Socio-Environmental Determinants of Cardiovascular Diseases

by

Antony Chum

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Geography
University of Toronto

© Copyright by Antony Chum 2012
Social-environmental Determinants of Cardiovascular Diseases

Antony Chum

Doctor of Philosophy

Department of Geography & Program in Planning

University of Toronto

2012

Abstract

Cardiovascular diseases (CVDs) are the leading cause of death and disability around the world. The purpose of this thesis is to investigate the impact of socio-environmental determinants of CVDs at the neighbourhood scale in order to inform actionable interventions, which may lead to large-scale reductions in preventable CVDs.

Drawing on 2411 surveys carried out in Toronto, Canada, this thesis employs multilevel models to estimate the magnitude of socio-environmental influences on the risk of CVD while adjusting for individual-level risk factors. To advance current research methodology, strategies and innovations were developed to 1) improve the characterization of neighbourhoods by empirically testing a full range of socio-environmental influences; 2) account for non-residential exposures by including a combined analysis of work and home contexts; 3) account for variations in the duration of exposure through the use of time-weighted models; 4) deal with problem of spatial data aggregation by developing and testing a novel method of neighbourhood zone design, and 5) account for the spatial scales of different socio-environmental determinants by modeling at multiple scales.
The thesis demonstrated that land use decisions are inextricably public health decisions. It found that living in neighbourhoods with inadequate access to food stores and areas for physical activity, burdened by violent crimes and fast food restaurants, and over-dependent on automobiles (leading to air pollution), with a high level of noise may significantly increase the risk of CVDs, over and above individual-level risks. The thesis also found that working in neighbourhoods that are socio-economically disadvantaged or have high-traffic may significantly increase CVD risk. The thesis developed and demonstrated novel methods to reduce the measurement error of neighbourhood exposures through 1) the use of “amoeba buffers” to improve neighbourhood zone design to better reflect participants’ local neighbourhoods and 2) the use of duration of exposure weights to adjust for individual differences in the time spent across different contexts. Finally, it found that the significance of socio-environmental factors depends on the scale of data aggregation; thus, investigation of multiple scales may be required to identify the relevant scale that matches the specific contextual factor in future research.
Acknowledgments

It would not have been possible to complete my PhD thesis without the help and support of my committee, family, friends, and colleagues, to only some of whom it is possible to give particular mention here.

I have been very fortunate with my supervisory committee and mentors. I would like to thank Alan for his great supervision. He never looked over my shoulder, and gave me room to develop as an independent researcher. Pat has been one of the greatest mentors I have had. It was her support and advice that kept me going at the most difficult times in this journey. Jim, Esme and Kathi were tremendously generous with their time and invaluable insights, for which I am very grateful. I would also like to thank Jack for his mentorship, guidance, and support throughout my entire time in graduate school. I thank Rosane for her support in biostatistics and patience.

Above all, I would like to thank Eddie for his personal and technical support, as well as his patience at all times. My parents and sister have given me their unequivocal support throughout, as always, for which my mere expression of thanks likewise does not suffice.

I had been fortunate to come across many good friends, without whom this journey would have been bleak and difficult. Special thanks go to Anita P for her humour and 11 years of friendship. I thank Sara for her ever-present support. Our trips abroad and weekend getaways have kept my years in graduate school fun and exciting. Valerie, Elyse, Juliana, Jen, Steven, Adam & Patrick I thank for the movies, dinners, concerts, plays, and trips that we enjoy together. I thank Zam for all those years of studying late with me at GGs and for being a great friend. I thank Shahrzad and Amir for believing in me especially when I had my doubts.

I have been very fortunate to be part of a research team at CRICH. I thank everyone I worked with at CRICH, and especially I thank my colleagues Janice and Anita M for their help with the data and their friendship. I also thank my fellow graduate students at
the Department of Geography and Planning for promoting a stimulating and welcoming academic environment.

I am tremendously grateful for the funding I received from the Social Science and humanities Research Council of Canada through their Joseph-Armand Bombardier Canada Graduate Scholarship program and the Aileen D Ross Fellowship program, as well as Munk Centre’s Comparative Program and Health and Society Program.

Antony
# Table of Contents

Introduction: The Purpose of Studying Neighbourhood Effects on Health .......................... 1
  Focus on Cardiovascular Health and Overview of the Study ........................................... 3
  Central Research Problems for the Study of Neighbourhood Effects on Health .............. 4
  Methodological approach of the thesis .............................................................................. 12
  Roadmap of the thesis ........................................................................................................ 15

Study 1: Socio-Environmental Determinants of Cardiovascular Diseases ....................... 26
  Literature Review .............................................................................................................. 27
  Methods ............................................................................................................................ 34
  Results .............................................................................................................................. 38
  Discussion .......................................................................................................................... 46

Study 2: Place effects on cardiovascular health: Do non-residential exposures and the duration of exposure make a difference? ................................................................. 60
  Literature Review .............................................................................................................. 62
  Methods ............................................................................................................................ 69
  Results .............................................................................................................................. 78
  Discussion .......................................................................................................................... 86

Study 3: Socio-environmental determinants of Cardiovascular Diseases: Do neighbourhood boundaries matter? ................................................................. 97
  Literature Review .............................................................................................................. 99
  Methods ............................................................................................................................ 111
  Results .............................................................................................................................. 119
  Discussion .......................................................................................................................... 130

Conclusion .............................................................................................................................. 142
  Section 1. Summary of findings and contribution to research and policy ......................... 142
  Section 2. Outstanding issues and next steps .................................................................. 154
List of Tables

Table 1.1: Self-Reported History of CVD Outcomes and Characteristics of Study Participants .........................................................................................................................40

Table 1.2: Multilevel Logistic Regressions of outcome 1 (MI only) and outcome2 (any CVDs) .........................................................................................................................................................44

Table 1.3: Multilevel Logistic Regression – Regression coefficients for neighbourhood level factors after adjusting for individual confounders and mediators (BMI and regular physical activity) ..............................................................................................................................................46

Table 2.1: Regression Coefficients for predictors of time spent at home in minutes (n=6175) ..............................................................76

Table 2.2: Descriptives of socio-demographic characteristics for entire cohort and participants with at least 1 CVD .........................................................................................................................................................79

Table 2.3: Unconditional Models: Intra-class Correlations and Model Fit Statistic represented by -2 restricted log pseudo-likelihood ........................................................................................................................................80

Table 2.4. Model Set #4. Including Residential and Workplace Socio-environmental variables and Individual level Controls, Unweighted vs. Time-Weighted Analysis ........................................................................................................83

Table 3.1: Selected literature linking socio-environmental factors to CVD outcome and risk factors: Analytical Strategies and Neighbourhood Boundaries Used ........................................................................................................108

Table 3.2: Self-Reported History of CVD Outcomes and Characteristics of Study Participants with $\chi^2$ test of differences for pairs of categorical variables ..................................................................................................120

Table 3.3: Correlations between neighbourhood socio-environmental predictors measured at six different neighbourhood scales ........................................................................................................................................121

Table 3.4: Comparison of Neighbourhood Definitions and Model Set #1 intra-class correlations .........................................................................................................................................................................123

Table 3.5: Multilevel logistic regression adjusting for all socio-environmental factors, estimates for 6 neighbourhood scale (95% confidence intervals) ........................................................................................................125

Table 3.6: Multilevel logistic regression adjusting for socio-environmental and individual factors, estimates for 6 neighbourhood scales (95% confidence intervals) ..................................................................................126

Table 3.7: Change in significant associations comparing 1) larger to smaller administrative boundaries and 2) administrative boundaries to respective amoeba buffers ..................................................................128
List of Figures

Figure 0.1: Amoeba buffer overlaid on administrative boundary (same as Figure 3) ............... 9

Figure 0.2: Food stores distributed across administrative boundary and the corresponding amoeba buffer (same as Figure 4) ................................................................. 10

Figure 1: Hypothetical administrative boundary and participant locations ....................... 110

Figure 2: Illustration of Neighbourhood Planning Area, Census Tract and Dissemination Area scales for measurement of the socio-environmental context ........................................... 113

Figure 3: Amoeba buffer overlaid on administrative boundary ....................................... 113

Figure 4: Food stores distributed across administrative boundary and the corresponding amoeba buffer ............................................................... 114
Socio-environmental determinants of Cardiovascular Diseases

Introduction Chapter

The Purpose of Studying Neighbourhood Effects on Health

Recent developments in health geography and social epidemiology play a major role in the recognition of ‘place’ as a significant factor underpinning health inequalities, given the uneven geographic development of risks associated to place as well as access to health promoting resources (Macintyre, Ellaway, & Cummins, 2002; S. J. Smith & Easterlow, 2005). The spatial turn in health research has engendered global comparisons of health inequalities (Shaw, Orford, & Brimblecombe, 2000; WHO, 1997, 2002) and research on national and regional variations in health outcomes (Griffiths & Fitzpatrick, 2001; Howe, 1986). At the local scale, a body of studies has emerged to establish the effects of neighbourhoods, independent of individual level risk factors, on a variety of health outcomes and behaviours including smoking (Kleinschmidt, Hills, & Elliott, 1995), cardiovascular health (Daniel, Moore, & Kestens, 2008), physical activity (Colabianchi et al., 2007), obesity (Ellaway, Anderson, & Macintyre, 1997), breast cancer (Barrett et al., 2008), low birthweight (Bell, Zimmerman, Almgren, Mayer, & Huebner, 2006), child accidents (Haynes, Jones, Reading, Daras, & Emond, 2008), general mortality (Yen & Kaplan, 1999), perceived health (Soobader & LeClere, 1999), mental health (Min-Ah, 2009), and domestic violence (O'Campo et al., 1995).

Investigations into how and the degree to which geography matters in the social production of health can offer an alternative perspective on health that has been largely dominated by a biomedical understanding of disease and pathology (Smyth, 2008). The social construction of place and its complex relationship with people and health also has a radical edge: it offers a critique of neoliberal politics, specifically, its tendency to blame individuals for their own misfortunes, and the social construction of diseases as simply a matter of personal responsibility or choice (Smyth, 1998; Sontag, 1989). Implicitly, this perspective shifts the onus of maintaining public health from individuals to governments and collectivities (Smith & Easterlow, 2005). Rather than solely relying on traditional individual-based health interventions, national governments, especially in the UK, have heavily invested in environmental and contextual interventions (such as community and neighbourhood based initiatives to tackle social determinants of health, e.g. UK’s Health
Action Zones) resulting in successful and mixed health outcomes (Bauld et al., 2005; O’Dwyer, Baum, Kavanagh, & Macdougall, 2007; M Stafford, Nazroo, Popay, & Whitehead, 2008; Thomson, Atkinson, Petticrew, & Kearns, 2006). There has also been a renewed interest of place effects on health by an increasing number of urban planners, designers and architects who strongly believe that reconnecting their respective fields with public health will help both fields to achieve their mission of social betterment by working collaboratively to address the health of urban populations (Corburn, 2004; Duhl & Sanchez, 1999; WHO City Action Group on Healthy Urban Planning, 2003). For example, in a report by Jackson and Kocktitzky for the Centers for Disease Control and Prevention (2009), they argue for the reintegration of land use planning and public health, explicitly linking transportation and land use planning as part of a comprehensive strategy to reduce cardiovascular and respiratory risks.

However, crafting and implementing place-based interventions require that we develop a better understand of the causal pathways that underlie neighbourhoods and health. The empirical studies in this thesis are designed to advance this understanding by developing and testing innovative strategies and methods to 1) improve how places are characterized, 2) account for non-residential exposures, 3) account for variations in the duration of exposure, 4) deal with the ‘modifiable areal unit problem’ (MAUP) where the zones that constitute neighbourhoods may not necessarily match the spatial extent of residents’ everyday life, and 5) account for the different spatial scales of different socio-environmental determinants of health. All five ‘gaps’ have been identified as major obstacles for the investigation of neighbourhood effects on health outcomes (Chaix, 2009; Chaix, Merlo, Evans, Leal, & Havard, 2009; Ana V. Diez-Roux, 2001; Kawachi & Berkman, 2003; Macintyre et al., 2002), and thus, by overcoming these obstacles, this project helps to improve research and the success of future targeted interventions.

This introductory chapter is organized into 4 sections: first, I will explain the rationale for the focus on cardiovascular health and introduce the *Neighbourhood Effects on Health and Wellbeing study* (NEHW), a large scale study in which this thesis was developed under. Second, I lay out five central research problems for the study of neighbourhood effects on health, and discuss the innovations and strategies that have been developed in my thesis to tackle these problems. Third, I discuss the method of
generalized multilevel modeling and explain why it is the method chosen in my empirical studies. Finally, I briefly outline each of the studies and discuss their collective theoretical and empirical contribution to research and how they may help to improve the success of place-based interventions.

1. **Focus on Cardiovascular Health and Overview of the Study**

   This thesis focuses on cardiovascular health for a number of reasons. First, cardiovascular diseases (CVD) are the leading cause of death and disability in the US, UK, Canada and in most countries around the world (American Heart Association, 2003; Bonow, Smaha, Smith, Mensah, & Lenfant, 2002), and a vast number of CVD cases are preventable. Second, while conventional cardiovascular disease (CVD) preventive strategies typically involve the modification of individual-level behavioural and biological risk factors (Canto & Iskandrian, 2003), such as diet modification, daily exercise, and weight management; the socio-environmental risks and resources that may impact CVD outcome such as the local food environment, opportunities for physical activities, neighbourhood sources of psychosocial stress, and environmental factors such as noise and air pollution are not typically considered. Careful consideration of the socio-environmental context relevant to CVD may be a productive way to shift some of our focus from treatment and prevention solely at the individual level to prevention for ‘sick populations’ (Rose, 2001), which may lead to large scale reductions in preventable CVD mortalities and morbidities.

   While a number of national and regional surveys contain individual level health information, most are not designed specifically to understand the impact of socio-environmental determinants of health at the neighbourhood scale. The Neighbourhood Effects on Health and Well-being (NEHW) project, based in Toronto, Canada, implements a cross-sectional survey designed specifically to understand the impact of neighbourhood level determinants, acting independently or interactively with individual level factors, on population health including CVD outcomes. Toronto is an ideal setting for this research because it is ethnically, economically, and socially diverse. Data were collected between Mar 2009 and June 2011. Data collection involved surveys of 2412 participants that are adults aged 25-65, randomly selected from 87 socio-economically
and demographically diverse census tracts. The survey response rate was 72%.

Neighbourhood information is collected from 1) participants, 2) commercial and administrative databases (e.g. DMTI Spatial Inc, 2011a, 2011b; Statistics Canada & Canadian Centre for Justice Statistics, 2010), and 3) municipal government and other state authorities (e.g. Toronto Public Health Inspection, 2011; Toronto Transportation Services, 2011).

Project NEHW obtained information on individual health status including the CVD outcomes relevant to this thesis. They include self-reported history of physician diagnosis of myocardial infarction (MI), angina, coronary heart disease (CHD), stroke, and congestive heart failure (CHF). Although the survey did not include clinical details, previous studies of the validity and reliability of self-reported conditions have suggested a high level of agreement with medical records for the conditions considered here (Bush, Miller, Golden, & Hale, 1989; Kehoe, Wu, Leske, & Chylack, 1994).

2. Central Research Problems for the Study of Neighbourhood Effects on Health

While the study of neighbourhood effects on health has been steadily gaining popularity over the past two decades across fields such as medical sociology, health geography and social epidemiology (Ana V. Diez-Roux, 2001; Macintyre et al., 2002; S. J. Smith & Easterlow, 2005; Smyth, 2008), researchers have identified some major obstacles that are common amongst many studies of neighbourhood effects on health, in particular for studies employing quantitative methods (Chaix, 2009; Chaix et al., 2009; Ana V. Diez-Roux, 2001; Kawachi & Berkman, 2003; Macintyre et al., 2002). They include: 1) underdeveloped characterization of neighbourhoods, 2) lack of consideration for non-residential exposures (also commonly termed the ‘residential trap’), 3) a lack of consideration for variations in the duration of exposure, 4) ignoring the ‘modifiable areal unit problem’ (MAUP) where the zones that constitute neighbourhoods may not necessarily match the spatial extent of residents’ everyday life, and 5) ignoring the fact that neighbourhood level exposures may operate in different spatial scales (also commonly termed the ‘local trap’). In the following, I will provide an account of how each of these common problems may hinder the study of neighbourhood effects on
health, and briefly discuss the innovations developed in the three studies in this thesis to tackle them.

_Underdeveloped characterization of neighbourhoods_  

While sophisticated biometric measures and individual-level attributes are commonly used in public health and epidemiological studies, socio-environmental constructions of ‘context’ that are equally important, such as place and neighbourhood, remain comparatively underdeveloped (Ana V. Diez-Roux, 2004). For example, in the field of CVD research, one key area of research looks at the independent impact of neighbourhood deprivation. It found that living in a deprived neighbourhood is associated with increased incidence of coronary heart diseases (A. V. Diez-Roux et al., 2001; Sundquist, Malmstrom, & Johansson, 2004; Winkleby, Sundquist, & Cubbin, 2007), increased incidence of myocardial infarction (Lovasi et al., 2008), increase in all cause and cardiovascular disease mortality (G. D. Smith, Hart, Watt, Hole, & Hawthorne, 1998), increase in coronary heart disease case fatality (Winkleby et al., 2007), and increase in risk factors such as smoking, physical inactivity, obesity, diabetes, and hypertension (Cubbin et al., 2006; Ellaway et al., 1997; Matheson, White, Moineddin, Dunn, & Glazier, 2010) after adjusting for individual level risks in the studies above. Indicators of neighbourhood deprivation are easily derived from national census data and may be seen as a proxy for a range of more specific neighbourhood features relevant to CVDs that are not directly measured; however, since it is the only information on the neighbourhood context, the lack of directly measured socio-environmental features remain a major limitation (Ana V. Diez-Roux, 2003; Sampson & Raudenbush, 2004). To overcome this gap, studies analyzing neighbourhood deprivation need to also include additional specific features of the built environment. This is important since neighbourhood deprivation can be acting as a proxy for a number of contextual factors like healthy food availability, so it is important to understand its unique contribution (i.e. strictly neighborhood deprivation effect and not as a proxy for resource availability) to CVD variance.

Plausible socio-environmental factors relevant to CVD risk including the local food environment, opportunities for physical activities, neighbourhood sources of
psychosocial stress, and environmental factors such as noise and air pollution will be considered in this thesis. Virtually no previous work has combined all these areas into an integrated approach to understand the impact of the built environment on CVDs. This is a valuable approach because 1) we can evaluate their relative influence on CVD to set policy priorities, 2) one-dimensional characterization of neighbourhoods, such as only using neighbourhood deprivation, does not do enough to clarify specific contextual factors, which may help to elucidate the specific causal pathways and provide opportunities for more effective interventions, and 3) by testing a full range of possible environmental factors affecting CVDs, we can begin to understand the built environment as a complex system where multiple contextual factors may interact with each other and also with individual level covariates.

In summary, while neighbourhood deprivation is well established as a social determinant of health in social epidemiology (Pickett & Pearl, 2001; Riva, Gauvin, & Barnett, 2007), less emphasis has been placed on the effects of a full range of other neighbourhood characteristics outside of theoretical papers and qualitative studies. By testing the theoretical pathways on a variety of neighbourhood characteristics, in addition to neighbourhood SES, the studies in this thesis will offer valuable insights for the implementation of population health interventions.

Lack of consideration for non-residential exposures (or the ‘residential trap’)

Research on human activity and travel patterns suggests that the geographic extent of everyday lives is not limited to residential neighbourhoods (Buliung & Kanaroglou, 2006; Gliebe & Koppelman, 2005; Law, 1999; Naess, 2006). It is thus a critical limitation that virtually all studies of place effects on health have focused solely on the residential environment, save for rare exceptions (Inagami, Cohen, & Finch, 2007; Muntaner et al., 2006). Chaix (2009) has termed this research gap the ‘residential trap’, because of the exclusive reliance on local residential environments and the systematic neglect of non-residential environments.

In order to tackle this empirical lacuna, the second empirical study of this thesis examines the impact of residential and workplace exposure (separately and simultaneously) on CVD risk. Since not all participants in a given residential
neighbourhood go to work in the same area, this situation is said to have a ‘cross-classified’ structure and is best modeled using cross-classified multilevel analysis (Goldstein, 2003; Sykes & Musterd, 2011). Using a cross-classified multilevel approach, I am able to estimate the influence of each context while controlling for effects of the other context. While in previous studies of place effects on CVD risk (Inagami et al., 2007), even when non-residential exposure was considered, the lack of a cross-classified approach did not allow researchers to understand the relative importance of each setting. This is because by treating each context as a ‘level’ in the cross-classified approach (rather than as a ‘variable’ in the case of the Inagami study), variance proportion values, most commonly represented by the intraclass correlation coefficients (ICC), are calculated for each level or setting (i.e. one for each level or context considered). This allows us to understand the total variance that is accounted for by observed and unobserved factors in each of the levels, which gives an indication of the relative importance of each setting. The information on the relative importance of home versus work contexts, as well as the investigation of specific socio-environmental resources and risks for CVD in each setting, may potentially advance knowledge for research and intervention because the analysis: 1) may identify potentially new pathways through which non-residential socio-environmental contexts influence CVD etiology, and 2) may demonstrate that the workplace context may be an important context to consider for inclusion in the next generation of studies of place effects on health.

Lack of consideration for the duration of exposure

In virtually all neighbourhood effects on health studies, little has been done to account for time spent in a context and how duration of exposure may modify place-health associations. A possible reason why some previous studies have failed to identify robust associations between residential contextual influences and health may be because the effects are masked by large variations in the unobserved duration of exposure; assuming that each person’s duration of exposure is equal when in fact they are not might lead to an underestimation of effects, because those who spend substantial amounts of time out for work/school/leisure/other activities will not be as influenced by residential factors compared to people who sit around at home all day. The implicit assumption of
equal duration in most studies leads to the underestimation of contextual effects and may even increase the risk of type 2 errors.

To account for the duration of exposure, a time-weighted analysis is incorporated into the cross-classified approach that combines the work and home exposures in the second empirical study of this thesis. The time-weighted analysis uses information provided by the participants on the average time they spend at work, and from this information I estimate the time they spend at home. Using the information on the amounts of time spent at work and home, I weigh the respective sets of residential and workplace exposures that include variables representing the levels of neighbourhood deprivation, the local food environment, opportunities for physical activities, neighbourhood sources of psychosocial stress, and environmental factors such as noise and air pollution in both contexts. The weighted analysis is then compared to an unweighted analysis to see if the time-weights improved model fit as well as the strength of the place-CVD associations, which would indicate that accounting for time improves the model, which may justify that duration of exposure should be an important area of consideration in future studies.

*Lack of Consideration for the ‘Modifiable Areal Unit Problem’ (MAUP) and neighbourhood zone design*

Empirical investigations into how places affect health commonly employ geographical areas as units of analysis; therefore, their results may be greatly influenced by the design of these areas – an effect known as the modifiable areal unit problem (MAUP). MAUP is a problem that comes into play whenever researchers employ spatial data aggregation, because the boundaries that comprise areal units are modifiable (Cockings & Martin, 2005; Haynes, Daras, Reading, & Jones, 2007). While administrative zones are very typically employed in multilevel and ecological studies of place effects on health as proxies for neighbourhoods, the use of these spatial units have received much criticism because they may not necessarily match the spatial extent of residents’ everyday life (Ana V. Diez-Roux, 2004; Pickett, Collins, Masi, & Wilkinson, 2005). The third empirical study in this thesis develops a method of using *amoeba buffers* to improve on administrative boundaries to more accurately assess the area of exposure
for residents. Typical census and administrative boundaries are approximately rectangular areas that artificially truncate the residents’ area of exposure for those that live near the edges, and amoeba buffers solve this problem by augmenting the shape of the neighbourhood.

The use of amoeba buffer (figure 0.1) changes the shape of the original administrative boundary into a shape resembling an amoeba organism, and may improve the measurement of the residents’ area of exposure by extending the zone into adjacent administrative boundaries where residents are near the edge of the original administrative boundary.

Figure 0.1: Amoeba buffer overlaid on administrative boundary
To illustrate how amoeba buffers augment neighbourhood contextual measures, the following is an example of how this study counted the number of food stores in the administrative versus amoeba neighbourhood as shown in figure 0.2:

*Figure 0.2: Food stores distributed across administrative boundary and the corresponding amoeba buffer*

![Diagram showing food stores and amoeba buffer](image)

If we only consider the administrative boundary, only food stores A and B would be included (thus, a value of 2 food stores is assigned to every resident within the neighbourhood). If we consider the amoeba boundary in which the same group of residents are within, only food stores B, C and D would be included (thus, a value of 3 food stores is assigned to every resident within the neighbourhood). This example demonstrates that amoeba buffers do not always consider ‘more’ risks or resources compared to the corresponding administrative boundary, but in fact may remove risks/resources that may be within the same administrative boundary but are not within 1 km of any study participants (for more details on the rationale and design of amoeba buffers, refer to Study #3).

In the third study of this thesis, I formally test whether amoeba buffers represent an improved mode of neighbourhood zone design by comparing the effect sizes of the contextual influences of CVD between multilevel models employing amoeba buffer versus models using the original administrative boundary. In addition to overcoming the ‘edge truncating’ effect of administrative boundaries, the amoeba buffers have the added
advantage of preserving the original clusters of residents to give an appropriate
estimation of the standard errors for generalized linear mixed models (compared to
methods that get rid of neighbourhoods all together, such as using Euclidean distance or
distance decay functions to proxy exposure, which ignores the clustering of residents).
Administrative and census boundaries will probably continue to be a popular choice for
the purposes of investigating neighbourhood effects, because of their convenience, and
they represent a common set of neighbourhood boundaries that researchers can use to
replicate analyses. This is why I think it is impractical to disregard administrative zones,
but a more pragmatic approach, one taken in this thesis, is to improve how they can better
represent the real/lived areas of exposure for residents.

*Lack of consideration for the spatial scales of neighbourhood level exposures (or the
‘local trap’)*

While many epidemiology and health geography studies use only a single set of
neighbourhood definitions, it is often acknowledged that a single scale may be inadequate
to investigate the spatial extent for different types of human activities (Macintyre et al.,
2002), which may operate at different spatial scales. The (incorrect) assumption that all
contextual influences of health may be studied using only one neighbourhood scale is
aptly termed the ‘local trap’ by Cummins (2007).

There are two main reasons why researchers should pay greater attention to the
spatial scales for contextual influences: first, studies using a single set of boundaries may
underestimate levels of health inequality, especially if boundaries define areas where
within-group variation in the health outcome (as well as contextual determinants of
health) is relatively larger than the between-group variation. In other words, if the
neighbourhood boundaries define areas that are internally heterogeneous in terms of
health, the analysis will not detect the full extent of the local area differences, since the
inequalities may be occurring at a different scale. Second, the area of effect for different
contextual influences on health may operate at different scales. For example, while the
availability of healthy food might be relevant to health at a relatively local scale, access
to a hospital for acute care would not be expected to be important at the same scale. Thus,
when we measure the characteristic of place for health ‘resources’, there should be two
distinct scales in which we measure access to healthy food versus access to acute care. If
scale is not taken into account, the analysis will tend to bias results towards a null finding: increasing the likelihood of type II error (M. Stafford, Duke-Williams, & Shelton, 2008).

In the third empirical study of this thesis, I will analyze neighbourhood effects on CVD risk at 3 different administrative scales ranging from the census dissemination areas, which typically only include a few city blocks (mean area = 0.18 km$^2$) up to Toronto neighbourhood planning areas, which are large multi-block areas (mean area=4.56km$^2$), roughly equivalent to typical US zip code areas. The results of the study let us understand the appropriate scale of analysis for a range of contextual influences of CVD risk, and if certain associations are only significant at certain scales, it would justify that multiple scales be used in future analysis. To do so is not to advocate that researchers ‘fish’ for significant associations by testing a range of scales, but to recognize that health-relevant risks and resources operate at multiple scales, and any single scale might mask relevant spatial variation if it is too large or excludes relevant socio-environmental features if it is too small.

3. Methodological approach of the thesis

Many statistical tests, such as ordinary least squares, logistics regression, and ANOVA, rely on the assumption of independent observations, such that the response of one person must not affect the response of another. This assumption makes these methods not suitable to analyze complex social relations where people should be understood as groups and the behaviour of one person can influence another (e.g. neighbours, friendship groups, students in the same class, etc.) or they may be simultaneously affected by a common context (e.g. they live in the same city or neighbourhood governed by a common set of social processes) (Raudenbush & Bryk, 2002). Multilevel modeling offers a new way of thinking about and analyzing social contextual variables as they affect groups of people. It does so by conceptualizing data as nested or hierarchical (Field, 2009), and disentangles the effects of neighbourhoods from those of individuals on health by simultaneously modeling the processes at different levels of a population hierarchy (Kawachi & Berkman, 2003; Raudenbush & Bryk, 2002). The form of the multilevel
model can be simply understood as an extension of the basic linear model expressed as follows:

\[ Y_i = \beta_0 + \beta_1 W_i + \varepsilon_i \]

Where \( \beta_0 \) is the intercept, \( \beta_1 \) is the slope for the variable \( W \) (a level one/individual level variable), and \( \varepsilon_i \) is the error term for the unique \( i^{th} \) case.

To add in a neighbourhood-level variable (level 2, expressed below as the subscript \( j^{th} \) case), we can do perform one of the three procedures: 1) add a random intercept (i.e. allowing intercepts to vary across neighbourhoods), or 2) add a random slope (i.e. allowing slopes to vary across neighbourhoods), or 3) allowing both intercepts and slopes to vary (Field, 2009).

