On the Associations Between Local Concentrations of Ambient Air Pollution, Neighbourhood-scale Deprivation, and Postmenopausal Breast Cancer in Montreal, Canada

Dan L. Crouse

Department of Geography
McGill University, Montreal

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

©Dan L. Crouse, 2009
Abstract

Urban air pollution is a complex mixture composed of hundreds of different solids, liquids, and gases. Some pollutants, such as nitrogen dioxide, are highly variable at a local scale, such that the variability in concentrations within cities can be greater than the variability between cities. Long-term exposures to air pollution have been associated with many health outcomes, and there is evidence that individuals of lower socioeconomic status may be more vulnerable to the effects of exposure. Associations between exposure to ambient air pollution and the incidence of breast cancer have not been investigated adequately despite findings from occupational studies that have suggested increased risks among women exposed to some contaminants found also in ambient pollution. Only about one-third of new cases of breast cancer are attributable to known risk factors; thus much breast cancer aetiology remains unexplained. The task of assigning exposure to air pollution to individuals is the principal challenge in studying associations between air pollution and health, due in part to the fact that levels of pollution vary over small areas. Thus, this dissertation develops and describes improved methods for modelling local-scale variations in concentrations of ambient air pollution. First, samples of nitrogen dioxide were collected at 129 locations across Montreal, Quebec, Canada on three occasions. A spatial model was developed that described approximately 80% of the variability in concentrations of nitrogen dioxide across Montreal. Next, spatial associations were identified at the neighbourhood scale between patterns of ambient air quality and indicators of social deprivation (e.g., percentage of people living alone) and material deprivation (e.g., low income households). Lastly, this dissertation demonstrates an association between the incidence of postmenopausal breast cancer and long-term exposure to concentrations of intra-urban air pollution. Women in the highest quartile of exposure to estimates of intra-urban air pollution had an increased risk of 2.33 (95% confidence interval 1.29 – 4.21) compared to those in the lowest quartile of exposure. This latter finding is among only a handful of studies that have investigated the possible link between the incidence of breast cancer and exposure to ambient air pollution.
Résumé

La pollution atmosphérique urbaine est un mélange complexe d’une centaine de différents polluants primaires et secondaires. Certains polluants, dont le dioxyde d'azote, sont présents en quantités très variables localement. Ainsi, la variabilité des concentrations à l'échelle intra-urbaine peut être plus prononcée que celle des concentrations à l'échelle inter-urbaine. Une exposition prolongée à la pollution atmosphérique a été associée à plusieurs effets néfastes sur la santé dont la réduction de l'espérance de vie et un taux de mortalité plus élevé chez les personnes atteintes de maladies cardio-pulmonaires et de cancer du poumon. Les liens entre l'exposition à la pollution atmosphérique et l'incidence du cancer du sein n'ont pas été étudiés adéquatement, malgré les conclusions d'études professionnelles qui indiquent une augmentation du risque de développer le cancer du sein lors de l'exposition à certains contaminants présent dans l'air ambiant. Environ seulement un tiers des nouveaux cas de cancer du sein sont imputables aux facteurs de risque connus. L'étiologie du cancer du sein demeure donc inexplicable. Le principal défi d'une étude concernant les liens entre la pollution atmosphérique et la santé est de quantifier l'exposition à la pollution atmosphérique ambiante au niveau de l'individu. Cela est dû en partie à la variabilité des concentrations de polluants à une échelle géographique réduite. Cette dissertation a pour but de (1) développer et décrire une méthodologie améliorée pour modéliser les variations locales des concentrations de pollution atmosphérique à petite échelle; (2) identifier des liens entre la qualité de l'air ambiant et des indicateurs de privation sociale et matérielle à l'échelle local pour la ville de Montréal (Québec, Canada); (3) démontrer le lien entre l'incidence du cancer post-ménopausique du sein et l'exposition prolongée à la pollution atmosphérique intra-urbaine. Les femmes faisant partie du plus haut quartile d'exposition à la pollution atmosphérique intra-urbaine sont plus à risque (2.33, intervalle de confiance à 95% 1.29 – 4.21) que celles faisant partie du quartile le plus bas. Cette dissertation figure parmi peu d'études examinant cette hypothèse.
Acknowledgements

I would not have been able to complete this dissertation without the support, guidance, and assistance of a number of people and organizations to whom I am very grateful.

First, I would like to thank Nancy Ross for her encouragement, vision, and brilliant insights and ideas that contributed so much to the development of this dissertation and to my development as an independent researcher. Thank you for being such a great teacher to me during the past four years.

Second, I’d like to thank Mark Goldberg for his dedication to my research, and to training me to be an independent researcher. I appreciate all the time you spent reading countless drafts, and all your helpful comments – even those in boldface, upper-case, red font – that have helped me to become a better writer. Thank you also for the many stimulating conversations over lunch at the Thomson House.

I’d also like to thank Profs Ian Strachan and Michael Jerrett for their ideas, suggestions, and advice, all of which contributed to improvements in my research. Very special thanks to Ian Haase for contributing his time, good humour, and navigational skills to my data collection activities. Thanks also to Marie-France Valois who helped me through some frustrating times with S-PLUS and to Hong Chen who helped me through some frustrating times with ArcGIS.

I am grateful to the Canadian Institutes of Health Research (CIHR), who provided the funding (Canada Graduate Scholarship) that enabled me to conduct my research, to travel to conferences and meetings to present my work, and to meet others in my field. I am grateful also to the Warren Fellowship, which provided me with additional financial support.

I would like to acknowledge the excellent assistance that I received from the administrative staff in the Geography Department at McGill, in particular
Maria Marcone, Pauline Nesbitt, Joseph Vacirca, and Nancy Secondo.

My experience as a PhD student at McGill would not have been complete without all the times spent not working on my PhD; the times spent outside of Burnside Hall. A very big thank you to my friends Chris, Brian, Tim, Ciara, Julia, Chantelle, Ian, and Jer for all the coffee breaks, trips to Super Sandwiches, and evenings discussing music, recipes, Scotch, the Habs, life in general, and, of course, our respective research projects. Here, I’d like to single-out Pete with an extra-special thank you for all your support and friendship during the past four years: it wouldn’t have been the same without you here.

I’d like to thank my family, in particular my parents and my sister, Melanie, for all their love and encouragement over the past four years. We didn’t get to spend as much time together as we would have liked to have, and I appreciate your understanding while I pursued this goal. I appreciate also the trips you made to Montreal to visit me whenever possible.

Lastly, I am grateful to my wife, Abby, for her support and patience while I completed this degree. Thank you for putting up with my distance and the evenings that I chose to spend with my laptop instead of with you. I couldn’t have made it through these past few years without you, and I wouldn’t have wanted to.
Contributions of Authors

This dissertation is a collection of three related manuscripts that have been, or will be, submitted to peer-reviewed journals for publication. My supervisors, Nancy Ross and Mark Goldberg, helped me conceptualize and set the goals and objectives of the dissertation research. I performed all of the spatial and statistical analyses. I wrote all three manuscripts, but received valuable intellectual and editorial comments from both supervisors, who are included as co-authors on all three. Additionally, Hong Chen, a PhD student in Epidemiology at McGill, contributed some geographic analyses to the methods presented in the third manuscript (which are acknowledged therein).
# Table of Contents

Abstract...........................................................................................................................ii  
Résumé...........................................................................................................................iii  
Acknowledgements.......................................................................................................iv  
Contributions of Authors..............................................................................................vi

**Chapter 1**  
Introduction......................................................................................................................1  
1.1 Problem Statement .................................................................................................1  
1.2 Research Objectives...............................................................................................4  
1.3 Dissertation Outline...............................................................................................5

**Chapter 2**  
Research Context and Framework............................................................................... 8  
2.1 Determinants of Health..........................................................................................8  
2.2 Health Inequalities.................................................................................................11  
2.3 Air Pollution and Health.......................................................................................15  
2.4 Chapter Summary..................................................................................................22

**Chapter 3**  
A Prediction-based Approach to Modelling Temporal and Spatial Variability of Traffic-related Air Pollution.................................................................................23  
3.1 Introduction............................................................................................................24  
3.2 Materials and Methods..........................................................................................27  
3.3 Results ...................................................................................................................33  
3.4 Discussion and Conclusions..................................................................................41

**Chapter 4**  
Double Burden of Social and Material Deprivation and High Concentrations of Air Pollution at the Neighbourhood-scale....................................................................45  
4.1 Introduction............................................................................................................46  
4.2 Context...................................................................................................................51  
4.3 Materials and Methods.........................................................................................55  
4.4 Results...................................................................................................................59  
4.5 Discussion and Conclusions..................................................................................65

**Chapter 5**  
Postmenopausal Breast Cancer is Associated with Exposure to Traffic-related Air Pollution in Montreal, Canada .................................................................................70
5.1 Introduction........................................................................................................... 71
5.2 Breast Cancer and Air Pollution......................................................................... 72
5.3 Materials and Methods....................................................................................... 74
5.4 Results................................................................................................................. 81
5.5 Discussion and Conclusions.............................................................................. 94

Chapter 6
Summary and Conclusions...................................................................................... 99

6.1 Key Findings and Contributions......................................................................... 99
6.2 Limitations of the Dissertation.......................................................................... 101
6.3 Directions for Further Research........................................................................ 103
References.............................................................................................................. 106
CHAPTER 1
INTRODUCTION

The purpose of this dissertation is to contribute to our understanding of how exposure to local concentrations of ambient air pollution influences health outcomes. The research develops and describes improved methods for modelling local-scale seasonal and spatial variations in concentrations of ambient air pollution; identifies spatial associations between patterns of ambient air quality and indicators of social and material deprivation at the neighbourhood scale; and, demonstrates an association between the incidence of postmenopausal breast cancer and long-term exposure to concentrations of intra-urban ambient air pollution. This latter finding is among only a handful of other studies that have investigated the hypothesis that the incidence of breast cancer may be associated with exposure to ambient concentrations of air pollution. This dissertation improves upon these previous studies by developing a more comprehensive prediction model of estimates of exposure, and by controlling for potential confounding associated with occupational exposures to other toxic substances. This research is based in Montreal, Quebec, Canada.

1.1 Problem Statement

It is easy to understand, on a superficial level, that everyone’s health is related to where they live. Where people live influences the kinds and numbers of health risks to which they might be exposed; the characteristics of social and physical environments that surround them; the accessibility (both financially and physically) of health care; and the occupations that may be available to them. In the context of health geography—that is, broadly, the study of variations in health status across space—health is considered a function of the characteristics of individuals (i.e., compositional factors) and of the environments in which they live and work (i.e.,
contextual factors) (Duncan et al., 1993; Macintyre et al., 1993; Duncan and Jones, 1995). Compositional factors include, for example, one's level of education, smoking and drinking behaviours, and genetic makeup. Contextual factors can include aspects of the social environment (e.g., characteristics of friends, family, co-workers) as well as of the physical environment (e.g., size and quality of one's home, climate, air quality or quality of drinking water). Population-level variations in health status can be observed between countries, between regions or provinces within countries, between cities, and between neighbourhoods in any given city. Not only can variations in health outcomes be observed at the neighbourhood scale, but so too can variations in the quality and characteristics of the physical environment. For example, ambient air pollution is highly variable at the local scale, which creates significant disparities in air quality between neighbourhoods in the same city (Briggs et al., 2000; Jerrett et al., 2001).

Short-term effects of air pollution on health have been observed consistently in studies conducted around the world, and Montreal is no exception (Goldberg et al., 2000; 2001a; 2001b; 2001c). Additionally, long-term exposures have been associated with reduced life expectancy, and both cardiopulmonary and lung cancer mortality (Pope et al., 2002). The task of assigning exposure to ambient air pollution to individuals is the principal challenge in studying associations between air pollution and health outcomes. The challenge relates to a number of issues, including the fact that people are mobile throughout the day, and the mixture of pollutants varies over small areas and between indoor and outdoor environments. It is important to note, that despite this limitation, two of the most important and influential studies of the effects long-term exposures on health (i.e., Dockery et al., 1993 and Pope et al., 2002) are cohort studies that assigned exposure to individuals based on the mean levels for entire cities.

The attribution of exposure to air pollution varies by the objectives and design of the study, as well as by available data. For example, in cohort studies that compare rates of disease between cities, all subjects in a given community may be assigned the same level of exposure, usually derived from a centralized monitoring site (e.g,
Dockery et al., 1993). In other studies it may be important to estimate personal levels of exposure, so subjects may wear personal monitoring devices for a period of time (e.g., Adgate et al., 2004). Currently, there is great interest in developing models to describe accurately spatial patterns of concentrations of intra-urban air pollution (Briggs 2005; Jerrett et al., 2005a; Hoek et al., 2008). Such spatial models can be used to estimate individual exposures to different pollutants and to link these exposures with health outcomes.

While there have been almost no studies to link air pollution with the incidence of breast cancer in humans, polycyclic aromatic hydrocarbons (PAHs), which are by-products of fuel combustion, and benzene, which is present in gasoline, are proven mammary carcinogens in rodents (US National Toxicology Program 1986; Maltoni et al., 1988; Huff et al., 1989). Benzene and PAHs, both of which are primary contaminants of traffic-related air pollution, are thus considered suspected risk factors for breast cancer (Labrèche and Goldberg, 1997; Sasco, 2003; Coyle, 2004).

Accepted risk factors for breast cancer are numerous and include, among others, family history of breast cancer, reproductive history, and alcohol consumption. Among Canadian women, breast cancer has the highest incidence of all cancers (22,900 estimated new cases in 2009) and is the second leading cause of cancer deaths (5,400 estimated deaths in 2009) (Canadian Cancer Society, 2009). Additionally, Canadian breast cancer incidence rates had been increasing by more than 1% per year since the 1960s for reasons that are not well understood, although these rates have begun to decline recently (Canadian Cancer Society, 2009). Breast cancer does not follow the traditional social gradient in Canada (i.e., incidence is not systematically higher among poorer women) unlike most other chronic and infectious diseases (Wilkins et al., 2002). Geographic location, however, is considered a strong predictor of breast cancer incidence, with consistent findings of higher rates in urban areas compared to rural areas, both in the U.S. (Reynolds et al., 2004; Hall et al., 2005) and in Canada (Bako et al., 1984). Only about one third of new cases of breast cancer are attributable to known risk factors, and thus, much breast cancer aetiology remains unexplained (Coyle, 2004). For these reasons, exploration of new
theories is warranted.

1.2 Research Objectives

This dissertation has three objectives:

1. Describe spatial and seasonal patterns of concentrations of intra-urban ambient air pollution in Montreal. The purpose of this objective was to develop a better understanding of the spatial patterns of air pollution across Montreal at a very local level, and to determine whether these spatial patterns are different at different times of the year. To achieve this objective, spatial models (raster surfaces) were developed using both recent and historical samples of indicators of traffic-related air pollution (i.e., samples of nitrogen dioxide (NO$_2$)) across Montreal. The historical observations provide insight into how spatial patterns and levels of air pollution have changed over the past few decades. Further, measurements sampled across different times of year allow for describing seasonal differences in both concentrations and in spatial patterns.

2. Identify spatial associations between concentrations of ambient air quality and indicators of social and material deprivation at the neighbourhood scale across Montreal. The purpose of this objective was to find out whether neighbourhoods in Montreal characterised by higher levels of deprivation were characterised also by higher levels of air pollution. This objective was achieved by using Statistics Canada Census tracts as proxies for neighbourhood boundaries overlaid on the air pollution surfaces. Associations were computed between indicators of deprivation and local concentrations of ambient pollution to identify whether or not some population groups (e.g., lower-income households) were at greater risk of exposure to higher levels of air pollution than others.

3. Determine whether the incidence of postmenopausal breast cancer is associated with exposure to local concentrations of air pollution after accounting for individual-level risk factors and occupational exposures. Very few studies have examined whether exposure to ambient pollution may be implicated in the aetiology of breast cancer despite findings from occupational studies that have shown associations
with exposure to some contaminants found also in ambient pollution. This objective was achieved with a hospital-based case control study design. Cases and controls in this cohort were all residents in Montreal and diagnosed with cancer between 1996 and 1997. All subjects completed a structured questionnaire with content related to personal risk factors, occupational exposures, and home address at time of diagnosis, which was used to geo-code (address match) subjects to the pollution surfaces described above. The case control dataset was collected in 1996-97 by other researchers for the purposes of a separate study. Odds ratios and 95% confidence intervals for risk of breast cancer associated with exposure to intra-urban concentrations of air pollution were estimated with unconditional logistic regression models.

1.3 Dissertation Outline

This dissertation consists of six chapters, based on three manuscripts. Two of the manuscripts have been published in the journals *Atmospheric Environment* and *Social Science and Medicine*, respectively. **Chapter Two** presents the research context and framework. This chapter provides a review of literature that this dissertation draws from and contributes to, namely health geography and environmental determinants of health; effects on health associated with exposure to air pollution; and, methods of modelling concentrations of intra-urban air pollution. Given that the manuscripts were maintained intentionally in their original format for this dissertation, there is some theoretical and conceptual material reviewed in this chapter that appears also in the introductory sections of the subsequent three manuscript chapters.

**Chapter Three** consists of the first manuscript and describes first the methods of collecting samples of concentrations of ambient NO₂ across Montreal on three separate occasions in 2005-2006. This chapter describes also the methods used for creating spatial surfaces in a geographic information system (GIS) of seasonal, as well as mean annual, concentrations of ambient NO₂. The method of spatial modelling used is land use regression, which is a statistical prediction method based
on associations between the observed concentrations of NO$_2$ and variables describing proximal and distal characteristics of land use and traffic. An original approach to land use regression modelling is presented, which predicts estimates of concentrations of NO$_2$ better than those found elsewhere in the literature. Although the sampling and modelling results demonstrate significant differences in mean concentrations of NO$_2$ between the three sampling periods, there was little spatial variability in the patterns observed across these periods.

The objective of Chapter Four is to investigate an important issue of environmental justice by identifying whether neighbourhoods in Montreal characterised by social and material deprivation have higher levels of exposure to ambient air pollution than do less deprived neighbourhoods. In this chapter some of the social and historical factors that contributed to producing associations between local indicators of deprivation and levels of air pollution, such as the long-standing presence of affluent inner-city neighbourhoods, are identified. In addition to identifying specific neighbourhoods that are characterised by a double burden of high levels of deprivation and high concentrations of ambient NO$_2$, this chapter illustrates also that due to the unique social geography in Montreal, there are exceptions to that pattern. For example, some of the most affluent neighbourhoods in Montreal are in the borough of Ville-Marie, which is located downtown in an area characterised by some of the heaviest concentrations of traffic and some of the highest levels of pollution on the whole Island.

Chapter Five includes the third manuscript, which builds upon the work developed in the first two manuscripts, and investigates the role of urban air pollution in the aetiology of postmenopausal breast cancer. Here, estimates of exposure to NO$_2$ from the pollution surfaces described in the first manuscript are assigned to cases and controls in an existing cancer cohort (Lenz et al., 2002; Labrèche et al., 2003). A method is presented for extrapolating the estimates of NO$_2$ collected in 2005-2006 back to levels that might have been present at the time of diagnosis (i.e., 1996) and ten years prior to this (i.e., 1985). The results of this study show an association between the incidence of postmenopausal breast cancer and
exposure to local concentrations of air pollution, after accounting for numerous individual-level risk factors and selected occupational exposures. For each increase in exposure of 5 parts per billion of NO$_2$ in 2005-2006 the fully-adjusted odds ratio for increased risk was 1.37 (95% confidence interval: 0.95 – 1.97). For estimated exposures in the 1980s and 1990s, the odds ratios varied between 1.16 and 1.32. Overall, there was about 1.5 to 2-fold increased risk for postmenopausal breast cancer among those women exposed to the highest concentrations of ambient NO$_2$ in Montreal as compared to those exposed to the lowest concentrations. The results presented in this chapter provide compelling evidence of an association between the incidence of breast cancer and exposure to traffic-related pollution, although further studies are needed to confirm whether the causal agent is indeed NO$_2$ or rather some other component of ambient air pollution.

**Chapter Six** concludes the dissertation. This chapter reviews the major findings, as well as the theoretical, substantive, and methodological contributions of the research. The key contributions include the development of an improved method for modelling intra-urban patterns of air pollution; the demonstration of evidence of environmental injustice across Montreal; and, evidence of a possible causal association between exposure to ambient air pollution and the incidence of postmenopausal breast cancer. This chapter includes a review of the limitations of the dissertation, and, finally, recommendations for further research are suggested.
This chapter provides a general research framework that synthesizes the relationships between the social environment, the physical environment, and how these interact with characteristics of individuals throughout the lifecourse to define inequalities in health. Next, the effects on health associated with exposure to ambient air pollution are discussed. Lastly, this chapter introduces challenges to, and progress in, methods for modelling concentrations of intra-urban ambient air pollution.

### 2.1 Determinants of Health

This dissertation follows a multi-disciplinary approach that incorporates theory and methods from health geography, environmental epidemiology, and social epidemiology. **Health geography** is the study of how where people live influences their opportunities for health, and involves the application of geographic information to the study of health outcomes and health inequalities. Health geography has both quantitative and qualitative branches, both of which examine inequalities. Quantitative health geography considers the geographical determinants and patterning of disease and health outcomes as well as the patterning of health care, while qualitative health geography considers relationships between health, health care, and place (Andrews and Moon, 2005). It could be argued that the former is concerned with health and *space*, while the latter is concerned with health and *place*.

**Environmental epidemiology** is defined as “the study of health effects on populations of exposure to physical, chemical, and biological agents external to the human body, and of immediate and remote social, economic, and cultural factors…related to these physical, chemical, and biological agents” (Last, 2001, 59). **Social epidemiology** is a
branch of epidemiology that examines the social distribution and social determinants of health states (Berkman and Kawachi, 2000). Clearly, there are theoretical and methodological overlaps between these three fields. For example, an important underlying tenet of all three—and the basis for the research approach described here—is that human health is a function of both compositional and contextual factors, as described in chapter one. Individuals are embedded in societies, and the risk of illness in individuals cannot be considered in isolation from the risk of illness in the population (Berkman and Kawachi, 2000).

