZ, Dragomir A, Cote R, et al. Impact of a better adherence to antihypertensive agents on cerebrovascular disease for primary prevention. Stroke 2009;40:213.

273. Engels EA, Brock MV, Chen J, et al. Elevated incidence of lung cancer among HIV-infected individuals. Journal of clinical oncology 2006;24:1383.

274. Din-Dzietham R, Liu Y, Bielo MV, et al. High blood pressure trends in children and adolescents in national surveys, 1963 to 2002. Circulation 2007;116:1488.

275. Hansen ML, Gunn PW, Kaelber DC. Underdiagnosis of hypertension in children and adolescents. JAMA: The Journal of The American Medical Association 2007;298:874.

276. Will JC, Galuska DA, Ford ES, et al. Cigarette smoking and diabetes mellitus: evidence of a positive association from a large prospective cohort study. International journal of epidemiology 2001;30:540.

277. Johnson D, Parker JD. Air pollution exposure and self-reported cardiovascular disease. Environmental research 2009;109:582-9.

278. Van Hee VC, Adar SD, Szpiro AA, et al. Exposure to traffic and left ventricular mass and function: the Multi-Ethnic Study of Atherosclerosis. American journal of respiratory and critical care medicine 2009;179:827.

279. Kelishadi R, Mirghaffari N, Poursafa P, et al. Lifestyle and environmental factors associated with inflammation, oxidative stress and insulin resistance in children. Atherosclerosis 2009;203:311-9.

280. Brook RD, Jerrett M, Brook JR, et al. The relationship between diabetes mellitus and traffic-related air pollution. Journal of Occupational and Environmental Medicine 2008;50:32.

281. Pershagen G. Transport Noise and Incidence of Cardiovascular Disease. Epidemiology 2009;20:S249.

282. Allen RW, Adar SD. Are both air pollution and noise driving adverse cardiovascular health effects from motor vehicles? Environ Res 2011;111:184-5.

283. Selander J, Nilsson ME, Bluhm G, et al. Long-term exposure to road traffic noise and myocardial infarction. Epidemiology 2009;20:272-9.

284. Huss A, Spoerri A, Egger M, et al. Aircraft noise, air pollution, and mortality from myocardial infarction. Epidemiology 2010;21:829-36.

285. Giles LV, Barn P, Knnzli N, et al. From Good Intentions to Proven Interventions: Effectiveness of Actions to Reduce the Health Impacts of Air Pollution. Environmental health perspectives 2011;119:29.

286. Knnzli N, McConnell R, Bates D, et al. Breathless in Los Angeles: the exhausting search for clean air. American journal of public health 2003;93:1494.