Case 1: random intercept, fixed slope

\[ Y_i = (\beta_0 + u_{0j}) + \beta_1 X_i + \varepsilon_i \]

Case 2: fixed intercept, random slope

\[ Y_i = \beta_0 + (\beta_1 + u_{0j}) X_i + \varepsilon_i \]

Case 3: random slope, random intercept

\[ Y_i = (\beta_0 + u_{0j}) + (\beta_1 + u_{0j}) X_i + \varepsilon_i \]

Random intercepts are included because they allow each neighbourhood context to vary. By adding in random slopes, we can also drop the assumption of homogeneity across regression slopes. In other words, in different contexts (or level-2 units), covariates can have vastly different relationships with the outcome variable. Thus, allowing for both random intercepts and slopes in models may make them more realistic. However, the addition of random slopes is not always computationally possible: if the number of level-1 units within level-2 groups are not sufficiently large (usually around 10, but may be further restricted by the degrees of freedom available), this will typically result in model non-convergence (Goldstein, 2003). Random intercepts are not affected by this restriction.
Lastly, by relating the level 1 unit (subscript i) to the level 2 unit (subscript j) for neighbourhood-level variable (X) and individual-level variable (W), we get the following with interactions (XW) and residuals accounted for at the individual, neighbourhood, and interactional levels:

\[
Y_i = \beta_{ij} + \beta_{1j}X_j + \beta_{2j}W_{ij} + \beta_{3j}X_jW_{ij} + \varepsilon_{1j} + \varepsilon_{2ij} + \varepsilon_{3j}W_{ij}
\]

**Generalizations:**

**Adding the logit link function**

Since my thesis analyzes the determinants of cardiovascular disease (CVD) outcomes, further generalization of the above linear model is made to accommodate the dependent variable that is a dichotomous variable for singular or combinations of CVD conditions (depending on the analysis), which include self-reported history of physician diagnosis of myocardial infarction (MI), angina, coronary heart disease (CHD), stroke, and congestive heart failure (CHF). Therefore, probability of the CVD outcome is equal to \( p_{ij} = \Pr(y_{ij}=1) \), where \( p_{ij} \) is modeled using a logit link function, and \( y_{ij} \) follows the Bernoulli distribution (Guo & Zhao, 2000). This can be summarized as

\[
\log\left[ \frac{p_{ij}}{1-p_{ij}} \right] = (\beta_0 + u_{0j}) + (\beta_1 + u_{0j})X_i + \varepsilon
\]

in a combined model that includes both random intercepts and random slope.
Adding cross-classification

While the above models are appropriate for simple grouping of participants nested within residential neighbourhoods, an additional generalization of the model is required for my analysis involving participants nested within residential neighbourhoods and workplaces. Since not all participants in a given residential neighbourhood go to work in the same area, this situation is said to have a ‘cross-classified’ structure and is best modeled using cross-classified multilevel analysis (Goldstein, 2003; Sykes & Musterd, 2011). Cross-classified models recognize the simultaneous membership of level-1 units (i.e. participants) in multiple higher and non-nested levels (i.e. residential and workplace neighbourhoods). Cross-classified models estimate the influence of each context while controlling for effects of the other context, which make it an appropriate strategy. Given the inclusion of two different level-2 contexts (i.e. home and work), which is notated by $j_1$ and $j_2$ (with respective random effects $u_{0j_1}^{(1)}$ and $u_{0j_2}^{(2)}$), the following summarizes the incorporation of cross-classification into a multilevel logit-linked model with random slope and random intercept:

$$\log\left[\frac{p_{ij_1j_2}}{1 - p_{ij_1j_2}}\right] = (\beta_0 + u_{0j_1}^{(1)} + u_{0j_2}^{(2)}) + (\beta_1 + u_{0j_1}^{(1)} + u_{0j_2}^{(2)})X_i(j_1,j_2) + \varepsilon_{0i(j_1,j_2)}$$

Conclusions

Roadmap of the thesis

In the following, I will first present a roadmap of the entire thesis, and then briefly discuss the research questions set out in each of the studies. Study #1-3 make up the empirical chapters of my thesis, which will be followed by a concluding chapter, where I summarize and the research and findings of the three empirical studies and discuss how they help to advance theory and research in a coherent way.

This thesis extends the scope of neighbourhood characteristics typically considered in social epidemiological studies beyond neighbourhood deprivation. It does so by investigating a full range of neighbourhood characteristics. Study #1, in addition to neighbourhood deprivation, examines characteristics of the local food environment,
opportunities for physical activities, neighbourhood sources of psychosocial stress, and environmental factors such as noise and air pollution. The study evaluates the association between CVDs and the multi-dimensional neighbourhood socio-environmental context, thus the study advances beyond a relatively simplistic characterizing of neighbourhoods, using neighbourhood deprivation as the only contextual factor as many other multilevel studies have done. The study helps to shed light on the specific causal pathways and provides opportunities for more effective and advanced interventions. Study #1 also discusses the results with regards to possibilities of developing novel CVD preventive strategies through modifications to the built environment (using specific examples of partnerships between public health practitioners and urban planners) to supplement conventional individual-based strategies.

Study #2 extends the understanding of place-health associations to non-residential exposures for CVDs. Bridging this research gap will help us to 1) identify new (non-residential) sites and potential opportunities for effective intervention, and 2) clarify the true association of residential neighbourhood on health by adjusting for exposure to non-residential neighbourhoods. In virtually all neighbourhood effects on health studies, little has been done to account for time spent across different contexts and how the duration of exposure may modify place-health associations. By excluding time in these models, there is an implicit assumption that people spend relatively similarly amounts of time at home and in other settings, which is likely not the case. In study #2, I compare the time-weighted analysis to an unweighted analysis, and I test the hypothesis that the time-weighted analysis would result in improved model fit and stronger regression coefficients (for already significant effects in the unweighted analysis). The study is important because it can identify potential ‘workplace’ pathways to CVD, and may demonstrate that the workplace context and duration of exposure are important elements to consider for inclusion in the next generation of studies of place effects on health.

Finally, study #3 investigates whether neighbourhood boundaries can significantly impact the estimates for a range of socio-environmental neighbourhood effects on CVDs. This study will help to identify the appropriate scale and zone for the spatial distribution of health-relevant neighbourhood characteristics for CVD outcomes, which will help to
define ‘neighbourhood’ in the context of CVD research and help to inform the spatial extent for place-based interventions. It presents and tests a novel method of defining neighbourhood boundaries that supplements administrative boundaries using buffers, which is termed amoeba buffers, to help reduce the measurement error in the level of exposure. Specifically, amoeba buffers deal with the problem that residents who are located near the edges of a given neighbourhood boundary have their area of exposure truncated by the boundary.

**Research Questions**

The specific research questions associated with each of the 3 studies are presented in the following table. These questions drive the method of investigation in each study as well as the respective discussion sections, which shed light on how the findings may be applicable to research, policy, and intervention.

<table>
<thead>
<tr>
<th>Study #1 research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Is neighbourhood-level socio-environmental context correlated to higher risk of myocardial infarction and other CVDs? Can socio-environmental factors explain the association between neighbourhood deprivation and CVD outcomes?</td>
</tr>
<tr>
<td>2) To what extent do socio-environmental characteristics, taken together, account for neighbourhood variation in CVD?</td>
</tr>
<tr>
<td>3) Does the association between the socio-environmental context and CVDs remain after controlling for individual level risk factors?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study #2 research questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) What is the relative importance of the residential context versus the workplace context for CVD risk?</td>
</tr>
<tr>
<td>2) Do residential and/or workplace socio-environmental contexts correlate to the risk of CVDs after adjustments for individual-level risk factors?</td>
</tr>
<tr>
<td>3) Does the time-weighted analysis improve model-fit and strengthen the effect sizes of possible associations between socio-environmental risks and CVD?</td>
</tr>
<tr>
<td>Study #3 research questions</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>1) Does the CVD variability explained at the neighbourhood level vary significantly depending on the neighbourhood scale/boundaries used?</td>
</tr>
<tr>
<td>2) Is the association between neighbourhood socio-environmental factors and CVDs stronger for smaller administrative zones compared to larger ones? Is the association between neighbourhood socio-environmental factors and CVDs stronger for the amoeba boundaries compared to administrative boundaries? Do these associations remain after controlling for individual risk factors?</td>
</tr>
<tr>
<td>3) What is the extent to which neighbourhood boundaries may impact studies that involve the use of multilevel models to explain the risk of CVDs? What theoretical and practical considerations should be taken into account when constructing neighbourhood-level variables with regards to the associated modifiable areal unit problem?</td>
</tr>
</tbody>
</table>

The concluding chapter of this thesis will address the following meta-question of the thesis: how does addressing the five central research problems for the study of neighbourhood effect on health (identified above) help to advance research, policy and interventions? I will also discuss a number of outstanding issues and ways to move forward with regards to knowledge translation and future research.

---

1 These five central research problems for the study of neighbourhood effects on health include: 1) underdeveloped characterization of neighbourhoods, 2) lack of consideration for non-residential exposures (also commonly termed the ‘residential trap’), 3) a lack of consideration for variations in the duration of exposure, 4) ignoring the ‘modifiable areal unit problem’ (MAUP) where the zones that constitute neighbourhoods may not necessarily match the spatial extent of residents’ everyday life, and 5) ignoring the fact that neighbourhood level exposures may operate in different spatial scales.
Reference List


Study 1: Socio-Environmental Determinants of Cardiovascular Diseases

Introduction

Conventional cardiovascular disease (CVD) preventive strategies typically involve the modification of individual-level behavioural and biological risk factors (Canto & Iskandrian, 2003). This is done in two ways: 1) public health education aimed at lifestyle change, which includes smoking cessation, diet modification (reduced sodium and saturated-fat intake), daily exercise, and weight management; and 2) Early detection of risk factors such as hypertension and hypercholesterolemia through medical interventions, e.g. antihypertensive and lipid-lowering drug therapies. Given that current CVD intervention strategies (and for many other chronic diseases) are framed as matters of individual choice and medical care, it may be surprising for many to find that public health interventions were historically grounded in the understanding that land use decisions and the built environment influence population health (Melosi, 2000).

In Western Europe and North America, public health and urban planning co-evolved as a consequence of late 19th-century efforts to reduce the harmful effects of rapid urbanization and industrialization, particularly infectious diseases, through improvements to the urban infrastructure to solve the problems of sanitation, water delivery, and waste disposal (Corburn, 2004; Melosi, 2000). It is widely recognized that the improvements developed in this period have led to dramatic increases in life expectancies across industrialized nations (Condran & Crimmins-Gardener, 1978). Despite the common historical origin linking public health and urban planning, the two fields have since diverged (Northridge, Sclar, & Biswas, 2003). Public health today draws on a primarily biomedical paradigm to inform individual risk factor modification. The field shifted to addressing individuals rather than environments because the former is easier for physicians to influence with their specialized training. On the other hand, urban planning in postwar America first promoted public infrastructure projects to boost economic development, and then turned its attention to suburban expansion, which led to an era of inner-city divestment and residential segregation (J. Jackson, 1985; Marshall, 2000). However, many researchers today are convinced that reconnecting urban planning and public health will help both fields to achieve their original mission of social
betterment and working collaboratively to address the health of urban populations (Corburn, 2004; Duhl & Sanchez, 1999; WHO City Action Group on Healthy Urban Planning, 2003). A report, by Jackson and Kocktitzky for the Centers for Disease Control and Prevention, argues for the reintegration of land use planning and public health, explicitly linking transportation and land use planning to public health outcomes such as increased cardiovascular and respiratory risks (2009).

Cardiovascular diseases are the leading cause of death and disability in the US, UK, Canada and in most countries around the world (American Heart Association, 2003; Bonow, Smaha, Smith, Mensah, & Lenfant, 2002), and this emerging global epidemic requires our immediate attention. The purpose of this paper is to investigate the causal pathways underlying the socio-environmental context and CVDs in order to inform collaborations between urban planning and public health to tackle this epidemic. This paper has 3 parts: first, it reviews the literature linking the residential environment and CVD outcomes, identifies the research gaps, and presents a map of the causal pathways of the possible place-health associations. Second, it presents an original study of neighbourhoods in Toronto, Canada to investigate some of these pathways. Third, it discusses the results with regards to possibilities of developing novel CVD preventive strategies through modifications to the built environment to supplement conventional strategies.

1. Review of literature: Neighbourhood socio-environmental context and Cardiovascular Risks

A growing body of evidence suggests that socio-environmental factors are likely to be important in shaping the distribution of CVDs (Ana V. Diez-Roux, 2003). One key area of research looks at the independent impact of neighbourhood deprivation. Studies have found that living in a deprived neighbourhood is associated with increased incidence of coronary heart diseases (A. V. Diez-Roux et al., 2001; Sundquist, Malmstrom, & Johansson, 2004; Winkleby, Sundquist, & Cubbin, 2007), increased incidence of myocardial infarction (Lovasi et al., 2008), increase in all cause and cardiovascular disease mortality (Smith, Hart, Watt, Hole, & Hawthorne, 1998), increase in coronary heart disease case fatality (Winkleby et al., 2007), and increase in risk factors such as
smoking, physical inactivity, obesity, diabetes, and hypertension (Cubbin et al., 2006; Ellaway, Anderson, & Macintyre, 1997; Matheson, White, Moineddin, Dunn, & Glazier, 2010) after adjusting for the effects of at least gender, age, and individual level socioeconomic status (SES) in the studies above. Indicators of neighbourhood deprivation are easily derived from national census data and may be seen as a proxy for a range of more specific neighbourhood features relevant to CVDs that are not directly measured; however, since it is the only information on the neighbourhood context in these studies, the inability of these studies to directly measure specific socio-environmental features remain a major limitation of this work (Ana V. Diez-Roux, 2003; R. Sampson & Raudenbush, 2004).

A number of studies investigate the impact of other contextual factors on cardiovascular risks (Chaix, 2009). These factors can be usefully grouped by the mechanism through which they impact CVDs: 1) diet, 2) physical activity, 3) psychosocial stress, and 4) air pollution and noise. I will discuss these in turn and provide an integrated map of the plausible causal pathways and their mediators to CVD outcomes.

i. Diet

A number of studies have shown that differences in diet may not be entirely dependent on individual characteristics such as SES, gender and ethnicity, but may be partly explained by contextual factors (Ana V. Diez-Roux, 2003; Laraia, Siega-Riz, Kaufman, & Jones, 2004; Mooney, 1990; Moore, Diez Roux, Nettleton, & Jacobs Jr., 2008; Morland, Wing, & Diez Roux, 2002). There are 2 important areas of inquiry with regards to the influence of the built environment on diet: 1) the spatial distribution of food and 2) the extent to which local food environment affect dietary choices.

First, the location and availability of stores selling healthy foods is an area of active research, and might also explain part of the reason why higher CVD rates are associated with neighbourhood deprivation. For example, in a study of four US states, Morland et al. (2002) found that there are 4 times more supermarkets in white neighborhoods compared to black neighborhoods. They also found that there are 3 times fewer places to consume alcoholic beverages in the wealthiest compared to the poorest neighborhoods. Other studies have also documented a similar trend that healthy foods are
less available in deprived neighbourhoods compared to more affluent ones (Mooney, 1990; Wechsler, Basch, & Shea, 1995), and the density of fast food restaurants is found to be higher in black and low-income neighbourhoods (Block, Scribner, & DeSalvo, 2004). However, it is worth noting that research in Glasgow did not find the presence of ‘food deserts’, where food stores were numerous in deprived as well as non-deprived neighbourhoods (Cummins, Curtis, Diez-Roux, & Macintyre, 2007). The authors note that their findings are in contrast to studies from US cities. They suggest that there is a relatively equal spatial distribution of retail food stores in British cities compared to the US, which could be due to the trans-Atlantic differences in the pattern of urban dwellings and rental markets.

Second, studies that relate specific features of the local food environment to the actual dietary behaviors and health outcomes of individuals also suggest there may be differences between the UK versus US. In a US study of 10,623 participants, Black Americans’ fruit and vegetable intake increased by 32% for each additional supermarket in the census tract, and for White Americans’ fruit and vegetable intake increased by 11% with the presence of 1 or more supermarkets (Morland, Wing, & Diez Roux, 2002). Similarly, using data from the US Multi-ethnic Study of Atherosclerosis, a study finds that participants with no supermarkets near their homes were 25-46% less likely to have a healthy diet than those with the most stores, after adjustment for age, sex, race/ethnicity, and socioeconomic indicators (Moore et al., 2008). In a study based in North Carolina, US, researchers also confirm that the proximity of food retail outlets independently influences the diet quality of pregnant women (Laraia et al., 2004). On the other hand, in Glasgow, Scotland, researchers found few statistically significant associations between proximity to food retail outlets and diet or obesity, for unadjusted or adjusted models, or when stratifying by gender, car ownership or employment (Macdonald, Ellaway, Ball, & Macintyre, 2011). These results are somewhat expected since in an earlier study Macintyre had established that Glasgow has a relatively equal spatial distribution of retail food stores. It should be noted that results from different studies are not consistent in this field of research. For example, Bonne-Heinonen et al., in a US longitudinal study of young and middle-aged adults (2011), found that while fast food consumption is independently associated with fast food availability, there was no detectable relationship
between food stores availability and diet quality as measured by fruit and vegetable intake.

**ii. Physical Activity**

Recent efforts in urban planning research have focused on how planners can use land use and design policies to increase transit use as well as active forms of transportation such as walking and bicycling (Handy, Boarnet, Ewing, & Killingsworth, 2002). This research is particularly relevant to public health because of the expert consensus that short daily episodes of moderate physical activity may be sufficient to produce cardiovascular and other health benefits (Pate et al., 1995). The built environmental factors affecting physical activity can be grouped into 1) land development patterns (e.g. density and land use) and 2) urban design factors. An early study from Washington State have found that population and employment density as well as mixed land use are positively related to transit use and walking for shopping and work-related trips (L. D. Frank & Pivo, 1994). Later studies have generally confirmed an independent association between the rate of active transportation and land use mix/density in other US urban areas (Cervero & Radisch, 1996; Friedman, Gordon, & Peers, 1994; Kitamura, Mokhtarian, & Laidet, 1997).

With regards to urban design, a number of studies have investigated elements of New Urbanism, an urban design movement which combines traditional (i.e. pre-automobile era) neighbourhood design and transit-oriented development to promote a number of goals such as to increase walkability, increase social and land use mix, and reduce urban sprawl (Grant, 2006), and its association to active transportation. In the San Francisco Bay Area, researchers found that pedestrian-oriented design (with improved sidewalks, street lighting, and planted strips) led to increase in active forms of travel after adjusting for family and individual characteristics (Cervero & Kockelman, 1997). Handy’s study of non-work trips also suggests that traditional neighbourhood design, characterized by the rectilinear street grid, may lead to reduced motorized forms of travel compared to post-WWII suburban style development, which are characterized by curvilinear street patterns and numerous cul-de-sacs (SL Handy, 1996). Despite these findings regarding the environmental determinants of travel behaviour, Diez-Roux (2003) notes that the study’s emphasis on predictors of travel behaviour, rather than physical
activity, limits its application to inform health interventions, since walking and cycling are only two activities amongst many other sports and physical activities.

Parallel to the planning literature, public health researchers have also studied environmental determinants of physical activities (Humpel, Owen, & Leslie, 2002). Humpel et al.’s 2002 comprehensive literature review, through an assessment of 19 quantitative studies that assessed the relationship between perceived and objectively determined environmental attributes and physical activity behavior, determined that physical environmental factors, in particular 1) accessibility of facilities, 2) opportunities for activity, and 3) aesthetic qualities of the area, are consistently associated with physical activity behaviour after adjusting for individual characteristics. Accessibility refers to the presence of recreational facilities and distances to them. Opportunities for activity include awareness of facilities and satisfaction with them, the availability of sidewalks and equipment, and residents’ perceptions regarding opportunities for physical activity in their area. Aesthetics include the extent to which the local area is perceived as pleasant, attractive, and friendly. They also found that safety (e.g. crime, perceived neighborhood safety, street lights, and traffic) and poor weather only had mild to moderate impact on physical activity relative to the former three factors.

iii. Psychosocial Stress

The neighbourhood context may affect psychosocial stressors, such as depression, hostility, anxiety, and the level of social support, and in turn they may affect the risk of CVDs (Everson-Rose & Lewis, 2005). First, I will discuss the mechanism linking psychosocial stress and CVD; then, I will discuss how the neighbourhood context may shape psychosocial stress.

A number of studies have investigated the pathophysiological mechanism linking psychosocial stress and CVD risk. It is generally argued that psychosocial factors may influence cardiovascular function and promote atherogenesis through the hypothalamic-pituitary-adrenal axis, which are activated in response to stress, depression and anxiety (Musselman, Evans, & Nemeroff, 1998). Chronic dysregulation of the hypothalamic-pituitary-adrenal axis, which occurs in depression and anxiety (Plotsky, Owens, & Nemeroff, 1998), can result in hormonal and neuroendocrine alterations, including hypercortisolemia or excess glucocorticoid secretion. Even small increases in
glucocorticoids sustained over time can contribute to hypertension, insulin resistance, visceral obesity, coagulation changes, and increased lipid levels, all of which are precursors to CVD (Chrousos & Gold, 1998). In a selective review of key epidemiological and clinical studies by Everson-Rose and Lewis (2005), they conclude that methodologically strong evidence indicates that depression, anxiety, and lack of social support do increase the risk of CVD in healthy populations.

Overall, there are few studies that look directly at neighbourhood influences on psychosocial stress as a mediator for clinical end-point CVDs. One such study, by Sundquist et al. (2006), examined two neighbourhood psychosocial stressors, namely neighbourhood rate of violent crimes and unemployment rate, and found that they were positively associated with coronary heart disease after controlling for individual level confounders. On the other hand, there are many studies that look at contextual influence on depression, and they might be useful here to shed light on plausible pathways. For example, Matheson et al.’s (2006) study in Canada suggests that residential instability (measured by indicators such as percentage living alone and percent home ownership) and neighbourhood deprivation are key neighbourhood-level psychosocial stressors that were associated with depression, after controlling for gender, age, education, marital and visible minority status and neighborhood-level ethnic diversity. In addition, in a study of New York City residents, Galea et al. (2005) found that residence in a neighbourhood characterized by poor quality built environment (i.e. housing disrepair) was associated with greater likelihood of past six month and lifetime depression in multilevel models adjusting for age, race/ethnicity, sex, and income and for neighbourhood level income. Taken together, these studies suggest that the neighbourhood context may impact levels of psychosocial stress, which may in turn have an impact on CVD risk.

iv. Air pollution and Noise

In a recent American Heart Association scientific statement (Brook et al., 2004), experts discuss how exposure to carbon monoxide, oxides of nitrogen, sulfur dioxide, ozone, lead, and coarse and fine particulate matter can lead to adverse CVD outcomes including hospitalization and mortality due to CVD (Dominici et al., 2006; Mann, Tager, & Lurmann, 2002; Pope, Burnett, & Thun, 2002; Pope, Burnett, & Thurston, 2004;
Samet, Dominici, & Curriero, 2000), myocardial infarction (Peters, Dockery, Muller, & Mittleman, 2001; Zanobetti & Schwartz, 2005), heart failure exacerbation, progression of atherosclerosis (Suwa et al., 2002), and decreased rate of survival after stroke (Maheswaran et al., 2010). Since motorized vehicular traffic is the source for a large number of these pollutants, closely related to this body of literature is the investigation of traffic exposure on CVD (Adar & Kaufman, 2007). There is now evidence to suggest that residential proximity to traffic is an independent factor for increased cardiovascular and stroke mortality rates (Finkelstein, Jerrett, & Sears, 2004) and myocardial infarction (Hoffmann et al., 2006; Tonne et al., 2007). While traffic counts are often used as a proxy for actual measures of chemicals that make up air pollution, which may be more expensive and harder to obtain due to field sampling requirements, Vedal (2009) raises the concern that traffic-related exposures such as noise may have a distinct impact on CVDs which should be studied separately. In fact, Stansfeld et al.’s review of the impact of noise on urban health summarized its negative impact on blood pressure and sleep quality and duration, proposes that further investigations of how noise may be an important contextual determinant of CVD should be undertaken.

Based on the review above, I identify two important research gaps that informed the empirical study undertaken here. First, studies using neighbourhood deprivation needs to include additional specific features of the built environment. This is important since neighbourhood deprivation can be a proxy for a number of contextual factors like healthy food availability, so it is important to understand its unique contribution (i.e. strictly neighborhood deprivation effect and not as a proxy for resource availability) to CVD variance. Second, while each of the mechanisms discussed above (diet, physical activity, psychosocial stress, and the environment) represent possible areas of collaboration for urban planners and public health professionals, no previous work has been done to combine all four areas into an integrated approach to understand the impact of the built environment on CVDs. This is a valuable approach because 1) we can evaluate their relative influence on CVD to set policy priorities, 2) one-dimensional characterization of neighbourhoods, such as only using neighbourhood deprivation, does not do enough to clarify specific contextual factors, which may help to elucidate the specific causal pathways and provide opportunities for more effective interventions, and 3) by testing a
full range of possible environmental factors affecting CVDs, we can begin to understand
the built environment as a complex system where multiple contextual factors may interact
with each other and also with individual level covariates. The analysis to follow aims to
answer the following research questions:

1. Is neighbourhood-level socio-environmental context correlated with higher risk of
myocardial infarction and other CVDs? Can socio-environmental factors explain
the association between neighbourhood deprivation and CVD outcomes?
2. To what extent do socio-environmental characteristics, taken together, account for
neighbourhood variation in CVD?
3. Does the association between the socio-environmental context and CVDs remain
after controlling for individual level risk factors?

2. Study of Neighbourhood environment and CVDs in Toronto, Canada
i. Methods

The Neighbourhood Effects on Health and Well-being (NEHW) project uses a
cross-sectional survey designed specifically to understand the impact of neighbourhood
level determinants, acting independently or interactively with individual level factors, on
population health including CVD outcomes. The study uses a three stage sampling
technique: first, 50 out of the total 140 neighbourhood planning areas (NPA) delineated
by the City of Toronto were randomly selected; second, from the 50 NPAs sampled (each
containing 2 to 10 census tracts (CT), median = 3), 2 CTs were randomly selected,
resulting in 100 randomly selected CTs. CTs are small, relatively stable geographic units
with populations between 2500 and 8000, and are used as a proxy for residential
neighbourhoods because they generally have a high degree of internal homogeneity with
regards to social and economic conditions (Matheson et al., 2006), and their boundaries
are bordered with topographic features and major roads.

Due to budget limitations, the final number of sampled CTs was reduced from
100 to 87, where some NPAs only had 1 CT selected rather than 2. Thorough checks
were performed to ensure that the remaining CTs were representative of the
socioeconomic profile of the City of Toronto. Third, households were randomly selected
within each sampled CT based on their residential address. The recruitment target was 30
households per CT, but the actual number of households recruited ranged from 9-31
households, and the goal of 30 households per CT was reached in 51 out of 87 (58.6%)
CT sampled. An individual was eligible to participate if they met all of the following criteria: 1) a resident of the selected household (one per household), 2) were between the ages of 25 and 64, 3) were able to communicate in English, and 4) had lived in the neighbourhood for at least 6 months. The response rate was 72%.

Data were collected between Mar 2009 and June 2011. In total, 2412 participants, representing 47 NPAs and 87 CTs, were included in the study’s sample. Data were obtained from in-person interviews using a computer-assisted personal interviewing program administered by interviewers trained by the Survey Research Unit (SRU) at the Centre for Research on Inner City Health (CRICH) at St Michael’s Hospital. Interviews were conducted at participants’ residences or in private interview rooms located at CRICH. All participants provided written informed consent at the time of their interview. The Research Ethics Board at St Michael’s hospital in Toronto, Canada provided ethics approval for this study.

ii. Data Sources

Outcomes

The CVD outcomes were self-reported history of physician diagnosis of myocardial infarction (MI), angina, coronary heart disease (CHD), stroke, and congestive heart failure (CHF). Although this survey did not include clinical details, previous studies of the validity and reliability of self-reported conditions have suggested a high level of agreement with medical records for most of the conditions considered here (Bush, Miller, Golden, & Hale, 1989; Kehoe, Wu, Leske, & Chylack, 1994). This study examined 2 outcome sets with increasing prevalence, because some of the CVD outcomes were rare events, and also for the purpose of checking the robustness of the findings. Outcome group 1 included participants with a self-reported history of MI. Outcome group 2 included participants with a reported history of MI, angina, CHD, CHF, or stroke.

Predictors

i. Individual level measures

The structured interview provided self-reported information on age and total after-tax household income (in thousands) as continuous variables, and gender, visible minority status (which includes persons, other than Aboriginal persons, who self-
identified as not white in race or colour), and education as categorical variables. While age and measures of individual SES are more typically entered into regression as categorical variables, this is not exclusively the case (Eldeirawi & Lipton, 2003).

While unmeasured individual-level variables has been critiqued as a problem for the study of neighbourhood effects (Oakes, 2004), many of them, such as health behaviours (e.g. exercise and diet) and medical history (e.g. diabetes, hypertension, and obesity), are intermediate endpoints (or mechanisms) that lie between socio-environmental exposure and CVD outcomes (Ana V. Diez-Roux, 2004). Thus, variables that represent sociodemographic risk factors believed not to be on the causal pathway between the socio-environmental context and CVD were chosen. Also, smoking, coded as two dummy variables 1) former smoker vs. non-smoker and 2) current smoker vs. non-smoker, is included as an individual level control variable.

**ii. Socio-environmental contextual measures**

The selection of these variables is guided by the previous research reviewed above linking contextual measures and CVD. For each of the 4 mechanisms through which the neighbourhood context may impact CVD outcomes (diet, physical activity, psychosocial stress, and air pollution & noise), I will discuss the specific variables used.

With regards to diet, two variables are included. Data on access to 1) food stores and supermarkets and 2) fast food restaurants were obtained from the City of Toronto Public Health food inspection reports (Toronto Public Health Inspection, 2011). Since all food-selling premises in the city are required to be regularly reviewed for licensing purposes, the database provides comprehensive and up-to-date information. Addresses were geocoded using GeoPinpoint v. 3.3 (DMTI Spatial Inc., Markham, On, Canada), and the data were imported into ArcGIS Editor 9.3 (ESRI, Redlands, CA, USA) for spatial data manipulation and analysis. In each CT, the number of 1) supermarkets and food stores and 2) fast food restaurants were summed and normalized by the total area (in km²) of the CT.

Information on the availability of parks and recreational facilities was included in order to consider the impact of socio-environmental context on physical activity. Land
use information from the CanMap Route Logistics annually updated geo-database (DMTI Spatial Inc, 2011) were used to calculate the total area in each CT used for parks or recreational facilities. This number is normalized by the total area of each CT, which results in the mean area used for parks and recreation per km².