Individual health outcomes are influenced by such compositional characteristics as personal behaviours, psychosocial experiences, and social status, in addition to genetic makeup (Figure 2.1). Psychosocial experiences may influence health directly, by causing biological changes that predispose one to illness; and indirectly, by influencing personal behaviours that are themselves determinants of risk (Brunner, 2000). The contextual factors presented in Figure 2.1 include: proximal features of one's local environment, including social relationships at home and work; physical characteristics of home and home neighbourhood, such as availability of parks and the quality of drinking water and indoor and outdoor air pollution; as well as more distal factors such as climate, culture, and political economy.
Although the influence of characteristics of the physical environment on health, such as exposure to polluted air or water, may seem intuitive, it is perhaps more difficult to understand the influence of social context on health. Berkman and Kawachi (2000) describe four mechanisms through which the social environment influences behaviours that in turn influence health, namely: shaping of norms; enforcing patterns of social control; providing opportunities to engage in certain behaviours; and, reducing/producing stress. For example, a study by Ross et al. (2005) reported that neighbourhood context in Canadian cities had a greater influence on individuals’ behaviours (e.g., smoking and drinking) than on health outcomes per se.

Arguably, a deeper social theoretical approach reflected in Figure 2.1 is that of a structurationist approach that recognizes the interplay between societal structures and human agency in influencing health outcomes. That is to say, health outcomes, which manifest locally, are influenced by both greater socio-political structures and environmental contexts, as well as by individual actions. Structuration theory
“acknowledges that structures shape social practices and actions, but that in turn, such practices and actions can create and recreate social structures” (Gatrell, 2002, 42). The relationship between structure and agency is immediately relevant to studies of neighbourhood effects on health: neighbourhoods are where individuals encounter social structures and where they live out their lifecourses (Bartley et al., 1998).

The theoretical standpoint is that neighbourhood contexts both constrain and enable individual health possibilities. For some, neighbourhoods can be a gateway to life-enhancing opportunities, including social and health services, as well as commercial and employment opportunities (Mayer and Jencks, 1989; Ross et al., 2002). For others, neighbourhoods may impose limitations and challenges to healthy living, and to social and professional development (e.g., through lack of safe recreation areas or available public transit options). Therefore, the quality of the physical environment, and the available resources located in or near the spaces where someone lives can reflect both what kind of life they live, as well as what kind of life they may be able to live (Smith, 2004).

Finally, as suggested in Figure 2.1, health status is influenced by exposures to all of the above-mentioned factors over the course of one’s life: from conception, through childhood and adulthood, to old age. The “lifecourse perspective” recognizes the importance of both time (i.e., some diseases have long latency periods) and of timing (i.e., early life exposures can influence outcomes later in life) (Lynch and Davey Smith, 2005). In this context, Macintyre and Ellaway (2003, 35) argue that “models of neighbourhood influences on health need to take into account both a temporal and spatial dimension [and that] it is important to consider the likely time lag between neighbourhood influences and their expression in health.” Consequently, the first objective of this dissertation is to describe current, as well as historical, patterns of concentrations of ambient air pollution.

2.2 Health Inequalities

The central purpose of health geography is to identify, explore, and explain
variations in health outcomes across space and between places. This discipline developed in response to a need to apply geographic knowledge and information to the investigation of health inequalities. The concept of health inequalities can be difficult to define. Of the many different definitions or conceptualizations of health inequalities, however, Whitehead’s (1990, 168) is among the most concise and accessible: “Equity in health implies that ideally everyone should have a fair opportunity to attain their full health potential and, more pragmatically, that no one should be disadvantaged from achieving this potential, if it can be avoided.” Irrespective of definition, measuring health inequalities requires three basic components: an indicator of health (e.g., life expectancy or incidence of disease); an indicator of group membership, (i.e., a way of categorizing people into comparison groups, such as by race or income); and a method for comparing the health indicator between the different groups (Braveman, 2006).

There are many different forms of health inequalities and ways of defining comparison groups. Around the world, women live longer than men, people of lower socioeconomic status or lower social position (SES) die younger than those of higher SES, and in Canada, those who live in rural areas have higher mortality rates than those who live in cities (Canadian Institute for Health Information, 2006). Different kinds of health inequalities have been observed within and between countries around the world and for numerous health outcomes. Furthermore, inequalities have been observed at a variety of different geographic scales. Clearly there are many forms of health inequality, most of which have their origins in differences in physical and social environments. The following sections identify some of those inequalities most relevant to the field of health geography.

2.2.1 Regional Inequalities

Some of the most striking health inequalities are those observed between places, such as between countries, in particular between developed and developing countries. Globally, for example, male life expectancy at birth ranges from less than 45 years in many African countries to over 77 years in more-developed countries such as Canada, Norway, and Japan (WHO, 2006a). Large-scale, regional variations within a
single nation exist but are less dramatic: male life expectancy varies by only a few years between Canadian provinces (Statistics Canada, 2001). At the local level, however, there is significant variability in health outcomes between cities and between neighbourhoods within the same city. Male life expectancy across the Island of Montreal, for example, varies by more than 13 years between different neighbourhoods (Direction de santé publique, 2002). Differences between urban and rural areas are another example of regional inequalities, as noted above.

It should be noted, however, that although these inequalities are described here as “regional,” they do not occur in the absence of underlying social inequalities. The disparities in life expectancy noted above are not due simply to where those places are located. These examples demonstrate how national-level measures of health and health inequalities (including life expectancy) can mask conditions that occur at regional and local scales. Furthermore, they demonstrate how different patterns can emerge from the same set of data as the geographic scale of analysis is changed. In the example described above, the difference between (but not within) Canadian regional and local-scale variability is due more to variations in population density, than to variations in socioeconomic or environmental characteristics. The following section, however, explores inequalities that are characterised specifically by variations in the socio-demographic and socioeconomic characteristics of populations.

2.2.2 Social Inequalities

Perhaps the most easily identifiable manifestations of health inequality are those that are related to social inequality. Social inequalities in health may be characterised by differences in health outcomes between groups or individuals with different levels of educational attainment, occupational class, income, or other measures of SES. Social inequalities can be pervasive throughout the lifecourse. For example, SES inequalities in mortality are present at birth, in childhood and adolescence, and through adult life to old age (Black, 1980).

The case of international disparities in life expectancy, presented above to describe regional inequalities, can also be described as social inequalities. In essence,
while some health inequalities are measured or identified by their regional manifestations, it is the underlying social characteristics of the regions that produce them. In other words, the regional distinction simply describes the inequalities, whereas the social factors actually play the strongest role in causing them. International differences in life expectancy are due to a combination of both physical environmental characteristics (e.g., climate, air and water quality, latitude, agricultural opportunities), and social conditions of people and places (e.g., politics, culture, overall wealth, distribution of wealth). This example demonstrates the complexity involved in teasing out uniquely geographical effects from social ones. The next section adds a third complicating dimension to the study of health inequalities.

### 2.2.3 Racial and Ethnic Inequalities

In addition to those health inequalities characterised broadly by their regional and socioeconomic manifestations, there are also those associated with race and ethnicity. For example, in the United States, blacks experience higher infant mortality rates, lower life expectancy (National Center for Health Statistics, 2005), lower self-reported health, and other disparities when compared to whites (Ferraro et al., 1997; Farmer and Farmer, 2005). The study of racial inequalities in health presents the formidable challenge to researchers of separating effects that are due to race/ethnicity from those associated with social inequalities.

This challenge is complicated further, especially in the US, by spatial patterns of residential segregation, whereby certain groups (e.g., blacks or low-income families) are concentrated into (usually) depressed neighbourhoods of a city. Typically, the groups who are segregated due to economic reasons are also those in the ethnic minority. Residential segregation tends to produce patterns of increasing poverty and decreasing social capital, neither of which bodes well for the health of residents (Massey and Fischer, 2000; Cattell, 2001; Jargowsky, 2003). As Leo et al. (1998, 6) explain, “when a metropolitan area is divided into neighbourhoods where poverty predominates and others where comfortable circumstances are the rule, it is inevitable that there will be a concentration of social problems in the poor areas.” Furthermore, it has been argued that it is in fact *racism* and discrimination, and not
race, that cause the racial/ethnic disparities in health (Krieger, 2000).

Despite the wide body of evidence appearing to demonstrate racial inequalities, however, Krieger et al. (2005) argue that in the United States, the contribution of socioeconomic inequalities in health to racial/ethnic health disparities (or vice-versa) is poorly documented and cannot readily be monitored, due to limitations in available data from public health surveillance systems. In practical terms, then, even though the examples described above are identified principally as social or ethnic inequalities in health, it may be counterproductive to consider them independently of their environmental contexts, that is, independently of geographic knowledge. It is therefore imperative for health geographers to disentangle the uniquely geographic aspects of health inequalities.

2.3 Air Pollution and Health

Outdoor air pollution consists of a complex mixture of substances, including sulphur dioxide (SO\textsubscript{2}), ozone (O\textsubscript{3}), nitrogen oxide (NO\textsubscript{x}), nitrogen dioxide (NO\textsubscript{2}), carbon monoxide (CO), carbon dioxide (CO\textsubscript{2}), particulates (i.e., mixture of solid and liquid droplets of varying sizes), and volatile organic compounds (VOCs). Air pollution does not respect administrative boundaries: “winds can transport pollutants long distances away from their source, adding to the levels of air pollution that are generated locally” (UNCSD 2005, 3). Vehicular traffic is a primary local contributor to urban air pollution through direct emissions of NO\textsubscript{x}, CO, CO\textsubscript{2}, SO\textsubscript{2}, VOCs, polycyclic aromatic hydrocarbons (PAHs), and particulates. In metropolitan Montreal, Canada, for example, 85% of emissions of NO\textsubscript{x} and 43% of VOCs have been attributed to transportation (King et al., 2005).

Although individual pollutants, such as ozone, nitrogen oxides, and particulates may each exert unique toxic effects on the respiratory and cardiovascular systems, they are all potent oxidants, either through direct effects on lipids and proteins or indirectly, through the activation intracellular oxidant pathways (Brunekreef and Holgate, 2002). Essentially, pollutants are inhaled directly into the throat and lungs and are absorbed through the skin. Particulate matter less than 10 µm in diameter
can be inhaled, and those less than 2.5 µm can be inhaled deep into the lungs (into the alveolar region where gas exchange occurs). Once pollutants enter the bloodstream they can be carried to the heart and to every other organ in the body. General mechanistic pathways to biological effects on health associated with exposure to particulate matter include, among others: rapid progression of cardiopulmonary disease; pulmonary and systemic oxidative stress; vascular alterations; modulated host defenses and immunity; and, lung damage and decline in lung function (Pope and Dockery, 2006).

Some of the effects on health associated with exposure to air pollution are observed shortly after exposure (e.g., headaches, nausea, upper-respiratory infections, daily hospitalizations) and others develop from continuous exposures throughout life, such as increased mortality from cardiovascular disease. A search in PubMed (April, 2009) for the keywords “air pollution” AND “health” produced over 9900 references in total, over 3000 in the last five years, and over 650 published in the last year alone. This is a huge and increasing body of literature, and therefore, this discussion attempts only to highlight some of the most important health-related findings associated with long-term exposure to traffic-related pollution in particular. For a comprehensive review of the air pollution and health literature see the articles by Brunekreef and Holgate (2002) or Samet and Krewski (2007).

Globally, the World Health Organization (2002) estimates that urban outdoor air pollution causes about 5% of cancers of the trachea, bronchus, and lung, and 2% of cardiopulmonary mortalities. Overall, the WHO attributes to air pollution about 1.4% of deaths and 0.8% of disability-adjusted life years, the bulk of which occurs in large urban areas in developing countries. Although the burden of increased risk to health may appear relatively low, it is important to recognize that the number of people exposed to air pollution in cities worldwide is huge. As such, the total number of people affected, or potentially affected, by exposure is substantial (Chen and Kan, 2008). More importantly, numerous epidemiological studies have shown that the pollutant concentrations associated with increased risks to health are quite low. In Canadian cities, where ambient pollution levels are relatively low compared
to those in large cities elsewhere in the world, Judek et al. (2004) estimated that approximately 5900 deaths were caused annually by combined short- and long-term exposure to air pollution. This latter study was based on residents in only eight Canadian cities and reported over 1,500 premature deaths in the city of Montreal alone. This number of premature deaths in eight major Canadian cities is more than the amount estimated nationally for each of breast cancer, prostate cancer, pancreatic cancer (Canadian Cancer Society, 2008), and motor vehicle accidents (Ramage-Morin, 2008).

2.3.1 Effects on Health Associated with Long-term Exposure to Air Pollution

Long-term exposure to air pollution throughout life is associated with increased mortality and morbidity from chronic diseases, including chronic respiratory disease, lung cancer, heart disease, and damage to the brain, nerves, liver, and kidneys. A number of case-control studies have reported associations between long-term exposures to combustion-related pollutants and lung cancer (Jedrychowski et al., 1990; Barbone et al., 1995) and childhood cancer (Feychting et al., 1998; Raaschou-Nielsen et al., 2001; Langholz et al., 2002). In a recent review of the effects of NO$_2$ on human health, Latza et al. (2008) found evidence that exposures to mean annual concentrations below 22 parts per billion (ppb) were associated with respiratory symptoms and diseases, hospital admissions, mortality, and otitis media.

Only a few cohort studies have investigated the effects on health associated with long-term exposure to air pollution. Most notably, two large, prospective cohort studies based in the US, namely the American Cancer Society (ACS) study (Pope et al., 1995; Pope et al., 2002) and the Harvard Six Cities study (Dockery et al., 1993), demonstrated robust and statistically significant associations between long-term exposure to relatively low concentrations of ambient pollution and mortality rates from cardiopulmonary diseases and lung cancer, after adjusting for smoking and other risk factors. In 2005(b), Jerrett et al. used a sub-cohort from the ACS study and showed a strong association (relative risk (RR) 1.11, 95% confidence interval (CI) 0.99-1.25) between all-cause mortality and an increase of 10 μg/m$^3$ of PM$_{2.5}$ in
Los Angeles, after controlling for ecologic confounder variables and 44 individual covariates. This latter study presented relative risks for premature mortality three times greater than Pope et al., 2002 reported in the national study and is especially important as it suggests that “the chronic health effects associated with within-city gradients in exposure to PM$_{2.5}$ may be even larger than previously reported across metropolitan areas” (Jerrett et al., 2005b, 727). In a separate re-analysis of the ACS study, Jerrett et al. (2009a) examined the risk to mortality associated with exposure to tropospheric (ground-level) ozone. Here, they found that the risk of dying from respiratory causes was relatively low (RR for increase of 10 ppb was 1.04 95% CI 1.01 to 1.07), however, the risk was more than three times as great in the areas with the highest concentrations compared to those with the lowest.

In Europe, a cohort study based in the Netherlands (Hoek et al., 2002) found similar results to these American studies with cardiopulmonary mortality related to both living near a major road (RR 1.95, 95% CI 1.09–3.52) and for an increase of 10 μg/m$^3$ of ambient black smoke (RR 1.34, 95% CI 0.68–2.64). A more recent follow-up of this Dutch study (Beelen et al., 2008) demonstrated again that traffic intensity on the nearest road was associated independently with mortality, however, the effect sizes were smaller (e.g., RR for cardiopulmonary mortality 1.07, 95% CI 0.98–1.15 for an increase of 10 μg/m$^3$ in concentrations of black smoke).

Canadian cohort studies based in Hamilton (Finkelstein et al., 2003) and in Toronto (Jerrett et al., 2009b) also found that local concentrations of traffic-related air pollution were important correlates of all-cause and circulatory mortality. Specifically, Finkelstein et al. (2003) followed over 5,000 people between 1985 and 1999, and assigned to them estimated neighbourhood-scale exposures to total suspended particulates and sulphur dioxide, based on observations from a relatively sparse network of fixed-site monitors. They found that subjects living in neighbourhoods characterised by the lowest incomes and highest levels of particulates had the substantially increased risk (odds ratio) of all-cause mortality of 2.62 (95% CI 1.67 – 4.13) compared to those living in neighbourhoods characterised by the highest incomes and lowest levels of particulates. In the case of Toronto,
Jerrett et al. (2009b) followed 2 360 people between 1992 and 2002 and assigned to them estimates of NO₂ from a land use regression model. Here, they found a more modest risk of all-cause mortality of 1.17 (95% CI 1.00 – 1.36) after controlling for several individual-level and ecologic variables. It should be noted, however, that in this latter study these authors based their exposure estimates on samples collected in 2002 and 2004 – thus beyond the point of their mortality follow-up – and made the assumption that levels and patterns of pollution had not changed significantly during the ten years of the study.

A recent systematic review conducted by Chen et al. (2008) concluded that long-term exposure to fine particles (PM_{2.5}) increases the risk of non-accidental mortality by 6% per a 10µg/m³ increase, independent of age, gender, and geographic region. This same review reported associations between an increase of 10µg/m³ in ambient NO₂ and incidence of mortality from respiratory diseases (pooled RR 1.16, 95% CI 1.06 - 1.26) and with cardiovascular mortality (pooled RR 1.02, 95% CI 0.98-1.07).

It must be acknowledged, however, that due to the co-locational association with other pollutants, it remains unclear to what extent the effects on health observed in any study are associated with any individual pollutant as opposed to any other, or to the complex mixture that makes up ambient air pollution (Pope and Dockery, 2006). In other words, 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 (Goldberg, 2007). Goldberg (2007) argues that the specific causal agent in this mixture cannot be identified. This implies that the measured or observed pollutants should be considered instead as markers for the mixture.

2.3.2 Health, Wealth, and Air Pollution

Air pollution does not affect all people equally. Individual responses to air pollution, both acute and chronic, depend on the pollutant to which one is exposed, the degree of exposure, one’s initial health status, genetics, and myriad other factors. Due to increased biological sensitivities and different exposure patterns, some
individuals, including young children, the elderly, and those who suffer from respiratory problems or cardiovascular disease, are more at risk for health problems related to air pollution (Chen and Kan, 2008).

Additionally, individuals of lower SES tend also to be at increased risk for health problems related to air pollution. Reviews by O’Neill et al. (2003) and Laurent et al. (2007) both provide evidence that air pollution contributes to creating or accentuating socioeconomic gradients in air pollution-related health outcomes and in premature mortality. Low educational attainment, which is one indicator of deprivation, has been shown to have an effect on the social gradient in air pollution-related health outcomes. For example, Pope et al. (2002) found that the association with particulate pollution was stronger for both cardiopulmonary and lung cancer mortality among individuals with lower levels of education. The fact that SES modifies the relationship between air pollution and health relates to both differential exposure and differences in existing health status. O’Neill et al. (2003) outline three possible mechanisms to explain how exposure to air pollution may contribute to greater effects on health among individuals of lower SES:

1) lower SES may increase susceptibility to air pollution-related health risks directly through increased levels of psychosocial stress, limited access to health care, or increased likelihood of living in lower quality housing;

2) some health conditions (e.g., asthma, diabetes, and cardiovascular diseases (Goldberg et al., 2006)), behaviours (e.g., smoking), and genetic traits that increase susceptibility to effects of air pollution are distributed differentially by SES; and,

3) populations with low SES may have more frequent or more intense exposures to air pollution than those with high SES due to environmental inequalities.

In this context, these authors argue that:

if both exposures and susceptibilities vary across socioeconomic gradients, these factors are likely to act together to influence the health response of groups classified by socioeconomic level. An air pollution epidemiology study that considers air pollution exposure, [SES] measures, and potentially other factors related to [SES] (disease status, sex, behaviour) must be based, implicitly
or explicitly, on a conceptual model that accounts for complex relationships among these factors (O’Neill et al., 2003, 1866).

The research described in this dissertation is an example of a study that strives to account specifically for, and to understand better, the complex relationships between SES and exposures to air pollution.

2.3.3 Spatial Modelling of Concentrations of Intra-urban Air Pollution

One of the most pressing problems in the investigation of the effects on health of ambient air pollution is the lack of high-quality data on personal exposures (Briggs, 2007). In earlier studies, investigators estimated variability in exposure to ambient air pollution between cities with data collected by relatively sparse networks of monitoring stations operated by government agencies (Dockery et al., 1993; Pope et al., 2002). More recently there has been growing interest in assessing exposure at the finer scale of neighbourhoods within the same metropolitan area (Jerrett et al., 2005a; Marshall et al., 2008).

A number of different techniques have been developed in the last decade to assess intra-urban exposure to air pollution, including, among others, dispersion models (Bellander et al., 2001), proximity-based assessments (Venn et al., 2000), different methods of geostatistical interpolation, such as kriging and inverse distance weighting (Jerrett et al., 2001; Marshall et al., 2008), and land use regression (Briggs et al., 2000) (see Jerrett et al., 2005a for a review of the techniques). Land use regression is a statistical prediction method that estimates in two-dimensional space, levels of pollution from measurements taken at specific locations within an urban area. It has proved to be more effective for describing spatial variability in levels of pollution than dispersion models and some methods of interpolation (Briggs et al., 1997; Briggs et al., 2000; Lebret et al., 2000; Hoek et al., 2001).

In land use regression, a spatially dense measurement campaign of levels of pollutants is conducted within a well-defined geographic area. The prediction model then incorporates land use, road and population densities, and characteristics of vehicular traffic. Land use regression models have been shown to explain between 50% and 80% of the spatial variability in particulate matter smaller than 2.5
micrometres (PM$_{2.5}$) and NO$_2$ in several European cities (Briggs et al., 1997; Briggs et al., 2000; Brauer et al., 2003; Rosenlund et al., 2008), American cities (Ross et al., 2006; Moore et al., 2007; Ross et al., 2007), and Canadian cities (Gilbert et al., 2005; Sahsuvaroglu et al., 2006; Henderson et al., 2007; Jerrett et al., 2007; Wheeler et al., 2008). Despite the development of land use regression models for a number of different cities using very similar data sources and methods, it has been argued that city-specific models are not readily transferable to other cities, given the inherent differences in meteorology, local topography and land use, and patterns of traffic between places (Briggs, 2007; Poplawski et al., 2009). Dense, locally-based measurements are required. Data related to traffic patterns and street networks are key components of land use regression given that automobiles and trucks are major contributors to air pollution. Despite the broad mixture of different pollutants originating from sources of traffic, NO$_2$ is recognized as a good indicator of all traffic-related pollution due to its demonstrated co-locational association with other pollutants in the mix (Nieuwenhuijsen, 2000; Brunekreef and Holgate, 2002; Beckerman et al., 2008; Wheeler et al., 2008).