Two potential contextual sources of psychosocial stress are captured in this study. First, the *incident-based uniform crime reporting database* (Statistics Canada & Canadian Centre for Justice Statistics, 2010) provided data on the number of violent crimes per CT for the year 2006, which includes the most serious violations in incidents of sexual assault, criminal harassment, uttering threats, minor/major assault, and robbery. These counts are divided by the total area for each CT, which gives the average number of violent crimes per km². Second, the percent of housing requiring major repairs in each CT was used as another source of psychosocial stress. Data on housing disrepair aggregated at the CT level is obtained from the Canada census 2006 long form questionnaire (Statistics Canada, 2008), which is a 20% sample of the total census population.

Traffic data obtained from the City of Toronto Transportation Services (Toronto Transportation Services, 2011) is used as a proxy for air pollution. Average weekday 24 hour volume of Toronto Street segments were recorded in the geo-database of a street map shapefile in ArcGIS 9.3, then 100-meter circular buffers were created around each participant’s home location. Using buffer analysis, those who resided within 100-meters of a major road, defined by the City of Toronto as having at least an average of 20,000 vehicles in a 24-hour period, were coded as exposed (1), those who do not have a major road within their 100-meter buffer were coded as *not exposed to high traffic* (0). This method of categorizing traffic exposure and the use of a 100-meter buffer zone has been used previously to investigate the effect of exposure to traffic on mortality rate advancement periods (Finkelstein et al., 2004).

Furthermore, data on exposure to noise is added to understand its contribution to CVD risk, and to ensure that the traffic data properly proxies air pollution by separating out the unique effect of noise. Participants were asked in the interview whether they agreed with the statement “the noise level where I live often disturbs me” with the answer chosen on a 5-point Likert scale from strongly agree to strongly disagree. Lastly,
neighbourhood level median after-tax household income from the Canada census 2006 as a proxy of neighbourhood level SES, since neighbourhood SES has been shown to have an independent effect on CVD risk.

iii. Statistical Analysis

The analysis began by evaluating the bivariate association between individual level and neighbourhood level predictors and self-reported history of outcome 1 (MI only) and outcome 2 (any of MI, angina, CHD, CHF, or stroke) using $\chi^2$ test for pairs of categorical variables and 1-way analysis of variance for unbalanced design pairs of continuous and categorical variables. Since the data has a 2-level structure, with individuals nested in neighbourhoods, multilevel logistic regression is used to account for the lack of independence within neighbourhood clusters (Raudenbush & Bryk, 2002). First, the null model (model #1) was used to analyze how much of the variance in each of the 2 outcome measures can be attributed to the census tract level. This is followed by a series of models: first, introducing individual level risks only (model #2), then model #3 which adjusts for only neighbourhood level factors without any individual level factors. Model #4 introduces all the neighbourhood level predictors without individual level risk factors, while model #5 includes individual level controls along with the neighbourhood level predictors. This “block-entry” method allowed us to understand 1) whether the neighbourhood SES was a unique predictor from (or whether it is proxy for) specific socio-environmental predictors, and 2) the effect of adjusting for individual level risks on the neighbourhood-context only model: a significant reduction in the estimates of the socio-environmental predictors’ effect sizes would suggest that the contextual effects are better explained by the clustering of individuals with similar socio-demographic risk profiles. Multilevel logistic regression was performed using SAS 9.3 in the GLIMMIX (generalized linear mixed model) procedure.

Results

Descriptive Analysis

Participants in Project NEHW were 58% female and 71% white (self-identified), with a mean age of 48 years (table 1.1) and a mean after-tax family income of $107,712.
Those who have CVDs have a mean age of 54 and a mean after-tax family income of $77,033: one-way ANOVA testing for difference in means between participants with CVD vs. those without are significant for both variables at p<0.001. While a minimum of six months of residence in their home was required for all participants, the mean number of years was 11.84 (range = 0.5-55 years, SD = 8.5), which mean that most participants may have lived long enough in their neighbourhood to be affected by possible socio-environmental factors. Overall, 46 (1.91%) reported a previous history of MI, and 108 (4.48%) reported a previous history any CVDs (any one of MI, angina, CHD, CHF, or stroke). Among participants who reported MI, 71.74% (33 out of 46) also reported a history of angina, CHD, CHF, or stroke. Bivariate associations between Outcome 1 & 2 and socio-demographic characteristics are presented in table 1.1.
Table 1.1: Self-Reported History of CVD Outcomes and Characteristics of Study Participants

<table>
<thead>
<tr>
<th></th>
<th>Outcome 1: MI only</th>
<th>Outcome 2: Any of MI, angina, CHD, CHF, or stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entire Cohort No. (%)</td>
<td>No. (%)</td>
</tr>
<tr>
<td>Participants</td>
<td>2412 (100)</td>
<td>46 (1.91)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-35</td>
<td>359 (14.88)</td>
<td>4 (1.11)</td>
</tr>
<tr>
<td>36-45</td>
<td>598 (24.79)</td>
<td>6 (1.00)</td>
</tr>
<tr>
<td>46-55</td>
<td>761 (31.55)</td>
<td>12 (1.58)</td>
</tr>
<tr>
<td>56-65</td>
<td>694 (28.77)</td>
<td>24 (3.46)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1390 (57.63)</td>
<td>12 (0.86)</td>
</tr>
<tr>
<td>Men</td>
<td>1022 (42.37)</td>
<td>34 (3.33)</td>
</tr>
<tr>
<td>Visible Minority Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Yes</td>
<td>833 (34.54)</td>
<td>14 (1.68)</td>
</tr>
<tr>
<td>2. No</td>
<td>1579 (65.46)</td>
<td>32 (2.03)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Less than High School</td>
<td>88 (3.65)</td>
<td>4 (4.55)</td>
</tr>
<tr>
<td>2. High School Complete</td>
<td>402 (16.67)</td>
<td>9 (2.24)</td>
</tr>
<tr>
<td>3. Diploma Complete</td>
<td>616 (25.54)</td>
<td>15 (2.44)</td>
</tr>
<tr>
<td>4. Completed Undergrad and above</td>
<td>1306 (54.15)</td>
<td>18 (1.38)</td>
</tr>
</tbody>
</table>

Note: MI = Myocardial infarction; CHD = coronary heart disease; CHF = congestive heart failure. For variables with more than 2 categories, P is for the $\chi^2$ test of differences across row in the percentage reporting the outcome. Otherwise P is for a 2-tailed test of difference between 2 groups in the proportion reporting each outcome.

Multilevel logistic regression

Table 1.2 reports a series of multilevel logistic regressions predicting outcome 1 (MI only) and outcome 2 (any CVDs). Starting with a fully unconditional ordered logit model of the dichotomous outcome of MI (outcome 1) and any CVD (outcome 2), shown under model 1 in table 1.2, the intraclass correlation is calculated to decompose variance in the predictors across individual level versus CT level of analysis. The intraclass correlation (ICC) for multilevel logistic regression is calculated as:

$$ICC = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2}$$
where $\sigma^2_{u}$ is the variance between neighbourhoods, and $\sigma^2_{e}$, defined by the constant $\pi^2 / 3$, is the variance between individuals (Guo & Zhao, 2000). The estimate of the intercept variance in the null model (model 1) is 0.13996 for MI (p<0.001) and 0.15620 for any CVDs (p<0.001), which results in an ICC of 0.0408 and 0.0453 respectively (see table 1.2). This suggests, on average, 4.08% of the variance in MI and 4.53% of the variance in any CVDs can be attributed to the neighbourhood level.

The result of the ICC answers research question #2: namely, CT level socio-environmental context accounts for about 4-5% of the variance of MI and CVD prevalence. However, there are a number of points to keep in mind when interpreting the ICC values. While most of the variability in CVDs is at the individual level, the values of these ICCs are similar to ICCs found in other multilevel studies (Snijders & Bosker, 1999), which rarely exceed 20%. In is important to note that 1) the size of the ICC does not rule out relatively large effects of neighbourhood level measures (Duncan & Raudenbush, 1999), and 2) two out of the eight socio-environmental predictors, namely living near a major road and noise, have a very small area of effect (up to 100 meters) such that the between group variance of both predictors are much smaller than within group variance at the CT level (1-way ANOVA has been performed to confirm this with p<0.999 and p=0.998 respectively). Thus, while noise and living near a major road are factors shaped by the built environment, they operate at a smaller scale than the CT, and the effect they have on CVD outcomes will not be captured by the ICC. In other words, the 4-5% ICC for MI and CVD capture only the effects of observed and unobserved factors that operate at the CT scale, but the variance of CVD explained by noise and traffic (which operate at a scale smaller than CT) are not captured. Thus, one might argue that ICC values underestimate the true total variance attributable to all socio-environmental factors, since not all factors operate at a single neighbourhood scale.

Model 2 reports the results of a regression predicting MI (outcome 1) and all CVD (outcome 2) based on demographic variables. Despite the inclusion of individual-level characteristics, the variance of the random intercept remains significant for both outcomes, which indicates that there are still unexplained variance at the CT level (p<0.001) after accounting for the way in which individuals (with similar characteristics) may cluster.
Model 3 reports results for outcomes 1 and 2 regressed on neighbourhood wealth, using the CT-level median after-tax household income as the proxy. The negative effect that neighbourhood wealth, using CT median after tax household income (in thousands), has on the likelihood of MI (OR = 0.979, p<0.05) and any CVDs (OR = 0.989, p<0.05), is consistent with the literature linking neighbourhood deprivation and CVDs. Also, the variance of the random intercept remains significant for outcome 1 and 2 (p<0.001), which means that neighbourhood wealth/deprivation alone cannot completely explain the pattern of CVDs at the CT level.

Model 4 introduces all the socio-environmental characteristics in addition to neighbourhood wealth. From this model, it is found that some neighbourhood-level socio-environmental factors, such as access to food stores, fast food, parks and recreation, violent crimes, living close to a major road and noise, are indeed correlated to a higher risk of myocardial infarction and other CVDs (answer to question #1). At this point, the intercept variance component is reduced to zero, meaning that the socio-environmental variables included here explained all the variance at the CT level. Changes in neighbourhood wealth across model 3 and 4 (for both outcomes) indicate that neighbourhood wealth is no longer significant after adjusting for socio-environmental variables (i.e. food stores, fast food, parks/recreational land uses, violent crimes, housing disrepair, living near a major road, and noise), which indicates that socio-environmental factors can explain the association between neighbourhood deprivation and CVD outcomes.

Model 5 presents the fully adjusted model including neighbourhood wealth, socio-environmental variables, as well as individual level risk factors. After adjusting for sociodemographic variables, it is found that each additional food store per km\(^2\) in the neighbourhood is associated with a reduction in the likelihood of both MI (OR = 0.955, p<0.05) and any CVDs (OR = 0.963, p<0.05); having at least 1 fast food restaurant in the neighbourhood is associated with an increase in the likelihood of MI (OR = 4.613, p<0.05) and any CVDs (OR = 2.212, p<0.05); each additional percent of area used for parks and recreation per km\(^2\) is associated with a reduced likelihood of MI (OR = 0.931, p<0.05) but not associated any CVDs (outcome 2); each additional violent crime per km\(^2\) in the neighbourhood is associated with an increase in the likelihood of MI (OR = 1.004,
p<0.05) and any CVDs (OR = 1.004, p<0.05); the level of housing disrepair is not significantly associated with MI or any CVD outcomes; living near a major road is associated with an increase likelihood of MI (OR = 3.671, p<0.001) and any CVDs (OR = 1.728, p<0.01); and participants who are disturbed by noise at home (strongly agree vs. strongly disagree) have an increased likelihood of MI (OR = 5.457, p<0.01) and any CVDs (OR = 2.868, p<0.01).
Table 1.2: Multilevel Logistic Regressions of outcome 1 (MI only) and outcome2 (any CVDs)

<table>
<thead>
<tr>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Intercept only (Null Model)</td>
<td>Adjusted for individual level risk factors only</td>
<td>Adjusted for neighbourhood level deprivation only</td>
<td>Adjusted for all neighbourhood level factors</td>
<td>Adjusted for all neighbourhood level factors and individual sociodemographic risk factors</td>
</tr>
<tr>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
</tr>
<tr>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
</tr>
</tbody>
</table>

*\(p<0.05\); **\(p<0.01\); ***\(p<0.00\)
In separate analyses (not shown), models were specified with random slopes for gender, age, individual level income, education, and visible minority status in order to determine whether these effects vary across neighbourhoods. The fixed effects of gender, age, income, and education were significant in the models reported above (table 1.2), but allowing the slopes to vary on these covariates revealed no significant variance components suggesting that gender, age, income and educational differences is constant across the CTs. Also, no significant cross-level interactions were found.

In summary, the analyses indicate that socio-environmental contexts, including the access to food, fast food restaurants, access to parks and recreational areas (for MI only), rate of violent crimes, living near high traffic and noise exert significant effects on MI and any CVDs, even after controlling for individual level risk factors (answer to research question #3). On the other hand, while housing disrepair was a possible pathway to CVD through its effect on psychosocial stress (Galea et al., 2005), this analysis did not find the association to be significant for either MI or any CVDs.

Checking the pathway: testing obesity and physical activity as mediators

A number of neighbourhood level factors such as the availability of food stores/supermarkets, fast food, and areas for parks and recreation have a significant effect on CVD or MI risk, where their effect is thought to be mediated by their impact on obesity and physical activity. In order to verify the theorized pathway, an additional model (model #6) is specified where the participants’ body mass index (BMI) and regular physical activity were adjusted. Since the American College of Sports Medicine and the American Heart Association recommend a minimum of 30 minutes of moderate-intensity aerobic (endurance) physical activity on five days each week to promote and maintain cardiovascular health (Haskell et al., 2007), a dichotomous variable is created for those who self-report to exercise for at least 2.5 hours (coded as 1) versus those who do not (coded as 0). If obesity and physical activity were mediators for the aforementioned neighbourhood factors, we would expect that their effects drop to null once the two mediators are added in model #6. See table 1.3 below for the model #6 regression coefficients of the neighbourhood level factors.
Table 1.3: Multilevel Logistic Regression – Regression coefficients for neighbourhood level factors after adjusting for individual confounders and mediators (BMI and regular physical activity)

<table>
<thead>
<tr>
<th>Model 6</th>
<th>‘Over-adjusted’ model: adjusting for all individual confounders, neighbourhood factors, and 2 mediators: BMI and engage in regular physical activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome 1: MI only Odds ratio (95% CI)</td>
<td>Outcome 2: any CVD Odds ratio (95% CI)</td>
</tr>
<tr>
<td>Total (n=2412)</td>
<td>1.000 (1.000-1.000)</td>
</tr>
<tr>
<td>Median after-tax household income of CT</td>
<td>0.998 (0.739-1.257)</td>
</tr>
<tr>
<td>Have at least 1 fast food restaurant in CT</td>
<td>4.613 (0.072-17.618)</td>
</tr>
<tr>
<td>Mean area used for parks and recreation per km2 in CT</td>
<td>1.386 (0.001-4.216)</td>
</tr>
<tr>
<td>Mean number of violent crimes per km2 in CT</td>
<td>1.004</td>
</tr>
<tr>
<td>% of housing in CT requiring major repairs</td>
<td>93.133 (&lt;0.001-100.000)***</td>
</tr>
<tr>
<td>Lives within 100 meters of a major road</td>
<td>3.471 (1.733-6.774)**</td>
</tr>
<tr>
<td>Disturbed by noise at home</td>
<td>5.455 (1.934-15.379)**</td>
</tr>
<tr>
<td>1. Strongly Agree</td>
<td>0.629 (0.181-2.195)</td>
</tr>
<tr>
<td>2. Agree</td>
<td>1.312 (0.441-3.894)</td>
</tr>
<tr>
<td>3. Neither agree or disagree</td>
<td>1.256 (0.531-2.972)</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001

Note that the model above should not be used to estimate neighbourhood contextual effects, because by including mediators into the model, it is technically over-adjusted. As expected, as seen in table 1.3 above, the 95% confidence interval for the odds ratios for the availability of food stores/supermarkets, fast food, and areas for parks and recreation now includes 1.00 and are no longer statistically significant once the mediators (BMI and regular exercise) were added the model; however, violent crimes, exposure to high traffic and noise continue to be significant because BMI and regular exercise are not in their theorized causal pathway. The findings in model #6 add credibility to the pathways theorized in this study.

Discussion

While health professionals have generally targeted individual lifestyle choices (such as diet and exercise) to modify health behaviors, land-use, architectural and urban
design decisions are rarely considered. The findings here suggest that the reintegration of urban planning and public health may be a productive way to shift some of our focus from treatment and prevention solely at the individual level to prevention for ‘sick populations’ (Rose, 2001), since a number of socio-environmental determinants (with significant effect on CVDs) are shaped by land-use decisions and the built environment.

Before discussing how to incorporate knowledge of the socio-environmental determinants to inform novel CVD preventive strategies, it is important to point out a number of limitations for this study. First, disease assessments were limited to self-reports; therefore, it is not possible to determine the clinical nature or severity of conditions, and undiagnosed cases may be differentially misreported. Second, neighbourhood boundaries were limited to census tracts, which may not provide the most relevant construct of neighbourhood, as is the case for the area effect of noise and living close to a high-traffic road. Third, it is also important to acknowledge that the relationships found here may not necessarily be causal, because of health-selected migration, that is, residents may choose the neighbourhoods in which they live on the basis of health-related characteristics (R. J. Sampson, Morenoff, & Gannon-Rowley, 2002). For example, a hypothetical case might exist where an individual’s CVD status leads to a loss of income-earning capacity, and as a result has to move from a relatively high-quality neighbourhood into a more disadvantaged neighbourhood. In order to overcome the said bias, a longitudinal study, with a sufficiently lengthy study period, is required to examine the change in CVD status over the study period so that the effects of health-selected migration can be controlled out. The cross-sectional design of this study simply does not allow for health-selection migration to be controlled. Despite these limitations, this exploratory study points to potentially new pathways through which the socio-environmental context may influence CVD etiology.

Land use decisions are inextricably public health decisions. The research here suggests that living in neighbourhoods with inadequate access to local food stores and areas for physical activity, burdened by an over-representation of violent crimes and fast food restaurants, and over-dependent on automobiles (leading to air and noise pollution) may have a significant impact on the risk of CVDs, over and above individual-level risks. The following briefly outlines some emerging planning, design and architectural
innovations, through key examples, that may be useful starting places to forge productive coalitions between health professionals, planners, architects, builders, urban designers, and transportation officials to ameliorate these socio-environmental risks.

With regards to improving access to healthy and affordable food, the Puget Sound Regional Food Policy Council in Washington, have been at the forefront of regional food planning that involves local partners from planning, public health, retail food sector, academic researchers, and community anti-hunger advocates since 2010 (Puget Sound Regional Council, 2011). Beginning with a regional food assessment of the Puget Sound region, the council has pushed to implement a number of best practices to promote food equity. For example, the use of planning and economic incentives to open grocery stores, farmers markets, and food carts in underserved communities, and providing municipal support to simplify the licensing process and reducing the cost of opening flexible low-cost, non-“brick and mortar”, forms of food access such as fresh fruit/vegetable carts and stands.

To support healthier choices in dining out venues, taking a draconian approach that some health researchers have suggested to restrict fast food restaurants in residential areas though zoning by-laws (Boone-Heinonen et al., 2011) will probably be unsuccessful for a number of reasons: 1) the risk of antagonizing private interests and the business community significantly undermines the political will to pass such a municipal bylaw, and 2) since zoning only applies to new development, the existing fast food restaurants already in place will not be affected. In central city areas where established businesses are already in place, changing zoning does little to shape the landscape of dining options. An innovative strategy implemented by the Rainier Valley Healthy Restaurant Initiative in Seattle Washington, in consultation with dietitians, collaborated with 16 restaurants to make changes in their recipes to reduce fat, sodium, and calories, and to increase fiber, fruit and vegetable servings (Rainier Valley Health Coalition, 2011). Scaling up such public health and private collaborations may be our best chance at improving healthier options in dining out venues.

Urban planners and designers have been looking at ways to shape our cities to reduce the dependence on motor vehicle use in order to decrease air pollution, traffic congestion and increase active forms of transportation for decades (Ewing & Cervero,
2001; L. Frank & Engelke, 2000; L. D. Frank & Pivo, 1994; Freudenburg, Galea, & Vlahov, 2006; Friedman et al., 1994; S Handy, 1996; Handy et al., 2002). The research presented here confirms the significant impact of traffic on cardiovascular health. Rather than taking current levels of reliance on automobiles for granted, the level of vehicular-use may one day be seen as a modifiable risk factor given innovative design and planning strategies. Research across US and European countries show that levels of non-motorized travel is, at least partially, attributable to higher levels of population and employment density, greater level of mixing land uses, better infrastructure for pedestrians and cyclists (L. Frank & Engelke, 2000). In a review of literature on urban form’s effect on mode choice, Frank and Engelke (2000) have identified consensus among researchers regarding the specific design features to reduce automobile traffic and promote non-motorized travel: 1) traffic calming, 2) pedestrian and cycling friendly street design, 3) increasing mixed use neighbourhood design, and 4) site design features like reduction in building setback. Health professionals can productively collaborate with urban designers and planners in order to implement and evaluative some of these strategies.

In addition to the finding of the association between noise of an elevated risk of MI and other CVDs, many of the aforementioned planning and design practices may lead to increased population density, and in turn may aggravate residential exposure to noise; therefore, site planning and architectural solutions to reduce exterior noise may be increasingly important. In the case of new developments, we can orientate residences away from sources of noise, use built and natural noise shields, create room arrangements to place noise sensitive rooms (sleep and rest areas) further from sources of exterior noise, and use building materials and techniques known to be acoustically isolating (US Department of Transportation, 2011). In the case of existing residences, soundproofing windows facing major roads, and sealing cracks and edges of walls may be the most economical solution.

With regards to the findings of the association between the rate of violent crimes and CVDs, one way for health professionals and planners to collaborate is to consider how environmental design may be used to reduce crime through the production of defensible spaces, a concept elaborated by Oscar Newman (1973) and later researchers (Cozens, Saville, & Hillier, 2005). By optimizing opportunities for surveillance (e.g.
improved street lighting and reducing refuge for the potential offender), clearly defining boundaries (and defining preferred use within such spaces), creating and maintaining a ‘positive image’, increasing active community participation (i.e. to engender positive social activities, encourage neighbours to take ownership of space, and take advantage of natural surveillance), proponents argue that environmental design is effective in reducing both crime and fear of crime in the community (Cozens et al., 2005). While definitive proof has not been demonstrated, design oriented crime prevention tools may be a pragmatic and cost-effective first step for cross-disciplinary collaboration to improve public health and CVD outcomes.

While this study may be exploratory in nature, employing a cross-sectional and observational method, the study has nonetheless successfully evaluated the association between CVDs and the multi-dimensional neighbourhood socio-environmental context, thus the study advanced beyond a relatively simplistic characterizing of neighbourhoods, using neighbourhood deprivation as the only contextual factor as many other multilevel studies have done. In doing so, this study helps to shed light on the specific causal pathways and provide opportunities for more effective and advanced interventions that are best carried out through cross-disciplinary partnerships between health professionals and urban planners.

In linking the study findings to the concrete and explicit development of innovative interventions, this study invites urban planners and designers to join the critical and ongoing discussion of how to tackle the global leading cause of death and disability, and for health professionals to rethink the built environment as a modifiable risk factor.
Reference List


Study 2: Place effects on cardiovascular health: Do non-residential exposures and the duration of exposure make a difference?

Research on human activity and travel patterns suggests that the geographic extent of everyday lives is not limited to residential neighbourhoods (Buliung & Kanaroglou, 2006; Gliebe & Koppelman, 2005; Law, 1999; Naess, 2006). It is thus a critical limitation that virtually all studies of place effects on health have focused solely on the residential environment, save for rare exceptions (Inagami, Cohen, & Finch, 2007; Muntaner et al., 2006). Chaix (2009) has termed this research gap the ‘residential trap’, because of the exclusive reliance on local residential environments and the systematic neglect of non-residential environments. The purpose of this study is to extend the understanding of place-health associations to non-residential exposures for cardiovascular diseases (CVD), a leading cause of death and disability in the US, UK, Canada and in most countries around the world (American Heart Association, 2003; Bonow, Smaha, Smith, Mensah, & Lenfant, 2002). Bridging this research gap is important because greater understanding of the geography of everyday lives can help us 1) identify new sites and potential opportunities for effective intervention, and 2) clarify the true association of residential neighbourhood on health by adjusting for exposure to non-residential neighbourhoods.

Multilevel investigations to examine residential neighbourhood-level context have noted consistency documenting the independent effects of neighbourhood context after adjusting for individual-level covariates on a variety of cardiovascular health outcomes including increased incidence of coronary heart diseases (A. V. Diez-Roux et al., 2001; Sundquist, Malmstrom, & Johansson, 2004; Winkleby, Sundquist, & Cubbin, 2007), increased incidence of myocardial infarction (Lovasi et al., 2008), increased all-cause and cardiovascular disease mortality (Smith, Hart, Watt, Hole, & Hawthorne, 1998), increased coronary heart disease case fatality (Winkleby et al., 2007), and increased risk factors such as smoking, physical inactivity, obesity, diabetes, and hypertension (Cubbin et al., 2006; Ellaway, Anderson, & Macintyre, 1997; Matheson, White, Mouneddin, Dunn, & Glazier, 2010). Yet for these studies, two important factors impede effective translation of research findings to policies for CVD interventions: 1) while the impact of
the neighbourhood of residence may be especially salient to those who are spatially segregated or socially isolated (Inagami et al., 2007), for others, the impact of non-residential environments such as school and work may be equally or more influential, thus these studies leave out major sources of contextual exposure; and 2) the use of neighbourhood socioeconomic status (SES) as a proxy for unmeasured environmental and psychosocial factors does not allow for the development or recommendation of policies aimed at modifying socio-environmental risk factors due to its unspecific nature. The present study overcomes both barriers by using specific socio-environmental risk factors across multiple environments to better understand the ‘real-life spatial trajectories’ and its benefits and risks for CVDs (Cummins, Curtis, Diez-Roux, & Macintyre, 2007).

Simultaneously the analysis of residential and non-residential environments presents a unique opportunity to investigate the duration of exposure in different settings throughout the day. In virtually all neighbourhood effects on health studies, little has been done to account for time spent across different contexts and how the duration of exposure may modify place-health associations. By excluding time in these models, there is an implicit assumption that people spend relatively similar amounts of time at home and in other settings, which is likely not the case. In the study to follow, I compare the time-weighted analysis to an unweighted analysis, and I hypothesize that the time-weighted analysis would result in improved model fit and stronger regression coefficients for already significant effects in the unweighted analysis. These improvements might result because I am able to properly modify the influence of socio-environmental factors according to an estimated time spent in the associated context.

This paper has three sections: first, I review the literature on possible pathways linking residential neighbourhood socio-environments and CVDs. I also discuss previous research on non-residential exposures (Inagami et al., 2007), and suggest ways to overcome research gaps. Second, I present an original study of neighbourhoods in Toronto, Canada to investigate the association between CVD risk and residential/non-residential socio-environmental contexts to address these gaps. Third, I discuss the results with regards to implications for future research, policy, and interventions.
i. Literature review of Neighbourhood effects on CVDs

In her literature review of neighbourhood effects on CVD outcomes, Diez-Roux (2003), states that socio-environmental factors are likely to be important in shaping the distribution of CVDs. After adjusting for individual level risk factors, studies have found that neighbourhood deprivation had an independent impact on a number of cardiovascular health outcomes including increased incidence of myocardial infarction (Lovasi et al., 2008), increased incidence of coronary heart diseases (A. V. Diez-Roux et al., 2001; Sundquist et al., 2004; Winkleby et al., 2007), increase in coronary heart disease case fatality (Winkleby et al., 2007), increase in all cause and cardiovascular disease mortality (Smith et al., 1998), and increase in risk factors such as smoking, physical inactivity, obesity, diabetes, and hypertension (Cubbin et al., 2006; Ellaway et al., 1997; Matheson et al., 2010). In the studies above, neighbourhood deprivation is perceived to be a proxy for a range of unmeasured neighbourhood features, but the lack of information on specific features diminishes their applicability to policy and intervention (Ana V. Diez-Roux, 2003; R. Sampson & Raudenbush, 2004).

In order to deepen our understanding of neighbourhood features relevant to CVD risk, we can look at the mechanisms through which place has been theorized to influence cardiovascular health (Chaix, 2009). There are four mechanisms that are most discussed in the literature (Ana V. Diez-Roux, 2003; Kawakami, Li, & Sundquist, 2011): 1) diet, 2) physical activity, 3) psychosocial stress, and 4) air pollution and noise. These four mechanisms are discussed below in turn.

1) Diet

Many studies show that an individual’s diet is not solely determined by SES, gender and ethnicity, but is also shaped by the local food environment (Boone-Heinonen et al., 2011; S Cummins & Macintyre, 1999; Ana V. Diez-Roux, 2003; Laraia, Siega-Riz, Kaufman, & Jones, 2004; Larson, Story, & Nelson, 2009; Mooney, 1990; Moore, Diez Roux, Nettleton, & Jacobs Jr., 2008; Morland, Wing, & Diez Roux, 2002; Morland, Wing, Diez Roux, & Poole, 2002; Wechsler, Basch, & Shea, 1995). Researchers have pursued two parallel lines of inquiry to understand the impact of the local food
environment on diet. The first deals with the spatial distribution of food, and the second deals with the extent to which the local food environment affect dietary choices.

Research on the spatial distribution of food and grocery stores suggests that healthy food availability might mediate the relationship between neighbourhood deprivation and CVD rates. For example, Morland et al (2002) in a US study found that there are 4 times more supermarkets in affluent white neighborhoods compared to poor black neighborhoods. Their study suggests that the negative health impact of living in poor neighbourhoods may be mediated by the inadequate access to healthy foods. The unequal spatial distribution of healthy food, where the level of accessibility is negatively correlated to neighbourhood deprivation, has also been documented across a number of cities in countries including the UK (Mooney, 1990) and the US (Wechsler et al., 1995). Moreover, there is evidence to suggest that the availability of unhealthy food is positively correlated with neighbourhood deprivation. For example, the density of fast food restaurants is found to be higher in low-income and predominately black neighbourhoods (Block, Scribner, & DeSalvo, 2004). Also, Morland et al (2002) found that there are 3 times fewer places to consume alcoholic beverages in the wealthiest compared to the poorest neighborhoods in a study that included four US states. While the (mostly American) research cited above has identified an association between availability of healthy food and neighbourhood deprivation, research in Glasgow found that food stores were numerous in deprived as well as non-deprived neighbourhoods (Cummins et al., 2007). In order to explain their null finding in contrast to the numerous American studies, they suggest that there is a relatively equal spatial distribution of retail food stores in the Scottish city compared to the US, which could be due to the difference between European versus American patterns of urban dwellings and rental markets.

Parallel to the research on the spatial distribution of healthy/unhealthy food, studies have also examined the extent to which the local food environment shapes an individual’s dietary habits and health outcomes. However, there are contradicting results between US and UK studies. For example, in a US study of the impact of supermarket availability on the diet of individuals (n=10,623), it was found that each additional supermarket in the census tract increased fruit and vegetable intake by 32% for Black Americans, while white Americans who live in census tracts with at least 1 supermarket
have their fruit and vegetable intake increase by 11% compared to those with no supermarkets (Morland, Wing, & Diez Roux, 2002). In another US study, Moore et al (2008) found that individuals with no supermarkets near their homes were 25-46% less likely to have a healthy diet compared to those with the most stores, after adjustment for individual level variables. Laraia et al. (2004) also found that proximity of retail food stores is positively associated with the diet quality of pregnant women in in North Carolina, US. However, not all US studies have consistently identified the association between local food environment and diet. For example, Bonne-Heinonen et al, (2011) found that while fast food consumption is independently associated with fast food availability in a US longitudinal study of young and middle-aged adults, the relationship between supermarket/grocery stores availability and diet quality and fruit and vegetable intake were inconclusive. The results above contrast with a study from Glasgow, Scotland (Macdonald, Ellaway, Ball, & Macintyre, 2011) where researchers found no association between proximity to food stores and diet or obesity. The null result is not unexpected because Glasgow has a relatively equal spatial distribution of retail food stores across the city as shown in the Cummins et al. (2007) study cited above.

2) Physical Activity

There is a growing interest to understand how planners can use land use and urban design to increase active forms of transportation such as walking and bicycling (Handy, Boarnet, Ewing, & Killingsworth, 2002). Research in this area has tremendous relevance for population health because studies have shown that even short daily episodes of moderate physical activity may be sufficient to produce cardiovascular and other health benefits (Pate et al., 1995). Researchers have investigated how 1) land development patterns (e.g. density and land use) and 2) urban design factors can promote or hinder active forms of transportation and levels of physical activity. In the following, I discuss the evidence for both dimensions respectively.

With regards to land development patterns, Frank and Pivo (1994), in their study of Washington state, have found that population and employment density as well as mixed land use are positively associated with transit use and walking for shopping and work-related trips. Other studies also have similar findings that confirm an independent
association between land use mix/density and the rate of active transportation and in other North American urban centres (Cervero & Radisch, 1996; Friedman, Gordon, & Peers, 1994; Kitamura, Mokhtarian, & Laidet, 1997).

Related to research on land development patterns, planners and architects have also explored the use of urban design to promote physical activity. Many of these efforts are unified by the ‘New Urbanism’ movement, an urban design movement which combines pre-automobile era neighbourhood design and transit-oriented development to promote a number of goals such as to increase active forms of transportation such as walking and cycling, increase social and land use mix, and reduce urban sprawl (Grant, 2006). Cervero and Kockelman (1997), in their research of the San Francisco Bay Area, found a number of design features to have a significant impact on increasing active transportation even after adjusting for family and individual characteristics. These features include improved sidewalks, street lighting, and planted strips. In addition, research on non-work trips suggests that traditional neighbourhood design, which features a rectilinear street grid, reduced motorized forms of travel compared to post-WWII suburban style development characterized by cul-de-sacs and curvilinear street patterns (Handy, 1996). While these studies demonstrate the impact of the built environment on travel behaviour, increasing the amount of active transportation at the population level may not be enough as a form of health intervention, since walking and cycling are only two activities amongst many other sports and physical activities (Ana V. Diez-Roux, 2003).

Outside of the fields of planning and architecture, public health researchers are also interested in the environmental determinants of physical activity. In a systematic review of quantitative studies that investigated the relationship between perceived as well as objectively determined environmental attributes and physical activity behaviour (Humpel, Owen, & Leslie, 2002), the authors found three factors that are consistently associated with physical activity behaviour after adjusting for individual characteristics. The factors include 1) accessibility of facilities, 2) opportunities for activity, and 3) aesthetic qualities of the area. The first factor refers to how easy it is to get to a recreational facility measured by distance. The second factor deals with residents’ perception of local opportunities for physical activities, which can include awareness and
satisfaction with sports facilities, sidewalks, parks, trails, and other amenities. The third factor deals with whether the residents’ neighbourhood is rated as pleasant, attractive, and friendly, with high rated areas more conducive to physical activities. Surprisingly, the review also noted that the presence of street lights, traffic, poor weather, safety (as measured by the number of crimes, perceived neighbourhood safety) only had mild impact on physical activity compared to the three main factors identified above.

3) Psychosocial Stress

The pathway linking neighbourhood psychosocial stressors to CVD outcomes is somewhat complex, and requires that we look at two bodies of research that have developed separately to hypothesize a possible pathway: 1) individual mental health and psychosocial stress as determinants of cardiovascular health, and 2) the neighbourhood socio-environmental determinants of psychosocial stress.

First, there is mounting evidence that individual psychosocial factors, such as depression, hostility, and anxiety, chronic/acute psychosocial stressors, and lack of social support are independently associated with the development and progression of CVDs (Everson-Rose & Lewis, 2005). Musselman et al. (1998) argue that a pathophysiological mechanism links psychosocial stress to CVD risk, because psychosocial factors can promote atherogenesis through the hypothalamic-pituitary-adrenal axis, which is activated in response to emotional distress, stress, and anxiety. This can have a negative impact on cardiovascular health because chronic dysregulation of the hypothalamic-pituitary-adrenal axis as a result of psychosocial stress can lead to neuroendocrine and hormonal changes, and hypercortisolemia (Plotsky, Owens, & Nemeroff, 1998). If sustained over time, even minor hypercortisolemia, or small increases in glucocorticoids, can increase the risk to CVDs, because it can contribute to insulin resistance, coagulation changes, increased lipid levels, visceral obesity, and hypertension (Chrousos & Gold, 1998). Everson-Rose and Lewis’s review (2005) of key epidemiological and clinical studies also concludes that there is now a scientific consensus that anxiety, depression, and lack of social support increase the risk of CVD holding other factors constant.

With regards to neighbourhood determinants of psychosocial stress, researchers have identified neighbourhood social disorganization, measured by neighbourhood rate of
violent crimes and unemployment rate, which is associated with coronary heart disease after controlling for individual level confounders (Sundquist et al., 2006). We can also look at the literature on the contextual influence on depression since it has been identified as a determinant of CVD in the literature above. In a Canadian study, researchers found that residential instability (measured by indicators such as percentage living alone and percent home ownership) and neighbourhood deprivation are key neighbourhood-level psychosocial stressors that were associated with depression, after controlling for age, gender, marital status, education, and visible minority status and neighborhood-level ethnic diversity (Matheson et al., 2006). In a New York City study (Galea, Ahern, Rudenstine, Wallace, & Vlahov, 2005), built environment (measured by poor quality/deteriorating structures and housing disrepair) was also found to be associated with greater likelihood of past six month and lifetime depression in multilevel models adjusting for age, race/ethnicity, sex, and income and for neighbourhood level income. These studies suggest that contextual factors such as residential instability and the built environment may have an effect on the risk of CVD through their impact on psychosocial stress or depression.

4) Air pollution and Noise

There is a scientific consensus that exposure to carbon monoxide, oxides of nitrogen, sulfur dioxide, ozone, lead, and coarse and fine particulate matter can lead to adverse CVD outcomes including hospitalization and mortality due to CVD (Brook et al., 2004; Dominici et al., 2006; Mann, Tager, & Lurmann, 2002; Pope, Burnett, & Thun, 2002; Pope, Burnett, & Thurston, 2004; Samet, Dominici, & Curriero, 2000), myocardial infarction (Peters, Dockery, Muller, & Mittleman, 2001; Zanobetti & Schwartz, 2005), heart failure exacerbation, progression of atherosclerosis (Suwa et al., 2002), and decreased rate of survival after stroke (Maheswaran et al., 2010).

The association between exposure to vehicular traffic and CVD risk is also noted in the literature because it is a major source for many of the aforementioned pollutants (Adar & Kaufman, 2007). A number of studies have that that living near high traffic is an independent contributor for increased rates of myocardial infarction (Hoffmann et al., 2006; Tonne et al., 2007), and cardiovascular and stroke mortality rates (Finkelstein,
Since measurements of air pollution can be expensive and difficult to obtain due to field sampling requirements, the above studies make use of traffic counts as a proxy for these measures. However, Vedal (2009) cautions the effects of traffic may be confounded by traffic noise, which may have its own independent impact on CVD risk that deserves to be studied separately. This point is further emphasized by Stansfeld et al.’s systematic review of the impact of noise on urban health. They found that noise may negatively impact blood pressure and sleep quality/duration. As such, the study below will include both traffic measures and noise measures as distinct predictors to estimate their independent contribution to CVD risk.

\textit{ii. Non-residential exposures and the importance of time-weighted analysis}

Inagami et al. (2007) introduced non-residential contexts (using neighbourhood level SES as a proxy for unobserved influences) to the study of the influence of small area influence on self-rated health, and found that 1) residence in disadvantaged neighbourhood was associated with poor self-rated health, and 2) greater exposure to less disadvantaged non-residential neighbourhoods was associated with improved self-rated health. It was an innovative study that overcame the ‘residential trap’; however, an important limitation was the lack of consideration for the relative amount of time spent in residential vs. non-residential settings. In virtually all neighbourhood effects on health studies, little has been done to account for time spent in a context and how duration of exposure may modify place-health associations. Similarly, in the Inagami study, residential and non-residential contexts were assumed to be equally influential for participants (i.e. they were unweighted); however, it is more likely that the amounts of time spent in various contexts differ between participants. In the study to follow, I compare the time-weighted analysis to an unweighted analysis, and hypothesize that the time-weighted analysis would result in improved model fit and stronger regression coefficients for already significant effects in the unweighted analysis. These improvements might result because the model is able to properly modify the influence of socio-environmental factors according to an estimated time spent in the associated context.
Inagami et al. did not use cross-classified models, which mean that the non-residential contexts are treated as a ‘variable’ rather than ‘levels’. In this way, they missed an important opportunity to understand the relative importance of each setting. In a cross-classified analysis, variance proportion values, represented by the intraclass correlation coefficient (ICC), are calculated for each level or setting (e.g. one for each level or context considered). This allows us to understand the total variance that is accounted for by observed and unobserved factors in each of the levels, which gives an indication of the relative importance of each setting. In the study to follow, I use cross-classified modeling, and restrict the focus to the two most influential settings for working-age adults (i.e. the target population of our study is working age adults from age 25-65), the work and home settings, so that I can estimate the relative importance of these two settings to answer the question of relative importance. Finally, Inagami et al. use neighbourhood SES as a proxy for unmeasured contextual factors, which I previously critiqued for its unspecific nature. The use of socio-environmental risk predictors in this study may shed light on the specific pathways that underlie what has generally been referred to as ‘neighbourhood disadvantage’.

Based on the research gaps identified above, this study seeks to answer the following research questions:

1) What is the relative importance of the residential context versus the workplace context for CVD risk?
2) Do residential and/or workplace socio-environmental contexts correlate to the risk of CVDs after adjustments for individual-level risk factors?
3) Does the time-weighted analysis improve model-fit and strengthen the effect sizes of possible associations between socio-environmental risks and CVD?

2. Methods

The Neighbourhood Effects on Health and Well-being (NEHW) project uses a cross-sectional survey designed specifically to understand the impact of neighbourhood level determinants, acting independently or interactively with individual level factors, on population health including CVD outcomes. The study uses a three stage sampling technique: first, 50 out of the total 140 neighbourhood planning areas (NPA) delineated
by the City of Toronto were randomly selected; second, from the 50 NPAs sampled (each containing 2 to 10 census tracts (CT), median = 3), 2 CTs were randomly selected, resulting in 100 randomly selected CTs. CTs are small, relatively stable geographic units with populations between 2500 and 8000, and are used as a proxy for residential neighbourhoods because they are constructed to maximize internal homogeneity with regards to social and economic conditions, and their boundaries are bordered with topographic features and major roads.

Due to budget limitations, the final number of sampled CTs was reduced from 100 to 87, where some NPAs only had 1 CT selected rather than 2. Thorough checks were performed to ensure that the remaining CTs were representative of the socioeconomic profile of the City of Toronto. Third, households were randomly selected within each sampled CT based on their residential address. The recruitment target was 30 households per CT, but the actual number of households recruited ranged from 9-31 households, and the goal of 30 households in 51 out of 87 (58.6%) CT sampled was reached. An individual is eligible to participate if they met all of the following criteria: 1) a resident of the selected household (one per household), 2) were between the ages of 25 and 64, 3) were able to communicate in English, and 4) had lived in the neighbourhood for at least 6 months. The response rate was 72%.

Data were collected between Mar 2009 and June 2011. In total, 2411 participants, representing 47 NPAs and 87 CTs (residential neighbourhoods), were included in the NEHW survey. Since this study is interested in the interaction between residential and workplace influences, only individuals who were employed or on leave are included. While full address information was available for each participant’s residential location, workplace locations were provided by 1681 out of the total 1705 (98.6%) participants who were employed or on temporary/parental leaves. In the process of geocoding these locations, 39 addresses were discovered to contain incorrect information and 16 participants worked outside of the Toronto census metropolitan area, and these participants were removed. This resulted in a usable sample of N=1626. It should be noted that 1444 people (88.8%) worked in the City of Toronto (downtown), and the remaining 182 participants are employed in the Toronto metropolitan area but outside the city core area.
Data were obtained from in-person interviews using a computer-assisted personal interviewing program administered by interviewers trained by the Survey Research Unit (SRU) at the Centre for Research on Inner City Health (CRICH) at St Michael’s Hospital. Interviews were conducted at participants’ residences or in private interview rooms located at CRICH. All participants provided written informed consent at the time of their interview. The Research Ethics Board at St Michael’s hospital in Toronto, Canada provided ethics approval for this study.

**Outcome**

The CVD outcome used is based on self-reported history of physician diagnosis of myocardial infarction (MI), angina, coronary heart disease (CHD), stroke, and congestive heart failure (CHF). Although this survey did not include clinical details, previous studies of the validity and reliability of self-reported conditions have suggested a high level of agreement with medical records for most of the conditions considered here (Bush, Miller, Golden, & Hale, 1989; Kehoe, Wu, Leske, & Chylack, 1994). Participants with a reported history for any of MI, angina, CHD, CHF, or stroke are coded as 1, others are coded as 0.

**Individual level predictors**

The structured interview provided self-reported information on age and total after-tax household income as continuous variables, and gender, visible minority status (which includes persons, other than Aboriginal persons, who are not white in race or colour), and education as categorical variables. While unmeasured individual-level variables has been critiqued as a problem for the study of neighbourhood effects (Oakes, 2004), many of them, such as health behaviours (e.g. exercise and diet) and medical history (e.g. diabetes, hypertension, and obesity), are intermediate endpoints (or mechanisms) that lie between socio-environmental exposure and CVD outcomes (Ana V. Diez-Roux, 2004). Thus, only variables that represent sociodemographic risk factors believed not to be on the causal pathway between the socio-environmental context and CVD were chosen. Smoking, coded as two dummy variables 1) former smoker vs. non-
smoker and 2) current smoker vs. non-smoker, was also included as an individual level risk factor.

**Socio-environmental contextual predictors**

The selection of these variables is guided by the previous research reviewed above linking contextual factors and CVD. For each of the 4 mechanisms through which the neighbourhood context may impact CVD outcomes (diet, physical activity, psychosocial stress, and air pollution & noise), I will discuss the specific variables used. Since virtually no research examined the possible pathways linking workplace socio-environmental context and CVD risk, I tested the effects of the same set of socio-environmental risk factors for both residential and workplace contexts.

With regards to diet, two variables are included. Data on access to 1) food stores and supermarkets and 2) fast food restaurants were obtained from an annually updated commercial geodatabase, Enhanced Points of Interest 2011 (DMTI Spatial Inc, 2011b), which contains all Canadian business locations categorized by business type and specialty. In each workplace and residential CT, the number of 1) supermarkets and food stores and 2) fast food restaurants were summed and normalized by the total area (in km$^2$) of the CT.

Information on the availability of parks and recreational facilities were included in order to consider the impact of socio-environmental context on physical activity. Land use information from the CanMap Route Logistics annually updated geo-database (DMTI Spatial Inc, 2011a) were used to calculate the total area in each residential and work CT used for parks or recreational facilities. This number is normalized by the total area of each CT, which results in the mean area used for parks and recreation per km$^2$.

Two potential contextual sources of psychosocial stress are captured for residential and work CTs. First, the *incident-based uniform crime reporting database* (Statistics Canada & Canadian Centre for Justice Statistics, 2010) provided data on the number of violent crimes per CT for the year 2006, which includes the most serious violations in incidents of sexual assault, criminal harassment, uttering threats, minor/major assault, and robbery. These counts are divided by the total area for each CT, which gives the average number of violent crimes per km$^2$. Second, I used the percent of
housing requiring major repairs in each CT as another source of psychosocial stress. Data on housing disrepair aggregated at the CT level is obtained from the Canada census 2006 long form questionnaire (Statistics Canada, 2008), which is a 20% sample of the total census population.

Traffic data obtained from the Toronto Transportation Services (2011) is used as a proxy for air pollution. Average weekday 24 hour volume of Toronto Street segments were recorded in the geo-database of a street map shapefile in ArcGIS 9.3, then 100-meters circular buffers were created around each participant’s home location. Using buffer analysis, those who resided within 100-meters of a major road, defined by the City of Toronto as having at least an average of 20,000 vehicles in a 24-hour period, were coded as exposed (1), those who do not have a major road within their 100-meter buffer were coded as not exposed (0). This method of categorizing traffic exposure and the use of a 100-meter buffer zone has been used previously to investigate the effect of exposure to traffic on mortality rate advancement periods (Finkelstein et al., 2004).

Furthermore, data on residential exposure to noise is added to understand its contribution to CVD risk, and to ensure that the traffic data properly proxies air pollution by separating out the unique effect of noise. Participants were asked in the interview whether they agreed with the statement “the noise level where I live often disturbs me” with the answer chosen on a 5-point Likert scale from strongly agree to strongly disagree. Noise information is not available for workplace locations.

Neighbourhood SES variable, for both residential and workplace, is included since neighbourhood deprivation has been shown to have an independent effect on CVDs. Neighbourhood level median after-tax household income from the Canada census 2006 is used as a proxy.

Finally, road network distance between each participant’s home and workplace is calculated using the Network Analyst tool in ArcGIS 9.3. This is a better estimation of the distance each participant travels from home to work compared to Euclidean distance, because it takes into account the distance traveled over the existing road network. The distance between home and work may capture a number of contextual influences such as additional exposure to air pollution on the road and a potential source of psychosocial stress that is the result of time spent in commute.
Time-Weighted Analysis

The time-weighted analysis used information provided by the participants on the time they spend at work. Each participant reported the number of hours worked in an average week (mean = 36.93, S.D. = 13.63). Since time spent at home was not a question asked on the NEHW survey, it is predicted based on the respondent’s time spent at work and their socio-demographic and lifestyle characteristics. Time-activity pattern is significantly influenced by demographic and lifestyle factors, and the strongest predictors consistently include age, gender, work status, marital/relationship status, and whether the participant has any children living at home (McCurdy & Graham, 2003; Schweizer et al., 2007; Yang et al., 2011). Using the Canadian General Social Survey 2005 (cycle 19) that included a section on time-use, I built a multivariate linear regression to model the association between time spent at home in minutes (dependent variable) and a number of socio-demographic predictors, including gender, age, income, marital status, have any children (under 19) at home, and education, which have been proven to be statistically significant determinants of time spent at home in the studies cited above. Also, time spent at work (in minutes) is included as a predictor. The regression coefficients from this model are then used to predict each participant’s time spent at home in the NEHW study.

The Canadian General Social Survey (GSS) 2005 is a routine cross-sectional survey conducted by Statistics Canada with a sample of 19,598 individuals over 15 years of age. The models below employ the time-use estimates reported in the one-day time use diary portion of the survey, which details the time spent on a wide variety of day-to-day activities and the locations where these activities occurred (e.g. at work, at home, etc.). More details on the methodology of the survey can be found in the cited document (Statistics Canada, 2005). While the full sample of the GSS is 19,598 individuals representative of all the provinces of Canada, the purpose of the models below is to create a predictive model for individuals with the characteristics of project NEHW participants (i.e. age 25-65, living in a city, and working full-time or part-time). To achieve this, 2489 cases were deleted for individuals under the age of 25, and 3589 cases were deleted for individuals over the age of 65. From the remaining 13519 cases, 204 cases were deleted for individuals who did not report their labour force status, and 3578 cases were deleted
for non-working adults. From the remaining 9736 respondents, 2292 cases were deleted for all individuals who do not live in a large urban center (i.e. census metropolitan areas), and a further 1269 participants are removed because of missing data, which means a final sample of 6175 is used in the following sub-analysis. The variables for duration of time spent at home and at work include responses for both weekdays and weekends (correctly proportioned to be representative of the typical 7 day week), which means that the models below have taken into account the weekday versus weekend differences for time spent at home. All the findings discussed next are significant (at least) at the level of \( p<0.05 \).

Using only time spent at work as a predictor of time spent at home (Model A in table 2.1 below), I find that each additional minute spent at work reduces the amount of time spent at home by 0.63 of a minute. This model with only a single predictor explained 35% of the variance in the outcome. Model B improves upon model A by including predictors that were previously identified as socio-demographic determinants of time spent at home. Model B finds that females (compared to males) spend 16.36 more minutes at home; each additional year in age increases the amount of time by 1.39 minutes; single people (compared to married or people in common-law relationships) spend 60.31 fewer minutes at home; having at least 1 child under the age of 19 increases the number of minutes at home by 24.67; and there is a significant trend that people with a lower household income spends more time at home compared to those with higher incomes (see Table 2.1 Model B for details). While education, measured by those who completed high school versus those who did not, seems to have an impact on the time spent at home in a bivariate situation (1-way ANOVA test of difference in means was significant at \( p<0.05 \)), the impact of education was not statistically significant in the multivariate linear model. It should also be noted that the inclusion of socio-demographic predictors only led to a modest improvement of the variance explained by 1.5%, where the adjusted \( r^2 \) square of Model B is 36.5%.
Table 2.1: Regression Coefficients for predictors of time spent at home in minutes (n=6175)

<table>
<thead>
<tr>
<th></th>
<th>Model A</th>
<th></th>
<th></th>
<th>Model B</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>Standard</td>
<td>P-value</td>
<td>Beta</td>
<td>Standard</td>
<td>P-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>1091.41</td>
<td>4.52</td>
<td>.000</td>
<td>1002.83</td>
<td>19.80</td>
<td>.000</td>
</tr>
<tr>
<td>Duration of time at work (in minutes)</td>
<td>-0.63</td>
<td>0.01</td>
<td>.000</td>
<td>-0.62</td>
<td>0.01</td>
<td>.000</td>
</tr>
<tr>
<td>Gender (female=1)</td>
<td>16.36</td>
<td>6.12</td>
<td>.008</td>
<td>16.36</td>
<td>6.12</td>
<td>.008</td>
</tr>
<tr>
<td>Age</td>
<td>1.39</td>
<td>0.31</td>
<td>.000</td>
<td>1.39</td>
<td>0.31</td>
<td>.000</td>
</tr>
<tr>
<td>Single (not married or in common law)</td>
<td>-60.31</td>
<td>7.27</td>
<td>.000</td>
<td>-60.31</td>
<td>7.27</td>
<td>.000</td>
</tr>
<tr>
<td>Have at least 1 child at home under 19</td>
<td>24.67</td>
<td>6.66</td>
<td>.000</td>
<td>24.67</td>
<td>6.66</td>
<td>.000</td>
</tr>
<tr>
<td>Household income under $19,999 vs those with over $100k</td>
<td>93.27</td>
<td>16.12</td>
<td>.000</td>
<td>93.27</td>
<td>16.12</td>
<td>.000</td>
</tr>
<tr>
<td>Household income $20k to $39,999 vs those with over $100k</td>
<td>61.53</td>
<td>10.35</td>
<td>.000</td>
<td>61.53</td>
<td>10.35</td>
<td>.000</td>
</tr>
<tr>
<td>Household income $40k to $59,999 vs those with over $100k</td>
<td>46.64</td>
<td>9.11</td>
<td>.000</td>
<td>46.64</td>
<td>9.11</td>
<td>.000</td>
</tr>
<tr>
<td>Household income $60k to $79,999 vs those with over $100k</td>
<td>32.30</td>
<td>4.41</td>
<td>.001</td>
<td>32.30</td>
<td>4.41</td>
<td>.001</td>
</tr>
<tr>
<td>Household income $80k to $99,999 vs those with over $100k</td>
<td>26.39</td>
<td>10.19</td>
<td>.010</td>
<td>26.39</td>
<td>10.19</td>
<td>.010</td>
</tr>
</tbody>
</table>

Model A has an adjusted r square of 0.35; Model B has an adjusted r square of 0.365

The coefficients in Model B are substituted into the data in the NEHW survey to predict the participants’ time spent at home, and from there, a ‘work’ weight is created to be assigned to all variables describing the work context, and ‘home’ weight is created to be assigned to all variables for the home context. For example, a single, 30 year old, female with no children, working 40 hours per week, with a household income of $50,000 would be predicted as follows:


\[
\text{Minutes spent at home per day} = 1002.83^a - 0.62(342.86)^b + 16.36^c + 1.39(30)^d - 60.31^e + 0^f + 46.64^g
\]

- \(a\) the intercept
- \(b\) the average minutes of work per day is 40 hours divided by 7 days, and then multiplied by 60 minutes, which results in 342.86 minutes per day averaged across both weekends and weekdays
- \(c\) 16.36 minutes are added for being female
- \(d\) 1.39 minutes are added for each additional year of age; therefore, 30 is multiplied by 1.39 because the participant is 30 years of age
- \(e\) she is single; therefore, she spends 60.31 fewer minutes at home
- \(f\) she has no children; therefore, no minutes are added for time spent at home
- \(g\) 46.64 minutes are added because her household income is $50,000

In the example above, the model estimates that 834.65 minutes (13.91 hours) are spent at home per day averaged across the entire week; therefore 97.37 hours per week are estimated for time spent at home. Since a week has 168 hours in total, a weight of 0.58 (i.e. 97.37/168) is assigned to home variables and a weight of 0.24 (i.e. 40/168) is assigned to work variables.

For categorical variables, weighing by time effectively changes what it actually measures. For example, while the unweighted variable is dichotomous (1 vs. 0) for living near high traffic, the time-weighted equivalent estimates the proportion of time in a week that a participant is exposed to high traffic near their residence. For the hypothetical female participant above with a home weight of 0.58, the living near high traffic variable is calculated as (0.58*1)*100%, which indicates that in 58% of her time, she is exposed to a high-traffic road.

There is one concern for using time spent at work versus (the estimation of) time spent at home to weigh the contextual variables: if these ratios were correlated to the risk of CVD, it might bias the estimation of contextual effects. One-way ANOVA revealed that this was not the case. The number of hours spent at work did not seem to differ between those with CVD versus those without (p=0.819).

**Statistical Analysis**

I began by evaluating the bivariate association between individual level and residential/workplace neighbourhood level predictors and self-reported history of CVDs using \(\chi^2\) test for pairs of categorical variables and 1-way analysis of variance for unbalanced design pairs of continuous and categorical variables.
The structure of the data is participants nested within residential neighbourhoods and workplaces. Since not all participants in a given residential neighbourhood go to work in the same area, this situation is said to have a ‘cross-classified’ structure and is best modeled using cross-classified multilevel analysis (Goldstein, 2003; Sykes & Musterd, 2011). Cross-classified models recognize the simultaneous membership of level-1 units (i.e. participants) in multiple higher and non-nested levels (i.e. residential and workplace neighbourhoods). Cross-classified models estimate the influence of each context while controlling for effects of the other context, which make it an appropriate strategy in this study.

The analysis begins (model set #1) with the estimation of unconditional or ‘null’ models for a) residential context only, b) workplace context only, and c) combined residential and workplace contexts. These models estimate the variance in CVD outcome that is attributed to each of the levels. Model set #2 introduces socio-environmental factors for a) residential context only, b) workplace context only, and c) combined residential and workplace contexts. Finally, model set #3 repeats the model set #2 but adjusts for individual level controls. For model set #2 and #3, separate analysis is performed for the time-weighted variables and unweighted variables to understand whether the former yield stronger and/or better fitted results. Note that multilevel models are used for a) residential context only and b) workplace context only models, since they have a simple nested structure with individual nested neighbourhoods; cross-classified models are necessary for all models that examined c) combined residential and workplace contexts. Multilevel and cross-classified logistic regressions were performed using SAS 9.3 in the GLIMMIX (generalized linear mixed model) procedure. All regression coefficients are reported in odds ratios (OR) with a minimum significance level of p<0.05.

Results

Descriptive Analysis

Participants in the study were 55% female and 29% identified as visible minority (self-identified as non-white and non-aboriginal), with a mean age of 47 years (table 2.2) and a mean after-tax family income of $113,982. Those who have CVDs have a mean
age of 52 and a mean after-tax family income of $61,531: one-way ANOVA testing for difference in means between participants with CVD versus those who are CVD-free are significant for both variables at p<0.05. While a minimum of six months of residence in their home was required for all participants, the mean number of years was 11.84 (range = 0.5-55 years, SD = 8.5), which means that most participants may have lived long enough in their neighbourhood to be affected by possible socio-environmental factors. Overall, 50 participants (3.1%) reported a previous history of any CVDs (any one of MI, angina, CHD, CHF, or stroke).