2.4 Chapter Summary

This chapter has provided an overview of the interrelated and overlapping literatures that inform this dissertation, namely those related to the relationships between characteristics of the social environment and the physical environment, and how these interact with characteristics of individuals to influence health inequalities. It was shown that many health inequalities have their origins in environmental inequalities. Furthermore, this chapter described the principal effects on health associated with exposure to traffic-related air pollution and how not all people are affected equally by exposure. Lastly, in order to understand how the link between exposure to air pollution and health may be investigated, this chapter highlighted current approaches and methods for modelling concentrations of intra-urban air pollution.
In this chapter I address the first objective of the dissertation, namely, to describe the spatial and seasonal patterns of concentrations of intra-urban air pollution across Montreal. This dissertation consists of three inter-related projects, all of which hinge on the spatial model of ambient air pollution described below. The work presented in this chapter draws from and contributes to literature related to air pollution measurement and modelling, and to methods for spatial modelling of continuous data. This manuscript has been published in the journal Atmospheric Environment as follows:


**Abstract**

Concentrations of traffic-related air pollution can be highly variable at the local scale and can have substantial seasonal variability. This study was designed to provide estimates of intra-urban concentrations of ambient nitrogen dioxide (NO$_2$) in Montreal, Canada, that would be used subsequently in health studies of chronic diseases and long-term exposures to traffic-related air pollution. We measured concentrations of NO$_2$ at 133 locations in Montreal with passive diffusion samplers.
in three seasons during 2005 and 2006. We then used land use regression, a proven statistical prediction method for describing spatial patterns of air pollution, to develop separate estimates of spatial variability across the city by regressing NO$_2$ against available land-use variables in each of these three periods. We also developed a “pooled” model across these sampling periods to provide an estimate of an annual average. Our modelling strategy was to develop a predictive model that maximized the model $R^2$. This strategy is different from other strategies whose goal is to identify causal relationships between predictors and concentrations of NO$_2$.

Observed concentrations of NO$_2$ ranged from 2.6 ppb to 31.5 ppb, with mean values of 12.6 ppb in December 2005, 14.0 ppb in May 2006, and 8.9 ppb in August 2006. The greatest variability was observed during May. Concentrations of NO$_2$ were highest downtown and near major highways, and they were lowest in the western part of the city. Our pooled model explained approximately 80% of the variability in concentrations of NO$_2$. Although there were differences in concentrations of NO$_2$ between the three sampling periods, we found that the spatial variability did not vary significantly across the three sampling periods and that the pooled model was representative of mean annual spatial patterns.

3.1 Introduction

One of the most pressing problems in the investigation of the effects on health of ambient air pollution is the lack of high quality data on personal exposures (Briggs, 2007). In earlier studies, investigators estimated variability in exposure to ambient air pollution between cities with data collected by government agencies (Dockery et al., 1993; Pope et al., 2002). More recently there has been growing interest in assessing exposure at the finer scale of neighbourhoods within the same metropolitan area (Jerrett et al., 2005a; Marshall et al., 2008).

3.1.1 Spatial and Seasonal Variability of Ambient Air Pollution

Concentrations of ambient air pollution can be highly variable at the local scale. For example, Hewitt (1991) observed, in Lancaster, UK, differences in average annual concentrations of ambient nitrogen dioxide (NO$_2$) of more than 50 $\mu$g/m$^3$. 
between sampling locations less than 50 m apart from each other. Indeed, studies have shown that spatial variability of ambient air pollution can be greater within cities than between cities (Briggs, 2000; Zhu et al., 2002; Jerrett et al., 2005b). Intra-urban variability in ambient air pollution may be caused by a variety of factors, including the mixing of pollutants, local wind patterns (Seaman, 2000), patterns of traffic, and land use.

In addition to high spatial variability, there is usually substantial intra-urban seasonal variability in concentrations of air pollution (Ackerman and Knox, 2003). Seasonal variability may occur in cities that experience seasonal differences in patterns of urban heating and volume of traffic that may also be related to changing weather conditions (Andreeescu and Frost, 1998; Environment Canada, 2004); hours of sunlight, temperature, wind speed and direction, and amount and type of precipitation all influence the diffusion and dispersion of ambient pollutants (McGregor, 1999). The presence of heavy clouds, for example, reduces the amount of incoming ultraviolet radiation, thus limiting photochemical reactions that produce secondary pollutants, such as ozone (Jacobson, 2002).

In Montreal, Quebec, over the last five years, data from fixed-site air pollution monitoring stations located across the city showed that mean concentrations of NO\textsubscript{2} varied between 12 ppb in summer months to over 23 ppb in winter months (Environment Canada, National Air Pollution Surveillance (NAPS) data, available: www.etc-cte.ec.gc.ca/napsstations/Default.aspx). Wheeler et al. (2008) and Jerrett et al. (2009b) also observed nearly two-fold differences in concentrations of NO\textsubscript{2} between seasons in Windsor and Toronto, Ontario. Thus, it seems clear when estimating ambient air pollution, especially for the purposes of estimating chronic health effects, that both spatial and seasonal variability must be considered.

### 3.1.2 Techniques for Modelling Intra-urban Air Pollution Patterns

A number of techniques have been developed in the last decade to assess intra-urban exposure to air pollution, including, among others, dispersion models (Bellander et al., 2001), proximity-based assessments (Venn et al., 2000), geostatistical
interpolation, such as kriging and inverse distance weighting (Jerrett et al., 2001; Marshall et al., 2008), and land use regression (Briggs et al., 2000) (see Jerrett et al., 2005a for a review of the techniques). Land use regression is a statistical prediction method that estimates in two-dimensional space, concentrations of pollution from measurements taken at specific locations within an urban area. It has proved to be more effective for describing spatial variability than dispersion models and methods of interpolation (Briggs et al., 1997; Briggs et al., 2000; Lebret et al., 2000; Hoek et al., 2001).

In land use regression, a spatially dense measurement campaign of concentrations of pollutants is conducted within a well-defined geographic area. The prediction model then incorporates land use, road and population densities, and characteristics of vehicular traffic. Land use regression models have been shown to explain between 50% and 80% of the spatial variability in fine particulate matter (particles with aerodynamic diameters under 2.5µ; PM$_{2.5}$) and NO$_2$ in several European cities (Briggs et al., 1997; Briggs et al., 2000; Brauer et al., 2003; Rosenlund et al., 2008), American cities (Ross et al., 2006; Moore et al., 2007; Ross et al., 2007), and Canadian cities (Gilbert et al., 2005; Sahsuvaroglu et al., 2006; Henderson et al., 2007; Jerrett et al., 2007; Wheeler et al., 2008). Despite the development of land use regression models for a number of different cities using very similar data sources and methods, it has been argued that city-specific models are not readily transferable to other cities, given the inherent differences in meteorology, local topography, land use, and patterns of traffic between places (Briggs 2007; Poplawski et al., 2009).

Data related to traffic patterns and street networks are key components of land use regression given that automobiles and trucks are major contributors to air pollution through direct emissions of nitrogen oxides (NO$_x$), carbon monoxide, carbon dioxide, sulphur dioxide, volatile organic compounds (VOCs), polycyclic aromatic hydrocarbons, and particulates. In metropolitan Montreal, for example, 85% of NO$_x$ emissions and 43% of VOCs have been attributed to transportation (King et al., 2005). Despite this broad mixture of pollutants originating from road traffic, NO$_2$ is recognized as a good indicator of traffic-related pollution due to its
demonstrated co-locational association with other pollutants (Nieuwenhuijsen, 2000; Brunekreef and Holgate, 2002; Beckerman et al., 2008; Wheeler et al., 2008).

3.1.3 Objectives

The objective of this study was to develop a land use regression model in Montreal for describing intra-urban spatial patterns of NO$_2$ across seasons. The model that we developed is intended to be used for long-term exposure assessment in studies of chronic disease, which require annual estimates of NO$_2$, as opposed to assessments of short-term exposures and acute health effects. Important local variations in ambient pollution were found in all of the previous studies that used land use regression. With the exception of a study by Wheeler et al. (2008), however, all previous models were based on measurements of pollution at one point in time. As we mentioned, concentrations of ambient NO$_2$ in urban areas usually vary across seasons (Ackerman and Knox, 2003). What remains unclear, however, is to what extent the spatial patterns of ambient NO$_2$ remain consistent across seasons. As we conducted measurements in different seasons, we were able to explore these temporal patterns and identify whether one model created using an annual average can describe adequately patterns of NO$_2$ in all parts of the city and at all times of the year. Our modelling strategy was predictive rather than explanatory, in that our primary goal was to create a model that could reliably predict concentrations of NO$_2$, rather than to identify specific causal relationships between individual predictor variables and concentrations of ambient NO$_2$.

3.2 Materials and Methods

Montreal is the second largest metropolitan area in Canada (population 3.6 million people), but our study is set on the Island of Montreal, which has a population of approximately 1.8 million people (Statistics Canada, 2006). Ambient air pollution has been shown to vary spatially (Gilbert et al., 2005). Montreal has generally lower concentrations of air pollution than Canada’s largest city, Toronto, and other large cities in the United States, including Chicago, New York, and Philadelphia (Ontario Ministry of the Environment, 2006). The city has a temperate
climate, with mean daily temperatures during 1971-2000 in January and July of approximately -10°C (minimum of -35°C) and 22°C (maximum of 35°C), respectively (Environment Canada, Online Climate data: www.climate.weatheroffice.ec.gc.ca/climate_normals/index_e.html).

3.2.1 Monitoring of Ambient Air Pollutants

We developed the land use regression model by first conducting a series of sampling campaigns throughout the study area to estimate integrated two-week concentrations of NO₂ at individual points in space. We measured concentrations of NO₂ using two-sided Ogawa passive samplers (Ogawa and Co., USA). We deployed the samplers in three periods: November/December 2005 (to capture concentrations in cold weather), April/May 2006 (“temperate” weather), and August 2006 (“hot” weather). The Ogawa samplers have not been tested in conditions below -10°C and there are no published studies of their functionality in colder temperatures. We thus deployed the samplers in December, rather than in January or February, before Montreal’s temperatures became too cold for normal operation of the devices.

The samplers were installed at a height of 2.5 m above ground and were attached to street light poles, hydro-electric poles, or parking signs, usually near the sidewalk of the closest road. The geographic coordinates of each sampling location were recorded with a Garmin eTrex Legend Cx global positioning system (accurate to between 5 m and 15 m).

Location and Frequency of Sampling

NO₂ has high spatial variability, and so a relatively dense sampling network was used. The locations of the samplers were selected with a population-weighted location-allocation model that placed samplers in areas likely to have high spatial variability in traffic-related pollution, and with high population densities (Kanaroglou et al., 2005). Approximately 20 samplers were added to capture concentrations in residential areas that appeared to be under-represented by the initial allocation of sampling sites. Results from previous studies suggested that the precision of land use regression models depends more on the variability of the land use characteristics
captured by the sampling network rather than the total number of sampling sites (Ryan and LeMasters, 2007). Thus, we located a total of 133 samplers across the Island of Montreal, including in residential areas, industrial areas, parks, near and away from the shoreline, next to major roads and highways, and in areas with relatively high and low population densities (Figure 3-1). The minimum distance between any two neighbouring samplers was approximately 100 m and the maximum distance was just over three kilometres. The number and density of samplers used here is slightly larger than in other studies by Lebret et al. (2000) in Amsterdam, Sahsuvaroglu et al. (2006) in Hamilton, and Jerrett et al. (2007) in Toronto.

![Montreal Study Area and Sampling Locations](image)

**Figure 3-1: Locations of Ogawa passive samplers in Montreal, Quebec, 2005-2006**

**Analysis of Samplers**

The Ogawa passive samplers use triethanolamine-impregnated filters as an absorbent, and diffusion draws air into the sampler where a reagent absorbs the NO₂. Sampling begins when the sampler is exposed to air and ends when the
sampler is placed in a closed, impermeable container. The samplers were analyzed at an Environment Canada laboratory using ion chromatography (Gilbert et al., 2003).

3.2.2 Land-use Regression Models and Spatial Modelling

We developed the land use regression model in three steps. First, we developed a set of spatial variables that described characteristics of land use and road densities within buffers of various radii surrounding each sampler location. Second, we developed regression models to determine the associations between these variables and the observed concentrations of NO$_2$ and, subsequently, to create a prediction equation to estimate concentrations of NO$_2$ at locations in which measurements were not made. Third, we computed a spatial map to show, visually and numerically, the predicted concentrations of NO$_2$.

Step 1: Creation of spatial variables

The spatial variables were created in ArcGIS 9.2 (Redlands, CA) with the land use and traffic data in the 2006 DMTI CanMap Streetfiles (Markham, ON). This dataset contains a street centreline road network with six road classifications and topographic information, including land-use classifications and building footprints. The accuracy of these data range from 10 m to less than one metre.

There are two kinds of GIS data, namely vector, which consist of points, lines, polygons, and their associated attributes and, raster, which consist of grid-like, continuous surfaces in which individual pixels each represent a single attribute. The first step in this analysis was to convert the source vector-format data — which describe land uses and roads as polygons and lines respectively — to binary, raster surfaces at a resolution of 5 m x 5 m cells. For example, one data layer would include cells assigned as “industrial” or “not industrial”, while another data layer would include cells assigned either as “highway” or “not highway”. This was repeated for two road categories and six land use categories. We created similar surfaces in which each cell was assigned either the traffic count of the underlying segment of road or a value of “non-road”.

30
Next, we summed the total area (i.e., count of 5 m x 5 m cells) within buffers of multiple radii around every other cell in the entire study area. In this way, we were able to select any 5 m x 5 m location and immediately identify the total area of each land use (and total length of major roads or total traffic count on all major roads) within various buffer distances of that cell. This process was repeated with buffers of 100 m, 300 m, 500 m, and 750 m.

We also used data from the 2001 Canadian census to describe patterns of population density. We converted census tract boundaries to population-weighted centroids and then used kernel density estimation (Bailly and Gatrell, 1995) to create a density surface. Additionally, we used data from the 2005 National Pollutant Release Inventory (NPRI), which is Canada's legislated inventory of pollutants released and disposed of by industrial, institutional, and commercial facilities. This database includes emissions and address information from all facilities that manufacture or process any of the NPRI-listed substances (which includes NO$_2$) (see www.ec.gc.ca/pdb/npri/npri_home_e.cfm for more specific information about reporting requirements).

We used traffic count data generated from a region-wide transportation model developed in TransCAD, which is a comprehensive, GIS-based, urban transportation modelling software. The travel demand data were based on the 1998 origin destination survey undertaken by the Agence Métropolitaine de Transport de Montréal. The street network data with link travel times were acquired through DMTI Spatial Inc. (Markham, ON) and contained approximately 135,000 bi-directional links. These data present a detailed representation of peak morning automobile flow (6 am to 9 am) on a link-by-link basis for both primary provincial highways and major urban streets.

Lastly, we calculated also for each cell the straight-line distance to the shoreline, to the nearest highway, and to the known point sources of NO$_2$, namely the 33 facilities in the NPRI. Overall, we produced 47 different variables.
Step 2: Statistical Modelling

We created four separate models using as dependent variables the three sets of two-week integrated concentrations of NO$_2$ from the three sampling periods (i.e., December, May, August), as well as the mean of these three periods. This mean value was meant to represent an approximate annual estimate at each location. We developed our regression models using the natural logarithm of NO$_2$ because the data were distributed lognormally.

Our approach to modelling was to identify the model that explained the most variability (as estimated from the uncorrected $R^2$) of the natural logarithm of NO$_2$. To meet this objective, we included all variables measured except those that were perfectly collinear. This modelling strategy is different from that of other investigators (Henderson et al., 2007; Jerrett et al., 2007; Ross et al., 2007) who generally attempted to develop parsimonious models by using various model selection procedures (e.g., forward stepwise).

Our strategy is justified by the objective of the study, namely to develop a model that maximized prediction. Moreover, we were not overly concerned with the interpretation of the individual regression coefficients, only the resulting linear predictor. We thus included variables measuring the same construct but with different buffer sizes. In addition, statistical significance and expected sign (+ or -) of individual coefficients were not criteria for removal from the model, as excluding non-significant variables only reduced the predictive power of the model. Another advantage of our procedure is that it is transparent and reproducible and used the same set of variables across the three seasons so that comparison with the average, or “pooled” model, is simplified. We applied standard regression diagnostics to identify possible outliers and to ensure that the models conformed to the assumptions of linear regression. Additionally, we conducted sensitivity analyses to assess how the models performed after removing randomly 15% of the observations.
Step 3: Spatial Modelling

The four multivariable models were processed into individual spatial surfaces by calculating the linear predictor of $\text{NO}_2$ as $\exp(\beta_0 + \beta_{1,x_1} + \ldots + \beta_{i,x_i})$, where $\beta_i$ are the estimated regression coefficients for the independent predicting variables, $x_i$. These predictions were computed in GIS using the 47 spatial data layers and each predicted pollution map was computed at a 5 m resolution. Functionally, this translated into calculating the predicted concentration of $\text{NO}_2$ at every 5 m x 5 m location in our study area based on the measured physical characteristics of its surrounding areas.

3.2.3 Comparisons of the models across the sampling periods

To assess whether the mean of $\text{NO}_2$ across our three sampling sessions could serve as an adequate proxy for a mean annual concentration, we estimated Pearson correlation coefficients of the observed concentrations at the 129 sampler locations between the three seasons. We also intersected 5,000 randomly generated points with each of the four predicted surfaces and estimated the Pearson correlation coefficients between these. This step enabled us to assess whether locations that had relatively high (or low) concentrations of $\text{NO}_2$ in one season had similarly ranked concentrations in other seasons.

3.3 Results

3.3.1 Environmental Sampling

We obtained valid observations from all three sampling periods at 130 locations. Samplers were stolen or damaged on at least one occasion at three locations and so these locations were excluded from the analysis. Additionally, data from one sampler were discarded due to atypical circumstances. This sampler had been placed near an intersection of two single-lane streets in a residential block. In May 2006, construction activity led to the redirection of traffic from a major artery onto this side street, which increased traffic and idling by vehicles that normally would not have used that street. That sampler measured unusually high concentrations of $\text{NO}_2$,
including the highest observed value (36 ppb) of any location over all three sampling sessions. Observed concentrations of NO$_2$ from the remaining 129 locations ranged from approximately 2.6 ppb to 31.5 ppb, with mean values of 12.6 ppb in the December 2005 sampling session, 14.0 ppb in May 2006, and 8.9 ppb in August 2006 (Table 3-1). The greatest variability was observed during the May 2006 sampling period.

Table 3-1: Results of the two-week sampling of NO$_2$ across three sampling periods at 129 sites in Montreal, 2005-2006

<table>
<thead>
<tr>
<th></th>
<th>December</th>
<th>May</th>
<th>August</th>
<th>Average of the three seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>12.6</td>
<td>14.0</td>
<td>8.9</td>
<td>11.9</td>
</tr>
<tr>
<td>Median</td>
<td>12.7</td>
<td>13.8</td>
<td>8.8</td>
<td>11.8</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>2.6</td>
<td>4.3</td>
<td>3.1</td>
<td>3.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>6.7</td>
<td>6.1</td>
<td>2.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Maximum</td>
<td>20.1</td>
<td>31.5</td>
<td>16.9</td>
<td>19.0</td>
</tr>
</tbody>
</table>

3.3.2 Model Selection and Mapping

The magnitude and direction (+/-) of the coefficients in each multivariable model vary slightly by season (Table 3-2). This is not unexpected, because weather affects relationships between different land uses and topographic features and local concentrations of pollution. For example, differences in concentrations of residential heating between winter and summer alters the size of the regression coefficient of the residential land use variables without altering the overall predicted spatial patterns of NO$_2$.

Of the four multivariable models, the pooled model created with the mean of the three observation periods had the highest $R^2$ (0.80), whereas the models for May and August had the lowest (0.72 for both). The $R^2$ for December was 0.77. These results suggest that the collection of land use and traffic predictors included in our models collectively explain approximately 70-80% of the variability of NO$_2$ in Montreal.