Table 2.2: Descriptives of socio-demographic characteristics for entire cohort and participants with at least 1 CVD

<table>
<thead>
<tr>
<th>Variable</th>
<th>Entire Cohort</th>
<th></th>
<th></th>
<th>Participants with CVDs</th>
<th></th>
<th></th>
<th>P for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>No. (% of total)</td>
<td>Mean</td>
<td>S.D.</td>
<td>No. (row %)</td>
<td></td>
</tr>
<tr>
<td>Participants (N=1627)</td>
<td>---</td>
<td>---</td>
<td>1627 (100%)</td>
<td>---</td>
<td>---</td>
<td>50 (3.1%)</td>
<td>---</td>
</tr>
<tr>
<td>Individual Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>47.34</td>
<td>19.13</td>
<td>---</td>
<td>52.31</td>
<td>8.38</td>
<td>---</td>
<td>0.0004</td>
</tr>
<tr>
<td>Male</td>
<td>---</td>
<td>---</td>
<td>738 (45.36%)</td>
<td>---</td>
<td>---</td>
<td>28 (3.79%)</td>
<td>0.1248</td>
</tr>
<tr>
<td>Female</td>
<td>---</td>
<td>---</td>
<td>889 (54.64%)</td>
<td>---</td>
<td>---</td>
<td>22 (3.47%)</td>
<td>0.1248</td>
</tr>
<tr>
<td>Education levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1296</td>
</tr>
<tr>
<td>1. Less than High School</td>
<td>---</td>
<td>---</td>
<td>39 (2.40%)</td>
<td>---</td>
<td>---</td>
<td>3 (7.69%)</td>
<td></td>
</tr>
<tr>
<td>2. High School Complete</td>
<td>---</td>
<td>---</td>
<td>242 (14.87%)</td>
<td>---</td>
<td>---</td>
<td>9 (3.72%)</td>
<td></td>
</tr>
<tr>
<td>3. Diploma Complete</td>
<td>---</td>
<td>---</td>
<td>412 (25.32%)</td>
<td>---</td>
<td>---</td>
<td>16 (3.88%)</td>
<td></td>
</tr>
<tr>
<td>4. Completed Undergrad and above</td>
<td>---</td>
<td>---</td>
<td>934 (57.41%)</td>
<td>---</td>
<td>---</td>
<td>22 (2.36%)</td>
<td></td>
</tr>
<tr>
<td>Visible Minority</td>
<td>---</td>
<td>---</td>
<td>465 (28.58%)</td>
<td>---</td>
<td>---</td>
<td>14 (0.86%)</td>
<td>0.8204</td>
</tr>
<tr>
<td>After-tax family income (in thousands)</td>
<td>113.98</td>
<td>87.57</td>
<td>---</td>
<td>89.48</td>
<td>61.53</td>
<td>---</td>
<td>0.0445</td>
</tr>
</tbody>
</table>

Multilevel and Cross-Classified Logistic Analysis

I began the analysis by running unconditional models and calculating the ICCs, in order to understand where the variation in CVD risk lies. The intraclass correlation (ICC) for multilevel logistic regression is calculated as:

$$ICC = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2}$$

where $\hat{\sigma}_u^2$ is the variance between neighbourhoods, and $\hat{\sigma}_e^2$, defined by the constant $\pi^2/3$, is the variance between individuals (Guo & Zhao, 2000). The estimate of the intercept variance in the null model (see table 2.2) are 0.13996 for the residential context.
(p<0.001), 0.0959 for the work context (p<0.001), and in the combined home and work model, they are 0.1561 and 0.0754 respectively (both at p<0.001). The ICCs are reported in table 2.3, and they suggest that, on average, 4.53% of the variance of CVDs can be attributed to the residential context alone, and 2.84% of the variance of CVDs can be attributed to the work context alone. However, as shown in model 1c, when I control for the home context, the variance explained by the work context is slightly reduced (down to 2.13%) and the home context is unaffected.

Table 2.3: Unconditional Models: Intra-class Correlations and Model Fit Statistic represented by -2 restricted log pseudo-likelihood

<table>
<thead>
<tr>
<th></th>
<th>Model 1a: Residential Context</th>
<th>Model 1b: Workplace Context</th>
<th>Model 1c: Work and Home as cross-classified level-2 units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept of the Covariance Parameter (home)</td>
<td>0.1562</td>
<td>---</td>
<td>0.1561</td>
</tr>
<tr>
<td>Intercept of the Covariance Parameter (work)</td>
<td>---</td>
<td>0.0959</td>
<td>0.07154</td>
</tr>
<tr>
<td>Intra-class Correlation (ICC) Level 2: Home neighbourhood</td>
<td>4.53%***</td>
<td>---</td>
<td>4.53%***</td>
</tr>
<tr>
<td>Intra-class Correlation (ICC) Level 2: Work neighbourhood</td>
<td>---</td>
<td>2.83%***</td>
<td>2.13%***</td>
</tr>
<tr>
<td>-2 Res Log Pseudo-Likelihood</td>
<td>14238.84</td>
<td>10306.59</td>
<td>10146.06</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001

The results of the ICC answer research question #1, namely the relative importance of the residential context versus the workplace context for CVD risk. With the work context’s ICC of 2.13% in model 1c, the home context (with an ICC of 4.53%) was about 2 times more important than the work context. The total combined ICC of 6.66% is the total variance explained by work and home contexts. While most of the variability in CVDs is at the individual level, the values of these ICCs are in the typical range, which rarely exceeds 20% (Snijders & Bosker, 1999). It is important to note that 1) the size of the ICC does not rule out relatively large effects of neighbourhood level measures (Duncan & Raudenbush, 1999), and 2) a number of socio-environmental predictors, namely i) distance between home and work, ii) living near high traffic, iii) working near high traffic, and iv) residential noise, have areas of effect that do not operate at the scale of CTs. For example, living near a major road and noise, have a very small area of effect (up to 100 meters) such that the between group variance of both predictors are much smaller than within group variance at the CT level (1-way ANOVA.
has been performed to confirm this with p<0.999 and p=0.998 respectively). Thus, while all four variables are important socio-environmental variables, the effects they have on CVD do not operate at the scale of CT, and their influence would not be captured by the ICC. The -2 restricted pseudo-likelihood values in the last row of table 2.2 were used to compare the fit and complexity of the models (SAS, 2011), with lower values indicating a better model fit. The cross-classified model revealed a clear improvement in the fit relative to the models in which just home (1a) or work (1b) was included, suggesting that both work and home contexts should be considered when predicting CVD risk.

Next, I ran model set #2 and #3, each set is further broken down into separate models for a) residential contextual variables only, b) workplace contextual variables only, and c) the combined context variables. Model set #2 included contextual variables only, and model set #3 added individual level control variables. Since the results were substantially similar (i.e. adding individual level controls did not make any significant variables in model set #2 insignificant in model set #3 and resulted in only modest reductions in beta coefficients), therefore, I only show and discuss results for the fully adjusted model (model set #3 - see table 2.4). For each of the three levels, separate estimates were obtained for the unweighted and time-weighted analyses.

Considering the residential context alone (3a), each additional food store per km$^2$ was associated with a reduction in the likelihood of CVD (OR=0.904, p<0.001) in the unweighted analysis. When I weigh the foodstore/km$^2$ variable by the estimated time spent at home for each participant, the average effect size was increased to OR=0.801, p<0.001. In the unweighted analysis, each additional crime per km$^2$ was associated with an increased likelihood of CVD at OR=1.009, p<0.001; however, accounting for time spent at home strengthened the effects of crime to OR=1.013, p<0.001. Living near high traffic (within 100m of a road with a daily average volume of greater than 20,000 vehicles) led to a heightened risk of CVD (OR=1.904, p<0.05) for the unweighted analysis. Comparing this value to the time-weighted analysis is somewhat tricky because the odds ratio of 1.031 (p<0.05) in the time-weighted model is associated with a 1% increase of the time spent living near high traffic. Based on the work hours provided and the model to predict time spent at home, an average of 61% of total time per week is spent at home, and 25% of total time per week is spent at work. Therefore, the exposure
to high traffic at home would be calculated as $1.013^{61}$ to account for the average time spent at home, which would result in an odds ratio of 2.20 ($p<0.001$): a stronger effect compared to the unweighted model. Being strongly disturbed by noise at home was associated a heightened risk of CVD (OR=7.4, $p<0.001$); since the time-weighted model changed the dichotomous variable to a continuous variable that measures the associated effect of 1% increase in the time exposed to ‘strongly disturbing’ noise, it has to be converted to a comparable value. Using the mean of 61% time spent at home, the time-weighted model would give an estimate of OR=9.73 (i.e. $1.038^{61}$) for individuals exposed to ‘strongly disturbing’ noise at home. Thus, it appears that the influence of noise is strengthened when it is weighted using time. The overall fit of the unweighted model was also improved by the time-weighted analysis as measured by the restricted log pseudo-likelihood, which improved from 11231.35 to 9762.49.
Table 2.4. Model Set #4. Including Residential and Workplace Socio-environmental variables and Individual level Controls, Unweighted vs. Time-Weighted Analysis

1) Unweighted Analysis

<table>
<thead>
<tr>
<th>Model Set</th>
<th>Residential Environment</th>
<th>Workplace Environment</th>
<th>Residential Workplace Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model ia:</td>
<td>Median after-tax household income (in thousands)</td>
<td>Median after-tax household income (in thousands)</td>
<td>Median after-tax household income (in thousands)</td>
</tr>
<tr>
<td>Model ib:</td>
<td>Number of foodstores + supermarkets per km²</td>
<td>Number of foodstores + supermarkets per km²</td>
<td>Number of foodstores + supermarkets per km²</td>
</tr>
<tr>
<td>Model ic:</td>
<td>% of housing in CT requiring major repairs</td>
<td>% of housing in CT requiring major repairs</td>
<td>% of housing in CT requiring major repairs</td>
</tr>
</tbody>
</table>

2) Time-weighted Analysis

<table>
<thead>
<tr>
<th>Model Set</th>
<th>Residential Environment</th>
<th>Workplace Environment</th>
<th>Residential Workplace Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model ia:</td>
<td>Median after-tax household income (in thousands)</td>
<td>Median after-tax household income (in thousands)</td>
<td>Median after-tax household income (in thousands)</td>
</tr>
<tr>
<td>Model ib:</td>
<td>Number of foodstores + supermarkets per km²</td>
<td>Number of foodstores + supermarkets per km²</td>
<td>Number of foodstores + supermarkets per km²</td>
</tr>
<tr>
<td>Model ic:</td>
<td>% of housing in CT requiring major repairs</td>
<td>% of housing in CT requiring major repairs</td>
<td>% of housing in CT requiring major repairs</td>
</tr>
</tbody>
</table>

* *p<0.05; **p<0.01; ***p<0.001
While food stores, fast food restaurants, and crimes were significantly associated with the risk of CVDs in the home context, none of these factors were associated with CVD risk in the work context. In the work context only models (3b), increased neighbourhood SES as measured by a $1000 increase in the average CT after-tax household income was associated with a decreased risk of CVD (OR=0.978, p<0.05); accounting for time spent at work strengthened the regression coefficient to OR=0.952 (p<0.05). Working close to a high traffic road heightened the risk of CVD (OR=1.955, p<0.01), which is, unsurprisingly, an effect size similar to living near high traffic in the unweighted model. When I account for the time spent at work, an additional 1% time spent working near high traffic was associated with an increased risk of CVD (OR=1.031, p<0.01); and thus, based on the average of 25% of total time spent at work per week, working near high traffic corresponds with an increased risk of CVD of OR=2.145 (p<0.01), which is a stronger effect than seen in the unweighted model. The overall fit of the unweighted model, again, was improved when we accounted for time spent at work as shown by the restricted log pseudo-likelihood, which was decreased from 10118.76 to 9679.42.

In the cross-classified analysis (3c), none of the significant contextual influences in the previous two sub-models (3a and 3b) became insignificant after controlling for the effects from both contexts; however, as expected, the effect sizes of many beta coefficients were slightly reduced. In the fully adjusted cross-classified (unweighted) model, each additional food store in the residential CT reduced the risk of CVD (OR=0.906, p<0.001), and in the time-weighted analysis this effect is stronger (OR=0.812, p<0.001). In the unweighted model, having at least 1 fast food restaurant in the home CT heightened CVD risk (OR=3.556, p<0.05), and the time-weighted model estimated an odds ratio of 1.024 (p<0.05) for each 1% time spent in a home CT with fast food available, which is approximately an OR of 4.249 (p<0.05) given the average time (61%) spent at home. Each additional violent crime increased the risk of CVD in the unweighted analysis (OR=1.009, p<0.001), and the time-weighted model estimated OR=1.065 (p<0.001). In the unweighted model, living near high traffic increased the risk of CVD (OR=1.801, p<0.05), accounting for time resulted in an OR of 1.012 (p<0.05) for each 1% increase in time spent living near high traffic, which is approximately an odds
ratio of 2.07 (p<0.05) for total average exposure. Strongly disturbed by noise at home increased the risk of CVD in the unweighted model (OR=6.248, p<0.001), and an OR of 1.036 for each 1% time spend in a noisy home (p<0.001), which is approximately an OR of 8.649 for total average exposure. Each additional $1000 in the median after-tax household income in the work CT decreased the risk of CVD in the unweighted model (OR=0.979, p<0.05), and the effects increased in the time-weighted model (OR=0.981, p<0.05). Finally, working close to high traffic increased the risk of CVD in the unweighted model (OR=1.748, p<0.01), and OR of 1.022 is estimated for each 1% increase in time spent working near high traffic, which is approximately an OR of 1.995 for total average exposure.

In the case of every significant contextual factor, adjusting for the duration of exposure at work and home strengthened the effect sizes of the contextual variables. The model fit was also better in the time-weighted model compared to the unweighted model, as shown by the restricted log pseudo-likelihood that decreased from 10012.08 to 9211.73. Thus, the answer is affirmative for both research question #2 and #3: significant residential and workplace socio-environmental influences are associated with CVD risk after controlling for individual-level risk factors, and the time-weighted analysis improve model-fit and strengthened the effect sizes of all possible associations between socio-environmental risks and CVD.

*Checking the pathway: testing obesity and diet quality as mediators*

Two contextual level factors and its impact on CVD risk, namely the availability of food stores/supermarkets and fast food in the home neighbourhood, are theorized to be mediated by obesity and poor diet. In order to verify the theorized pathway, an additional model (not shown) is specified where the participants’ body mass index (BMI) and diet quality (proxied using times per month eating fruits and vegetables) were adjusted in addition to the individual level confounders used in model 3. When the two mediators (BMI and diet quality) were added to the multilevel cross-classified models (i.e. model 3c), in both the cases of the time-weight model and the unweighted model, the effect of the availability of 1) food stores/supermarkets and 2) fast food in the home neighbourhood became statistically non-significant (i.e. the 95% confidence interval of
the OR for both variables included 1.00 in the new models). In the meantime, all other significant factors (as found in model 3c), including violent crimes, traffic and noise, remained significant in this new model (at a minimum of p<0.05). This finding adds credibility to the theory of BMI and diet quality as mediating factors in the association between the availability of food stores/supermarkets and fast food (in the home neighbourhood) and CVD risk.

**Discussion**

While numerous studies have analyzed the effects of neighbourhoods on cardiovascular health (Chaix, 2009; Diez-Roux, 2003), few have analyzed the influence of specific socio-environmental factors (instead using neighbourhood SES as a proxy), and virtually no studies have looked at non-residential contextual factors and accounted for the time-spent in different contexts. This study is an original contribution to the literature not only because of the specific contextual pathways identified, but I have shown that that the inclusion of non-residential exposure and accounting for duration of exposure in different contexts can significantly improve model fit and the strength of place-CVD associations. Traditional models assume everyone spends the same amount of time at home, which is an unrealistic assumption; and the home is assumed to be the only place of influence, which is also unrealistic.

To the assertion that there has been a “residential trap” in studies of neighbourhood, I wholeheartedly agree, and would add that duration of exposure has also been systematically neglected. Even in studies that solely look at residential contextual exposures, the modeling can be significantly improved by accounting for how much time people spend in the residential context, as demonstrated by model 3a. One reason why some previous studies have failed to identify robust associations between residential contextual influences and health may be because the effects are masked by large variations in the (unobserved) duration of exposure; assuming that each person’s duration of exposure is equal when in fact they are not might lead to an underestimation of effects, because those who spend substantial amounts of time out for work/school/leisure will not be as influenced by residential factors compared to people who sit around at home all
day. The implicit assumption of equal duration in most studies leads to the underestimation of contextual effects and may even increase the chance of type 2 errors.

My exploration of workplace socio-environmental variables was based on theories of how the residential context would potentially impact CVD risk, thus the characteristics investigated might not be the most appropriate or influential workplace factors for CVD. This is demonstrated by the remaining work context ICC found in model 3b and 3c (table 2.4); as contextual variables get entered into the model the ICC is reduced, and the remaining ICC represents the unaccounted for variance in CVD that still remains in the neighbourhood level. Since models 3b and 3c have already been adjusted for individual-level risks, the remaining significant ICCs in model 3b and 3c can be interpreted as unexplained level-2 CVD variance, which is potentially explainable if appropriate workplace socio-environmental factors are identified. The significant association between neighbourhood SES in the work context and CVD is also likely a proxy for unobserved socio-contextual factors, which are beyond the scope of this paper. Further theorization and exploration of what contextual workplace factors are important for CVDs will be an important next step for CVD research and intervention.

There are a number of limitations for this study. First, disease assessments were limited to self-reports; therefore, it is not possible to determine the clinical nature or severity of conditions, and undiagnosed cases may be differentially misreported. Second, neighbourhood boundaries were limited to census tracts, which may not provide the most relevant construct of neighbourhood, as is the case for the area effect of noise and living close to a high-traffic road. Third, it is also important to acknowledge that the relationships we find here may not necessarily be causal, because of health-selected migration, that is, residents may choose the neighbourhoods in which they live or work on the basis of health-related characteristics (R. J. Sampson, Morenoff, & Gannon-Rowley, 2002). Fourth, while the duration of work time was reported, the time spent at home was an estimation based on the hours worked, and may be inaccurate for those that spend many hours doing leisure or other activities, more information on where and how time is spent throughout the day would significantly improve the measurement for the duration of exposure and increase the accuracy of estimates of contextual effects.
Despite these limitations, this exploratory study points to potentially new pathways through which different socio-environmental contexts may influence CVD etiology, and demonstrates that the workplace context and duration of exposure are important elements to consider for inclusion in the next generation of studies of place effects on health. Inclusion of non-residential locations/characteristics and duration of exposure variables (in the form of how time is spent throughout the day) into national community health surveys or other national/regional surveys will likely have the most influence on whether they are regularly considered in future research and public health interventions. The ‘residential trap’ has been structurally reinforced by the abundance of information about the home place, such as from the census and household surveys, but with the increased spatial mobility of urban lives in developed nations, information on one’s home place alone may no longer be as relevant as it once was. As social scientists, population health researchers and practitioners, it is important that we ensure population information accurately reflect the complexities of everyday lives so that the best available evidence is used to inform public policies and public health interventions. I urge that you will consider making a case for the inclusion of non-residential location/characteristics as well as duration of exposure questions in routine surveys in your area given their significance demonstrated by this study.
Reference List


Study 3: Socio-environmental determinants of Cardiovascular Diseases: Do
neighbourhood boundaries matter?

Introductions

Recent developments in health geography and social epidemiology play a major role in the recognition of ‘place’ as a significant factor underpinning health inequalities, given the uneven geographic development of risks associated to ‘place’ as well as access to health promoting resources (Macintyre, Ellaway, & Cummins, 2002; S. J. Smith & Easterlow, 2005). At the neighbourhood level, a body of multilevel studies have emerged to establish the effects of neighbourhoods on a variety of cardiovascular health outcomes including increased incidence of coronary heart diseases (A. V. Diez-Roux et al., 2001; Sundquist, Malmstrom, & Johansson, 2004; Winkleby, Sundquist, & Cubbin, 2007), increased incidence of myocardial infarction (Lovasi et al., 2008), increase in all cause and cardiovascular disease mortality (G. D. Smith, Hart, Watt, Hole, & Hawthorne, 1998), increase in coronary heart disease case fatality (Winkleby et al., 2007), and increase in risk factors such as smoking, physical inactivity, obesity, diabetes, and hypertension (Cubbin et al., 2006; Ellaway, Anderson, & Macintyre, 1997; Matheson, White, Mowinckel, Dunn, & Glazier, 2010), even after adjusting for individual level risk factors.

These studies have an important feature in common: they employ geographical areas as units of analysis; therefore, their results may be greatly influenced by the design of these areas – an effect known as the modifiable areal unit problem (MAUP). While many epidemiology and health geography studies use only a single set of neighbourhood definitions, it is often acknowledged that a single scale may be inadequate to investigate the spatial extent for different types of human activities (Macintyre et al., 2002), which may operate at different spatial scales. There are two main reasons why researchers should pay greater attention to boundary definitions: first, studies using a single set of boundaries may underestimate levels of health inequality, especially if boundaries define areas where within-group variation in the health outcome (as well as determinants of health) is relatively larger than the between-group variation. In other words, if the neighbourhood boundaries define areas that are internally heterogeneous in terms of
health, the analysis will not detect the full extent of the local area differences, since the inequalities may be occurring at a different scale. Second, if the areas used are vastly different from the spatial extent experienced in resident’s everyday life, then this will lead to measurement error in the level of exposure. This will tend to bias results towards a null finding: increasing the likelihood of type II error (Stafford, Duke-Williams, & Shelton, 2008).

To date, while a number of recent studies address the modifiable areal unit problem (MAUP) in the study of neighbourhood effects on child accidents (Haynes, Jones, Reading, Daras, & Emond, 2008), obesity, alcohol intake, smoking, walking and self-rated health (Stafford et al., 2008), walking for leisure or errands (Oliver, Schuurman, & Hall, 2007), and mental health (Proper, Jones, & Bolster, 2005), this is a relatively new area of research, and the use of multiple sets of neighbourhood boundaries within a single study is the exception rather than the rule. Moreover, while a study has investigated the MAUP in relation to the effect of neighbourhood socio-economic status on myocardial infarction (Lovasi et al., 2008), no studies have investigated MAUP in relation to the effects of specific neighbourhood socio-environmental characteristics on CVDs. These specific characteristics include measures of the local food environment, opportunities for physical activities, and other possible contextual determinants of CVD such as neighbourhood sources of psychosocial stress, noise, and air pollution, all of which are included in this study. Given the fact that CVDs are the leading cause of death and disability in the US, UK, Canada and in most countries around the world (American Heart Association, 2003; Bonow, Smaha, Smith, Mensah, & Lenfant, 2002), investigating multiple scales through which place characteristics may affect CVDs can help to reduce the chance that the ‘wrong’ scale may mask significant effects. The purpose of this study is to investigate whether neighbourhood boundaries can significantly impact the estimates for a range of socio-environmental neighbourhood effects on CVDs. This study will help to identify the appropriate scale and zone for the spatial distribution of health-relevant neighbourhood characteristics for CVD outcomes, which will help to define ‘neighbourhood’ in the context of CVD research and help to inform the spatial extent for place-based interventions. I also present and test a novel method of defining neighbourhood boundaries that supplements administrative
boundaries using buffers, which is termed *amoeba buffers*, to help reduce the measurement error in the level of exposure. More specifically, it deals with the problem that residents who are located near the edges of a given neighbourhood boundary have their area of exposure truncated by the boundary. I hypothesize that the association between neighbourhood socio-environmental factors (which is defined as the local food environment, local opportunities for physical activity, neighbourhood sources of psychosocial stress, and environmental factors including traffic and noise) and CVDs would be stronger for the amoeba boundaries compared to typical administrative boundaries.

This paper has three sections: first, the literature on possible pathways linking neighbourhood socio-environments and CVDs is reviewed. The problems concerning the use of neighbourhood boundaries are discussed, and a novel method of amoeba boundaries to help improve the measurement of local socio-environmental exposures is proposed. Second, an original study of neighbourhoods in Toronto, Canada is presented to investigate some of the pathways linking neighbourhoods and health using a range of neighbourhood boundaries. Third, I discuss the results with regards to implications for future research, policy, and interventions.

Section 1:

i. Literature review of Neighbourhood effects on CVDs

Socio-environmental factors are likely to be important in shaping the distribution of CVDs (Ana V. Diez-Roux, 2003). In particular, the effects of neighbourhood deprivation on CVDs have been extensively studied and this work has generally confirmed the negative impact of neighborhood deprivation on cardiovascular health. For example, researchers found that neighbourhood deprivation is associated with increase in all cause and cardiovascular disease mortality (G. D. Smith et al., 1998), increase incidence of myocardial infarction (Lovasi et al., 2008), increase in coronary heart diseases incidences (A. V. Diez-Roux et al., 2001; Sundquist et al., 2004; Winkleby et al., 2007), increase in coronary heart disease case fatality (Winkleby et al., 2007), and increase in a number of risk factors such as smoking, physical inactivity, obesity,
diabetes, and hypertension (Cubbin et al., 2006; Ellaway et al., 1997; Matheson et al., 2010) after adjusting for individual level confounders.

While the research on the impact of neighbourhood deprivation have provided new insights on the importance of ‘place’ for health, it is not known whether a) neighbourhood deprivation alone increases CVD risk (i.e. given that all other factors are held constant), or b) neighbourhood deprivation is acting as a proxy for a number of other unobserved neighbourhood level factors to increase CVD risk. In order to understand the true causal pathway, it is necessary to include other neighbourhood level socio-environmental features into future models (Ana V. Diez-Roux, 2003). In the following, I will explore the specific socio-environmental features that have been theorized as plausible risks for CVDs before modeling them (alongside neighbourhood deprivation) in section 2. The work on the contextual influence on CVD risk is voluminous (Chaix, 2009), but it can be summarized by the mechanism through which they impact CVD risk, which includes diet, physical activity, psychosocial stress, and lastly, air pollution and noise.

1) Diet

A number of studies now suggest that the quality of diet is not entirely determined by individual factors such as SES, gender and ethnicity, but may be by the locational and environmental factors (Ana V. Diez-Roux, 2003; Laraia, Siega-Riz, Kaufman, & Jones, 2004; Mooney, 1990; Moore, Diez Roux, Nettleton, & Jacobs Jr., 2008; Morland, Wing, & Diez Roux, 2002). Researchers have approached studies on the impact the ‘local food environment’ from two different angles. The first angle looks at the spatial distribution of healthy and unhealthy food to determine whether there is an unequal spatial distribution that can impact access. The second angle tries to investigate whether (and how much) individuals’ diet choices are shaped by their local food environment. Both areas of research are discussed respectively in the following.

Studies on the uneven spatial distribution of food suggest that the availability of healthy food is scare in deprived neighbourhood, which may in turn drive the spatially uneven rates of CVDs (Mooney, 1990; Wechsler, Basch, & Shea, 1995). For instance, in Morland et al’s (2002) US study, the researchers found that there are 400% more
supermarkets in white neighborhoods versus black neighborhoods. With regards to
‘unhealthy’ foods, they found that there are 300% more locations to purchase and
consume alcohol in the poorest neighbourhoods compared to the wealthiest
neighbourhoods. Numerous other studies also found that the availability of healthy food
(usually measured by number of supermarkets and food stores selling fresh produce) is
negatively correlated with neighbourhood deprivation (Mooney, 1990; Wechsler et al.,
1995). Adding to the burden of the lack of healthy food in deprived neighbourhoods,
Block et al (2004) have also found that fast food and unhealthy take-out restaurants are
more prevalent in low-income and black neighbourhoods. However, the geographical
distribution of healthy/unhealthy food is not the same across every city. In a study of
Glasgow, Scotland, researchers found that the distribution of food stores were similar
across deprived as well as non-deprived neighbourhoods (Cummins, Curtis, Diez-Roux,
& Macintyre, 2007). Researchers of the Glasgow study notes that their findings reveals
that the research based in the US may not be applicable in the UK, where the latter enjoy
a more equitable distribution of food stores across both wealthy and poor
neighbourhoods.

It follows that if there is an unequal spatial distribution of healthy and unhealthy
food (at least in the North American context), the next question to ask is how much the
local food environment actually shape individuals’ diet and in turn their health outcomes.
There is evidence to show that high diet quality is associated with the availability of
healthy food. In a study based on the US Multi-ethnic Study of Atherosclerosis,
researchers found that participants with no supermarkets near their homes were 25-46%
less likely to have a healthy diet than those with the most stores, after adjusting for
individual level confounders (Moore et al., 2008). Further support for this association is
given in a study by Morland, Wing and Diez Roux (2002), where Black Americans’ fruit
and vegetable intake increased by 32% for each additional supermarket in the census
tract, and for White Americans’ fruit and vegetable intake increased by 11% with the
presence of 1 or more supermarkets (n=10,623). In addition, a study of prenatal dietary
quality in North Carolina, US, also confirm that the proximity of food retail outlets
independently influences the diet quality of pregnant women (Laraia et al., 2004).
While a substantial body of evidence show that the local food environment may impact individuals’ diet quality, the results are not always consistent. A study in Glasgow, Scotland found no significant associations between proximity to food stores and diet or obesity in unadjusted or adjusted models, or when stratifying by gender, car ownership or employment (Macdonald, Ellaway, Ball, & Macintyre, 2011). However, given the earlier finding that there is no significant difference in the number of food stores across Glasgow neighbourhoods (Cummins et al., 2007), it seems that the lack of variation in the number of food stores across neighbourhoods may have led to this particular finding. Also, in a US longitudinal study of adults (Boone-Heinonen et al., 2011), researchers did not find an association between food store availability and diet quality measured by vegetable and fruit intake; however, fast food consumption was found to be associated with the availability of fast food restaurants in the local area.

2) Physical Activity

The research and practice of urban planners and architects, through land use planning and urban design, to increase the amount of walking and biking (i.e. active transportation) is a good place to explore how the environment may influence the level of physical activity and in turn CVD outcomes (Handy, Boarnet, Ewing, & Killingsworth, 2002). Efforts to promote active transportation can be an effective public health intervention for CVD because there is a general scientific consensus that even short daily episodes of moderate physical activity may be sufficient to produce cardiovascular and other health benefits (Pate et al., 1995). Both land development patterns, in terms of density and land use, as well as urban design can influence physical activity.

With regard to land development patterns, a study based in Washington State (Frank & Pivo, 1994) has identified a number of development patterns that seem to promote walking for shopping and work-related trips. They include high population and employment density, and a high level of mixed land-use. Other studies also have similar findings that land-use mixing and high densities promote active transportation (Cervero & Radisch, 1996; Friedman, Gordon, & Peers, 1994; Kitamura, Mokhtarian, & Laidet, 1997).
Urban design also seems to influence the level of active transportation. New Urbanism is an urban design movement which combines traditional (i.e. pre-automobile era) neighbourhood design and transit-oriented development to promote a number of goals such as to increase walkability, increase social and land use mix, and reduce urban sprawl (Grant, 2006). There are studies on the use of New Urbanism design elements to promote physical activity that may be instructive here. In a study of the San Francisco Bay Area, researchers found that pedestrian-oriented design, which features widened sidewalks, street lighting, and planted strips, have a significant impact on increasing the level of active transportation even after adjusting for individual and family characteristics (Cervero & Kockelman, 1997). In a study of the impact of traditional neighbourhood design that features rectilinear street grid, Handy (1996) finds that traditional design reduces motorized forms of travel compared to development in the style of the post-WWII suburb, which are characterized by the use of cul-de-sacs and curvilinear street patterns.

In addition to the research on the built-environmental determinants of active transportation, public health researchers have also studied the environmental determinants of physical activities in general (Humpel, Owen, & Leslie, 2002). In a systematic literature review, Humpel and Owen (2002) have reviewed quantitative studies that assessed the association between environmental characteristics (both perceived and objectively determined) and the level of physical activity. They found three contextual factors that are consistently associated with physical activity levels even after adjusting for individual level characteristics; they include 1) the accessibility of facilities, 2) the opportunities for activity, and 3) the aesthetic qualities of place. Accessibility is generally measured by the distances (in Euclidean or Network distances) to facilities that promote physical activities, which can include swimming pool, running track, or tennis courts that can be used by the public. ‘Opportunities for activity’ deals with the residents’ perceptions regarding opportunities for physical activity in their area. It includes their level of awareness of the facilities and their satisfaction with them. It can also refer to the quality of sidewalks and equipment in facilities, but the underlying factor for this category is that measures of ‘opportunities’ are subjectively perceived. Finally, places that are perceived as pleasant, attractive, and friendly also have an impact on the
level of physical activity. Somewhat surprising is that the fact that safety (e.g. crime, perceived neighborhood safety, street lights, and traffic) and poor weather only had a mild influence on physical activity compared to the former three factors.

3) Psychosocial Stress

To establish the plausible pathway linking neighbourhood context and CVD risk as mediated by psychosocial stress, we can look at two bodies of research. First, it is necessary to establish that psychosocial stress is a determinant of cardiovascular health. Second, we need to look at socio-environmental factors that can promote psychosocial stress. Each linkage is discussed in turn in the following.

Recent research now show that psychosocial wellbeing, which are partly indicated by depression, hostility, and anxiety, chronic/acute psychosocial stressors, and lack of social support are independently associated with the development and progression of CVDs (Everson-Rose & Lewis, 2005). The key pathophysiological mechanism is believed to be the hypothalamic-pituitary-adrenal axis, which are activated in response to stress, depression and anxiety, and in turn negatively impact cardiovascular function and promote atherogenesis (Musselman, Evans, & Nemeroff, 1998). In other words, depression and anxiety can lead to the chronic dysregulation of the hypothalamic-pituitary-adrenal axis, which in turn leads to hormonal and neuroendocrine alterations, including hypercortisolemia or excess glucocorticoid secretion (Plotsky, Owens, & Nemeroff, 1998). This is harmful because even a small increase in glucocorticoid levels sustained over time can contribute to hypertension, visceral obesity, coagulation changes, insulin resistance, and increased lipid levels, all of which increase the risk of CVDs (Chrousos & Gold, 1998). A literature review of epidemiological and clinical studies (Everson-Rose & Lewis, 2005) concludes that there is now a clear scientific consensus that psychosocial wellbeing, which can be measured in terms of (the lack of) depression, anxiety, and levels of social support, is one of the key determinants of cardiovascular health.

It is possible that neighbourhood level factors may influence psychosocial wellbeing and in turn have an impact on CVD risk, but few studies look directly at
neighbourhood influences on psychosocial wellbeing as a mediator for clinical end-point CVD. One such study was carried out by Sundquist et al. (2006): they found that two neighbourhood psychosocial stressors, namely neighbourhood rate of violent crimes and unemployment rate, were positively correlated with coronary heart disease even after adjusting for age, gender, income and other individual level confounders. While studies that look at the pathway between neighbourhood context and CVD risk as mediated by psychosocial wellbeing are rare, we can look at a number of studies of neighbourhood influences on depression and anxiety to shed light on the plausible pathway. One such study is Matheson et al.’s (2006) Canadian study of the influence of neighbourhoods on the risk of depression. They found that neighbourhood deprivation and residential instability (measured by the percent living alone and percent home ownership) and are key psychosocial stressors that were positively correlated with depression even after controlling for individual level confounders as well as neighborhood-level ethnic diversity. With regards to the built environment, Galea et al.’s (2005) study of New York City residents found that living in a neighbourhood with poor quality built environment (i.e. housing disrepair) was associated with higher risk of past six month and lifetime depression in multilevel models adjusting for individual level confounders and neighbourhood deprivation. Given the emerging understand that neighbourhood context can impact psychosocial wellbeing, and the scientific consensus on the harmful effect of psychosocial stress on cardiovascular health, it is probable that psychosocial stress is a key mechanism for neighbourhoods to influence CVD risk.

4) Air pollution and Noise

Most recently, exposure to air pollution has been confirmed to heighten the risk of CVDs. In a recent statement released by the American Health Association (Brook et al., 2004), researchers drew on epidemiological and clinical evidence to demonstrate that exposure to sulfur dioxide, lead, carbon monoxide, oxides of nitrogen, ozone, and coarse and fine particulate matter (commonly known as PM$^{10}$ and PM$^{2.5}$ respectively) increased the risk of CVDs including myocardial infarction (Peters, Dockery, Muller, & Mittleman, 2001; Zanobetti & Schwartz, 2005), heart failure exacerbation, progression of atherosclerosis (Suwa et al., 2002), hospitalization and CVD-related mortality (Dominici et al., 2006; Mann, Tager, & Lurmann, 2002; Pope, Burnett, & Thun, 2002; Pope,
Burnett, & Thurston, 2004; Samet, Dominici, & Curriero, 2000), and decreased rate of survival after stroke (Maheswaran et al., 2010). Given the fact that vehicular traffic accounts for a large variation of these pollutants, exposure to traffic is also studied for its impact on CVD risk (Adar & Kaufman, 2007). A number of studies now suggest that living in close proximity to traffic can significantly heighten the risk of myocardial infarction (Hoffmann et al., 2006; Tonne et al., 2007) and increase the number of cardiovascular and stroke mortalities (Finkelstein, Jerrett, & Sears, 2004).

Traffic counts can also be used as a proxy for air quality measures, since the latter are more expensive and difficult to obtain due to field sampling requirements, or may only be freely available through government owned permanent monitoring stations but the data is often at a spatial resolution that is too sparse for modeling at the intra-urban level. However, Vedal (2009) cautions that a portion of the effect of traffic vis-à-vis air quality may be confounded by the impact of traffic noise on CVD risk. As such, if a separate measure of noise is omitted from models, we may be overestimating the burden of traffic on CVD risk qua air pollution. Stansfeld et al.’s systematic review on the impact of noise as a determinant of health in urban areas have identified its harmful impact on blood pressure and sleep quality and duration, which suggests that further investigation of how noise can contribution to CVD risk should be specifically studied. By including both noise and traffic into the CVD risk models (section 2), this study will help to correct possible overestimation of the impact of traffic (vis-à-vis air quality) and investigate the association between noise and CVD risk.

ii. Placing Neighbourhood boundaries: methods and challenges

Given the proliferation of studies exploring neighbourhood effects on CVDs, there remains a critical concern over methods for the identification and delineation of study neighborhoods. Table 3.1 summarizes selected studies linking neighbourhood factors to CVDs and related risk factors, as well as the boundaries used. All but one study use only a single set of neighbourhood boundaries, namely Lovasi et al. (2008). Their study found an association between socioeconomic context--myocardial infarction (MI), but no neighbourhood definition (i.e. 1 km buffer, block group, census tract, and ZIP code) led to consistently stronger associations with MI. However, the authors only
looked at the effects of neighbourhood deprivation, and not a range of specific socio-environmental features. Furthermore, the lack of similar studies outside of the US restricts the generalizability of this study to other countries. On the other hand, outside of the field of cardiovascular health, the risk of accidents to preschool children have been found to vary significantly between different sets of neighbourhood boundaries in a study of Southwest England, with stronger effects observed in areas with populations less than 4000 (Haynes et al., 2008).

It should be noted that neighbourhood boundaries are often chosen due to convenience, because relevant data comes pre-aggregated and access to micro-data for re-aggregation to other scales is unfeasible. Mostly commonly, the neighbourhood deprivation—CVD association was studied at the level of census block group and census tract, which makes it easy to retrieve census based pre-aggregated socioeconomic data. On the other hand, air pollution was often studied at the city level (Pope et al., 2002; Samet et al., 2000; Zanobetti & Schwartz, 2005), not because it is more appropriate than other scales, but because of the lack of data at finer spatial resolutions.

Proximity to health resources or risks (measured in Euclidean or network distances) was used in some studies rather than defined neighbourhood boundaries (Boone-Heinonen et al., 2011; Laraia et al., 2004; Macdonald et al., 2011), these are typically studies looking at the proximity of food resources. While the use of distance as a proxy for exposure condition eliminates the need for predefined boundaries, there is a concern that by ignoring the clustering of residents in neighbourhoods (i.e. ignoring the intra-neighbourhood correlation) we risk violating the assumption of independence which can lead to incorrect estimation of the standard errors of model parameters in non-hierarchical models and possibly draw incorrect inference on the effect sizes of neighbourhood characteristics (Pickett & Pearl, 2001). For this reason, I choose to not take the ‘distance’ approach in the study to follow, and instead use multiple neighbourhood scales to group participants, and take a hierarchical regression approach to account for the clustering of residents in neighbourhoods.
In studies of neighbourhood socio-environmental determinants of CVDs, different scales are used without consensus as to which is the most relevant. This is a concern because size matters for boundary selection. For neighbourhood boundaries that are too large, this can mask relevant spatial variation and also result in large scale residual confounding. However, boundaries that are too small effectively excludes relevant socio-environmental features that can affect CVDs and related health behaviours. Thus, measuring multiple scales can help researchers established the plausibility of mediators, reduce the chance of using a wrong scale that masks significant effects, interpret studies testing similar hypothesis at different scales, and build knowledge on relevant scales for neighbourhood influences on CVD to inform place-based interventions.
There has been a valid concern that the use of administrative boundaries or GIS generated buffers might be vastly different from the spatial extent experienced in residents’ everyday life. In a UK study of the differences between computer generated neighbourhoods (Haynes, Daras, Reading, & Jones, 2007), the authors found that residents’ perceived neighbourhoods could be explained by socio-economic conditions at the local scale of enumeration districts with populations of about 500. Other studies have also suggested that perceived and measured features are highly correlated within small areas, especially within the 1 km buffer areas (M. Duncan & Mummery, 2005; Moudon et al., 2006). This suggests that neighbourhoods with meaning for residents are 1) much smaller than those typically investigated in neighbourhood effects on health studies and 2) may be reasonably represented by small administrative zones. Thus, for the analysis to follow, I hypothesize that smaller administrative zones might represent the area of effects for socio-environmental factors relevant to CVDs better than larger ones.

Finally, the location of each participant within the neighbourhood can be an important, and often neglected, source of variation in levels of exposure for individuals. Figure 1 is a hypothetical neighbourhood boundary with points representing locations of 7 participants. While participant A, close to the center of the neighbourhood, may have an area of exposure approximated by the boundary chosen, participant B and C located at the edges are equally exposed to the effects of adjacent neighbourhoods, which means that their area of exposure might be inappropriately truncated by the chosen boundary.
None of the studies reviewed here have addressed this issue, and my study will attempt to resolve this problem by adding a buffer to the entire group of participants to form an amoeba buffer that respects the original neighbourhood grouping, but also will use the characteristics of adjacent neighbourhoods to inform the area of exposure (see figure 3, discussed below). Based on the review and discussion above, the analysis to follow will address the following research questions:

1. Does the CVD variability explained at the neighbourhood level vary significantly depending on the neighbourhood scale/boundaries used?

2. Is the association between neighbourhood socio-environmental factors and CVDs stronger for smaller administrative zones compared to larger ones? Is the association between neighbourhood socio-environmental factors and CVDs stronger for the amoeba boundaries compared to administrative boundaries? Do these associations remain after controlling for individual sociodemographic characteristics, including gender, income, education level, age, and visible minority status?

3. What is the extent to which neighbourhood boundaries may impact studies that involve the use of multilevel models to explain the risk of CVDs? What theoretical and practical considerations should be taken into account when constructing neighbourhood-level variables with regards to the associated modifiable areal unit problem?

Question number 2 sets up two formal hypotheses that will be tested in the statistical analysis section, namely 1) associations between contextual factors and CVD risk are
stronger/more significant in smaller (compared to bigger) administrative zones, and 2) the use of amoeba buffers (compared to their respective administrative boundary) result in stronger/more significant associations between contextual factors are CVD risk, after adjusting for individual level covariates. Rationale, design and construction of the amoeba buffers will be more thoroughly discussed in the section titled *Neighbourhood boundaries investigated*.

**Section 2:**

*Methods*

The Neighbourhood Effects on Health and Well-being (NEHW) project implements a cross-sectional survey designed specifically to understand the impact of neighbourhood level determinants, acting independently or interactively with individual level factors, on population health including CVD outcomes. The study uses a three stage sampling technique: first, 50 out of the total 140 neighbourhood planning areas (NPA) delineated by the City of Toronto were randomly selected; second, from the 50 NPAs sampled each containing 2 to 10 census tracts (CT), median = 3), 2 CTs were randomly selected, resulting in 100 randomly selected CTs.

Due to budget limitations, the final number of sampled CTs was reduced from 100 to 87, where some NPAs only had 1 CT selected rather than 2. Thorough checks were performed to ensure that the remaining participants were representative of the socioeconomic profile of the City of Toronto. Third, households within each sampled CT were randomly selected based on their residential address. The recruitment target was 30 households per CT, but the actual number of households recruited ranged from 9-31 households, and the goal of 30 households in 51 out of 87 (58.6%) CT sampled was reached. An individual is eligible to participate if they met all of the following criteria: 1) a resident of the selected household (one per household), 2) were between the ages of 25 and 64, 3) were able to communicate in English, and 4) had lived in the neighbourhood for at least 6 months. The response rate was 72%.

Data were collected between Mar 2009 and June 2011. In total, 2412 participants, representing 50 NPAs and 87 CTs, were included in the study’s sample. Data were obtained from in-person interviews using a computer-assisted personal interviewing program administered by interviewers trained by the Survey Research Unit (SRU) at the
Centre for Research on Inner City Health (CRICH) at St Michael’s Hospital. Interviews were conducted at participants’ residences or in private interview rooms located at CRICH. All participants provided written informed consent at the time of their interview. The Research Ethics Board at St Michael’s hospital in Toronto, Canada provided ethics approval for this study.

**Neighbourhood boundaries investigated**

In this analysis, six different neighbourhood boundaries were used: 1) Toronto neighbourhood planning areas (NPAs), 2) census tracts (CTs), 3) dissemination areas (DAs), 4) 1km amoeba buffers respecting NPA groupings, 5) 1km amoeba buffers respecting CT groupings, and 6) 1km amoeba buffers respecting DA groupings. NPAs were created by the City of Toronto in 2002 to delineate Toronto’s neighbourhoods and provide information at a meaningful geographic level for government and community agencies (City of Toronto, 2002). It is an important scale to consider for the purposes of place based interventions because its boundaries coincide with existing Public Health neighbourhood planning areas, and is used as service boundaries by Toronto community agencies. Their population ranges from 10,000 to approximately 70,000. Each NPA is comprised of at least two census tracts, and respect existing boundaries such as rivers, major roads streets/highways as well as census boundaries. CTs are smaller, relatively stable geographic units with populations between 2,500 and 8,000, and are used very typically as a proxy for residential neighbourhoods in many studies because they are constructed to maximize internal homogeneity with regards to social and economic conditions, and their boundaries are bordered with topographic features and major roads/highways. DAs are the smallest geographic area for which all census data are disseminated and have a population from 400 to 700. DA boundaries follow roads and other features (such as railways, water features, power transmission lines), where these features also form part of the boundaries of census tracts. Figure 2 illustrates the relationship between NPA, CT and DA.
In order to deal with the problem of mismeasured area of exposure for participants living near the edges of these administrative boundaries, 1 km amoeba buffers were created that respect the original neighbourhood grouping, but can extend the resident’s area of exposure into adjacent administrative boundaries (see figure 3). To create the amoeba buffers, participants’ addresses were geocoded using GeoPinpoint v. 3.3 (DMTI Spatial Inc., Markham, On, Canada) to assign a latitude and longitude coordinate for each participant and the data were imported into ArcGIS Editor 9.3 (ESRI, Redlands, CA, USA). A 1 km circular buffer was created for each participant, but the boundaries that overlapped with other participants within the same administrative boundary were dissolved so that all the participants from the originating administrative boundary are kept as a group. This effectively changed the shape of the original administrative boundary into a shape resembling an amoeba organism.

Figure 3: Amoeba buffer overlaid on administrative boundary
Ameba buffers were created for all 3 scales of administrative boundaries. A 1 km buffer was chosen because previous studies have demonstrated neighbourhood environment within 1 km of resident’s home to have an impact on health behaviour such as walking (Oliver et al., 2007), as well as studies that have demonstrated that the 1 km buffer areas are closely associated with residents’ perceived neighbourhood boundaries (M. Duncan & Mummery, 2005; Moudon et al., 2006).

To illustrate how amoeba buffers augment neighbourhood contextual measures, the following is an example of how this study counted the number of food stores in the administrative versus amoeba neighbourhood as shown in figure 4:

*Figure 4: Food stores distributed across administrative boundary and the corresponding amoeba buffer*

If we only consider the administrative boundary, only food stores A and B would be included (thus, a value of 2 food stores is assigned to every resident within the neighbourhood). If we consider the amoeba boundary in which the same group of residents are within, only food stores B, C and D would be included (thus, a value of 3 food stores is assigned to every resident within the neighbourhood). This example demonstrates that amoeba buffers do not always consider ‘more’ risks or resources compared to the corresponding administrative boundary, but in fact may remove risks/resources that may be within the same administrative boundary but are not within 1 km of any study participants. Finally, it should be noted that when the final food stores
variable is created, this study normalizes the number of food stores within each zone by the total area of that zone measured in km\(^2\), since this procedure controls for the differing size of zones as a factor for food stores availability by calculating the food store density.

When amoeba buffers were constructed for DAs, 24% of the total area covered by DA amoeba buffers were overlapped by other amoebas, this value was reduced to 16% for CT amoebas, and 9% for NPA buffers. This overlapping of amoeba buffers is not a concern for the analytical approach taken in this study because 1) each amoeba is a unique clustering of residents (i.e. no participant is a member of more than one amoeba group within a particular scale of analysis), 2) each amoeba represents a geographic neighbourhood with a unique combination of socio-environmental risks and resource, and 3) no pair of amoebas are perfectly overlapped. For example, it does not matter that a food store is within (shared by) 2 overlapping amoebas; in fact, this is a more realistic assumption than the typical single scale study (i.e. where once a resource is assigned to a particular neighbourhood, it is not assignable to other neighbourhoods) since more than one neighbourhood can (and do) share access to the same resources.

**Outcome**

The CVD outcome used is based on self-reported history of physician diagnosis of myocardial infarction (MI), angina, coronary heart disease (CHD), stroke, and congestive heart failure (CHF). Although this survey did not include clinical details, previous studies of the validity and reliability of self-reported conditions have suggested a high level of agreement with medical records for most of the conditions considered here (Bush, Miller, Golden, & Hale, 1989; Kehoe, Wu, Leske, & Chylack, 1994). Participants with a reported history for any of MI, angina, CHD, CHF, or stroke are coded as 1, others are coded as 0.

**Individual level predictors**

The structured interview provided self-reported information on age and total after-tax household income as continuous variables, and gender, visible minority status (which includes persons, other than Aboriginal persons, who are not white in race or colour), and education as categorical variables. While unmeasured individual-level
variables has been critiqued as a problem for the study of neighbourhood effects (Oakes, 2004), many of them, such as health behaviours (e.g. exercise and diet) and medical history (e.g. diabetes, hypertension, and obesity), are intermediate endpoints (or mechanisms) that lie between socio-environmental exposure and CVD outcomes (Ana V. Diez-Roux, 2004). Thus, control variables that represent sociodemographic risk factors believed not to be on the causal pathway between the socio-environmental context and CVD were chosen. Also, smoking, coded as two dummy variables 1) former smoker vs. non-smoker and 2) current smoker vs. non-smoker, was entered into the fully adjusted model as an additional control variable.

*Socio-environmental contextual predictors*

The selection of these variables is guided by the previous research reviewed above linking socio-environmental context and CVDs. For each of the mechanisms through which the neighbourhood context may impact CVD outcomes (diet, physical activity, psychosocial stress, and air pollution & noise), I will discuss the specific variables used.

With regards to diet, two variables are included. Data on access to 1) food stores and supermarkets and 2) fast food restaurants were obtained from the City of Toronto Public Health food inspection reports (Toronto Public Health Inspection, 2011). Since all food-selling premises in the city are required to be regularly reviewed for licensing purposes, the database provides comprehensive and up-to-date information. Addresses were geocoded using GeoPinpoint v. 3.3 (DMTI Spatial Inc., Markham, On, Canada), and the data were imported into ArcGIS Editor 9.3 (ESRI, Redlands, CA, USA) for spatial data manipulation. For each of the six neighbourhood boundaries, the number of 1) supermarkets and food stores and 2) fast food restaurants were summed and normalized by the total area in km².

Information on the availability of parks and recreational facilities were included in order to consider the impact of the socio-environmental context on physical activity. Land use information from the CanMap Route Logistics annually updated geo-database (DMTI Spatial Inc, 2011) were used to calculate the total area in each of the six
neighbourhood boundaries used for parks or recreational facilities. This number is normalized by the total area of each neighbourhood, which results in percent area used for parks and recreation for each of the six boundaries.

Two potential contextual sources of psychosocial stress are captured in the study. First, the *incident-based uniform crime reporting database* (Statistics Canada & Canadian Centre for Justice Statistics, 2010) provided data on the number of violent crimes per CT for the year 2006, which includes the most serious violations in incidents of sexual assault, criminal harassment, uttering threats, minor/major assault, and robbery. For CTs and NPAs (which are constituted by multiple CTs), it was a straightforward process of taking the totals counts for each area and dividing them by the total area for each neighbourhood, which gives the average number of violent crimes per km$^2$. For amoeba boundaries, approximate crime counts per buffer were calculated as weighted averages (weighted by the area) using the overlay feature in ArcGIS 9.3. For example, an amoeba buffer might have 25% of their total area in census tract A and 75% of their total area in census tract B, the estimated total violent crime would be $0.25 \times \text{(crime of CT A)} + 0.75 \times \text{(crime of CT B)}$. Since the data were pre-aggregated to the CT scale, the coarse spatial resolution does not allow the crime predictor to vary between the DA and the CT scale. Thus, there is a data limitation: multiple DAs nested within the same CTs get the same *violent crimes per km$^2$* as their parent CT. However, the amoeba boundaries grouped by DAs do vary from their parent CTs because their buffers often extend beyond boundaries of their parent CTs, thus the values were modified by adjacent areas through the weighted average method discussed above. Second, the percent of housing requiring major repairs is used in each neighbourhood as another source of psychosocial stress. Data on housing disrepair aggregated to both the CT and DA levels are obtained from the Canada census 2006 long form questionnaire (Statistics Canada, 2008), which is a 20% sample of the total census population. For NPAs, the percent requiring major repairs is obtained by summing all households requiring major repairs and dividing it by the total number of households in the NPA – all calculated using information from the CT level. The weighted average method is taken again for the three levels of amoeba buffers.
Traffic data obtained from the City of Toronto Transportation Services (Toronto Transportation Services, 2011) is used as a proxy for air pollution. Average weekday 24 hour volume of Toronto Street segments were recorded in the geo-database of a street map shapefile in ArcGIS 9.3, then 100-meters circular buffers were created around each participant’s home location. Using buffer analysis, those who resided within 100-meters of a major road, defined by the City of Toronto as having at least an average of 20,000 vehicles in a 24-hour period, were coded as exposed (1), those who do not have a major road within their 100-meter buffer were coded as not exposed (0). This method of categorizing traffic exposure and the use of a 100-meter buffer zone has been used previously to investigate the effect of exposure to traffic on mortality rate advancement periods (Finkelstein et al., 2004).

Furthermore, data on exposure to noise is added to understand its contribution to CVD risk, and to ensure that the traffic data properly proxies air pollution by separating out the unique effect of noise. Participants were asked in the interview whether they agreed with the statement “the noise level where I live often disturbs me” with the answer chosen on a 5-point Likert scale from strongly agree to strongly disagree. Both noise and traffic are entered into hierarchical models as individual level covariates, because they depend on factors such as the home’s location within the neighbourhood and the level of noise isolation of the residential structure.

Lastly, neighbourhood level median after-tax household income from the Canada census 2006 (neighbourhood wealth) is included since neighbourhood SES has been shown to have an independent effect on CVDs. Both DA and CT level data for median after-tax household income are available, and NPA medians were calculated using CT information. Median incomes for the 3 levels of amoeba buffers were obtained using the weighted average method described above.

iii. Statistical Analysis

The analysis began by evaluating the bivariate association between individual level and neighbourhood level predictors and self-reported history of CVDs using $\chi^2$ test for pairs of categorical variables and 1-way analysis of variance for unbalanced design
pairs of continuous and categorical variables. Descriptive summaries for the six
neighbourhood boundaries were calculated (i.e. number of areas, participants per area,
and area sizes in km\(^2\)), and Pearson correlations were obtained between socio-
environmental characteristics measured at the six neighbourhood scales.

Since the data has a 2-level structure, with individuals nested in neighbourhoods,
multilevel logistic regression is used to account for the lack of independence within
neighbourhood clusters (Raudenbush & Bryk, 2002). First, the null model (model #1) is
used to analyze the variation in CVD risk attributable to observed and unobserved
contextual factors across the six different neighbourhood scales through the intraclass
correlation (ICC). These ICCs can be interpreted as the maximum proportion of variation
in CVD risk explained at each neighbourhood scale. Model #2 enters all the
neighbourhood socio-environmental factors and neighbourhood deprivation as predictors
of the risk of CVDs (model #2) across the six neighbourhood boundaries. Finally, in the
last model (model #3), individual level controls are entered along with the neighbourhood
level predictors. This block-entry method allowed us to understand the effect of adjusting
for individual level socio-demographic risk on the neighbourhood-context only model, a
reduction in the estimates of the socio-environmental predictors’ effect sizes (from model
#2 to #3) would suggest that the contextual effects are better explained by the clustering
of individuals with similar socio-demographic risk profiles. Multilevel logistic regression
was performed using SAS 9.3 in the GLIMMIX (generalized linear mixed model)
procedure.

Section 3: Results

Descriptive Analysis

Participants in Project NEHW were 58% female and 34.54% visible minority,
with a mean age of 48 years and a mean after-tax family income of $107,712. Those who
have CVDs have a mean age of 54 and a mean after-tax family income of $77,033: one-
way ANOVA testing for difference in means between participants with CVD vs. those
without are significant for both variables at p<0.001. While a minimum of six months of
residence in their home was required for all participants, the mean number of years was
11.84 (range = 0.5-55 years, SD = 5.5), which means that most participants may have lived long enough in their neighbourhoods to be affected by possible socio-environmental factors. Overall, 108 (4.48%) reported a previous history of any CVDs (any one of MI, angina, CHD, CHF, or stroke). Bivariate associations between having any CVDs (dichotomous outcome) and other participants’ socio-demographic characteristics are presented in table 3.2.

Table 3.2: Self-Reported History of CVD Outcomes and Characteristics of Study Participants with $\chi^2$ test of differences for pairs of categorical variables

<table>
<thead>
<tr>
<th>Participants with any of MI, angina, CHD, CHF, or stroke</th>
<th>Entire Cohort No. (%)</th>
<th>No. (%)</th>
<th>P for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>2412 (100)</td>
<td>108 (4.48)</td>
<td></td>
</tr>
<tr>
<td>25-35</td>
<td>359 (14.88)</td>
<td>5 (1.34)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>36-45</td>
<td>598 (24.79)</td>
<td>15 (2.51)</td>
<td>0.0414</td>
</tr>
<tr>
<td>46-55</td>
<td>761 (31.55)</td>
<td>30 (3.94)</td>
<td>0.4741</td>
</tr>
<tr>
<td>56-65</td>
<td>694 (28.77)</td>
<td>58 (8.36)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>1390 (57.63)</td>
<td>52 (3.74)</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>1022 (42.37)</td>
<td>56 (5.48)</td>
<td></td>
</tr>
<tr>
<td>Visible Minority Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Yes</td>
<td>833 (34.54)</td>
<td>28 (4.01)</td>
<td></td>
</tr>
<tr>
<td>2. No</td>
<td>1579 (65.46)</td>
<td>80 (4.76)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1. Less than High School</td>
<td>88 (3.65)</td>
<td>13 (14.77)</td>
<td></td>
</tr>
<tr>
<td>2. High School Complete</td>
<td>402 (16.67)</td>
<td>20 (4.98)</td>
<td></td>
</tr>
<tr>
<td>3. Diploma Complete</td>
<td>616 (25.54)</td>
<td>34 (5.52)</td>
<td></td>
</tr>
<tr>
<td>4. Completed Undergrad and above</td>
<td>1306 (54.15)</td>
<td>41 (3.14)</td>
<td></td>
</tr>
</tbody>
</table>

Note: MI = Myocardial infarction; CHD = coronary heart disease; CHF = congestive heart failure. For variables with more than 2 categories, P is for the $\chi^2$ test of differences across row in the percentage reporting the outcome. Otherwise P is for a 2-tailed test of difference between 2 groups in the proportion reporting each outcome.