Diagnostics and Model Validation
We examined several different diagnostic statistics and graphical plots to ensure validity of each of the four models. Specifically, we tested for homogeneity of variance in the residuals, normality of the residuals, and autocorrelation in the residuals. We found no heteroscedasticity or autocorrelation and we also found that the residuals had a reasonably normal distribution, all of which demonstrate that the models did not violate any of the assumptions of multiple regression. We also sought to identify influential or outlier cases by examining Cook’s $D$ and by inspecting plots of the observed versus the predicted values. Each model produced consistent predictions with almost no important outliers (Figure 3-2). We produced also a plot of the predicted values versus observed mean annual values (mean daily values during our total sampling period) at the nine fixed-site NAPS stations for which data were available (Figure 3-3). This plot too, showed good agreement between our predictions and the observations at the fixed sites across Montreal.
Table 3-2: Results of multivariable regression models, Montreal, 2005-2006
(*NPRI= National Pollutant Release Inventory Facility; N/A = variable excluded due to extreme multicollinearity)

<table>
<thead>
<tr>
<th>Variable</th>
<th>December 2005</th>
<th>May 2005</th>
<th>August 2005</th>
<th>Average of the three periods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlled land use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of units</td>
<td>Beta</td>
<td>Coef.</td>
<td>Standard</td>
<td>Coef.</td>
</tr>
<tr>
<td>Parks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Industrial land use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of units</td>
<td>Beta</td>
<td>Coef.</td>
<td>Standard</td>
<td>Coef.</td>
</tr>
<tr>
<td>Residential land use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of units</td>
<td>Beta</td>
<td>Coef.</td>
<td>Standard</td>
<td>Coef.</td>
</tr>
<tr>
<td>Density of buildings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number of units</td>
<td>Beta</td>
<td>Coef.</td>
<td>Standard</td>
<td>Coef.</td>
</tr>
<tr>
<td>Length of major road (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trafic on primary road</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trafic on major road</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to shoreline (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance to NTRI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

36
Figure 3-2: Comparison of observed (Ogawa) and predicted concentrations of NO₂ (n = 129) across three sampling periods, Montreal, 2005-2006

Multicollinearity is a problem that limits the ability to separate and assess the partial effects of correlated independent variables, but it does not hinder the ability to assess their joint effects. As such, we did not test for multicollinearity in our process of model evaluation (other than removing variables identified as being perfectly collinear, as explained above). Although high correlation between two or
more predictor variables may affect the sign and magnitude of the individual regression coefficients and their standard errors, it does not violate any of the assumptions of multiple regression, and it does not affect overall prediction (Mason and Perreault Jr. 1991; Allison, 1999).

As a further test of validity, we recomputed each model with a random selection of 85% of the observations and were able to produce comparable results to those achieved from the full dataset. Additionally, we created a surface map for the model of NO$_2$ based on the average from the three sampling periods and compared the predicted estimates at the 15% of locations that had been excluded from the model with our observed concentrations at those locations. Here we found Pearson correlation coefficients of ~0.9 between the observed and predicted values.

Surface maps of predicted concentrations of NO$_2$ from each model show very similar spatial patterns when categorized into groups of relatively high and low values, despite differences in mean concentrations between the sampling periods (Figure 3-4). Moreover, the surface maps have strong face-validity, with the highest predicted concentrations of NO$_2$ appearing along highway corridors and in the
downtown of the city (observed more easily in Figure 3-5). As well, the areas with the highest predicted concentrations appear to be the among most densely populated areas (see Figure 3-1) compared to the less densely-populated east and west ends of Montreal.

![Figure 3-4: Surface maps predicted concentrations of NO\textsubscript{2} across from sampling periods and their average, Montreal, 2005-2006](image)

### 3.3.3 Comparison of Seasonal and Spatial Variability

We explored the variability of spatial patterns of both observed and predicted concentrations of NO\textsubscript{2} across seasons. In both cases, we found strong positive Pearson correlation coefficients (i.e., 0.73 to 0.81) between the values in each season (Table 3-3). These correlations demonstrated that the locations characterized by relatively high and low concentrations of NO\textsubscript{2} remained consistent across the sampling periods, so that the spatial variability did not vary appreciably according to sampling period.
Figure 3-5: Surface map of predicted concentrations of NO$_2$ based on the average of the three sampling periods, Montreal, 2005-2006

Table 3-3: Pearson correlation coefficients of observed and predicted concentrations of NO$_2$ between sampling periods in Montreal, 2005-2006

<table>
<thead>
<tr>
<th>Observed concentrations at 129 sites</th>
<th>December</th>
<th>May</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>December</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>0.73</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>0.76</td>
<td>0.73</td>
<td>1.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predicted concentrations at 5000 random points</th>
<th>December</th>
<th>May</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td>December</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>August</td>
<td>0.81</td>
<td>0.80</td>
<td>1.00</td>
</tr>
</tbody>
</table>
3.4 Discussion and Conclusions

3.4.1 Principal findings

We showed that there were considerable differences in mean, integrated two-week concentrations of ambient NO$_2$ between the three sampling periods. Additionally, we found that there was significantly more variability in concentrations of NO$_2$ in temperate weather as compared to “cold” and “hot” weather. We observed also that the spatial variability in ambient NO$_2$ did not vary by sampling period, suggesting that one land use regression model based on the average of the three sampling periods may be used as a reasonable proxy for an annual estimate.

The spatial variability of concentrations of NO$_2$ was related directly to characteristics of local geography, namely population density and patterns of land use, vegetation, open space, and roads, and traffic. In this context, some of the coefficients in our models are different (in size and direction) from those reported by researchers working in other cities. Many other coefficients, however, (e.g., population density, industrial land use within 750m, distance to shoreline, traffic density within ~500m) are comparable to those obtained in models by other authors (Gilbert et al., 2005; Ross et al., 2006; Sahsuvaroglu et al., 2006; Henderson et al., 2007; Jerrett et al., 2007). The magnitude of coefficients from land use regression models is related to the set of variables included in the models, as well as to the absence of those variables that have been excluded from the models. Some modelling strategies, such as forward selection, will remove variables that are competing for the same variability through p-value selection criteria, whereas our procedure was based on retaining all variables that explained variability. Our models produced some of the highest predictions in the literature, and use of forward variable selection procedures with our data yielded concentrations of prediction similar to those found in the literature (data not shown).

3.4.2 Strengths and Weaknesses

Our model of estimated mean annual concentrations of NO$_2$ across Montreal reflects progress over the model described by Gilbert et al. (2005). In that study,
fewer factors for predicting concentrations of NO$_2$ were used, and the model was
based on samples of NO$_2$ from approximately half as many locations (67) as were
used here (129). Five of the seven variables included in that earlier study were
related to roads and traffic, whereas the present model is based on numerous
characteristics of the landscape, including area of commercial and industrial space,
building density, distance to the shoreline, and proximity of known point sources of
NO$_2$, in addition to density of roads and traffic. It could be argued, then, that the
present model may perform more reliably than the earlier model in areas of the city
characterised by more varied land uses and at increased distances from major
highways and expressways.

We believe that our variable selection procedure is optimal for the purpose of
producing the most highly predictive model, is reproducible, and is well suited for the
task of assigning estimates of ambient exposures at the intra-urban scale. If,
however, the goal of our study had been to inform policy for mitigating air pollution,
for identifying specific patterns and densities of land use that tend to be associated
with higher concentrations of pollution, or for determining independent predictors
of air quality, then another model selection procedure would have been used.

Gilbert et al. (2005) recommended incorporating data on industrial point sources
to improve model performance, which was achieved here with the NPRI data. It
should be acknowledged, however, that the NPRI dataset that we used to identify
point sources of pollution is likely to be incomplete. Although it is mandatory under
Canadian Law for industrial facilities to report to the NPRI, coverage may not be
complete because smaller industrial facilities may not be aware or able to provide
these data. In addition, the data provided to NPRI are not based usually on actual
measurements, but rather are engineering calculations. We also did not consider
wind direction or speed, both of which may influence the extent and direction of
dispersion of pollution from these facilities. Nevertheless, the associations between
NPRI point sources and NO$_2$ suggest that it is important to consider the specific
industrial landscape and presence of point source emitters when modelling ambient
pollution.
The use of estimates of traffic counts is a strength of this study. We found that the predictive power of our models (as measured by the $R^2$) improved by approximately 2-4% after adding the traffic data. Although these traffic data might be expected to necessarily improve model performance, in practice, their value-added to prediction in land use regression has been mixed. For example, Ross et al. (2006) and Sahsuvaroglu et al. (2006) both found that traffic estimates in San Diego County, CA, and Hamilton, ON, respectively, contributed significantly to the improvement of the models. On the other hand, Henderson et al. (2007) found that variables describing the length of roads proved to be as effective as variables describing vehicle density for predicting pollutant concentrations in Vancouver. Rosenlund et al. (2008) also found that variables based on estimates of traffic counts in Rome did not improve their model significantly. It is possible that the inconsistent findings related to the importance of traffic counts in previous land use regression models may have been due to competition of correlated variables for explaining variance. Estimates of traffic counts can be difficult or expensive to collect, and may be unavailable in some cities, but our findings suggest that these data will improve the predictive power of the model.

### 3.4.3 Implications

The results of this study showed that although concentrations of ambient NO$_2$ in Montreal vary throughout the year (due, perhaps, to seasonal changes in traffic volume, urban heating, and weather), the spatial patterns of these concentrations do not. It has already been shown that the principal source of ambient NO$_2$ in Montreal is exhaust from vehicular traffic (King et al., 2005). The high correlations between both the seasonal observations and the seasonal estimates of NO$_2$, as shown in Table 3-3, suggest that patterns of vehicular traffic in Montreal likely remain relatively consistent across seasons. The key implication of this finding is that we can conclude that use of the mean of the observations from the three sampling periods does not mask underlying spatial patterns. This lack of spatial variation between seasons suggests also that the spatial patterns visible in the final land use regression surface map are generally representative of the spatial patterns.
throughout the year, despite variations in actual concentrations across seasons.
In this chapter I address the second objective of the dissertation, namely, identifying spatial associations between concentrations of ambient air quality and indicators of social and material deprivation at the neighbourhood scale across Montreal. In the previous chapter the methods of data collection and spatial modelling of local concentrations of NO$_2$ were described. The research presented in this chapter builds upon that work by looking more closely at how the spatial patterns of NO$_2$ relate to the underlying patterns of population distribution across the Island of Montreal. Here, neighbourhood boundaries are overlaid on the pollution surface to explore whether certain neighbourhoods and population groups are exposed systematically to higher levels of pollution than others. The work presented in this chapter draws on and contributes to literature from environmental justice, health geography, and social geography. This manuscript has been published in the journal Social Science and Medicine as follows:


Abstract

Some neighbourhoods in urban areas are characterised by concentrations of
socially and materially deprived populations. Additionally, levels of ambient air
pollution in a city can be variable at the local scale and can create disparities in air
quality between neighbourhoods. Socioeconomic and physical characteristics of
neighbourhood environments can affect the health and well-being of local residents.
In this paper we identify whether neighbourhoods in Montreal, Canada, characterised
by social and material deprivation have higher levels of ambient air pollution than do
others.

We collected two-week integrated samples of nitrogen dioxide (NO$_2$) at 133 sites
in Montreal during three seasons between 2005 and 2006. We used these data in a
geographic information system, along with data describing characteristics of land
use, roads, and traffic, to create a spatial model of predicted mean annual
concentrations of NO$_2$ across Montreal. Next, we collected neighbourhood
socioeconomic information for 501 census tracts and overlaid their boundaries on
the pollution surface. We calculated Pearson correlation coefficients and 95%
confidence intervals (CI) between neighbourhood-level indicators of deprivation and
levels of ambient NO$_2$.

We found associations between concentrations of NO$_2$ and neighbourhood-level
indicators of material deprivation, including median household income ($r=-0.38$,
95% CI: -0.45 to -0.30), and with indicators of social deprivation, including
proportion of people living alone ($r=0.46$, 95% CI: 0.39 to 0.53). We identified
specific neighbourhoods that were characterised by a double burden of high levels of
deprivation and high concentrations of ambient NO$_2$. Because of the unique social
geography in Montreal, we found that not all deprived neighbourhoods had high
levels of pollution and that some affluent neighbourhoods in the downtown core
had high levels. Our results underscore the importance of considering social
tontexts in interpreting general associations between social and environmental risks
to population health.

4.1 Introduction

The concept of the social gradient in health, namely that individuals of lower
socioeconomic position tend to have poorer health than more advantaged individuals, is well established (Wilkins et al., 2002; Marmot, 2004; Lynch et al., 2006; Wilkins et al., 2008). Social gradients in health outcomes have been observed for all-cause mortality and for most chronic diseases across the developed world, regardless of how socioeconomic status (SES) or deprivation are measured (Wilkinson and Marmot, 2003). Certain characteristics of both physical and social environments influence gradients in health.

4.1.1 Spatial Variability of Air Pollution and Associated Health Effects

Ambient air pollution is an example of an environmental hazard that may influence gradients in health and that can be highly variable at local scales. For example, Hewitt (1991) observed in Lancaster, UK, differences in average annual levels of ambient nitrogen dioxide (NO$_2$) of more than 50 $\mu$g/m$^3$ between sampling locations less than 50 m apart from each other. Numerous studies have shown that spatial variability of levels of ambient air pollution can be greater between different neighbourhoods within cities than those between cities (Briggs et al., 2000; Zhu et al., 2002; Jerrett et al., 2005b). Intra-urban variability in ambient air pollution may be caused by a variety of factors, including how pollutants mix locally in the air, and factors that affect dispersion of pollution (e.g., local wind patterns (Seaman, 2000)), land use, and patterns of traffic.

Vehicular traffic is a primary local contributor to urban air pollution through direct emissions of nitrogen oxides (NO$_x$), carbon monoxide, carbon dioxide, sulphur dioxide, volatile organic compounds (VOCs), polycyclic aromatic hydrocarbons, and particulates. In metropolitan Montreal, Canada, for example, 85% of emissions of NOx and 43% of VOCs have been attributed to transportation (King et al., 2005). Nitrogen dioxide is recognized as a good indicator of traffic-related pollution due to its demonstrated co-locational association with other traffic-related pollutants (Nieuwenhuijsen, 2000; Brunekreef and Holgate, 2002; Beckerman et al., 2008; Wheeler et al., 2008). In a recent review of the effects of NO$_2$ on human health, Latza et al. (2008) found evidence that exposure to mean annual levels below 22 ppb were associated with respiratory symptoms and diseases, hospital
admissions, mortality, and otitis media. In a separate systematic review Chen et al. (2008) reported associations between an increase of $10 \mu g/m^3$ in ambient NO$_2$ and incidence or mortality from respiratory diseases (pooled relative risk (RR) = 1.16, 95% CI: 1.06 - 1.26) and with cardiovascular mortality (RR = 1.02, 95% CI: 0.98-1.07). It must be acknowledged, however, that due to the co-locational association with other pollutants, it remains unclear to what extent the health effects observed are indeed associated with exposure to NO$_2$ as opposed to the other pollutants or the complex mixture that makes up ambient air (Pope and Dockery 2006).

Exposure to ambient air pollution is associated with a wide variety of negative health outcomes. Some of these health effects are acute, which are observed shortly after exposure (e.g., headaches, nausea, upper-respiratory infections, daily hospitalizations) and others are chronic, or long-term, which develop from continuous exposures throughout life. Several large cohort studies have shown positive associations between long-term exposure to air pollution and mortality rates from cardiopulmonary diseases and lung cancer, after accounting for smoking and other risk factors (Dockery et al., 1993; Pope et al., 1995; Pope et al., 2002). A systematic review conducted by Chen et al. (2008) has shown, for example, that long-term exposure to fine particles (PM$_{2.5}$) increases the risk of non-accidental mortality by 6% per a $10\mu g/m^3$ increase, independent of age, gender, and geographic region. Judek et al. (2004) estimated that approximately 5,900 deaths were caused annually by combined short- and long-term exposure to air pollution in eight Canadian cities, including over 1,500 in the city of Montreal alone. This number of premature deaths in eight major Canadian cities is more than the amount estimated nationally for each of breast cancer, prostate cancer, pancreatic cancer (Canadian Cancer Society, 2008), and motor vehicle accidents (Ramage-Morin, 2008).

4.1.2 Double Burden of Deprivation and Exposure to Air Pollution

The concept of deprivation – as distinct from poverty, which refers more specifically to a lack of financial resources – emerged in Britain in the 1980s from a long tradition of analyzing social inequalities in health (Pampalon and Raymond, 2000). Townsend (1987, 125) defined deprivation as “a state of observable and
demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs.” Some individuals or households may exhibit multiple dimensions of deprivation, such as suffering from a low income, as well as lacking in participation in typical roles, relationships, and responsibilities of membership in society.

Townsend (1987) differentiated between two main forms of deprivation, namely, material (i.e., lacking in basic material goods and conveniences of modern life, such as a safe place to live, adequate diet, and basic household amenities) and social (i.e., lacking in social relationships with members of one’s family, community, or workplace). Both of these forms of deprivation have important implications for public health. Pampalon and Raymond (2000) provide a brief review of the links between living in deprived neighbourhoods and mortality and with morbidity from a wide variety of health outcomes.

Low educational attainment, which is one indicator of deprivation, has been shown to have an effect on the social gradient in health. For example, Pope et al. (2002) found that the association with particulate pollution was stronger for both cardiopulmonary and lung cancer mortality among individuals with lower levels of education. As reviewed by O’Neill et al. (2003) and Laurent et al. (2007), there is evidence that air pollution contributes to creating or accentuating socioeconomic gradients in air pollution-related health outcomes and in premature mortality. O’Neill et al. (2003) outline three possible mechanisms to explain how exposure to air pollution may contribute to greater health effects among individuals of lower SES:

1) lower SES may increase susceptibility to air pollution-related health risks directly through increased levels of psychosocial stress, limited access to health care, or increased likelihood of living in lower quality housing;

2) some health conditions (e.g., asthma, diabetes, and cardiovascular diseases (Goldberg et al., 2006)), behaviours (e.g., smoking), and genetic traits that increase susceptibility to effects of air pollution are distributed differentially by SES; and,
3) populations with low SES may have more frequent or more intense exposures to air pollution than those with high SES due to environmental inequalities.

Differentials in exposure by SES are perhaps the least studied of the three mechanisms outlined above through which exposure to air pollution may contribute to the social gradient in air pollution-related health outcomes. Ambient exposure to air pollution at the household-level is an important component of daily exposure. The major factors that determine where people choose to live include the accessibility and availability of services and amenities, the proximity to work or school, and affordability of housing. In this context, land use restrictions and real estate costs are important factors leading to unequal distributions of exposure by SES. For example, industrial facilities that produce pollution are located typically away from affluent neighbourhoods because of zoning restrictions, the higher costs of land, and prevailing winds. Individuals that are constrained financially face limited choices of where to live and they may reside near sources of pollution, including near roads with high traffic density, industrial facilities, waste disposal facilities, or airports (Perlin et al., 1999; Gunier et al., 2003). In affluent households, however, people have a greater ability to avoid living in close proximity to undesirable areas.

There is increasing evidence that residents in some neighbourhoods in urban areas in North America and Europe face the double burden of lower SES and elevated exposure to air pollution (Jerrett et al., 2001; Naess et al., 2007; Premji et al., 2007; Havard et al., 2009). In their comprehensive analysis of environmental inequity in Great Britain, Briggs et al. (2008) found generally strong, positive associations between low SES and local levels of ambient air pollution. In some cases, however, these authors found non-linear, J-shaped, and U-shaped associations, thus highlighting the fact that simple correlations may mask some of the complexities in the associations observed. In all of these studies mentioned above, however, the local social geographical contexts that contributed to producing patterns of social and environmental inequity were not described.
4.1.3 Objectives

In this paper, we identify whether neighbourhoods in Montreal characterised by socially and materially deprived populations are associated with greater exposures to ambient air pollution. Additionally, we attempt to identify some of the social and historical factors that have contributed to producing the observed spatial patterns. Lastly, we describe the implications to environmental justice among those living in areas characterised by both deprivation and higher levels of pollution. The present paper responds to recent calls in the literature of environmental epidemiology (ONeill et al., 2003; Laurent et al., 2007), public health (Premji et al., 2007), and health geography (Asthana et al., 2002; Cutchin, 2007) to conduct research on the combined influence of socioeconomic and environmental risks to the health of populations.

4.2 Context

Montreal, Quebec, is the second largest metropolitan area in Canada (population 3.6 million people (Statistics Canada, 2006)), but we limited our investigation to the Island of Montreal (population 1.8 million) because of limitations in data on ambient air pollution.

The Setting

The Island of Montreal is bounded by the St. Lawrence River on the south and the Rivière des Prairies on the north. Located centrally in the city is a small mountain (Mount Royal), which is comprised mostly by a large park. Montreal has a temperate climate, with 1971-2000 mean daily temperatures in January and July of approximately -10°C and 22°C, respectively, although daily temperatures may reach as low as -35°C and as high as 35°C (Environment Canada, Online Climate data: www.climate.weatheroffice.ec.gc.ca/climate_normals/index_e.html).

Major sources of traffic in Montreal include several expressways that cross through the city, most notably, the Autoroute Métropolitaine (provincial highway number 40), which extends along the length of the Island (Figure 4-1). There are
also three large bridges that connect the downtown core to the south shore and several smaller bridges that connect to the north shore. The bulk of industrial activity, including several large petrochemical facilities, is located in the eastern part of the city.

Montreal has lower regional concentrations of NO$_2$ (mean annual concentration of 18 ppb) than Canada’s largest city, Toronto (23 ppb), and other large cities in the United States, including New York (28 ppb), Philadelphia (24 ppb), and Chicago (24 ppb) (Ontario Ministry of the Environment, 2006). Previous studies have shown that local levels of ambient air pollution vary spatially across the Island (Gilbert et al., 2005) and in our previous paper (Crouse et al., 2009a) we reported that observed concentrations of NO$_2$ from 129 sampling locations across the Island ranged from approximately 2.6 ppb to 31.5 ppb. A study in Toronto that followed similar sampling methods found that concentrations of NO$_2$ ranged from 17.4 to 61.1 ppb (Jerrett et al., 2007).

![Figure 4-1: Context map of selected boroughs and neighbourhoods in Montreal](image-url)
Social Geography

The Island of Montreal is fragmented into 19 large boroughs that are each composed of multiple distinct neighbourhoods and communities. The highest densities of population are concentrated around the centre of the Island in boroughs such as Le Plateau and Côtes-des-Neiges, compared to areas in the east and west (Figure 3-2). The social geography of Montreal reveals residential patterns—most notably the presence of socially and materially privileged inner-city communities—that have persisted for over a century (Ames, 1897; Ley, 1993). Historically, the division of English and French groups has been central to the differentiation of Montreal’s social space (Foggin and Polese, 1977). Generally, the east of the Island and the former industrial areas to the southwest of downtown have been long characterised as working-class Francophone districts, whereas the wealthier inner-city neighbourhoods were dominated by Anglophone populations.