In order to compare the socio-environmental variables across the 6 neighbourhood scales to understand their robustness with regards to neighbourhood scaling, Pearson correlations are calculated and all variables are found to be significant at the p<0.001 level (table 3.3). Socio-environmental characteristics were correlated across different neighbourhood definitions, with correlations ranging from 0.38 to 0.98.
Table 3.3: Correlations between neighbourhood socio-environmental predictors measured at six different neighbourhood scales (N=2411)

<table>
<thead>
<tr>
<th></th>
<th>NPA</th>
<th>Amoeba buffer: NPA</th>
<th>CT</th>
<th>Amoeba buffer: CT</th>
<th>DA</th>
<th>Amoeba buffer: DA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median Household income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.79</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.86</td>
<td>0.65</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.79</td>
<td>0.97</td>
<td>0.70</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.60</td>
<td>0.48</td>
<td>0.66</td>
<td>0.51</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.77</td>
<td>0.89</td>
<td>0.71</td>
<td>0.91</td>
<td>0.55</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Number of food stores per km2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.71</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.75</td>
<td>0.60</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.71</td>
<td>0.97</td>
<td>0.61</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.51</td>
<td>0.39</td>
<td>0.54</td>
<td>0.39</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.68</td>
<td>0.93</td>
<td>0.60</td>
<td>0.96</td>
<td>0.38</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Number of fast food restaurants per km2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.66</td>
<td>0.63</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.75</td>
<td>0.98</td>
<td>0.67</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.39</td>
<td>0.39</td>
<td>0.46</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.71</td>
<td>0.90</td>
<td>0.66</td>
<td>0.92</td>
<td>0.46</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>% area used for parks/recreation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.76</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.79</td>
<td>0.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.76</td>
<td>0.94</td>
<td>0.79</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.52</td>
<td>0.44</td>
<td>0.61</td>
<td>0.47</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.75</td>
<td>0.86</td>
<td>0.79</td>
<td>0.91</td>
<td>0.56</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>Violent Crimes per km2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.76</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.75</td>
<td>0.69</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.76</td>
<td>0.98</td>
<td>0.69</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.26</td>
<td>0.69</td>
<td>1.00</td>
<td>0.69</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.73</td>
<td>0.96</td>
<td>0.69</td>
<td>0.98</td>
<td>0.69</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>% housing requiring major repairs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPA</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: NPA</td>
<td>0.78</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.70</td>
<td>0.62</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: CT</td>
<td>0.78</td>
<td>0.97</td>
<td>0.69</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA</td>
<td>0.39</td>
<td>0.34</td>
<td>0.53</td>
<td>0.38</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Amoeba buffer: DA</td>
<td>0.72</td>
<td>0.81</td>
<td>0.68</td>
<td>0.85</td>
<td>0.40</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients are shown above, all are significantly different than zero ($p<0.001$); N=2411 participants.
Violent crimes per km$^2$ had the highest correlations across neighbourhood scales (with a mean correlation of 0.75 across all possible pairs of neighbourhood boundaries); however, this result needs to be interpreted with caution since the crime data were pre-aggregated to the CT level. More specifically, no new information was gained by moving into a higher spatial resolution (from CT $\rightarrow$ DA), since values are simply duplicated from CT to the DA level (artificially inflating the correlation between CT vs. DA violent crime per km$^2$ to 1.00). Without data at a higher spatial resolution (e.g. point level crime data or crimes aggregated to the DA level), it is impossible to understand the impact of neighbourhood scaling on the areal exposure of violent crime when moving below the CT level. Disregarding violent crimes due to data limitations, median household income (mean correlation = 0.72) and percent used for parks and recreation (mean correlation = 0.71) have the second and third highest correlations respectively across neighbourhood scales. This suggests that both variables are robust (relative to other variables chosen in this analysis) to the effects of neighbourhood scaling. On the other hand, percent housing requiring major repairs (mean correlation = 0.64) and number of food stores per km$^2$ (mean correlation = 0.65) had the lowest correlations across neighbourhood scales, and may be more sensitive to neighborhood scaling effects. For each socio-environmental characteristic, it should also be highlighted that correlations between all possible pairs of amoeba buffers derived from the three scales (NPA, CT and DA) have consistently the highest correlations (ranging from 0.81-0.98), which may suggest that the 3 amoeba scales are more similar to each other than to the 3 administrative boundaries.

**Multilevel logistic regression**

In the following, the results are presented for 3 sets of multilevel logistic regressions predicting individual CVD status. Starting with a fully unconditional ordered logit model of the dichotomous outcome of CVD status (model set # 1), the intraclass correlation is calculated to decompose variance in the predictors across individual level versus CT level of analysis. The intraclass correlation (ICC) for multilevel logistic regression is calculated as:

$$ ICC = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2} $$
where $\sigma^2_w$ is the variance between neighbourhoods, and $\sigma^2_e$, defined by the constant $\pi^2 / 3$, is the variance between individuals (Guo & Zhao, 2000). The estimates of the intercept variance in the null models (model set #1) are 0.1181 for NPA level, 0.1562 for CT level, and 0.2346 for DA level, all significant at $p<0.001$ (see table 3.3). This suggests, on average, 3.47%, 4.53%, and 6.66% of the variance in CVD status can be attributed to the NPA, CT, and DA levels respectively. Note that the amoeba boundaries do not give different ICC results than their ‘parent’ administrative boundary (e.g. CT scale versus amoeba buffer respecting CT boundary): while the grouping of residents are preserved from the respective administrative boundaries, only the areas of exposure for the groups are modified (see figure 3). The results of these ICCs suggest a positive response to research question #1: the amount of CVD variability explained at the neighbourhood level is significantly related to the neighbourhood boundaries chosen; moreover, the percentage of variability explained is negatively correlated with the physical size of the neighbourhood boundary (before adjusting for neighbourhood or individual level predictors).

**Table 3.4: Comparison of Neighbourhood Definitions and Model Set #1 intra-class correlations**

<table>
<thead>
<tr>
<th>Neighbourhoods</th>
<th>Number of areas with participants</th>
<th>Mean (range) number of participants per area</th>
<th>Mean area in km² (range)</th>
<th>ICC for CVD status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neighbourhood Planning Area (NPA)</td>
<td>47</td>
<td>51.30 (23-61)</td>
<td>4.5646 (0.4206-16.2716)</td>
<td>3.47%***</td>
</tr>
<tr>
<td>Amoeba buffer respecting NPA groupings</td>
<td>47</td>
<td>51.30 (23-61)</td>
<td>8.6601 (3.2174-14.8031)</td>
<td>3.47%***</td>
</tr>
<tr>
<td>Census Tract (CT)</td>
<td>87</td>
<td>27.46 (8-31)</td>
<td>1.0382 (0.1937-4.5565)</td>
<td>4.53%***</td>
</tr>
<tr>
<td>Amoeba buffer respecting CT groupings</td>
<td>87</td>
<td>27.46 (8-31)</td>
<td>6.4424 (4.2153-10.8672)</td>
<td>4.53%***</td>
</tr>
<tr>
<td>Dissemination Area (DA)</td>
<td>493</td>
<td>4.89 (1-30)</td>
<td>0.1750 (0.0138-6.6199)</td>
<td>6.66%***</td>
</tr>
<tr>
<td>Amoeba buffer respecting DA groupings</td>
<td>493</td>
<td>4.89 (1-30)</td>
<td>3.8109 (3.1412-7.7798)</td>
<td>6.66%***</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001

While most of the variability in CVDs is at the individual level, the values of these ICCs are in the typical range, which rarely exceeds 20% (Snijders & Bosker, 1999). It is important to note that 1) the size of the ICC does not rule out relatively large effects of neighbourhood level measures (G. J. Duncan & Raudenbush, 1999), and 2) two out of the eight socio-environmental predictors, namely living near a major road and noise,
have a very small area of effect (up to 100 meters) such that the between group variance of both predictors are much smaller than within group variance across all 6 scales (1-way ANOVA has been performed to confirm this with at least $p<0.95$). Thus, while noise and living near a major road are factors shaped by the built environment, they operate at a smaller scale than the neighbourhood, and the effect they have on CVD outcomes will not be captured by the ICC.

In a separate set of models not shown here, only socio-demographic control variables (age, gender, after-tax family income, and visible minority status) were added to predict the risk of CVD status, and the random intercepts were retained. Despite the inclusion of individual level controls, the variance of the random intercept remains significant for CVD status at NPA, CT and DA scales: remaining ICCs are 1.97% ($p<0.01$), 3.28% ($p<0.001$), and 5.34% ($p<0.001$) respectively. These results indicate that there are unexplained variances at all neighbourhood scales even after accounting for the way in which individuals with similar characteristics may cluster.

Model set #2 reports the results of a set of regressions predicting CVD status using all neighbourhood level predictors (see table 3.4). From this model, we learn that some neighbourhood-level socio-environmental factors, such as access to food stores, fast food, violent crimes, living close to a major road, and noise, are indeed correlated to a higher risk of CVDs, but the strength and significance of these effects varied across neighbourhood scale used.
Table 3.5: Multilevel logistic regression adjusting for all socio-environmental factors, estimates for 6 neighbourhood scale (95% confidence intervals)

<table>
<thead>
<tr>
<th>Model 2</th>
<th>Adjusted for socio-environmental factors only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NPA</td>
</tr>
<tr>
<td>Total (n=2411)</td>
<td>0.992 (0.973-1.012)</td>
</tr>
</tbody>
</table>

|                     | Median after-tax household income (in neighbourhood) | 1.003 (0.952-1.057) | 0.991 (0.916-1.073) | 0.962 (0.937-0.987)** | 0.872 (0.837-0.907)** | 0.998 (0.990-1.007) | 0.874 (0.842-0.906)** |
|                     | Number of foodstores + supermarkets per km2        | 1.021 (0.967-1.078) | 0.975 (0.917-1.036) | 1.005 (0.985-1.025) | 1.096 (1.056-1.136)** | 1.011 (1.002-1.020)** | 1.813 (1.263-2.363)** |
|                     | Number of fastfood restaurants per km2             | 7.476 (4.456-38.363) | 3.270 (1.229-46.444) | 1.643 (+0.001-999.999) | 1.072 (+0.001-9.9999) | 1.347 (0.490-3.706) | 1.078 (+0.001-999.999) |
|                     | % area used for parks and recreation                | 1.001 (0.995-1.006) | 1.003 (0.992-1.015) | 1.005 (1.002-1.007)** | 1.004 (1.002-1.007)** | 1.005 (1.002-1.007)** | 1.006 (1.004-1.008)** *** |
|                     | Mean number of violent crimes per km2              | 0.002 (<0.001-736.398) | 0.001 (<0.001-6.234) | 3.586 (<0.001-999.999) | <0.001 (<0.001-58.340) | 0.352 (0.010-10.598) | <0.001 (<0.001-21.970) |
|                     | 1. Yes                                           | --                | --                | --                | --                | --                | --                |
|                     | 2. No (reference)                                 | --                | --                | --                | --                | --                | --                |
|                     | 1. Strongly Agree                                 | 0.642 (0.285-1.446) | 0.629 (0.279-1.419) | 0.572 (0.003-113.848) | 0.636 (0.282-1.435) | 0.596 (0.263-1.346) | 0.622 (0.276-1.403) |
|                     | 2. Agree                                          | 1.625 (0.867-3.043) | 1.613 (0.862-3.017) | 1.549 (0.026-91.071) | 1.621 (0.865-3.038) | 1.520 (0.811-2.851) | 1.582 (0.844-2.967) |
|                     | 3. Neither agree or disagree                      | 1.146 (0.694-1.890) | 1.122 (0.680-1.850) | 1.120 (0.043-29.243) | 1.105 (0.670-1.824) | 1.092 (0.661-1.803) | 1.106 (0.670-1.825) |
|                     | 4. Disagree                                       | --                | --                | --                | --                | --                | --                |
|                     | 5. Strongly Disagree (reference)                  | --                | --                | --                | --                | --                | --                |
| Interceptor of the covariance parameter              | 0.1059 | 0.1295 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Intraclass correlation (ICC)                         | 0.031  | 0.038  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

*p<0.05; **p<0.01; ***p<0.001

In should be noted that results in model set #3 were fairly consistent with those in model set #2, which means that the associations found in model set #2 were not confounded by individual socio-demographic risk or the clustering of individuals with similar characteristics (see table 3.6). Since, the results are very similar between model set #2 and #3, I will discuss the results with regards to the fully adjusted model set #3 only (see table 3.6).
Table 3.6: Multilevel logistic regression adjusting for socio-environmental and individual factors, estimates for 6 neighbourhood scales (95% confidence intervals)

<table>
<thead>
<tr>
<th>Model 3 (final)</th>
<th>Adjusted for neighbourhood factors and individual level controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NPA</td>
</tr>
<tr>
<td>Total (n=2412)</td>
<td></td>
</tr>
<tr>
<td>Socio-environmental variables</td>
<td></td>
</tr>
<tr>
<td>Median after-tax household income in thousands</td>
<td>0.996</td>
</tr>
<tr>
<td>Number of foodstores + supermarkets per km2</td>
<td>1.013</td>
</tr>
<tr>
<td>Number of fast food restaurants per km2</td>
<td>1.025</td>
</tr>
<tr>
<td>Mean area used for parks and recreation per km2</td>
<td>8.453</td>
</tr>
<tr>
<td>Number of violent crimes per km2</td>
<td>0.999</td>
</tr>
<tr>
<td>% of housing requiring major repairs</td>
<td>0.002</td>
</tr>
<tr>
<td>Lives within 100 meters of a major road</td>
<td>1.777</td>
</tr>
<tr>
<td>2. No (ref)</td>
<td>---</td>
</tr>
<tr>
<td>Disturbed by noise at home</td>
<td>3.483</td>
</tr>
<tr>
<td>2. Agree</td>
<td>0.601</td>
</tr>
<tr>
<td>3. Neither agree or disagree</td>
<td>1.609</td>
</tr>
<tr>
<td>4. Disagree</td>
<td>1.114</td>
</tr>
<tr>
<td>5. Strongly Disagree (ref)</td>
<td>---</td>
</tr>
<tr>
<td>Individual level Variables</td>
<td></td>
</tr>
<tr>
<td>Age (in years)</td>
<td>1.066</td>
</tr>
<tr>
<td>Male</td>
<td>1.042-1.091***</td>
</tr>
<tr>
<td>Visible Minority</td>
<td>0.797</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
</tr>
<tr>
<td>2. High School Completed</td>
<td>1.463</td>
</tr>
<tr>
<td>3. Diploma Completed</td>
<td>1.546</td>
</tr>
<tr>
<td>4. Undergrad Completed (ref)</td>
<td>---</td>
</tr>
<tr>
<td>Tobacco use (daily average)</td>
<td></td>
</tr>
<tr>
<td>1. Current smoker vs. non-smoker</td>
<td>1.529</td>
</tr>
<tr>
<td>2. Former smoker vs. non-smoker</td>
<td>1.001</td>
</tr>
<tr>
<td>After tax Family Income in thousands</td>
<td>0.933</td>
</tr>
</tbody>
</table>

*p<0.05; **p<0.01; ***p<0.001
Model set #3 consists of a set of regressions predicting CVD status with adjustments for both socio-environmental predictors and individual socio-demographic controls across all six neighbourhood scales. In the following, I evaluate the associations across neighbourhood scale to answer research question #2, namely whether 1) associations are stronger/more significant in smaller administrative zones and 2) whether the use of amoeba buffers (compared to their respective administrative boundary) resulted in stronger/more significant associations, after adjusting for individual level covariates. See table 3.6 for a summary of results.

The NPA administrative boundary, and NPA amoeba buffer, scales consistently were not able to detect associations between any level 2 socio-environmental predictors and CVD status. While noise and living near traffic show up significant in the NPA and NPA amoeba buffer scales in table 3.5 and 3.6, these variables are effectively level 1 socio-environmental predictors (because individuals have unique values unrelated to other participants in their neighbourhood), and thus, their association with CVD status has little to do with neighbourhood scaling. This suggests that NPA zones and NPA amoebas may not be appropriate scales to analyze the association between socio-environmental context and CVD. This is presumably the result of the large areas they cover, with a mean area of 4.7 km² for NPA and 8.7 km² for NPA amoebas, which have masked relevant spatial variations. Moving down to the CT scale resulted in a two more significant associations: notably, an additional food store per CT is associated with the reduction in the likelihood of CVD (OR = 0.962, p<0.01) and an additional violent crime per km² in the CT is associated with an increase in the likelihood of CVD (OR = 1.005, p<0.001). This supports the ‘smaller administrative neighbourhood boundaries = stronger/more significant associations’ hypothesis.

Adding the amoeba buffers to the CTs, another association was discovered: each additional fast food restaurant is associated with an increase in the likelihood of CVD (OR = 1.01, p<0.05). This suggests that the location of residents within neighbourhood boundaries matter, and the measurement of the area of exposure for fast food restaurant
was improved using the amoeba buffer technique. Moreover, the negative association between food stores and CVD becomes slightly stronger when amoeba buffers are employed: moving from OR = 0.962 to OR = 0.873, both at p<0.01.

Table 3.7: Change in significant associations comparing 1) larger to smaller administrative boundaries and 2) administrative boundaries to respective amoeba buffers

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Change in significant associations between socio-environmental factors and CVD (at least p&lt;0.05)</th>
<th>Support Hypothesis 1: Smaller administrative stronger</th>
<th>Support Hypothesis 2: Amoeba buffers stronger</th>
</tr>
</thead>
</table>
| **Comparing bigger administrative scale and a smaller one**
| NPA → CT   | Gained (+) food stores per km$^2$ association to CVD | Yes | --- |
|            | Gained (+) violent Crimes per km$^2$ association to CVD | Yes | --- |
| CT → DA    | Gained (+) fast food restaurants per km$^2$ association to CVD | Yes | --- |
|            | Loss (-) food stores per km$^2$ association to CVD | No | --- |
| **Comparing administrative scale and the respective amoeba buffer scale**
| NPA → NPA amoeba buffers | No change | --- | No |
| CT → CT amoeba buffer | Gained (+) fast food restaurants per km$^2$ association to CVD | --- | Yes |
|                        | Stronger association for food stores per km$^2$ to CVD | --- | Yes |
| DA → DA amoeba buffer  | Gained (+) food stores per km$^2$ association to CVD | --- | Yes |
|                        | Stronger association for fast food restaurants per km$^2$ to CVD | --- | Yes |

Comparing the CT scale to the DA scale, the association between fast food restaurants and CVD is significant at the latter scale (OR = 1.01 for each additional restaurant per km$^2$, p<0.05), but not the former scale. This suggests that a more restricted area of effect to measure exposure may be appropriate for understanding access to fast food restaurants, and supports the ‘smaller administrative zone = stronger/more significant associations’ hypothesis. However, comparing the CT scale to the DA scale, the number of food stores per km$^2$ was significant at the former scale (OR = 0.962, p<0.01) but not in the latter. This does not support the ‘smaller administrative zone = stronger/more significant associations’ hypothesis.
Comparing DA zones to DA amoeba buffers, the association between additional food stores per km\(^2\) and decreased risk of CVD is detectable for the DA amoebas (OR = 0.873, p<0.01), but the same association was not found using DA zones. DA amoebas (compared to DA zones) also strengthened the association for between additional fast food restaurants per km\(^2\) and increased CVD risk from OR = 1.01 to OR = 1.81 (both at p<0.05). Both associations support the hypothesis that amoeba buffers (compared to its respective administrative boundary) produced stronger or more significant associations between socio-environmental factors and CVD status.

Overall, I am confident that the technique of amoeba buffering, in the case of median size (e.g. CT) and small size (e.g. DA) administrative boundaries, improved neighbourhood definitions by reducing the measurement error of exposure for a range of socio-environmental factors. On the other hand, smaller administrative boundaries do not consistently lead to stronger or more significant associations between socio-environment and CVD. Both medium and small administrative boundaries produced some significant associations, but large area boundaries (such as the NPA with a mean area of 4.6 km\(^2\)) are not an appropriate scale for studying socio-environment—CVD associations.

*Checking the pathway: testing obesity and diet quality as mediators*

There are two contextual level factors with significant impact on CVD risk, namely the availability of food stores/supermarkets and fast food, which are most likely mediated by obesity and poor diet. In order to verify the theorized pathway, an additional model (not shown) is specified where the participants’ body mass index (BMI) and diet quality (proxied using times per month eating fruits and vegetables) were adjusted in addition to the individual level confounders used in model set #3. When the two mediators (BMI and diet quality) were added, the effect for food stores/supermarkets and fast food restaurants became statistically non-significant. In other words, the 95% confidence interval of the odds ratio for both variables included 1.00 across every single neighbourhood definition tested above. In the meantime, all other significant factors (as found in model set #3), such as the number of violent crimes, traffic and noise retained their significant effect on CVD risk (at a minimum of p<0.05). This finding adds
credibility to the theory that BMI and diet quality are mediating factors in the association between the availability of food stores/supermarkets and fast food and CVD risk.

**Discussion**

While a new body of research now looks at how the characteristics of places affect health (Kawachi & Berkman, 2003), most researchers have tested place-effects based on a single set of administrative or census areas rather than using multiple scales. Since results vary significantly depending on the scale used, I recommend that several scales (at least small and medium sized boundaries) are used simultaneously for analysis. To do so is not to advocate that researchers ‘fish’ for significant associations by testing a range of scales, but to recognize that health-relevant risks and resources operate at multiple scales, and any single scale might mask relevant spatial variation if it is too large or excludes relevant socio-environmental features if it is too small. For example, while the number of fast food restaurants was not associated with the risk of CVD at the CT level, moving further down to DA scale produced a significant association, which means CTs were large enough to mask the spatial variation of fast food restaurants relevant to CVD risk. The ‘availability of fast food’ seems to be determined by an area more local than the size of a CT (which is more typically used in multilevel studies of health outcomes); a plausible mechanism behind this is that people might not want to walk very far to go eat fast food. The evidence here seems to suggest that fast food restaurants that are situated further than a resident’s home DA may have little impact on end-point CVD outcomes; however, further research is still necessary to understand how fast food restaurants of varying distances may impact diet. On the other hand, while the number of food stores was associated with CVD risk at the CT level, moving down to the DA scale eliminated the association, which means that food stores have a sufficiently large area of effect, and moving the scale of measurement to the DA level excludes too many food stores from the areas they affect. One plausible mechanism behind this is that people are willing to travel farther (at least beyond their home DAs) to access food stores. If a study only looks at food stores within a person’s DA, it will not have fully considered the local food environment. Due to these complex scalar effects, it is not safe to simply say that
smaller boundaries will produce stronger or more significant associations between socio-environmental characteristics and CVD risk.

While smaller boundaries are not necessarily better able to represent different socio-environmental characteristics, it should also be noted that there is a tendency for smaller areas to capture more between neighbourhood variations (table 3.3), before and after controlling for individual level covariates; therefore, clustering of CVDs appeared to be most marked at more local scales, but this local clustering effect does not necessarily coincide with the scale appropriate for measuring all health-relevant risks and resources. More importantly, this study goes to show that the size of the ICC does not capture all the variation of a given health outcome (in this case, CVD) explained by all neighbourhood contextual factors, because all contextual factors do not operate at a single scale. Many researchers have critiqued the field of ‘neighbourhood effects of health’ because the between area variance of a health outcome (most often measured using intraclass correlation, ICC) have been relatively small compared to the variance attributable to individual characteristics (Ellen & Turner, 1997; Sampson, Morenoff, & Gannon-Rowley, 2002), and very commonly, the value of the ICC derived from an unconditional or semi-adjusted model has been interpreted as percent of total (maximum) variance explainable by all observed and unobserved contextual factors (Inagami, Cohen, & Finch, 2007). However, this analysis shows that the value of the ICC is highly dependent on the scale of analysis (where smaller zones yielded higher ICCs); moreover, different contextual factors operate at differing scales and cannot be captured by an ICC calculated at a single scale alone. For example, noise and traffic impacted an area of exposure immediately adjacent to the home; the impact of fast food restaurants operated at the DA scale; and access to food stores operated at the CT scale. Thus, when the ‘between area variance’ is calculated at a single scale only, that value should not be interpreted as the maximum variance of a health outcome attributable to all neighbourhood factors, but simply the variance attributable to contextual factors that operate at the particular scale under investigation.

Despite their limitations, administrative and census boundaries will continue to be a popular choice for the purposes of investigating neighbourhood effects, because of their
ease of use, convenience, and they represent a common set of neighbourhood boundaries that researchers can use to replicate analyses. This study has demonstrated a method of using amoeba buffers to improve on administrative boundaries to more accurately assess the area of exposure for residents. Both CT amoeba buffers and DA amoeba buffers are relatively robust scales because they produced very similar results (see table 3.6), and unlike the CT and DA scales, they had no problems detecting associations of food stores or fast food to CVD risk. This study has not only demonstrated that the size of the boundary matters, but the shape matters as well: typical census and administrative boundaries are rectangular areas that artificially truncate the residents’ area of exposure for those that live near the edges, and amoeba buffers solves this problem by augmenting the shape of the neighbourhood. These buffers have the added advantage of preserving the original grouping to account for the clustering of residents to give an appropriate estimation of the standard errors for generalized linear mixed models. Lastly, in cases where the original administrative boundary is sufficiently large, amoeba buffers help to eliminate parts of the administrative boundary that are in reality quite far from residents sampled. Removing these areas help to improve neighbourhood definitions and highlight relevant spatial variations. While the effect sizes of the socio-environmental determinants in the study seem to be small, this is partly due to unit of measurement used: for example, at the DA amoeba buffer scale, while each additional violent crime is only associated with an increased odds ratio of 1.005 (p<0.001), for residents living in neighbourhoods with number of violent crimes 1 standard deviation above the mean would receive an odds ratio of $1.005^{56} = 1.32$.

NPAs have been adopted by the City of Toronto and community agencies as the neighbourhood scale for social planning and service delivery, but it is important to keep in mind that these areas may be neighbourhoods in name only, since they are generally too large to be immediately relevant for the residents that live in them, and their utility as a spatial scale for measuring exposure failed to find important associations to CVD risk that can be consistently identified at more local scales. They are similar in size and population to the US ZIP codes, which may also be an inappropriate scale to analyze the impact of socio-environmental context on CVDs. It is recommended that the City of Toronto and community agencies regularly using NPAs as the geographic scale of
analysis also consider using more local scales, especially for the purposes of evaluating programs for CVD interventions.

There are a number of limitations in this study. First, disease assessments were limited to self-reports; therefore, it is not possible to determine the clinical nature or severity of conditions, and undiagnosed cases may be differentially misreported. Second, it is also important to acknowledge that the relationships found here may not necessarily be causal, because of health-selected migration, that is, residents may choose the neighbourhoods in which they live on the basis of health-related characteristics (Sampson et al., 2002). Third, the study considered only 6 possible neighbourhood definitions; thus, the modifiable areal unit problem may not be explored comprehensively. Instead, the study compared neighbourhood boundaries that might be practically considered as alternatives in similar studies to analyze neighbourhood effects on CVD risk. The alternative scales chosen allowed us to simultaneously consider how shape and size of the boundary can impact model results. Despite these limitations, this exploratory study points to previously unexplored pathways through which the socio-environmental context may influence CVD etiology, it identified a number of scales that may be appropriate for this kind of study, and demonstrated amoeba buffers as a novel method of defining neighbourhood boundaries to help reduce measurement error.

The results of this study have important implications for researchers studying health with the aid of data for areal units, and indeed for any studies that employ ecological data. It is recommended that the following guidelines are considered in studies using aggregated data: first, the scale of analysis should not be taken for granted, especially when the topic under discussion is related to spatial processes, as is the case of neighbourhood effects on CVDs. The availability and relative ease of use of GIS software allows researchers to experiment with a range of spatial scales, and by aggregating spatial data in different ways we can begin to understand how different health-relevant risks and resources may have different areas of effect and should be measured at certain scales and not others. Second, amoeba boundaries can be used to improve neighbourhood definitions and may help produce more statistically consistent results compared to administrative boundaries.
Reference List


Concluding Chapter

The studies in this thesis demonstrate that an individual’s health status depends not only on the biological characteristics of individuals but also on the environments where people live and work, and these environments are in turn shaped by the economic and social policies driving differentiation, and shaping the features of residential (and non-residential) environments. By definition, advantaged neighborhoods offer safer and less stressful environments with better access to health-promoting resources compared to disadvantaged neighbourhoods. It would be shocking to learn that such contexts did not in any way have an impact on health. The unique contribution of this thesis points to the magnitude, mechanism, and mutability: what are the sizes of the effects, what are the specific pathways to health, and how might such information be used to improve and inform future research, interventions, and public policy?

The purpose of this concluding chapter is to summarize and demonstrate that the orientation, design, and findings of my three previous empirical studies have helped to advance theory and research in a coherent way. I will do so in the first section where I summarize the overall contribution of this thesis with regards to how my findings can be used to advance research and policy and help to fill existing research gaps in the field of place effects on health. The second section reflects on a number of outstanding issues such as 1) compensatory mechanisms, 2) causality, and 3) moving forward with exploring contextual influences at broader scales.