Figure 4-2: Population Density across Montreal, 2001
In the 1970s and early 1980s, proximity to regional amenities, such as McGill University and Mount Royal and the extensive parkland and existing pattern of elite districts around it, were important influences on continued social upgrading and gentrification around Montreal’s inner city (Ley, 1996). In their study of the 1971 Census, Foggin and Polese (1977) identified that the wealthiest neighbourhoods in the city were those closest to the centre, namely those located along the north, west, and southern slopes of Mount Royal. Ten years later, eight of the 10 census tracts reporting the highest mean household incomes were still located in this “horseshoe” of old inner-suburbs around the mountain, including Westmount, Outremont, and Mont Royal (Ley, 1993). In the late 1980s, the process of social upgrading began to extend beyond these original elite districts located to the north and west of the Mount Royal, into more-deprived, former industrial neighbourhoods to the south and east, such as St. Henri, Pointe Sainte-Charles, and Le Plateau (Ley, 1996).

Additionally, the far west end of the Island has long been home to low-density affluent Anglophone communities. The pattern of wealthy, Anglophone communities located around the city centre, and more-deprived Francophone communities located to the south and east persist, however, these patterns are less defined now than they were in the past.

In contrast to many other North American cities, Montreal’s downtown core has preserved its status as the city’s commercial, employment, and entertainment focal point (Séguin and Germain, 2000; Collin et al., 2003). This preservation of a vital city core has been possible in part because of a relatively slow population growth over the past 40 years that allowed for the relative constancy of the urban and social fabric (Séguin and Germain, 2000). The vitality of the downtown is due also to the presence of three of the city’s four universities (i.e., McGill, Concordia, Université du Québec à Montréal (U.Q.A.M.)) and because it is the central destination for all three principal modes of public transportation, namely the métro (subway), the bus station, and the train station, with the latter two both accommodating suburban-commuter as well as inter-city vehicles.

In Montreal, social status and ethnicity are not related directly in the way that they
are in many other large North American cities. Despite the presence of many distinct ethnic and multi-ethnic neighbourhoods, and concentrations of poverty in many neighbourhoods (Hatfield, 1997), there are no true ethnic-ghettos nor ghettos of socially-deprived groups as compared to what might be observed in large American cities (Walks & Bourne, 2006). Moreover, immigrants in Montreal have had high rates of economic success and there is no direct correlation between social or financial deprivation and either immigrant status (Séguin and Germain, 2000) or visible minority status (Bauder and Sharpe, 2002).

Regardless of the absence of clearly defined social or ethnic ghettos, Montreal has a relatively unequal distribution of income, at least by Canadian standards (Ross et al., 2000) and is one of the most economically segregated cities in Canada (Ross et al., 2002). Studies by Langlois and Kitchen (2001) and Apparicio et al., (2007) both identified the former industrial neighbourhoods of Pointe St-Charles, St.-Henri, and Lachine, to the southwest of downtown; Mercier-Hochelaga-Maisonneuve to the east; and Parc-Extension and Montréal Nord to the north, as among the most deprived areas on the Island.

The social geography of Montreal is mirrored in its patterns of health disparities. Life expectancy among men varies by more than 13 years between different regions of the city (Direction de santé publique, 2002), so there is ample evidence in Montreal of strong variations in health status at small geographic scales. The endurance of affluent residents in downtown neighbourhoods near major traffic routes, along with the fact that deprived areas are not confined to the central city, implies that elevated exposures to air pollution extends across social boundaries in Montreal. All of these geographical and historical factors described above demonstrate that Montreal presents a complex setting in which to examine relationships between neighbourhood-scale social- and physical-environmental characteristics.

### 4.3 Materials and Methods

This study has three broad methodological steps. First, we created a spatial
surface of predicted annual levels of NO$_2$ across Montreal (see Crouse et al., 2009a). Second, we described socioeconomic characteristics of populations at the neighbourhood scale. Third, we then combined these first two steps through a simple correlation analysis to identify the associations between local patterns of ambient air pollution and the underlying socioeconomic conditions.

4.3.1 Data

Air Pollution Data

In our previous paper (see Crouse et al., 2009a) we described the environmental sampling of ambient NO$_2$. Briefly, we measured levels of NO$_2$ using two-sided Ogawa passive samplers (Ogawa and Co., USA). These diffusion samplers, which provided estimates of mean daily concentrations during two-week sampling periods, were deployed at 133 locations across Montreal on three different occasions between 2005 and 2006. These samplers use triethanolamine-impregnated filters as an absorbent, and were analyzed at an Environment Canada laboratory using ion chromatography (Gilbert et al., 2003). We obtained valid observations from all three sampling periods at 129 locations. Observed levels of NO$_2$ in this dataset ranged from approximately 2.6 ppb to 31.5 ppb, with mean values of 12.6 ppb in the December sampling session, 14.0 ppb in May, and 8.9 ppb in August.

Socioeconomic Data

We used Statistics Canada 2001 census data aggregated to the census tract level to describe socioeconomic characteristics of Montreal's populations. Canadian census tract boundaries are determined by a committee of local specialists (including urban planners, health and social workers, and educators) working in conjunction with Statistics Canada. Census tract boundaries are defined with the intention of delineating neighbourhoods whose residents are relatively homogeneous in terms of socio-economic characteristics, economic status, and social living conditions, and are based on the physical characteristics of the landscape (for more information about the rules and methods for delineating census tracts, see the Statistics Canada data dictionary: www12.statcan.ca/english/census01/Products/Reference/dict/index.htm).
Typically, census tracts have populations of about 2,500 to 8,000 people. Ross et al. (2004) showed that census tracts are effective proxies for natural neighbourhoods in studies of contextual effects on health. Hence, we deemed that census tracts were the most appropriate unit for our analysis, and we use the terms ‘neighbourhoods’ and ‘census tracts’ synonymously throughout our discussions. It is important to recognize that each borough in Montreal is composed of multiple census tracts.

This dataset included 501 census tracts, after 14 small tracts were removed by Statistics Canada because of concerns that individuals in these census tracts could be identified. Many of the census tracts that were removed comprise large parks and few dwellings. The area of some census tracts on the Island of Montreal exceeded 20km², while many smaller tracts are less than 0.5km². The total populations of the 501 census tracts used here ranged from 293 to 9,288.

4.3.2 Step 1: Spatial Modelling of Air Pollution

We used the mean levels of NO₂ from our three sampling sessions in a land use regression model to create a spatial surface of predicted annual levels across Montreal. This surface was computed at a 5m x 5m resolution. Land use regression is a statistical prediction method that estimates levels of pollution at all points in a specific geographic region from measurements taken elsewhere (Jerrett et al., 2007). The estimates are not based on interpolation but rather on variables describing characteristics of the land uses, and road and traffic densities of surrounding areas. Functionally, this translates into calculating the predicted level of mean annual NO₂ at every 5m x 5m location based on the physical characteristics of the surrounding areas. Land use regression has proved to be more effective for describing spatial variability in pollution levels than dispersion models and some methods of interpolation (Briggs et al., 1997; Briggs et al., 2000; Lebret et al., 2000; Hoek, et al., 2001).

We developed the land use regression model by including 47 independent variables (describing characteristics of land uses, and densities of roads, traffic, buildings and population). We applied standard regression diagnostics to identify
possible outliers, evidence of extreme multicollinearity between independent variables, and to ensure that the model conformed to the assumptions of linear regression. Next, we calculated both the mean and median levels of predicted NO$_2$ for the areas bounded by each of the 501 census tracts. We used ArcMap 9.2 (ESRI, Inc.) for these computations.

4.3.3 Step 2: Describing Neighbourhood Socioeconomic Conditions

We chose 14 different variables to describe multiple dimensions of the socioeconomic characteristics of the populations in Montreal's neighbourhoods. In keeping with Townsend's (1987) differentiation between two main forms of deprivation, we included three variables that described conditions of material deprivation and four that described conditions of social deprivation. The selected variables have been used routinely as individual indicators of deprivation or as components of indices of deprivation in studies linking deprivation with negative health outcomes (Eames et al., 1993; Eachus et al., 1999; Pampalon and Raymond, 2000; Curtis et al., 2006; Dibben et al., 2006; Baumann et al., 2007). Some variables were transformed to normalize their distributions (i.e., unemployment rate and median household income, to their logarithms, and, percentage of recent immigrants to its square root). Given our knowledge of Montreal's social geography—for example, the presence of affluent neighbourhoods in the downtown core—we hypothesized that we may not see linear associations between the indicators of deprivation and elevated exposure to air pollution.

4.3.4 Step 3: Linking Socioeconomic Characteristics with Air Quality

The next step was to juxtapose the two main datasets, namely the air pollution surface and the spatial distributions of the socioeconomic variables. First, we examined scatter-plots of mean neighbourhood concentrations of NO$_2$ and each variable to ensure linear relationships between them. Next, we computed bivariable Pearson correlation coefficients between both the mean and median levels of predicted NO$_2$ in each neighbourhood and each of the 14 socioeconomic variables. We calculated 95% confidence intervals for these correlation coefficients.
4.4 Results

4.4.1 Spatial Patterns of Ambient Levels of NO$_2$

The results from the land use regression models showed that variables that described the density of buildings, population, traffic, and highways were the strongest predictors of NO$_2$ (univariable $R^2$ varied between 0.02 and 0.32 among these four variables), while variables describing total area of parkland and residential space were less predictive (univariable $R^2$ of these variables varied between 0.01 and 0.07). The full multivariable model describing mean annual NO$_2$ had an $R^2$ of 0.80 (Crouse et al., 2009a). This result suggests that the collection of 47 land use and traffic predictors included in our model collectively explained approximately 80% of the variability in NO$_2$ in Montreal.

The visual results of this prediction model showed that the highest concentrations of NO$_2$ were along highway corridors and in the downtown of the city. Generally, the more densely-populated areas of the city were among those with the highest predicted levels; low levels were found in the east and west ends of the Island and over Parc Mount Royal, where there is no industry or dwellings and few roads (Figure 4-3). Most of the west Island was characterized by mean annual levels of approximately 4-10 ppb in contrast to large sections of the boroughs of Anjou, St. Laurent, Côtes-des-Neiges, and the downtown core, where predicted levels were in the range of 20-35 ppb. The highest concentrations were found in the area around the intersection of the Autoroute Métropolitaine and the Autoroute Décarie (provincial highway 15), which are two of the busiest expressways in the city.
4.4.2 Neighbourhood Socioeconomic Conditions

The spatial patterns of socioeconomic status across Montreal were mostly consistent with the historical patterns identified by earlier studies, namely enclaves of affluence in the West Island and around Parc Mount Royal (e.g., Outremont, Westmount, Mont Royal), and pockets of deprivation in Montréal Nord, Mercier-Hochelaga-Maisonneuve, and the area south of Ville-Marie (Figure 4-4).
In terms of indicators of material deprivation, we found that across Montreal the mean neighbourhood-level proportion of adults without a high school diploma was 51.1% (ranging from 5.7% in one neighbourhood on the border of Côtes-des-Neiges and Outremont, to 96.8% in Lachine) (Table 4-1). Median neighbourhood-level household income was $CDN 39,509 (ranging from about $13,000 in one neighbourhood in Ville-Marie to about $213,000 in Westmount). In the borough of Ville-Marie there were two neighbourhoods where approximately two-thirds of households had total income below $20,000, yet another neighbourhood where no households had incomes below $20,000.

Among the indicators of social deprivation, the mean neighbourhood-level unemployment rate was 8.9% (ranging from nil in neighbourhoods in Westmount, Ville-Marie, and the West Island, to 46% elsewhere in Ville-Marie) and the mean neighbourhood-level proportion of lone-parent families was 21.5% (ranging from nil
in a few neighbourhoods in the Old Port area of Ville-Marie to 51.3% in a different
neighbourhood at the northern end of that same borough). The mean percentage of
people living alone was 18.9 (ranging from less than 2.0% in several West Island
communities to 52.4% in one neighbourhood in Ville-Marie).

4.4.3 Spatial Associations between Socioeconomic and Environmental
Risks

All but two of the variables (proportion of lone-parent families and proportion
of adults separated, divorced, or widowed) in our dataset had a linear relationship
with mean neighbourhood concentrations of NO$_2$. We therefore deemed it invalid
to consider simple correlations across the entire range of these two variables. In the
case of the latter variable, there appeared to be a positive association with levels of
NO$_2$ in neighbourhoods with less than 16% of adults separated, divorced, or
widowed, but a negative association in those with more than 16% (Figure 4-5). We
found Pearson correlation coefficients greater than 0.32 for two of the three
variables describing material deprivation and two of the four variables describing
social deprivation (Table 4-1).

![Figure 4-5: Scatter-plot of percentage of adults separated, divorced, or widowed versus median ambient concentrations of NO$_2$ (n = 501 neighbourhoods, LOESS curve using a span of 80%)](image)
Table 4-1: Variables describing socioeconomic characteristics of Montreal neighbourhoods (n = 501 census tracts)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Median</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Pearson Correlation with Mean NO2</th>
<th>95% C.I. Pearson Correlation</th>
<th>95% C.I. Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density</td>
<td>total population/km²</td>
<td>7.074</td>
<td>8.219</td>
<td>3.513</td>
<td>124</td>
<td>66.783</td>
<td>-0.44</td>
<td>0.37</td>
<td>0.51</td>
</tr>
<tr>
<td>Children</td>
<td>% population aged 14 and younger</td>
<td>15.2</td>
<td>15.1</td>
<td>5.0</td>
<td>1.9</td>
<td>32.3</td>
<td>-0.34</td>
<td>-0.42</td>
<td>-0.26</td>
</tr>
<tr>
<td>Seniors</td>
<td>% population aged 65 and older</td>
<td>13.7</td>
<td>15.0</td>
<td>7.0</td>
<td>3.5</td>
<td>52.3</td>
<td>0.14</td>
<td>0.23</td>
<td>0.06</td>
</tr>
<tr>
<td>Older neighbourhoods</td>
<td>% dwellings constructed prior to 1961</td>
<td>56.8</td>
<td>51.0</td>
<td>22.7</td>
<td>0.0</td>
<td>97.7</td>
<td>0.40</td>
<td>0.33</td>
<td>0.47</td>
</tr>
<tr>
<td>Recent immigrants</td>
<td>% population who immigrated between 1996 &amp; 2001</td>
<td>3.9</td>
<td>5.2</td>
<td>4.7</td>
<td>0.0</td>
<td>25.8</td>
<td>0.33</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>Visible minorities</td>
<td>% population who identify as visible minority</td>
<td>15.5</td>
<td>19.3</td>
<td>15.0</td>
<td>0.4</td>
<td>82.4</td>
<td>0.18</td>
<td>0.09</td>
<td>0.26</td>
</tr>
<tr>
<td>High income</td>
<td>% households with total income above $90,000</td>
<td>8.6</td>
<td>12.9</td>
<td>12.8</td>
<td>0.0</td>
<td>82.2</td>
<td>-0.27</td>
<td>-0.35</td>
<td>-0.19</td>
</tr>
<tr>
<td><strong>Material Deprivation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low education</td>
<td>% population aged 20+ without high school diploma</td>
<td>57.5</td>
<td>51.3</td>
<td>19.8</td>
<td>5.7</td>
<td>96.8</td>
<td>-0.17</td>
<td>-0.26</td>
<td>-0.09</td>
</tr>
<tr>
<td>Low income</td>
<td>% households with total income below $20,000</td>
<td>28.5</td>
<td>27.8</td>
<td>12.3</td>
<td>0.0</td>
<td>68.6</td>
<td>0.18</td>
<td>0.09</td>
<td>0.45</td>
</tr>
<tr>
<td>Median household income</td>
<td>dollars ($)</td>
<td>54,376</td>
<td>59,509</td>
<td>15,684</td>
<td>12,900</td>
<td>215,111</td>
<td>-0.36</td>
<td>-0.43</td>
<td>-0.28</td>
</tr>
<tr>
<td><strong>Social Desirability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>% population aged 15+ unemployed</td>
<td>7.9</td>
<td>6.9</td>
<td>4.6</td>
<td>0.0</td>
<td>46.0</td>
<td>0.30</td>
<td>0.22</td>
<td>0.38</td>
</tr>
<tr>
<td>Lone parents</td>
<td>% families that are lone parent families</td>
<td>21.3</td>
<td>21.5</td>
<td>7.8</td>
<td>0.0</td>
<td>51.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Living alone</td>
<td>% population living alone</td>
<td>38.3</td>
<td>38.0</td>
<td>9.3</td>
<td>1.0</td>
<td>53.4</td>
<td>0.43</td>
<td>0.36</td>
<td>0.50</td>
</tr>
<tr>
<td>Separated, divorced, widowed</td>
<td>% population aged 15+ separated, divorced, or widowed</td>
<td>19.7</td>
<td>19.6</td>
<td>9.0</td>
<td>7.2</td>
<td>59.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Air Pollution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean predicted NO2</td>
<td>parts per billion</td>
<td>11.7</td>
<td>11.7</td>
<td>2.1</td>
<td>6.2</td>
<td>17.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Median predicted NO2</td>
<td>parts per billion</td>
<td>11.6</td>
<td>11.7</td>
<td>2.1</td>
<td>6.1</td>
<td>16.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

63
The associations between deprivation and concentrations of NO$_2$ were not all in the expected direction. For example, neighbourhoods with low levels of education, as measured by the percentage of adults who have not completed high school, were found to be inversely associated with mean levels of NO$_2$ ($r = -0.15$, 95% CI: $-0.24$ – $-0.07$). In other words, neighbourhoods with high proportions of less-educated individuals were associated with lower exposures to air pollution, and vice-versa. We hypothesized that this inverse association could be explained in part by the presence of three universities (i.e., McGill, Concordia, U.Q.A.M.) and their associated “student ghettos” (i.e., concentrations of students residences), which are all located in the dense, traffic-heavy areas of the city. We therefore conducted a sensitivity analysis by excluding from analysis all neighbourhoods where less than 60% of the population aged 15-24 were students (either part-time or full-time). This reduced sample of neighbourhoods ($n= 131$) produced an even stronger inverse association with NO$_2$ exposure ($r = -0.44$, 95% CI: $-0.57$ – $-0.29$) (Figure 4-6). The prevalence of many older, first-generation immigrants, who have low levels of educational attainment – but may not necessarily be materially or socially deprived – living in the less-dense, eastern boroughs of the Island may contribute to this inverse association. This finding is, however, one example of how elevated exposure to air pollution appears to extend across boundaries of social status in Montreal.

We identified four neighbourhoods in the boroughs of Saint-Laurent, Mont-Royal, and Westmount that were also representative of elevated concentrations of air pollution coincident with conditions of higher SES. These neighbourhoods were characterised by having more than 26% (double the Montreal average) of households with mean annual incomes greater than $90,000; fewer than one quarter (less than half the Montreal average) of adults had not completed high school; and mean annual levels of NO$_2$ exceeded 23 ppb, which is double the Montreal average and in excess of the WHO air quality guidelines for annual exposures (WHO, 2006b). These neighbourhoods are located near the centre of the Island near some of the heaviest concentrations of busy roads.
Figure 4-6: Scatter-plot of the percentage of adults without a high school diploma versus mean ambient concentrations of NO$_2$ (n = 131 neighbourhoods where fewer than 60% of people aged 15 to 24 are students)

On the other hand, much of the borough of Montréal Nord is characterised by high levels of deprivation and yet has among the lowest levels of air pollution on the Island (mean annual levels not exceeding 12 ppb). Levels of NO$_2$ are relatively low there because of that borough’s distance from major expressways, industry, and the denser urban core of the city, and because of its close proximity to the Rivière des Prairies. Lastly, three neighbourhoods – two in Parc-Extension, and one in Lachine – were particularly representative of the double burden of deprivation and elevated concentrations of air pollution. In each of these three neighbourhoods more than 40% of households have earnings less than $CDN 20,000 per year; more than 20% of adults are separated, divorced, or widowed; and, mean annual levels of NO$_2$ exceed 23 ppb.

4.5 Discussion and Conclusions

4.5.1 Principal findings

The principal finding of this study is that while overall patterns fit within
generally accepted ideas of a triple burden of social, material, and environmental deprivation, there are many examples where excess exposures to air pollution cross social and economic boundaries in Montreal. A secondary contribution of this paper is the identification of specific neighbourhoods in Montreal where physical and social environmental characteristics may present cumulative risks to population health. Specifically, we found that neighbourhoods with higher proportions of unemployed adults, visible minorities, low-income households, and individuals living alone, were all positively associated with higher exposures to ambient NO$_2$ although clear exceptions existed. Those correlations cannot be used to infer causality, but simply to identify statistical associations between the different area-wide socioeconomic variables and exposure.

The fact that we found no linear association between neighbourhood concentrations of NO$_2$ and percentage of adults separated, divorced, or widowed is somewhat surprising given the linear association observed between air pollution and living alone ($r=0.46$, 95% CI: 0.39 – 0.53). One might have expected these two variables to reflect similar patterns. The conjugal experience in the province of Quebec, however, is quite different from that in the rest of Canada (Statistics Canada, 2002). Data from the 2002 Statistics Canada General Social Survey found that more than 70% of Quebec women chose common-law relationships to start their conjugal lives as opposed to marriage, compared to only 34% who did so elsewhere in Canada. For these reasons, some variables related to family or relationship status may not have the same associations with deprivation in Quebec as they might in other parts of Canada.

4.5.2 Strengths and weaknesses

The fact that we included a variety of indicators of deprivation and that we made use of high-quality estimates of concentrations of ambient NO$_2$ at the relatively fine scale of the neighbourhood are key strengths of this study. In addition, the land use regression model that we used (Crouse et al., 2009a) to create the spatial surface of mean annual concentrations of NO$_2$ explained approximately 80% of the spatial variability.
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 (Goldberg, 2007). Furthermore, people are mobile throughout the day and the mixture of pollutants varies over small areas and between indoor and outdoor environments. In light of this, we acknowledge that our analysis does not estimate personal exposure to air pollution. Total personal exposure depends on a complex number of factors and behaviours, including occupational exposures, amount of time spent indoors and outdoors, exposure to tobacco smoke (passive and active), and indoor home characteristics.

**Choice of Areal Units**

The choice of boundaries used in studies of ecologic effects can be contentious and at times be problematic due to a number of issues. One issue in this context is the modifiable unit problem (MAUP) (Openshaw, 1983), which arises in situations when the boundaries delineating a set of units (e.g., neighbourhoods) are essentially arbitrary, and hence modifiable. As such, associations observed between aggregate-level data may be artefacts of the choice of boundaries, and may not be present between units that are delineated differently or at other scales of aggregation. Researchers are encouraged to use the smallest units available for analysis because larger scales of spatial aggregation tend to exhibit smaller variances in the observed phenomena “as aggregated means regress to the regional averages” (Baden et al., 2007, 165).

A second challenge related to choosing areal units for studies of ecologic effects at the neighbourhood scale in particular relates to the fact that the units used should be representative of natural neighbourhoods, be ecologically meaningful (Pickett and Pearl, 2001), and in some way define communities (Willms, 2001). In this case, we had the option of choosing between dissemination areas, which are the smallest unit available, and census tracts, which are larger than dissemination areas but are defined explicitly with the intention of delineating meaningful neighbourhood boundaries (as
described earlier). We performed sensitivity analyses to compare the associations between levels of air pollution and some of our indicators of deprivation between both census tracts and dissemination areas (results not shown). The associations were similar, though slightly stronger at the larger scale of the census tracts. We felt that the implication of this finding, based on our explanations above, is that the census tract level analysis provided not only stronger, but also more meaningful associations than those at the dissemination area level and, so this is the unit that we chose for our study. We acknowledge, however, that like all ecologic analyses, the effect of the modifiable areal unit problem limits our results.

4.5.3 Implications

The implications of the results of this study are that an examination of local socio-geographic contexts is required when looking at environmental equity, since not all cities have the same patterns of land use or of population. Although the identification of overall associations between neighbourhood-scale indicators of deprivation and levels of ambient air pollution are important, it is the neighbourhoods that are the exceptions and in opposition to these overall associations that are perhaps the most interesting. An added benefit of local analyses and the identification of specific neighbourhoods characterised by a double burden of deprivation and elevated exposures to air pollution is the opportunity to provide specific policy-relevant information for local governments. For example, ambient exposures at the household-level represent one of the least-modifiable components of exposure in the short term. Thus, local estimates of concentrations of ambient exposure are useful for guiding the development of strategies to control and regulate pollution, and they provide important information from a social justice perspective. The choice of where to live and quality of housing is constrained for many lower-income households. Lower-income households are the least able to take action towards modifying exposure. For example, in addition to being more restricted in terms of where to live, lower-income individuals – including many students and seniors – likely reside in one of the 50% of Montreal homes without air conditioning, and may be living in poorer quality housing that is more open to
infiltration of pollutants.

Furthermore, future studies could extend the work presented here by combining exposure data with data describing individual-level health outcomes by SES. Longitudinal and hierarchical study designs that incorporate environmental exposures, neighbourhood-level characteristics, individual characteristics, and time-activity patterns could better investigate causal pathways between differential levels of exposure to ambient pollution and health outcomes.
CHAPTER 5

POSTMENOPAUSAL BREAST CANCER IS ASSOCIATED WITH EXPOSURE TO TRAFFIC-RELATED AIR POLLUTION IN MONTREAL, CANADA

This chapter consists of the third manuscript of the dissertation. In this chapter I address the third and final objective of the dissertation, namely, to determine whether the incidence of postmenopausal breast cancer is associated with exposure to local concentrations of air pollution after accounting for individual-level risk factors and occupational exposures. This chapter builds upon the results, methods, and knowledge acquired through the research presented in the previous two chapters. Whereas the previous chapters described the spatial patterns of local concentrations of air pollution across Montreal and how those patterns related to underlying population patterns, this chapter goes a step further by exploring how these spatial patterns are related to actual health outcomes in the city, namely the incidence of postmenopausal breast cancer. The work presented here is an example of interdisciplinary research that bridges geography, environmental measurement, epidemiology, and public health. This chapter draws from and contributes to literature related to environmental epidemiology, breast cancer aetiology, and methods of exposure assessment.

Abstract

The incidence of breast cancer has been increasing in many countries over the past few decades, and only about of 30% of new cases can be explained by known risk factors. Findings from occupational studies have shown associations between the incidence of breast cancer and exposure to some contaminants found also in
ambient air. As traffic-related air pollution is ubiquitous in urban areas, it is hypothesized that there may be an association between exposure and the incidence of postmenopausal breast cancer.

Between 1996 and 1997 a population-based case-control study of postmenopausal breast cancer was conducted in Montreal, Quebec. Cases were 556 women, age 50-75 years, with incident malignant breast cancer and controls were 613 women with other malignant cancers, frequency-matched for age, date of diagnosis, and hospital. Concentrations of nitrogen dioxide (NO\textsubscript{2}) were measured at 129 locations across Montreal on three occasions in 2005 and 2006. Using land use, road, and traffic data, a land use regression model was developed to predict concentrations of NO\textsubscript{2} at all points in the city. Additionally, using historical data from fixed-site monitors, a method was developed to back-extrapolate the estimates of exposure to the 1980s and 1990s. Exposure estimates were linked to each subject's residence at time of diagnosis. Unconditional logistic regression models were adjusted for all accepted risk factors, selected occupational exposures, and neighbourhood-level socioeconomic characteristics to compute odds ratios (ORs) and 95% confidence intervals (CI).

Odds ratios were increased for the accepted risk factors for breast cancer (e.g., family history, benign breast disease, education, age at menarche). The response function for NO\textsubscript{2} was linear. For each increase of 5 parts per billion of NO\textsubscript{2} in 2006 the OR\textsubscript{5ppb} was 1.37 (95% CI 0.95 – 1.97). For estimated exposures in the 1980s and 1990s, the OR\textsubscript{5ppb} varied between 1.16 and 1.32. Odds ratios ranged between 1.75 and 2.33 when comparing in each time period those in the highest quartile of exposure relative to those in the lowest. Although the results provide compelling evidence of an association between the incidence of breast cancer and exposure to traffic-related pollution, further studies are needed to confirm whether the causal agent is indeed NO\textsubscript{2} or rather some other component of ambient air pollution.

5.1 Introduction

Among women, breast cancer has the highest incidence rate of all cancers and is
the second leading cause of death from cancer in both Canada (Canadian Cancer Society, 2009) and the United States (American Cancer Society, 2009). Rates of incidence in Canada, as in most developed countries, had been increasing by more than 1% per year since the 1960s for reasons that are not well understood, although rates appear to have begun declining recently (Canadian Cancer Society, 2009). Accepted risk factors for breast cancer are numerous and include genetic mutations, family history of breast cancer, aspects of reproductive history, and lifestyle factors such as alcohol consumption and weight gain. Only about one-third of new cases of breast cancer are attributable to known risk factors, and thus, much of the aetiology remains unexplained (Rockhill et al., 1998; Coyle, 2004).

Wilkins et al. (2002) showed that breast cancer does not follow the traditional social gradient in Canada (i.e., incidence is not systematically higher among less-affluent women), unlike most other chronic and infectious diseases. In fact, studies conducted in the United States and elsewhere have shown that women with higher levels of education and higher income have an increased risk of developing breast cancer (Kelsey and Bernstein, 1996; Heck and Pamuk 1997), independent of other accepted risk factors. Similar results have been reported in ecologic studies using community-level indicators of socioeconomic status (Gorey et al., 1998; Prehn and West, 1998; Mackillop et al., 2000; Yost et al., 2001). Lastly, geographic location has been associated with patterns and clusters in the incidence of breast cancer (Mandal et al., 2009), and there have been consistent findings of higher rates in urban areas compared to rural areas, both in Canada (Bako et al., 1984) and in the United States (Nasca et al., 1992; Reynolds et al., 2004; Hall et al., 2005).

5.2 Breast Cancer and Air Pollution

Local vehicular traffic is the primary contributor to air pollution in urban areas, although there is usually a large regional component arising from long-range transport of pollutants from sources upwind (Akimoto, 2003; Galvez, 2007). An important component of local sources are vehicular emissions, which include gases, particles, volatile organic compounds, and polycyclic aromatic hydrocarbons (PAHs),
many of which are potential carcinogens. Benzene, for example, is present in gasoline and is a known human carcinogen (IARC 1987; US Department of Health and Human Services, 2005). Benzene has been shown to cause mammary carcinomas in rodents (US National Toxicology Program 1986; Maltoni et al., 1988; Huff et al., 1989). PAHs are lipophilic, which means that they are dissolved in fat and can therefore reach elevated concentrations in breast tissue and promote carcinogenesis in the cells of the breast (Morris and Seifter, 1992; Labrèche and Goldberg, 1997). In a review of the estrogenic and antiestrogenic activity of PAHs, Santodonato (1997, 842) reported that, with regards to aetiology of human breast cancer, current scientific literature “provides persuasive evidence for the hypothesis that certain carcinogenic PAHs produce a unique duality of pathologic effects encompassing both genotoxic and non-genotoxic components.”

A few studies have shown associations between the incidence of breast cancer and occupational exposure to benzene and PAHs (Petralia et al., 1999; Hansen, 2000; Gammon et al., 2002). Given that these same pollutants are present in vehicular exhaust and thus present in urban air pollution, it is plausible that traffic-related exposures may contribute to the incidence of breast cancer. Following this hypothesis, one study conducted in Nassau and Suffolk Counties, New York State, (Lewis-Michel et al., 1996) suggested a possible increased risk of breast cancer among postmenopausal women living near areas characterised by high traffic (adjusted odds ratio (OR) 1.29 and 95% confidence interval (CI), 0.77-2.15). Bonner et al. (2005) presented the first study to examine the association between risk of breast cancer and exposure to estimates of traffic-related air pollution as estimated by interpolated surfaces of total suspended particulates from fixed-site monitors. These authors used a case-control study of women in Erie and Niagara Counties, New York State, and found that among postmenopausal women, early life exposures to relatively high concentrations of particulates were associated with an increased risk of developing breast cancer (OR 2.42, 95% CI 0.97-6.09) compared with exposure to relatively low concentrations. In a second analysis of this study, estimates of residential exposures to benzo[a]pyrene were derived from a traffic
emissions model (Nie et al., 2007). Here, higher levels of exposure at the time when a woman first gave birth were associated with increased risk (OR 2.57, 95% CI 1.16-5.69) compared to those exposed at lower concentrations. Their findings were limited to non-smokers and they found no evidence that exposure at other periods of time was associated with increased risk.

The present study evaluates the hypothesis that exposure to air pollution may be implicated in the etiology of postmenopausal breast cancer. The objective of this study is to determine whether the incidence of postmenopausal breast cancer is associated with long-term exposure to intra-urban concentrations of nitrogen dioxide (NO\textsubscript{2}) – a marker for traffic-related pollution – after accounting for personal risk factors and occupational exposures.

5.3  Materials and Methods

5.3.1  Study Site Description

This study is set in Montreal, which is the second largest metropolitan area in Canada. The greater Montreal area has a population over 3.6 million, but our study is restricted to the Island of Montreal, which has a population of approximately 1.8 million people (Statistics Canada, 2006). Several busy expressways extend along the length of the Island and several large bridges connect the Island to the urban communities on both the south and north shores. Heavy industry, including a few large petrochemical facilities, a copper refinery, and a natural gas processing plant, is located in the eastern part of the Island. Previous studies have shown that levels of intra-urban air pollution vary spatially across the study area (Gilbert et al., 2005; Crouse et al., 2009a).

5.3.2  Design of the Epidemiological Study

This study uses a standard, hospital-based case-control design (Wacholder et al., 1992) by which incident cases of postmenopausal breast cancer were identified from hospitals and control subjects comprised women diagnosed with other selected types of cancer. The ethics committees of all participating hospitals and affiliated universities approved the research protocol, and informed consent was obtained
from participating subjects.

Details of the study were described in previous papers (Lenz et al., 2002; Labrèche et al., 2003; Labrèche et al., accepted). Briefly, the target population comprised postmenopausal women, aged 50-75 years at the time of diagnosis who were residents of the greater Montreal area in 1996 and 1997. Eligible case subjects were diagnosed between 1996 and 1997 with primary, malignant breast cancer (International Classification of Diseases, 9th revision, code 174) that was confirmed histologically. Cases were identified from all of the 18 hospitals in the region that treated breast cancer, thus ensuring almost complete coverage of the target population. Control subjects were selected at the same time and from the same hospitals as the cases, and had one of 32 other selected sites of incident, histologically-confirmed cancers. The controls were approximately frequency-matched to the cases by age and by hospital. Cancer controls were used in this study to minimize the potential for differential recall bias. It should be noted that the cancer sites among the control group are not known or suspected to be associated with ambient pollution.

Postmenopausal status was classified according to the World Health Organisation definition: women over the age of 50 who ceased menstruation naturally in the 12 months prior to interview or who ceased menstruating following a bilateral oophorectomy (WHO, 1981); women who were still menstruating but had used hormonal replacement therapy to alleviate symptoms of menopause four or more years before diagnosis; and, women who had a simple hysterectomy without oophorectomy and reported using hormonal replacement therapy.

One to three months after diagnosis, participants completed a structured questionnaire with content related to occupational history and other personal risk factors including, among others, reproductive history, educational attainment, family history of breast cancer, age at menarche, smoking and alcohol consumption, body mass index, and home address (and duration at that address) at the time of diagnosis. GeoPinpoint V6.4 (DMTI Spatial, Markham, ON) with ArcGIS 9.3 geographic information system (GIS) (ESRI, Redlands, CA) was used to geocode the address of
5.3.3 Estimating Occupational Exposures

Occupational exposures were estimated using a standard methodology (Gérin et al., 1985; Siemiatycki, 1990). Interviewers used a structured set of general and job-specific questionnaires (e.g., for hairdressers, nurses, textile workers) and were trained to probe for details regarding each occupation (and tasks performed) that the subject ever had in her working lifetime (Stewart et al., 1998).

The jobs provided by each participant were translated by a team of industrial hygienists and chemists into a set of exposure indices for about 300 chemical and physical agents. For each agent judged to be present, the team coded physical aspect, average duration of exposure (in hours) in a working day, percent of working days exposed during the period, confidence that there was actual exposure to each agent using a 4-point ordinal scale (probably no occupational exposure, and “low”, “medium”, and “high” confidence of exposure), and level of intensity. Intensity was assessed on a rank-ordered scale of low, medium and high intensity. The categories of intensity for selected substances were benchmarked using published occupational measurements, and exposure levels comparable to those observed in the general environment were deemed to be “unexposed”. For the present study, indices were included for exposure to four different substances that may be associated with breast cancer (Labrèche et al., accepted; Table 5-1). The indices were limited to exposures before the age of 36 years (the period during which breast tissue may be more susceptible to exogenous insults, as female breast cells continue to develop until that age (Anderson, 1999)).

5.3.4 Neighbourhood Contextual Confounders

Neighbourhood deprivation may be a confounding factor in the association between exposure to air pollution and the incidence of breast cancer because deprived populations often live in areas that are characterised by higher rates of air pollution (Jerrett et al., 2001; Premji et al., 2007; Crouse et al., 2009b; Havard et al., 2009). To estimate indicators of deprivation, Statistics Canada 1996 census data
were aggregated to the census tract level to describe socioeconomic characteristics of Montreal’s populations. Census tract boundaries are defined to delineate neighbourhoods whose residents are relatively homogeneous in terms of socioeconomic characteristics, economic status, and social living conditions, and the physical characteristics of the landscape. Variables describing median household income and percentage of adults who did not complete high school were compiled for the 350 census tracts that included the addresses of the cases and controls.

5.3.5 Assessment of Exposure to Traffic-related Air Pollution in Montreal

A dense sampling program of traffic-related air pollution was conducted during three periods in 2005 and 2006. The methods and findings were described previously (Crouse et al., 2009a). Briefly, the purpose of this sampling was to provide at a fine geographical scale, estimates of annual averages of NO$_2$. NO$_2$ is recognized as a good indicator of traffic-related pollution due to its co-locational association with other pollutants (Nieuwenhuijsen, 2000; Brunekreef and Holgate, 2002; Beckerman et al., 2008; Wheeler et al., 2008). In this case, any association observed between breast cancer and exposure may reflect an association with NO$_2$ specifically, or NO$_2$ may simply be acting as a marker for the causal agent or combination of causal agents found in the complex mixture of urban air pollution.

The locations of the samplers were selected using a population-weighted location-allocation model that placed samplers in areas likely to have high spatial variability in traffic-related pollution and high population densities (Kanaroglou et al., 2005). Approximately 20 samplers were added to capture concentrations in residential areas that appeared to be under-represented by the initial allocation scheme. Samplers were deployed at 133 locations across the Island of Montreal on three different occasions between 2005 and 2006: once in summer, once in winter, and once in spring. The sampling devices were two-sided Ogawa passive diffusion samplers that make use of triethanolamine-impregnated filters as an absorbent. They were analyzed at an Environment Canada laboratory using ion chromatography (Gilbert et al., 2003). The samplers were deployed in the field for two-week sampling periods. Valid observations were obtained from all three sampling periods at 129 locations.
Land Use Regression Modelling

In order to provide estimates of concentrations at locations where there were no samplers, a land use regression model was used to predicted concentrations of mean annual NO$_2$ at a resolution of 5 m across the Island of Montreal. Land use regression is based on a linear regression of the natural logarithm of NO$_2$ on land use and traffic-related variables, and thus provides a prediction model that can be visualized spatially in a GIS. Land use regression has been shown to be more effective for describing spatial variability than dispersion models and other spatial methods of interpolation (Briggs et al., 1997; Briggs et al., 2000; Lebret et al., 2000; Hoek et al., 2001). The goal of the modelling strategy was to maximize prediction, as measured by the model $R^2$. As described in Crouse et al. (2009a), the model included 47 predictor variables representing lengths of roads, traffic counts, population density, and area of different land uses across the study area and obtained an $R^2$ of 0.80.

Historical Estimates of Concentrations of NO$_2$

The purpose of the land use regression model in the context of this study was to link the exposure surface with addresses of cases and controls ascertained at the time of the study (i.e., 1996). It is possible that, if patterns of traffic and land use have not changed dramatically over time, this surface, based on data collected in 2005-06, may in fact represent the spatial distribution of pollution 10 years earlier. No previous studies in Montreal or elsewhere have evaluated whether spatial patterns of urban air pollution vary dramatically over time. It had been established, however, from measurements at Environment Canada's National Air Pollution Surveillance (NAPS) network of 13 fixed-site monitoring stations in Montreal (Figure 5-1), that mean regional concentrations of NO$_2$ have declined by about 4 parts per billion (ppb) over the past 20 years. Moreover, assuming that air pollution is implicated in the aetiology of breast cancer, the induction period may be quite long. Thus, in collaboration with Chen (Chen et al., 2009) a method was developed to extrapolate the land use regression model as far back in time as possible.
Making use of the daily concentrations of NO$_2$ from the 13 NAPS stations, it was possible to back-extrapolate the land use regression surface to 1996 and to 1985. The method used was as follows: mean annual concentrations of NO$_2$ were computed at 13 of these stations for each of 1985, 1996, and 2006. These estimates of NO$_2$ were then used to create inverse-distance weighted (IDW) spatial interpolation surfaces covering the Island of Montreal for each of the three periods. These spatial surfaces, which are based on observations at only 13 locations, were generated at 5 m resolution and describe the general spatial patterns across Montreal, but at lower accuracy than does the land use regression model.

The 1985-based interpolated surface was divided (i.e., grid cell by grid cell) by the 2006-based interpolated surface to calculate the ratio of the older concentrations and patterns of NO$_2$ to the newer ones. Then the land use regression model was multiplied, cell by cell, by this ratio to produce a re-weighted spatial surface of
estimated mean annual NO$_2$ for 1985. This step was repeated with the interpolated surface for 1996. Lastly, a fourth surface was created that was meant to describe the mean estimates of NO$_2$ during the ten-year period preceding diagnosis among the cases and controls, namely the period between 1985 and 1996. For this purpose, a final surface was created by calculating the mean of the re-weighted 1985 and 1996 surfaces. To assess the spatial variability between the four surfaces, 1000 randomly generated points were sampled on each one, and Pearson correlation coefficients were estimated between them.

**Extending the Geographic Extent of the Exposure Surfaces**

It was known in advance that the case control cohort included subjects from the greater Montreal region, and thus, that some subjects would have home addresses located outside the range of the pollution surfaces, which were restricted to the extent of the Island. Therefore, effort was made to create an additional pollution surface extended to the geographic extent of the greater Montreal region (i.e., the census metropolitan area) that could be used in a sensitivity analysis. Thus, the same predictor variables that were used in the original land use regression model (Crouse et al., 2009a) were used to generate estimates of concentrations of NO$_2$ for 2005-2006 for the whole Montreal region.

**5.3.6 Statistical Analysis**

Unconditional logistic regression was used in S-PLUS v5 to estimate odds ratios and associated 95% confidence intervals. In order not to lose subjects because of missing values of continuous non-occupational covariates, these potential confounding variables were modelled as categorical functions, whereby missing values were assigned to a “missing” category. Ultimately, only age at diagnosis and the two neighbourhood covariates were treated continuously, and the best fitting function was linear.

Both age-adjusted and fully-adjusted models are presented. The latter models included variables accepted to be associated with the incidence of breast cancer: age at diagnosis; family history; education; ethnicity; age at bilateral oophorectomy; age at
menarche; age at first full-term pregnancy; alcohol drinking; smoking; body mass index; and proxy respondent status. In addition, factors whose causal association with breast cancer is still uncertain were included as covariates: oral contraceptive use; smoking; total duration of breast-feeding; neighbourhood ecologic covariates, and selected occupational exposures before age 36 years (i.e., organic solvents with reactive metabolites, extremely low frequency magnetic fields, carbon monoxide, PAHs from petroleum) (Labrèche et al., 2003; Labrèche et al., 2010). Additionally, the models were adjusted for the hospital where the subjects were diagnosed because it was a frequency-matching variable. Standard regression diagnostics were applied to identify possible influential subjects and to ensure that the models did not violate the assumptions of logistic regression.

Associations for NO$_2$ are presented as odds ratios that were computed as a linear function in the logistic model. These odds ratios are presented for an increase in exposure to NO$_2$ of 5 ppb and for an increase across the inter-quartile range, as evaluated at each time period. (The inter-quartile range is the difference between the 75th percentile and the 25th percentile and provides a more robust comparison between the exposure surfaces given that they have different distributions). Odds ratios were estimated also for quartiles of the distribution of NO$_2$ amongst cases and controls. Additionally, odds ratios were computed with a reduced sample of the cohort including only those subjects who had been residents at the same address for ten years or more prior to diagnosis. This latter subset includes those subjects for whom there is the greatest confidence of long-term exposures at those locations.

5.4 Results

5.4.1 Exposure Surfaces for the Island of Montreal

The four exposure surfaces reflect the trend of decreasing regional concentrations of NO$_2$ over time, with the highest mean value observed in the surface for 1985 (i.e., 20.1 ppb) and the lowest in 2006 (i.e., 10.8 ppb) (Table 5-2). Additionally, the different surfaces reflected a narrowing of distributions of exposure over time. Although the ranges of concentrations of NO$_2$ varied between
the four exposure surfaces, they each exhibited similar spatial patterns (Figures 5-2, 5-3, 5-4, 5-5). Furthermore, there were strong positive Pearson correlation coefficients between the randomly sampled locations in each season (i.e., 0.97 – 0.99). These strong correlations suggest that, according to these models, the spatial patterns of NO₂ did not vary significantly during the 20-year period between 1985 and 2006.
Table 5-1: Distribution of risk factors and age-adjusted odds ratios (ORs) and 95% confidence intervals (CI), postmenopausal breast cancer, Montreal, Canada, n = 799, (NE = not estimated)

<table>
<thead>
<tr>
<th>Variables (reference category)</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal Risk Factors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother or sister with breast cancer (no)</td>
<td>210   54.8</td>
<td>279 67.1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>83    21.7</td>
<td>47 11.3</td>
<td>2.36</td>
<td>1.58 – 3.52</td>
</tr>
<tr>
<td>No for mother, no sisters</td>
<td>79    20.6</td>
<td>76 18.3</td>
<td>1.40</td>
<td>0.97 – 2.01</td>
</tr>
<tr>
<td>Missing</td>
<td>11    2.9</td>
<td>14 3.4</td>
<td>1.13</td>
<td>0.50 – 2.55</td>
</tr>
<tr>
<td>Oophorectomy (Never had an ovary removed)</td>
<td>276   72.1</td>
<td>194 46.6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Only one ovary removed</td>
<td>27    7</td>
<td>30 7.2</td>
<td>0.61</td>
<td>0.35 – 1.07</td>
</tr>
<tr>
<td>Age at bilateral oophorectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;44</td>
<td>26    6.8</td>
<td>34 8.2</td>
<td>0.55</td>
<td>0.32 – 0.95</td>
</tr>
<tr>
<td>45-49</td>
<td>19    5</td>
<td>15 3.6</td>
<td>0.89</td>
<td>0.44 – 1.79</td>
</tr>
<tr>
<td>50-54</td>
<td>16    4.2</td>
<td>26 6.3</td>
<td>0.39</td>
<td>0.20 – 0.76</td>
</tr>
<tr>
<td>≥55</td>
<td>17    4.4</td>
<td>106 25.5</td>
<td>0.12</td>
<td>0.07 – 0.20</td>
</tr>
<tr>
<td>Missing</td>
<td>2     0.5</td>
<td>11 2.6</td>
<td>0.13</td>
<td>0.03 – 0.59</td>
</tr>
<tr>
<td>Years of education (≤7)</td>
<td>107   27.9</td>
<td>154 37</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8-10</td>
<td>78    20.4</td>
<td>107 25.7</td>
<td>1.05</td>
<td>0.72 – 1.54</td>
</tr>
<tr>
<td>11-17</td>
<td>178   46.5</td>
<td>138 33.2</td>
<td>1.82</td>
<td>1.30 – 2.53</td>
</tr>
<tr>
<td>≥18</td>
<td>20    5.2</td>
<td>17 4.1</td>
<td>1.58</td>
<td>0.78 – 3.17</td>
</tr>
<tr>
<td>Ethnicity (French)</td>
<td>232   60.6</td>
<td>219 52.6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>English &amp; Others</td>
<td>98    25.6</td>
<td>110 26.4</td>
<td>0.86</td>
<td>0.62 – 1.20</td>
</tr>
<tr>
<td>Jewish &amp; Italian</td>
<td>53    13.8</td>
<td>87 20.9</td>
<td>0.58</td>
<td>0.39 – 0.86</td>
</tr>
<tr>
<td>Age at menarche (≥16)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-15</td>
<td>97    25.3</td>
<td>109 26.2</td>
<td>1.47</td>
<td>0.82 – 2.62</td>
</tr>
<tr>
<td>13</td>
<td>110   28.7</td>
<td>100 24</td>
<td>1.82</td>
<td>1.02 – 3.24</td>
</tr>
<tr>
<td>12</td>
<td>78    20.4</td>
<td>98 23.6</td>
<td>1.28</td>
<td>0.71 – 2.30</td>
</tr>
<tr>
<td>≤11</td>
<td>74    19.3</td>
<td>70 16.8</td>
<td>1.65</td>
<td>0.90 – 3.02</td>
</tr>
</tbody>
</table>
Table 5-1, Continued

<table>
<thead>
<tr>
<th>Variables (reference category)</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Age at first full-term pregnancy (never pregnant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;18</td>
<td>15</td>
<td>3.9</td>
<td>12</td>
<td>2.9</td>
</tr>
<tr>
<td>≥18-26</td>
<td>166</td>
<td>43.3</td>
<td>214</td>
<td>51.4</td>
</tr>
<tr>
<td>&gt;26-30</td>
<td>59</td>
<td>15.4</td>
<td>56</td>
<td>13.5</td>
</tr>
<tr>
<td>&gt;30</td>
<td>46</td>
<td>12</td>
<td>36</td>
<td>8.7</td>
</tr>
<tr>
<td>Pregnant but never full-term</td>
<td>14</td>
<td>3.7</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>Breastfeeding, cumulative duration in weeks (0)</td>
<td>296</td>
<td>77.3</td>
<td>303</td>
<td>72.8</td>
</tr>
<tr>
<td>&gt;0-80</td>
<td>69</td>
<td>18</td>
<td>77</td>
<td>18.5</td>
</tr>
<tr>
<td>&gt;80</td>
<td>18</td>
<td>4.7</td>
<td>36</td>
<td>8.7</td>
</tr>
<tr>
<td>Years of oral contraception use (never)</td>
<td>248</td>
<td>64.8</td>
<td>290</td>
<td>69.7</td>
</tr>
<tr>
<td>&lt;1</td>
<td>42</td>
<td>11</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>≥1</td>
<td>93</td>
<td>24.3</td>
<td>101</td>
<td>24.3</td>
</tr>
<tr>
<td>Hormone replacement therapy, duration in months (0)</td>
<td>170</td>
<td>44.4</td>
<td>236</td>
<td>56.7</td>
</tr>
<tr>
<td>1-19</td>
<td>53</td>
<td>13.8</td>
<td>57</td>
<td>13.7</td>
</tr>
<tr>
<td>20-44</td>
<td>29</td>
<td>7.6</td>
<td>34</td>
<td>8.2</td>
</tr>
<tr>
<td>45-74</td>
<td>35</td>
<td>9.1</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>75-99</td>
<td>16</td>
<td>4.2</td>
<td>13</td>
<td>3.1</td>
</tr>
<tr>
<td>≥100</td>
<td>80</td>
<td>20.9</td>
<td>55</td>
<td>13.2</td>
</tr>
<tr>
<td>Body mass index (&lt;21)</td>
<td>45</td>
<td>11.7</td>
<td>54</td>
<td>13</td>
</tr>
<tr>
<td>&gt;21-23.5</td>
<td>75</td>
<td>19.6</td>
<td>96</td>
<td>23.1</td>
</tr>
<tr>
<td>&gt;23.5-27</td>
<td>132</td>
<td>34.5</td>
<td>112</td>
<td>26.9</td>
</tr>
<tr>
<td>&gt;27-35</td>
<td>116</td>
<td>30.3</td>
<td>126</td>
<td>30.3</td>
</tr>
<tr>
<td>&gt;35</td>
<td>15</td>
<td>3.9</td>
<td>28</td>
<td>6.7</td>
</tr>
<tr>
<td>Tobacco exposure (none)</td>
<td>45</td>
<td>11.7</td>
<td>52</td>
<td>12.5</td>
</tr>
<tr>
<td>Exposure to environmental tobacco smoke only</td>
<td>160</td>
<td>41.8</td>
<td>148</td>
<td>35.6</td>
</tr>
<tr>
<td>Active smoker with or without exposure to environmental tobacco smoke</td>
<td>176</td>
<td>46</td>
<td>209</td>
<td>50.2</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>0.5</td>
<td>7</td>
<td>1.7</td>
</tr>
<tr>
<td>Respondent (self)</td>
<td>362</td>
<td>94.5</td>
<td>362</td>
<td>87</td>
</tr>
<tr>
<td>Proxy</td>
<td>21</td>
<td>5.5</td>
<td>54</td>
<td>13</td>
</tr>
</tbody>
</table>
### Table 5-1, Continued

<table>
<thead>
<tr>
<th>Variables (reference category)</th>
<th>Cases</th>
<th>Controls</th>
<th>OR</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcohol status (never drinker)</td>
<td>193</td>
<td>229</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Former drinker</td>
<td>50</td>
<td>55</td>
<td>1.01</td>
<td>0.66 – 1.57</td>
</tr>
<tr>
<td>Infrequent drinker</td>
<td>53</td>
<td>53</td>
<td>1.15</td>
<td>0.75 – 1.76</td>
</tr>
<tr>
<td>Current drinker</td>
<td>87</td>
<td>79</td>
<td>1.26</td>
<td>0.88 – 1.81</td>
</tr>
<tr>
<td>Benign breast disease (no) (OR not age-adjusted)</td>
<td>197</td>
<td>333</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>185</td>
<td>83</td>
<td>3.77</td>
<td>2.75 – 5.15</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0</td>
<td>NE</td>
<td>NE</td>
</tr>
</tbody>
</table>

**Occupational Exposures**

<table>
<thead>
<tr>
<th>Solvents with reactive metabolites (not exposed)</th>
<th>338</th>
<th>374</th>
<th>1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-substantial (5 years)</td>
<td>5</td>
<td>4</td>
<td>1.35</td>
<td>0.36 – 5.09</td>
</tr>
<tr>
<td>Substantial (5 years)</td>
<td>6</td>
<td>5</td>
<td>1.26</td>
<td>0.38 – 4.17</td>
</tr>
<tr>
<td>Exposed only at R=1</td>
<td>12</td>
<td>9</td>
<td>1.43</td>
<td>0.60 – 3.45</td>
</tr>
<tr>
<td>Others</td>
<td>22</td>
<td>24</td>
<td>0.99</td>
<td>0.55 – 1.81</td>
</tr>
<tr>
<td>Extremely low frequency magnetic fields (not exposed)</td>
<td>81</td>
<td>102</td>
<td>24.5</td>
<td>1</td>
</tr>
<tr>
<td>Non-substantial (5 years)</td>
<td>107</td>
<td>96</td>
<td>1.30</td>
<td>0.86 – 1.97</td>
</tr>
<tr>
<td>Substantial (5 years)</td>
<td>48</td>
<td>44</td>
<td>1.30</td>
<td>0.78 – 2.17</td>
</tr>
<tr>
<td>Exposed only at R=1</td>
<td>8</td>
<td>18</td>
<td>0.57</td>
<td>0.24 – 1.37</td>
</tr>
<tr>
<td>Others</td>
<td>139</td>
<td>156</td>
<td>1.06</td>
<td>0.73 – 1.55</td>
</tr>
<tr>
<td>Carbon Monoxide (not exposed)</td>
<td>299</td>
<td>339</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Non-substantial (5 years)</td>
<td>36</td>
<td>21</td>
<td>1.87</td>
<td>1.07 - 3.27</td>
</tr>
<tr>
<td>Substantial (5 years)</td>
<td>0</td>
<td>0</td>
<td>NE</td>
<td>NE</td>
</tr>
<tr>
<td>Exposed only at R=1</td>
<td>1</td>
<td>1</td>
<td>1.18</td>
<td>0.07 - 19.10</td>
</tr>
<tr>
<td>Others</td>
<td>47</td>
<td>55</td>
<td>1.32</td>
<td>0.62 - 1.43</td>
</tr>
<tr>
<td>Polycyclic Aromatic Hydrocarbons from Petroleum (not exposed)</td>
<td>354</td>
<td>386</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Non-substantial (5 years)</td>
<td>9</td>
<td>4</td>
<td>2.34</td>
<td>0.72 - 7.63</td>
</tr>
<tr>
<td>Substantial (5 years)</td>
<td>2</td>
<td>2</td>
<td>1.07</td>
<td>0.15 - 7.69</td>
</tr>
<tr>
<td>Exposed only at R=1</td>
<td>2</td>
<td>8</td>
<td>0.28</td>
<td>0.64 - 1.22</td>
</tr>
<tr>
<td>Others</td>
<td>16</td>
<td>16</td>
<td>1.03</td>
<td>0.51 - 2.10</td>
</tr>
</tbody>
</table>
Table 5-1, Continued

<table>
<thead>
<tr>
<th>Ecologic Covariates, 1996</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median household income in Canadian dollars</td>
<td>243 34,239</td>
<td>251 34,436</td>
</tr>
</tbody>
</table>

| Percentage of adults (aged 15 years and older) without high school diploma | 243 32.6 | 251 32.8 | 1.00 | 0.99 – 1.01 |

Table 5-2: Distributions of concentrations of NO$_2$ in 4 different exposure surfaces, Island of Montreal, Canada

<table>
<thead>
<tr>
<th>Model Year</th>
<th>Minimum</th>
<th>25th Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>75th Quartile</th>
<th>Maximum</th>
<th>Inter-quartile Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>4.2</td>
<td>8.9</td>
<td>10.3</td>
<td>10.8</td>
<td>12.2</td>
<td>35.9</td>
<td>3.3</td>
</tr>
<tr>
<td>1996</td>
<td>5.9</td>
<td>12.9</td>
<td>15.2</td>
<td>15.7</td>
<td>18.0</td>
<td>44.5</td>
<td>5.1</td>
</tr>
<tr>
<td>1985</td>
<td>7.9</td>
<td>16.5</td>
<td>19.4</td>
<td>20.1</td>
<td>23.1</td>
<td>66.8</td>
<td>6.6</td>
</tr>
<tr>
<td>1985-1996</td>
<td>7.0</td>
<td>14.8</td>
<td>17.2</td>
<td>17.9</td>
<td>20.5</td>
<td>55.6</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Figure 5-2: Surface map of mean annual concentrations of NO$_2$ in Montreal, 2005-2006
Figure 5-3: Surface map of mean annual concentrations of NO$_2$ in Montreal, 1996

Figure 5-4: Surface map of mean annual concentrations of NO$_2$ in Montreal, 1985
5.4.2 General Characteristics of the Breast Cancer Dataset

A total of 1,631 subjects were potentially eligible for this study. Interviews were conducted among 608 cases and 667 control subjects, thus obtaining response rates of 81.1% for cases and 75.7% for controls. Of these 1,275 participants, 106 were deemed to be premenopausal and were therefore removed from the study. A further 79 participants had incomplete or inaccurate address information. Finally, 291 subjects were geocoded but resided off the Island of Montreal. Thus, the present study is based on 799 subjects: 383 cases and 416 controls (Figure 5-6); as described below, however, an additional analysis was conducted with the 291 subjects who resided off the Island.

The most frequent sites of cancer in the control series were colon (21.6%), uterus (19.0%), ovaries (9.1%), rectum (6.7%), bladder (6.0%), and kidneys (5.8%). The analyses suggested associations for the generally accepted risk factors for postmenopausal breast cancer (i.e., family history, benign breast disease, education,
age at menarche, duration of hormonal replacement therapy). There was weak evidence of association with the number of full-term pregnancies, cumulative duration of breastfeeding, and body mass index, and there was little evidence of an association for oral contraceptive use. There were essentially no differences in characteristics between the home neighbourhoods of the cases and controls. For example, median neighbourhood household income and percentage of adults without high school diplomas was $34,239 and 32.6 respectively among cases, compared to $34,436 and 32.8 among the controls. The distributions and age-adjusted odds ratios for all risk factors are presented in Table 5-1.

Figure 5-6: Spatial distribution of the cases and controls across Montreal

5.4.3 Associations Between Postmenopausal Breast Cancer and Exposure to Air Pollution

Through the use of natural cubic splines and visual inspection it was identified that the response function for NO$_2$ was linear. Odds ratios were computed using each of the four models of exposure (Table 5-3). In the age-adjusted models, the odds ratios for an increase of 5 ppb of NO$_2$ (OR$_{5ppb}$) ranged from a low of 1.05
(95% CI 0.91 – 1.22) using the model extrapolated to 1985 to a high of 1.15 (95% CI 0.89 – 1.48) using exposure estimates from 2006. In the fully-adjusted models, the OR\textsubscript{5ppb} ranged from a low of 1.16 (95% CI 0.95 – 1.44) using the model extrapolated to 1985 to a high of 1.37 (95% CI 0.95 – 1.97) in 2006.

Table 5-3: Associations between exposure to NO\textsubscript{2} and postmenopausal breast cancer, Island of Montreal, Canada (n = 799)

<table>
<thead>
<tr>
<th>Exposure surface</th>
<th>Age-adjusted OR\textsubscript{5ppb}</th>
<th>95% CI</th>
<th>Fully-adjusted OR\textsubscript{5ppb}</th>
<th>95% CI</th>
<th>Fully-adjusted OR\textsubscript{inter-quartile range}</th>
<th>95% CI</th>
<th>Fully-adjusted OR\textsubscript{Q4 vs. Q1}</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>1.15</td>
<td>0.89-1.48</td>
<td>1.37</td>
<td>0.95-1.67</td>
<td>1.23</td>
<td>0.97-1.56</td>
<td>1.87</td>
<td>1.05-3.33</td>
</tr>
<tr>
<td>1996</td>
<td>1.10</td>
<td>0.91-1.32</td>
<td>1.32</td>
<td>1.01-1.73</td>
<td>1.33</td>
<td>1.01-1.75</td>
<td>2.33</td>
<td>1.29-4.21</td>
</tr>
<tr>
<td>Mean of 1996 and 1985</td>
<td>1.07</td>
<td>0.91-1.27</td>
<td>1.23</td>
<td>0.97-1.46</td>
<td>1.26</td>
<td>0.97-1.66</td>
<td>1.82</td>
<td>1.03-3.32</td>
</tr>
<tr>
<td>1985</td>
<td>1.05</td>
<td>0.91-1.22</td>
<td>1.16</td>
<td>0.95-1.44</td>
<td>1.23</td>
<td>0.93-1.64</td>
<td>1.75</td>
<td>1.40-3.11</td>
</tr>
</tbody>
</table>

* adjusted for hospital and all personal risk factors, occupational exposures, and neighborhood ecologic covariates listed in Table 5-1

* inter-quartile range presented in Table 5-2

* comparing those in the highest quartile of exposure (Q4) relative to those in the lowest quartile of exposure (Q1)

Given that each of the exposure surfaces had different distributions and mean concentrations of NO\textsubscript{2}, odds ratios associated with increase in exposure by inter-quartile range (IQR) were computed also, with the IQR computed separately for each surface. The ORs calculated per IQR were less variable between exposure periods compared to those computed per 5 ppb. Here, the estimates of risk in fully-adjusted models varied between 1.23 (95% CI 0.93 – 1.61) in 1985 and 1.33 (1.01 – 1.75) in 1996.

Additionally, the estimates of risk of breast cancer were computed by comparing those subjects in the highest quartile of exposure relative to those in the lowest quartile of exposure. These models produced the highest estimates of risk and included statistically significant associations. In the case of the model extrapolated to 1996, the fully-adjusted estimate of risk was 2.33 (95% CI 1.29 – 4.21). Similarly, the increased risk associated with exposure to estimates in 2005-06 was 1.87 (95% CI 1.05 – 3.33).

Finally, estimates of risk of breast cancer were computed among only those women in the study who had been residents at the same address for at least ten years prior to diagnosis. This reduced cohort of 408 subjects comprised 195 cases and 213 controls. Here the odds ratios were slightly larger than those produced with the full cohort, although the ranges of the 95% confidence intervals were significantly
wider (Table 5-4). For example, in the fully-adjusted models, the OR_{5ppb} ranged from a low of 1.23 (95% CI 0.87 – 1.75) using estimates for 1985 to a high of 1.53 (95% CI 0.82 – 2.85) in 2006. By comparing those subjects in the highest quartiles of exposure in 1996 relative to those in the lowest, the estimate of risk was 4.19 (95% CI 1.39 – 12.67). The higher odds ratios among these models perhaps reflect a dose-response relationship with exposures over time.

Table 5-4: Associations between exposure to NO\textsubscript{2} and postmenopausal breast cancer, Island of Montreal, Canada: including only subjects who were residents at the same address for at least ten years prior to diagnosis (n = 408)

<table>
<thead>
<tr>
<th>Exposure surface</th>
<th>Age-adjusted\textsuperscript{a}</th>
<th>Fully-adjusted\textsuperscript{b}</th>
<th>Fully-adjusted\textsuperscript{c}</th>
<th>Fully-adjusted\textsuperscript{d}</th>
</tr>
</thead>
<tbody>
<tr>
<td>per 5ppb OR &amp; 95% CI</td>
<td>per 5ppb OR &amp; 95% CI</td>
<td>inter-quartile range &amp; 95% CI</td>
<td>Q4 vs. Q1 &amp; 95% CI</td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>1.08 (0.75-1.56)</td>
<td>1.53 (0.62-2.85)</td>
<td>1.32 (0.98-2.00)</td>
<td>2.47 (0.88-7.29)</td>
</tr>
<tr>
<td>1996</td>
<td>1.02 (0.78-1.34)</td>
<td>1.35 (0.84-2.17)</td>
<td>1.36 (0.84-2.20)</td>
<td>4.19 (1.39-12.67)</td>
</tr>
<tr>
<td>Near of 1996 and 1985</td>
<td>1.04 (0.82-1.32)</td>
<td>1.29 (0.86-1.93)</td>
<td>1.33 (0.84-2.12)</td>
<td>2.62 (0.72-9.63)</td>
</tr>
<tr>
<td>1985</td>
<td>1.04 (0.85-1.29)</td>
<td>1.23 (0.87-1.75)</td>
<td>1.32 (0.83-2.09)</td>
<td>2.21 (0.76-6.45)</td>
</tr>
</tbody>
</table>

\textsuperscript{a} adjusted for hospital and all personal risk factors, occupational exposures, and neighbourhood ecologic covariates listed in Tables 5-1

\textsuperscript{b} inter-quartile range presented in Table 5-2

\textsuperscript{c} comparing those in the highest quartile of exposure (Q4) relative to those in the lowest quartile of exposure (Q1)

5.4.4 Model Diagnostics

For the fully-adjusted models, effort was made to identify outliers and cases that may influence the overall fit and performance of the models. Hosmer and Lemeshow (2000) and Menard (2002) both recommend plotting leverage values, squared Pearson residuals, and squared Studentized residuals against the predicted values to detect outliers and to identify cases for which the model fits poorly. In Figure 5-7, one case is identified that lies significantly away from the rest of the data points (i.e., case 2487). In Figures 5-8 and 5-9, three separate cases appear to be poorly fit by the model (i.e., cases 2976, 1257, 2093). The plots show that aside from these few exceptions the model fits reasonably well. As shown in Table 5-5, however, there are no substantial changes in model fit or in estimated odds ratios when any of the cases are deleted. Further inspection of the data suggest that although the characteristics of these four cases may be unusual, they are not implausible, and thus, all four were retained.
Figure 5-7: Scatter plot of Pregibon leverage vs. the estimated probability from the fully-adjusted model using estimates of NO$_2$ from 2005-2006, n = 799

Figure 5-8: Scatter plot of Studentized residuals squared vs. the estimated probability from the fully-adjusted model using estimates of NO$_2$ from 2005-2006, n = 799
Figure 5-9: Scatter plot of Pearson residuals squared vs. the estimated probability from the fully-adjusted model using estimates of NO\textsubscript{2} from 2005-2006, n = 799

Table 5-5: Model diagnostics for fully-adjusted model using estimates of NO\textsubscript{2} in 2005-06 with and without selected cases deleted, Island of Montreal, Canada

<table>
<thead>
<tr>
<th></th>
<th>Full dataset</th>
<th>2487</th>
<th>2976</th>
<th>1257</th>
<th>2093</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Goodness of fit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2 Log likelihood</td>
<td>819.30</td>
<td>819.30</td>
<td>812.49</td>
<td>813.32</td>
<td>813.48</td>
</tr>
<tr>
<td>Chi-square</td>
<td>286.99</td>
<td>285.52</td>
<td>292.33</td>
<td>291.49</td>
<td>291.33</td>
</tr>
<tr>
<td>Hosmer &amp; Lemeshow test</td>
<td>4.26</td>
<td>4.26</td>
<td>3.15</td>
<td>3.45</td>
<td>4.08</td>
</tr>
<tr>
<td><strong>Pseudo R\textsuperscript{2}</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cox &amp; Snell</td>
<td>0.30</td>
<td>0.30</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>Nagelkerke</td>
<td>0.40</td>
<td>0.40</td>
<td>0.41</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>Odds Ratio</strong></td>
<td>1.32</td>
<td>1.32</td>
<td>1.34</td>
<td>1.36</td>
<td>1.34</td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.95–1.85)</td>
<td>(0.95–1.85)</td>
<td>(0.95–1.87)</td>
<td>(0.97–1.91)</td>
<td>(0.96–1.88)</td>
</tr>
</tbody>
</table>
5.4.5 Subjects Residing Off the Island

Given that 291 subjects (i.e., 148 cases and 143 controls) resided outside the geographic extent of the original pollution surface, concentrations of NO$_2$ were predicted additionally across the full geographic extent of the greater Montreal region for 2005-2006. In this way, these additional subjects could be included in sensitivity analyses. First, there was no association with the incidence of breast cancer and exposure to NO$_2$ in the fully-adjusted model that included only the 291 subjects living not on the Island (OR$_{5\text{ppb}}$ 0.46; 95% CI 0.18 – 1.16). Second, by including these subjects in the full cohort (n = 1090), the fully-adjusted odds ratio per 5 ppb was 1.13 (95% CI 0.86 – 1.48). This latter estimate is significantly lower than that shown with the cohort that included only subjects living on the Island.

5.5 Discussion and Conclusions

5.5.1 Principal Findings

The principal finding of this study is the identification of a possible causal connection between exposure to traffic-related air pollution and the incidence of postmenopausal breast cancer in Montreal. Similar estimates of risk were associated with the inter-quartile range of exposure to NO$_2$ using estimates of exposure from models based on four different time periods. Moreover, estimates of risk were of similar magnitude to those computed with the full cohort when only those women who had been residing at the same address for at least one decade prior to diagnosis were considered. It should be noted that the estimates of exposure from each of the extrapolated historical surfaces have different limitations and advantages over each other related to temporal appropriateness and accuracy. It is difficult to determine which of these exposure surfaces provides the single most appropriate estimates of exposure for the purposes of this study. Nonetheless, the fact that a number of different approaches were used and the estimates of risk in each case were of similar magnitude provide compelling evidence of a true association between exposure to intra-urban concentrations of air pollution and the incidence of postmenopausal breast cancer. Although the specific odds ratios vary among the different estimates.
of exposure, there does appear to be an increased risk of approximately 25 percent for every increase of 5 ppb in exposure to NO$_2$, regardless of the method used. Alternatively, there is about 1.75 to 2-fold increased risk for postmenopausal breast cancer among those women exposed to the highest concentrations of ambient NO$_2$ in Montreal as compared to those exposed to the lowest concentrations.

### 5.5.2 Strengths and Limitations

The quality of the datasets used in the analyses, in particular the case-control dataset and the air pollution exposure estimates, is a strength of this study. The response rates for participation in the study among both cases and controls were exceptionally high, thus limiting the likelihood of selection bias. Given that controls were selected among other cancer patients reduces the possibility of recall bias affecting responses to the questionnaires. Furthermore, there is no possibility for recall bias in the exposure of interest, and the logistic models were adjusted for potential confounding for many personal, as well as occupational, risk factors. As noted earlier (and as published elsewhere, see Lenz et al., 2002; Labrèche et al., 2003), there were associations—of similar magnitude to those found in other studies—for most accepted risk factors for postmenopausal breast cancer, which further validates the reliability of this dataset. The land use regression exposure surfaces produced some of the best predictions in the literature. Additionally, the exposure surfaces were computed at a very high resolution (i.e., 5m), thus providing very local estimates of concentrations of NO$_2$.

There is less confidence in the accuracy of the estimates of NO$_2$ predicted for the regions off the Island, however. Some areas are suburban and/or semi-rural, and so they do not reflect the same mix of land uses, and densities of traffic and population present on the Island. Areas further away from the city are increasingly less-representative of the land use and traffic patterns found there; thus the estimates of NO$_2$ also are increasingly less reliable at greater distances from the city. Furthermore, no samples of concentrations of NO$_2$ were collected at anywhere off the Island. As has been shown by Jerrett et al. (2005a) and Poplawski et al. (2009), the transfer of land use regression models to locations characterised by dissimilar
land use patterns and densities of population can lead to incorrect predictions of concentrations of NO$_2$. In this context, the attenuated estimate of association reported in the sensitivity analysis with the off-Island subjects is likely a reflection of inaccuracies.

Two limitations of this study are related to population mobility: people do not necessarily live in the same home over the course of their lifetime, nor do they spend all of their time at home. While it is true that some subjects may work or spend their days away from home, a study by Leech et al. (2002) found that Canadian adults spend on average ~67-68% of their time at home (indoors and outdoors combined). Older, retired adults may spend even more time at home than employed, younger adults. As such, this study adopts the home address as the exposure proxy with the limitations that this choice necessarily entails.

High correlations have been observed between ambient (outdoor) and personal (indoor) exposures to PM$_{10}$ and PM$_{2.5}$ (Brunekreef et al., 2005). This relationship, however, is less clear in the case of ultrafine particles (i.e., PM$_{0.1}$) (Pekkanen and Kulmala, 2004) and NO$_2$ (Quackenboss et al., 1986; Rijnders et al., 2001). In light of this, it is acknowledged that the present analysis provides only partial information on personal exposure to air pollution. Total personal exposure relates to a complex number of factors and behaviours, including daily activity patterns and amount of time spent indoors and outdoors, among others.

Given the inherent imprecision associated with geocoded addresses and other geographic data, the risk estimates presented here are almost certainly subject to non-differential misclassification bias. That is, some subjects may be placed in incorrect exposure categories due to misplaced address points and/or inaccuracies and imprecision in the spatial datasets. This problem leads to an equal likelihood among cases and controls of being assigned inaccurate estimates of exposure, which tends to bias measures of association toward the null. It is thus likely that the results of this study underestimate the true estimates of risk of postmenopausal breast cancer associated with exposure to air pollution in this population.
5.5.3 Implications and Unanswered Questions

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 (Goldberg, 2007). Components of air pollution are made up of a variety of pollutants that derive mostly from fixed and moving combustion sources and from other non-anthropogenic sources. Moreover, there can be extensive spatial variability within a given city. Goldberg (2007) has argued that it is likely that the causal agent in this mixture cannot be identified. This implies that individual pollutants should be considered as markers for the mixture. As described earlier, the association between breast cancer and NO\(_2\) may reflect a true causal association, or NO\(_2\) may be a marker for the causal agent or combination of causal agents. Furthermore, it remains unclear whether exposure to air pollution may be an initiator or a promoter of breast cancer.

In summary, this study has demonstrated a possible connection between exposures to traffic-related air pollution and the incidence of postmenopausal breast cancer, in a city (Montreal) that has relatively low regional concentrations of pollution by international standards. These findings are similar to those reported in the two other studies that have examined the hypothesis that breast cancer may be associated with environmental exposures to air pollution (i.e., Bonner et al., 2005; Nie et al., 2007). The results presented here differ somewhat, however, from those studies. Among postmenopausal women, those authors found relatively strong associations between early life exposure to particulates (as markers of air pollution) and developing breast cancer, however, they found no associations with exposures 10 and 20 years prior to diagnosis. This study was unable to assess associations with early life exposures, nor does it show what age periods may be critical in the induction of cancer. The analysis of occupational exposures (Labrèche et al. 2010) suggests that exposures to some compounds (such as PAHs) before the age of 36 may be more important. If this is the case, then it is possible that the risks observed here are attenuated. Studies are needed to verify whether these results do represent
true associations or whether they are due to chance or to undetected bias. If these associations are indeed verified, additional studies should explore potential critical periods of exposure to air pollution in relation to the development of breast cancer.
CHAPTER 6

SUMMARY AND CONCLUSIONS

This dissertation has examined associations between spatial patterns of traffic-related air pollution in Montreal, and both conditions of neighbourhood deprivation and the incidence of postmenopausal breast cancer. This chapter concludes the dissertation by reviewing key findings and contributions of the dissertation, acknowledging some of the limitations of the research, and by discussing directions for future research.

6.1 Key Findings and Contributions

This dissertation contributes new methods and substantive contributions to academic scholarship in health geography, environmental epidemiology, spatial modelling of ambient pollution, and breast cancer aetiology. The key findings are shaped by the three research objectives of this dissertation, namely to:

1. Describe spatial and seasonal patterns of concentrations of intra-urban ambient air pollution in Montreal;

2. Identify spatial associations between concentrations of ambient air quality and indicators of social and material deprivation at the neighbourhood scale across Montreal; and,

3. Determine whether the incidence of postmenopausal breast cancer is associated with exposure to local concentrations of air pollution after accounting for individual-level risk factors and occupational exposures.

The first manuscript addressed the first objective and described the sampling of air pollution in Montreal, which consisted of collecting samples in three seasons at 129 locations across the Island. Previous studies had used considerably less-dense
networks of samplers and almost none had examined seasonal patterns of variability. This study is among the first to demonstrate the spatial variability of intra-urban concentrations of NO$_2$ across seasons. In this case there were significant differences in mean concentrations of NO$_2$ between the three seasons, however, the spatial patterns did not vary significantly. Furthermore, there was significantly more variability in mean concentrations of NO$_2$ in temperate weather as compared to “cold” and “hot” weather. Lastly, this study presented a new, prediction-based approach to modelling concentrations of NO$_2$ that achieved superior results (e.g., $R^2 = 0.80$) to those achieved by other researchers using different approaches.

The second manuscript addressed the second research objective and examined spatial associations between concentrations of ambient NO$_2$ and conditions of material and social deprivation at the neighbourhood scale across Montreal. Overall, the general patterns reflected a triple burden of social, material, and environmental deprivation for many residents of Montreal. For example, there were statistically significant associations between neighbourhood-scale concentrations of air pollution and unemployment rate (Pearson's $r = 0.32$, 95% CI 0.23 – 0.39) and the percentage of low-income households ($r = 0.40$, 95% CI 0.32 – 0.47). Moreover, this second manuscript makes a unique geographical contribution to the social/environmental justice literature by examining the specific social geographical characteristics that help us understand these findings. Because of the unique social geography and historical residential patterns in Montreal there were several neighbourhoods where excess exposures to air pollution crossed social and economic boundaries. The results of this second manuscript underscore the importance of considering social contexts in interpreting general associations between social and environmental risks to population health.

The third manuscript addressed the third research objective and extended the work presented in the first two manuscripts by assigning exposures from the pollution surface to subjects in an existing cohort of breast cancer cases (n=383) and controls (n=416). Most importantly, the third manuscript identified a possible causal connection between the incidence of postmenopausal breast cancer and exposure to
intra-urban concentrations of NO\(_2\) across Montreal, after controlling for personal risk factors and occupational exposures to other substances. For each increase of 5 ppb of NO\(_2\) in 2006 the OR was 1.37 (95% CI 0.95 – 1.97). By comparing subjects in the highest quartiles of exposure relative to those in the lowest quartile of exposure using estimates back-extrapolated to 1996, the fully-adjusted estimate of risk was 2.33 (95% CI 1.29 – 4.21). These associations were demonstrated using a variety of historical estimates of exposure in a city (i.e., Montreal) that is characterized by relatively low levels of pollution compared to other Canadian and international cities of comparable size. The strengths of the research presented in the third manuscript relate to the high quality of the population-based case control dataset (i.e., it included almost every woman diagnosed with breast cancer in Montreal between 1996 and 1997 and included numerous personal and occupational covariates known to be associated with breast cancer risk) and to the high quality of the estimates of exposures (i.e., estimates produced from a spatial model with a very high level of accuracy and computed at a 5 m resolution). Only a few other studies have examined the possible link between breast cancer and exposure to ambient pollution (e.g., Lewis-Michel et al., 1996; Bonner et al., 2005; Nie et al., 2007); the present study improves upon these earlier studies by controlling for more personal and occupational risk factors, and by using estimates of exposure from a prediction model based not only on traffic patterns, but also on land use patterns, population density, and locations of known point sources of pollution. Overall, the results of this dissertation contribute further evidence to the increasing body of literature that demonstrates the potential harmful effects to human health associated with exposure to urban air pollution.

6.2 Limitations of the Dissertation

There are two key limitations of this research and they relate broadly to the issue of accuracy of exposure estimates. Assigning exposure to ambient air pollution to individuals is a major challenge in epidemiological studies. Thus, first, it must be acknowledged that the estimates of mean annual NO\(_2\) generated through the prediction surfaces are in no way proxies for complete personal exposures. Total
personal exposure relates to a complex number of factors and behaviours, including daily activity patterns, quality of housing, and amount of time spent indoors and outdoors, among others. This limitation has possible implications for the findings presented in both chapters four and five. Specifically, some residents in Montreal may live in an area characterised by relatively low levels of pollution, yet may spend their days working or studying in a location characterised by especially high levels of pollution, or vice versa. Similarly, most people do not live out their whole lives at the same residence. A related point is that the case control dataset was lacking complete lifetime addresses of each subject. Ideally, addresses (and exposure estimates) would have been available for each subject at key points in life, such as menarche, first birth, and start of menopause, to help understand whether exposure at different periods is more important to the development of breast cancer than at others.

A second limitation of this research relates to one of its most notable strengths, namely the high level of detail and context-specificness of the land use regression model. A consequence of the high level of attention to local conditions is the inherent lack of generalizability that that level of singularity entails. That is, the strength of the prediction model and of the ultimate estimates of local concentrations of NO$_2$ is due to the fact that the model is based on 47 different variables describing land use, traffic, and population patterns specific to Montreal. By including 47 locally-specific predictor variables, the model produced estimates of NO$_2$ with exceptionally high levels of accuracy. A perhaps adverse consequence of this high level of precision calibrated to the Montreal landscape is that the model is not readily transferable to other locations beyond the geographic boundaries of the Island. For example, this model could not be replicated easily in say, Quebec City, without putting considerable effort into compiling numerous datasets describing specific traffic and land use characteristics of Quebec City. This issue relates directly to why geography – the description of the patterns of people and land use in Montreal – is central to the development of the land use regression model. In this case, context is everything. The land use regression model described in this dissertation is not a general prediction model that can be used to estimate levels of
pollution in any city, but rather is a model that can be used to estimate levels of pollution in Montreal, *exclusively.*

Conversely, a simpler, more parsimonious land use regression model based on only a few predictor variables might possibly be transferred more easily to another location, but the trade-off would be a reduction in the accuracy and precision of the predicted estimates of NO$_2$. In the case of this dissertation, this limitation resulted in a reduced sample size for the case control study. The original breast cancer dataset included 291 subjects from the greater Montreal area (extending beyond the immediate boundaries of the Island), who had to be excluded due a lack of confidence in the predicted estimates of NO$_2$ off the Island.

6.3 **Directions for Further Research**

The key areas for future research relate to developing a better understanding of what the key findings in the study really mean and to validating further the estimates of exposure. As described above, NO$_2$ is simply one of many constituent parts of the complex mixture that is urban air pollution. Studies conducted in other cities have shown that NO$_2$ is highly spatially correlated with other traffic-related pollutants (Nieuwenhuijsen, 2000; Beckerman et al., 2008; Wheeler et al., 2008). As such, it remains unclear whether the associations observed between NO$_2$ and the incidence of breast cancer described in chapter five are in fact due to exposure to NO$_2$ or rather to exposure to some other co-locational pollutant or to some combination of pollutants. In this context it would be worthwhile to sample and model other pollutants (e.g., sulphur dioxide, benzene, fine particles) to test whether they do in fact share the same spatial patterns in Montreal. It is possible that unique characteristics of the Montreal landscape (e.g., characteristics related to topography, location of industrial activity, climate) affect the distribution, dispersion, and concentrations of different pollutants. For example, the presence of petrochemical facilities in the east of Montreal may contribute significantly to the patterns of different pollutants.

This dissertation described a method for back-extrapolating estimates of NO$_2$
collected in 2005-06 to what they might have been 10 and 20 years earlier, in 1996 and in 1985, respectively. A similar method could be used to extrapolate these estimates forwards in time, for use in other epidemiological studies for which subjects are followed over time. For example, estimates of exposure covering an approximate 40-year period (i.e., 1985-2025) could potentially be generated from this dataset and assigned to existing cohorts or used in new studies. It can be very time-consuming and expensive to conduct a large-scale, multi-season sampling campaign, such as was described earlier. As such, there is immense value in the possibility of extrapolating estimates accurately into the future, thus preventing the need for repeating the full sampling campaign. The method of extrapolation could be evaluated, however, by performing a five- or ten-year follow-up of the sampling of NO$_2$ at some of the same locations sampled in 2005-06. In this way, the patterns of NO$_2$ estimated through forward-extrapolation could be compared with those from sampled observations. Likewise, samples collected for a five or 10-year follow-up could be back-extrapolated and compared with the samples collected in 2005-06. The data collected and the research methods developed for this dissertation can be used to address a number of problems and unanswered questions related to understanding how exposure to local concentrations of air pollution affect health outcomes in Montreal.

The overall impact of the research presented in this dissertation is the contribution to the body of evidence showing associations between exposure to air pollution and negative health outcomes. There is a need for continued, well-designed and well-executed scientific studies linking health outcomes to exposure to air pollution in order for this important topic to remain on the policy agenda of both local and national governments. 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. There is, in fact, strong evidence that reductions in local concentrations of pollution have been effective in improving life expectancy and reducing mortality rates in relatively short amounts of time (Clancy et al., 2002; Laden et al., 2006; Pope et al., 2009). The
most recent of those studies (Pope et al., 2009) presented compelling evidence that reductions in fine particulate matter contributed to an increase of approximately 15% in life expectancy, based on data from 211 counties in 51 cities across the US. These studies highlight the benefit of pollution control strategies, and the importance of continued research into the associations between health and exposure to air pollution.
References


Canadian Institute for Health Information (CIHI). 2006. How Healthy are Rural Canadians? An Assessment of Their Status and Health Determinants. Ottawa: CIHI.


109


Ley DF. 1993. Past elites and present gentry: neighbourhoods of privilege and the inner city. In: The Changing Social Geography of Canadian Cities (Bourne LS,


Santodonato J. 1997. Review of the estrogenic and antiestrogenic activity of