Section 1. Summary of findings and contribution to research and policy

The findings and contributions of my thesis can be grouped into three main categories: i) improving the conceptualization and measurement of neighbourhoods, ii) incorporating the real-life temporal/spatial trajectory into the study design, and iii) incorporating critical spatial thinking in the analysis of contextual effects on health. I will discuss each one in turn and explain how they may help to advance research, policy, and interventions.
i. Improving the conceptualization and measurement of neighbourhoods:

The majority of research on the contextual influences on cardiovascular diseases (CVD) to date, including studies reviewed by recent comprehensive literature reviews on the topic (Chaix, 2009; Ana V. Diez-Roux, 2003), focuses on the general impact of neighbourhood deprivation, but rarely analyzes specific environmental factors. Virtually no studies have empirically tested a full range of neighbourhood influences on CVD that are mediated by diet, physical activity, psychosocial stress, as well as neighbourhood factors that may have direct ‘non-conscious’ impacts on CVD risk (i.e. its impact of health is not mediated by health behaviour), which include air pollution and noise. The first major contribution of this thesis is the use of a full range of specific contextual influences to attempt to comprehensively understand the socio-environmental context of CVD risk. In the following, there are 2 main research and policy implications of taking this approach to the conceptualization of neighbourhoods: 1) shifting the focus from neighbourhood deprivation to specific socio-environmental factors as key contextual exposures; and 2) identification of specific-environmental exposures to diversify intervention strategies.

Shifting the focus from neighbourhood deprivation to socio-environmental exposures as key contextual exposures

Many studies have found robust associations between neighbourhood deprivation and CVD outcomes even after controlling for individual level risks (A. V. Diez-Roux et al., 2001; Sundquist, Malmstrom, & Johansson, 2004; Winkleby, Sundquist, & Cubbin, 2007). However, an important question regarding the mechanism that undergirds the association remains unanswered: is neighbourhood deprivation a proxy for other (unmeasured) contextual risks or does living in a deprived neighbourhood, holding all else equal, have its own unique impact? The results discussed in Study #1, in which both the neighbourhood SES and specific environmental factors were included in multi-level models estimating CVD, supports the first interpretation of the role played by neighbourhood SES. While neighbourhood SES alone was significantly associated with CVD risk (OR = 0.979 per $1,000 increase in median neighbourhood family income, p<0.05), neighbourhood SES was no longer significantly associated with CVD after
adjusting for specific environmental predictors including measures of the local food environment, availability of parks and recreational facilities, crimes, living near a major road, and noise. These results suggest that specific environmental predictors are more proximal determinants of CVD than neighbourhood SES, and the socio-environmental context possibly confounds the previously established neighbourhood deprivation--CVD association. Arguably, the results justifies that we shift our focus on neighbourhood deprivation to socio-environmental exposures as ‘key contextual predictors’ of health.

Admittedly, neighbourhood socio-economic characteristics may have other unobserved impacts on CVD through its impact on ‘individual-level variables’. An important theoretical quagmire for multilevel study of contextual influences is that individual SES is consistently seen to be uninfluenced by context, in other words, researchers have chosen to ignore contextual influences on individual income, employment status, profession, and education, treating them strictly as confounders of neighbourhood effects. Given the research on the spatial extent of labour markets disadvantage especially in the fields of regional studies and economic geography (Beatty, Fothergill, & MacMillan, 2000; Green & Owen, 1998; MackKay, 1999), individual incomes and professions are, arguably, partly shaped by where you live. Thus, individual SES may in fact be partly a mediator, and simultaneously a confounder, for CVD risk. The degree to which individual SES is spatially structured by the lack of local opportunities and area deprivation can mediate the association between neighbourhood deprivation and health (i.e. it is in the causal pathway of the contextual association). The degree to which individual SES is a product of agency or chance: individual SES can also confound the association between place and health. By agency or chance, I mean how an individual’s profession or income may also be influenced by choices and actions of the individual. Since previous studies that have identified a relationship between neighbourhood and CVD have controlled for individual SES, treating it as a confounder of the neighbourhood deprivation effect (A. V. Diez-Roux et al., 2001; Sundquist et al., 2004; Winkleby et al., 2007), I would interpret that in most cases, the neighbourhood deprivation measure is only acting as a proxy for other unobserved neighbourhood exposures. This is not to say that neighbourhood deprivation alone cannot influence CVD risk, but the typical approach of treating individual SES strictly as a confounder of
contextual effects, eliminates the possibility of identifying the neighbourhood deprivation → individual SES → CVD risk causal pathway. Future work should take into consideration that individual SES may be spatially structured (e.g. an individual’s education may be influenced by quality of local school or their income may be partly influenced by the local labour market), socially structured (e.g. influenced by race, parent’s SES, etc.) and also influenced by choice or chance. By controlling for individual SES, arguably, we may be over-controlling the effect size of context, because individual SES is partially determined by context. Therefore, the inclusion of individual SES as a confounder for the relationship between neighbourhood deprivation and health may be underestimating the effects of context. One way to overcome this problem, taken by study #1, is to model the effects of socio-environmental factors on CVD risk, which are not mediated by individual SES. The results in study #1 support that researchers include specific environmental factors, rather than simply neighbourhood deprivation, in multilevel studies of contextual effects on health.

*Diversifying interventions*

Conventional CVD preventive strategies typically involve the modification of individual-level behavioural and biological risk factors (Canto & Iskandrian, 2003). This is done in two ways: 1) public health education aimed at lifestyle change, which includes smoking cessation, diet modification, daily exercise, and weight management; and 2) early detection of risk factors such as hypertension and hypercholesterolemia through medical interventions, e.g. antihypertensive and lipid-lowering drug therapies. However, researchers have argued that these strategies alone may not be doing enough to address CVD risks at the population level (Chaix, 2009; Ana V. Diez-Roux, 2003; Rose, 2001). An important purpose of my thesis is to identify, and provide evidence to justify, place-based strategies aimed at prevention for ‘sick populations’ (Rose, 2001) to complement the conventional individualistic strategies, since a number of socio-environmental determinants with significant effect on CVDs are shaped by land-use decisions and the built environment. In the following, first, I will identify and deal with some common misconceptions about place-based interventions; then, I will discuss the specific contributions from this thesis that may help to advance place-based interventions.
Researchers often refer to the interest in the connection between place and public health with a sense of novelty, coloured by phrases like ‘recent interest’, ‘new paradigm’ and ‘an explosion of interest’ (Ana V. Diez-Roux, 2001, 2003), in some cases dating the starting point of this interest to the mid-1990s (S Macintyre, 2007). However, closer examination of the history of public health reveals a different story. In Western Europe and North America, public health and urban planning co-evolved as a consequence of late 19th century efforts to reduce the harmful effects of rapid urbanization and industrialization, particularly infectious diseases, through improvements to the urban environment and infrastructure to solve the problems of sanitation, water delivery, and waste disposal (Corburn, 2004; Melosi, 2000). It is widely recognized that the improvements developed in this period have led to dramatic increases in life expectancies across industrialized nations (Condran & Crimmins-Gardener, 1978). Thus, the interest in the connection between place and health is not new, but what is new, and somewhat surprising, is that some researchers are pitting individual-based interventions against ‘place-based interventions’ in a zero-sum approach (Andersson & Musterd, 2005; Dorling, 2001; McCulloch, 2001; Oreopoulos, 2008). For example, McCulloch (2001) in the context of the UK, and similarly Andersson & Musterd in the context of Western Europe (2005), advocate policies targeted at individuals (e.g. universal redistributive payments) over policies targeted at places (e.g. area-based regeneration) to deal with social and health inequalities. They advocate policies targeting individuals because the results of many multi-level studies found only moderate neighbourhood effects, and the individually targeted strategies are thought to reach many more disadvantaged populations. There is an important, and often neglected, dimension to this argument that relates to the distinction between ‘people policies’ and ‘place policies’, in which the previous researchers may be setting up a false dichotomy: for example, it would be hypocritical to advocate healthy diet in a public health education campaign without also addressing access to healthy foods with regards to its affordability as well local availability. Since access to many health-promoting resources are inherently place-based (for example, access to safe and affordable housing, availability of services, healthy food, access to clean air), any theoretical attempt to divorce people from the places they live in is theoretically and logically insufficient. In other words, ‘people policies’ and ‘place
policies’ must work together synergistically for effective interventions: changes to the socio-environmental context should be aligned to facilitate the effects of strategies that focus on individuals, and individual-based strategies can increase the awareness of the need for environmental change and may enhance support for this change.

The unique contribution of this thesis is the identification of specific socio-environmental factors that can be targeted to reduce CVD risk through place-based interventions. While previous studies have focused on the CVD risk of living in deprived neighbourhoods (A. V. Diez-Roux et al., 2001; Sundquist et al., 2004; Winkleby et al., 2007), this thesis clarifies the dimensions of this deprivation as 1) the lack of access to healthy food, 2) exposure to unhealthy fast food, 3) exposure to violent crimes, 4) exposure to air pollution and 5) exposure to noise, which are contextual factors that significantly increase CVD risk. With this information, we can move on to design innovative interventions to address these contextual risks. The specific recommendations suggested in this thesis include:

1) Planning and economic incentives may be used to encourage grocery stores, farmers markets, and food carts to serve in underserved communities, and provide municipal support to simplify the licensing process and reduce the cost of opening flexible low-cost, non-“brick and mortar”, forms of food access such as fresh fruit/vegetable carts and stands;
2) Restaurants may be incentivized to make changes in their recipes to reduce fat, sodium, and calories, and to increase fiber, fruit and vegetable servings;
3) Environmental design may be used to reduce crime using strategies such as improving surveillance (e.g. improving street lighting, reducing refuge for the potential offender) and encouraging neighbours to take ownership of space;
4) Urban design features may be used to reduce automobile traffic and promote non-motorized travel such as traffic calming, pedestrian and cycling friendly street design, increasing mixed use neighbourhood design, and site design features like reduction in building setback;
5) Site planning and architectural solutions may be used to reduce exterior noise: orientating residences away from sources of noise, using built and natural noise shields, creating room arrangements to place noise sensitive rooms (sleep and rest areas) further from sources of exterior noise, and using building materials and techniques known to be acoustically isolating, soundproofing windows facing major roads, and sealing cracks and edges of walls.

The above is not an exhaustive list of possible contextual interventions, but the findings of this thesis may stimulate an ongoing debate regarding what would be effective.
contextual interventions given the limits of feasibility and cost-effectiveness, which may be important factors for many cash-strapped and over-burdened municipalities. Evaluative and intervention-based studies to weigh the benefits and costs associated with these strategies would be an important next step towards implementation.

ii. Incorporating the temporal and spatial trajectory of everyday life into the study design

The second major area of contribution of my thesis is related to method development, which incorporates the temporal and spatial trajectory of daily life into the analysis of contextual effects. Specifically, this thesis tested original methods to address two empirical lacunae commonly found across studies of neighbourhood effects on health: 1) the implicit (and incorrect) assumption that duration of residential exposures are equivalent between participants, and 2) the lack of consideration of non-residential exposures save for rare exceptions (Inagami, Cohen, & Finch, 2007; Muntaner et al., 2006). To date, there are no other studies that analyzed the combined residential and non-residential socio-environmental impacts on CVD risk while accounting for durations of exposure. In the following, I will first review the specific findings related to the accounting of i) duration of exposure and ii) non-residential exposures; and then I will discuss how they may help to advance research and policy.

Admittedly, while the accounting for duration of exposure in study #2 was not perfect since detailed accounts of where and how time is spent throughout the day are not available, the estimations based on the weekly average work hours nonetheless provided weights that improved the effect sizes of every single significant contextual factor as well as improved the model fit when I compared the time-weighted models to the respective unweighted models. On the other hand, the inclusion of the workplace context into a cross-classified analysis revealed that the workplace is indeed a significant context, but the maximum variation of CVDs explained by the work context is only approximately half of what can be explained by the residential context. Moreover, the contextual factors that are significantly associated with CVD risk in the residential context are not significant in the work context (e.g. access to food store, fast food, and exposure to crime) with the exception of exposure to high traffic that is used as a proxy for air pollution. Neighbourhood-level SES in the work context was significantly associated
with CVD risk (OR: 0.981 per $1000 increase in the median household income in the workplace census tract, p<0.05) even after adjusting for residential contextual factors and individual-level risk factors, which may suggest that there may be unidentified contextual influences in the work neighbourhood. Apart from the effects of air pollution, further theoretical and empirical developments are needed to uncover other specific socio-environmental factors associated with CVD risk in the work context.

My thesis, thus, provides innovative methods to tackle two central research problems in the study of neighbourhood effects on health, namely the 1) ‘residential trap’ (Chaix, 2009), which refers to the exclusive reliance on residential environments and the systematic neglect of non-residential environments, and 2) the lack of consideration for duration of exposure in previous studies. One important implication for the interpretation of previous studies is that contextual effects may be masked by large variations in the (unobserved) duration of exposure: assuming that each person’s duration of exposure is equal when in fact they are not might lead to an underestimation of effects, because those who spend substantial amounts of time out for work/school/leisure will not be as influenced by residential factors compared to people who sit around at home all day. The implicit assumption of equal duration of exposure in most studies leads to the underestimation of contextual effects and may even increase the chance of type 2 errors. In order to more accurately assess the size of contextual effects, future studies should include duration of exposure weights where possible.

A second important implication of these findings relates to critiques of contextual effects research: critics often cite that mild to moderate neighbourhood effects, as well as the ‘low’ total variance of health explainable at the contextual level relative to individual-level variance, as reasons why place-based intervention may be unjustifiable (Andersson & Musterd, 2005; Dorling, 2001; McCulloch, 2001; Oreopoulos, 2008). However, since the research often cited in these critiques do not account for duration of exposure or non-residential exposure, it is reasonable to believe that the evidence-base of these critiques may be systematic underestimation of contextual effects. For instance, when I included the work context along with the residential context in the cross-classified analysis, the total variance explainable at the contextual level was increased by almost 40% (see study
Since this study only accounted for two contexts, total variance explainable at the contextual level may yet be higher if we develop the means to account for influences from additional contexts, and thus fully consider the temporal and spatial trajectory of everyday life. Where, and how, time is spent need to be accounted for before concluding that contextual effects only have a weak effect.

### iii. Incorporating critical spatial thinking into the analysis of contextual effects on health

While there has been a proliferation in the consideration of place and geography across the social sciences, public health and social epidemiology (Goodchild & Janelle, 2010; Kawachi & Berkman, 2003; Warf & Arias, 2009), more nuanced elements of geographical analysis including methodological considerations of spatial scale and spatial zone design have been relatively underappreciated even as place has gained recognition as one of the principle axis of social analysis (Nakaya, 2000; Purcell & Brown, 2005; Zhang & Kukadia, 2005). For instance, while many epidemiology and health geography studies use only a single set of neighbourhood definitions, it is often acknowledged that a single scale may be inadequate to investigate the spatial extent for different types of human activities (Sally Macintyre, Ellaway, & Cummins, 2002), which may operate at different spatial scales. The (incorrect) assumption that all contextual influences of health may be studied using only one neighbourhood scale is aptly termed the ‘local trap’ by Cummins (2007). On the other hand, while the use of administrative zones as proxies for neighbourhoods in ecological and multilevel studies of contextual effects has received much criticism because these zones may not necessarily match the spatial extent of residents’ everyday life (Ana V. Diez-Roux, 2004; Pickett, Collins, Masi, & Wilkinson, 2005), few researchers have attempted to improve the neighbourhood zone design to more accurately access levels of contextual exposure. The third major contribution of my thesis involves incorporating critical spatial thinking into the analysis of contextual effects on health by i) investigating the scalar dimension of contextual influences and ii) developing and testing a novel method of designing neighbourhood zones to more accurately access levels of contextual exposure. In the following, I will first review my major findings, and then I will discuss how these findings may be a unique contribution to research and policies.
Most research to date has paid scant attention to issues of the relevant scale in studying the effects of residential environments on health. In study #3, I analyzed neighbourhood effects on CVD risk at 3 different administrative scales including the (smallest) census dissemination areas (DA), which typically only include a few city blocks (mean area = 0.18 km$^2$), to (medium sized) census tracts (CT), which are the most commonly used proxy for residential neighbourhoods (mean area = 1.042 km$^2$), and up to the (largest) Toronto neighbourhood planning areas (NPA), which are large multi-block areas (mean area=4.56km$^2$) roughly equivalent to typical US zip code areas. While no significant association between contextual variables and CVD risk was found at the NPA scale, certain significant relationships existed at the CT scale but not at the DA scale and vice versa. Specifically, the association between fast food restaurants and CVD is significant at the DA scale (OR = 1.01 for each additional restaurant per km$^2$, p<0.05), but not at the CT scale$^1$. On the other hand, the availability of food stores was significantly associated with CVD risk at the CT scale (OR = 0.962, p<0.01), but not at the DA scale. These findings suggest that the issue of scale in the study of contextual effect can be complex, and investigation of multiple scales may be required to identify the relevant scale that matches the specific contextual factor in question.

With regards to designing and testing a novel method to reduce the measurement error of contextual exposures, study #3 deals specifically with the problem of edge participants of administrative boundaries who have their area of exposure artificially truncated. A method of amoeba buffering was devised to augment the shape of the neighbourhood so that every participant within a given administrative boundary is fully enveloped: with a minimum of 1 km distance to the amoeba edge for the outermost participant (see study #3 on the rationale of size and details of construction). In order to test the effectiveness of these buffers, I compared the effect sizes of contextual variables between the amoeba buffers constructed at the DA, CT, and NPA scales with their respective administrative boundaries. In the case of median size (e.g. CT) and small size (e.g. DA) administrative boundaries, that amoeba buffers (compared to its respective administrative boundary) produced stronger associations between socio-environmental

$^1$ In study #3, the full linear combination (continuous variable) is used to characterize the availability of fast food restaurants. The full linear combination measure of fast food restaurants was only significantly related to CVD risk at the DA scale and not at the CT scale.
factors and CVD risk, which suggests that amoeba buffers may improve neighbourhood definitions by reducing the measurement error of exposure.

While the scalar dimension of specific socio-environmental influences on CVD risk has been theorized (Ana V. Diez-Roux, 2003), virtually no studies have empirically tested it, since one of the only multi-scalar studies of contextual influence on CVD risk limited the scope of their analysis to neighbourhood SES (Lovasi et al., 2008). Thus, study #3 is a unique contribution to research in 2 ways: 1) it is the first to design and test a new method to reduce the measurement error of contextual exposures by carefully considering the location of edge participants, and 2) it is the first multi-scalar multilevel study testing specific socio-environmental factors on CVD risk. The findings are important because it justifies that 1) multiple scales are used in future studies since not all contextual associations with health are detectable at all scales, and 2) standard administrative boundaries may underestimate effect sizes of contextual influences by producing measurement error for edge participants, and amoeba buffers is an innovative and relatively convenient way to get around this problem.

Many researchers have critiqued the field of ‘neighbourhood effects of health’ because the between area variance of a health outcome (most often measured using intraclass correlation, ICC) have been relatively small compared to the variance attributable to individual characteristics (Ellen & Turner, 1997; Sampson, Morenoff, & Gannon-Rowley, 2002), and very commonly, the value of the ICC derived from an unconditional or semi-adjusted model has been interpreted as percent of total variance explainable by all observed and unobserved contextual factors (Inagami et al., 2007). Some researchers have gone on use this as the primary reason for dismissing all place-based interventions (Dorling, 2001; McCulloch, 2001). This study shows that variance measured at a single neighbourhood scale, indicated by ICC value, does not capture all the variation of a given health outcome explained by all neighbourhood contextual factors. This is because contextual factors do not operate at a single scale. For example, noise and traffic impacted an area of exposure immediately adjacent to the home; the impact of fast food restaurants operated at the DA scale; and access to food stores operated at the CT scale. Thus, when the ‘between area variance’ is calculated at a single scale only, that value should not be interpreted as the maximum variance of a health
outcome attributable to all neighbourhood factors, but only as the variance attributable to contextual factors that operate at the particular scale under investigation. Study #3 also shows that the value of the ICC is highly dependent on the scale of analysis, where smaller zones yielded higher ICCs. Since virtually all studies have only used a single spatial scale to arrive at what researchers consider to be the variance explainable by neighbourhood influences, it is reasonable to conclude that this is yet another way contextual effects have been grossly underestimated across most studies.

By incorporating critical spatial thinking into the field of neighbourhood effects on health, I have helped to advance the understanding of ‘neighbourhood’ and ‘place’ by introducing the notion of scale as well as zone. It is my hope that the work that has begun here will help to challenge the interpretations of previous research into contextual effects on health by the way of pointing out the previously missing pieces, which includes the lack of consideration for 1) duration of exposure, 2) non-residential exposures, 3) the problem of single scale studies, and 4) measurement error of contextual influence that is the result of the ‘edge-participant’ effect: all of which may lead to the underestimation of contextual effects, and thus, possibly resulting in erroneous conclusions drawn about the importance and relevance of place-based interventions. However, this thesis is, by no means, a project only based on criticism or theoretical deconstruction. For each of the central research problems that have been identified by leading researchers in the field of neighbourhood effects on health (Chaix, 2009; Chaix, Merlo, Evans, Leal, & Havard, 2009; Ana V. Diez-Roux, 2001; Kawachi & Berkman, 2003; Sally Macintyre et al., 2002), namely 1) underdeveloped characterization of neighbourhoods, 2) lack of consideration for non-residential exposures (also commonly termed the ‘residential trap’), 3) a lack of consideration for variations in the duration of exposure, 4) ignoring the ‘modifiable areal unit problem’ (MAUP) where the zones that constitute neighbourhoods may not necessarily match the spatial extent of residents’ everyday life, and 5) ignoring the fact that neighbourhood level exposures may operate in different spatial scales (also commonly termed the ‘local trap’), I have either designed or introduced (as well as tested) novel techniques to deal with these problems. These techniques include, in matching order to the problems, 1) the use of specific socio-environmental predictors, 2) cross-classified analysis to introduce workplace contextual influences, 3) time-weighted
analysis to account for duration of exposure, 4) amoeba buffers to improve neighbourhood zone design to more accurately access areas of exposure, and 5) multi-scalar analysis to account for exposures operating at different spatial scales to tackle each of the central research problems in the study of neighbourhood effects on CVD, but most of which are also applicable to the study of neighbourhood effects on health more generally. Finally, in addressing the research gaps that were identified in this thesis, it was only through work that transcend disciplinary boundaries, bringing together theory and research in the fields of urban planning, sociology, geography, behavioral science, environmental studies and social epidemiology that made it possible to tackle some of the central research problems in the study of neighbourhood contextual effects on health. Further work taking a trans-disciplinary approach will be necessary to help translate this research into actionable interventions. In the next section of this concluding chapter, I will highlight a number of outstanding issues such as 1) compensatory mechanisms, 2) causality, and finally 3) moving forward by exploring contextual influences at broader scales.

Section 2. Outstanding issues and next steps

Compensatory mechanisms and mitigating factors

While I have identified a number of socio-environmental factors that significantly impact CVD risk, these contextual influences may be modified by the fact that individuals may mitigate deficiencies in their residential (or non-residential) neighbourhoods through diverse mechanisms. For instance, participants who live in areas with few recreational facilities may still be physically active because they have access to a gym at work. Participants that live in areas with no access to food stores may have their groceries delivered. On the other hand, neighbourhoods that have no fast food restaurants may be targeted by advertisements for fast food delivery. Relationships can sometimes be complex and the influence of contextual factors may be modified in unexpected directions. For example, in a US study by Ross (2000), the author expected that residents from poor neighbourhood would be less likely to walk and exercise, but the results were not consistent with expectation as those who lived in poor neighbourhood were less likely to drive and more likely to walk (even after controlling for gender, individual poverty,
household income, education, race, marital status and age), because of better access to public transportation and mixed land use also encouraged walking. Consideration of these compensatory and mitigating factors, as well as investigating whether these effects may differ by country, can help us better understand the true contextual effects on health as well as design intervention strategies that take these complexities into account. Detailed accounts of how and where time is spent throughout the day can help to elaborate on these nuances, as will qualitative studies that explore how residents may use coping strategies to tackle contextual challenges.

_Causality_

This study, like the majority of other studies linking place to CVD, has been observational. Observational design has to contend with the issue of residential selection, and thus unable to draw a conclusive causal link between predictor and outcome. In other words, participants may be selected into different types of residential areas based on individual-level characteristics that may themselves be related to CVD risk, which would result in a case of reverse causation. Although we can attempt to account for these individual-level differences by statistically controlling for individual-level variables, the possibility of selection bias remains due to study design.

The limitations of observational studies have led some researchers to call for increasing the use of randomized community trials in the study of the effects of residential environments on health to overcome selection bias. Although these quasi-experimental approaches would help to solve many of the problems of observational studies, they have their own set of methodological and ideological problems. I will explain my position to follow: first, by briefly outlining their history, how they work, and then offering my critique.

Quasi-experimental design studies based on housing mobility programs have been argued to be the strongest evidence available of neighbourhood effects on health outcomes (Acevedo-Garcia et al., 2004). Following the well-known Gautreaux program, a US federal court-ordered racial desegregation program in Chicago in the late 1970s (Goering, 1996), the Moving to Opportunity program (MTO), the most well-known of
these housing mobility programs, included a randomized controlled trial to understand
the independent effect of moving households from high-poverty to lower-poverty
neighbourhoods in five US cities between the years of 1994-2006 (Goering, 2005;
Jackson et al., 2009; J.R. Kling, Leibman, Katz, & Sanbonmatsu, 2004). This
experimental strategy divides participants into three groups (Jackson et al., 2009): 1) a
control group that does not relocate and remains in public housing, 2) a comparison
group that has access to regular Section 8 vouchers with the choice to relocate to any
neighbourhood, and 3) an experimental group that receives a restricted housing voucher
that can only be used in a low-poverty neighbourhood. While health was not originally
anticipated as an outcome of the MTO project, it is now an important discourse and
rationale behind similar housing-mobility programs and poverty deconcentration policies
(Acevedo-Garcia et al., 2004; Briggs, 2005; J.R. Kling et al., 2004; Leventhal & Brooks-
Gunn, 2003). A number of reviews summarize some positive and mixed socio-economic
and health outcomes of those who are relocated (Acevedo-Garcia et al., 2004; Jackson et
al., 2009). They are primarily associated with mental health improvements among adults
and reduced exposure to crime in the experimental group (Goering, Feins, & Richardson,
2002; J.R. Kling et al., 2004; Jeffrey R Kling, Ludwig, & Katz, 2005; Leventhal &

MTO, seen as a research program and social experiment, has had an immense
impact on our policy imagination. Numerous studies based on the program have had the
ability to overcome the problem of selection bias, and thus many researchers see them as
representing a promising future direction for neighbourhood effects on health research
(Acevedo-Garcia et al., 2004; Briggs, 2005; Sampson et al., 2002). However, this
experimental literature is also intrinsically linked to certain assumptions about the
concentration of poverty and questions around whether its deconcentration can improve
health outcomes (Chum, 2011). While these studies investigate whether housing mobility
programs can improve the health of the families involved, it is important to emphasize
that they do not indicate whether poverty deconcentration is the best policy for reducing
health inequalities. Since poverty dispersal is the primary treatment considered in these
studies, other possible social interventions and policies aimed at broader scales that may
be more health promoting or socially just, are not considered within this research
framework. In other words, randomized community trials inspired by the MTO model do not test the effectiveness of modifying contextual risk factors, but the effectiveness of ‘moving people out of poor neighbourhoods’ as a form of treatment.

Since the results of my thesis show that neighbourhood deprivation may be a proxy for more proximal contextual determinants of health, interventions aimed directly at modifying these proximal contextual risk factors such as exposure to crime, noise, air pollution/traffic, access to healthy food may be a more effective strategy. Evaluation of the behavioral and health effects of interventions that modify the built environment (including the five specific recommendations that I discussed above) is likely to yield more useful and relevant information compared to randomized community trials based on the MTO model, because the results are directly useable by urban planners, designers and architects who are already involved in shaping our built environment. The results from these studies will help to connect their respective fields with public health relevance, and may help to mainstream health equity concerns in broader cross-disciplinary and intersectoral debates.

Finally, longitudinal studies can be used to deal with the problems of selection bias and reverse causation because we can track the temporal string of events to understand the extent to which causes are followed by effects. However, there is a danger that the past is used to control out context in these studies (Dorling, 2001), especially if the time between the cycles is not sufficiently long enough for the context to influence health status. In other words, since longitudinal studies are concerned with the change in health status between time 1 (T1) and time 2 (T2) in a simple 2 cycle study, controlling for health status before T1, in the case where the time between T1 and T2 is not long enough, is likely to actually control out the influence of past context. Obviously, ‘long enough’ is up for debate, but the incubation periods of chronic diseases can take up to decades, making new investigations unfeasible.

**Moving forward: exploring contextual influences at broader scales**

While none of the socio-environmental influences that I tested were significantly associated with CVD at the NPA scale (highest scale in my study), it certainly does not
rule out the possibility of unobserved contextual influences at sub-municipal, municipal and regional scales, which may include a different set of contextual determinants of health than the ones tested in this study. These broader scales are often neglected in multilevel studies of place effects on health (Sally Macintyre et al., 2002). However, I have begun to theorize plausible influences at these scales in my discussion in *Health and Place* (Chum, 2011), where I articulated a theory of competitive urban structure that shapes neighbourhood resources and opportunities, and in turn structuring health outcomes. The competitive urban structure may exert influences on health through the interrelated process of municipal fragmentation, exclusionary politics, and municipal competition. In the following, I will discuss the method of influence (mechanism) for some of these higher scalar factors, and then I will discuss how we should proceed with regards to research and policy.

A high number of municipalities per household (i.e. municipal fragmentation) have been argued to increase consumer choice (Altshuler, Morrill, Wolman, Mitchell, & Committee on Improving the Future of US Cities Through improved Metropolitan Area Governance, 1999; Tiebout, 1956), where each offers a unique combination of public goods and levels of taxation. However, most services and amenities provided by the local government (such as clean water, safety, and education) are fundamental to a minimum level of well-being for all. Moreover, this local government fragmentation creates a situation where municipalities behave like firms in a competitive market economy. They struggle over scarce resources through the process of exclusionary land use planning and compete with other municipalities for tax income (Luce, 1998). Winners and losers inevitably emerge in this competition, and elucidation of these struggles can tell us much about the structure that shapes our patchwork of affluence and deprivation in our cities today.

Since a large proportion of local revenue is derived from property taxes typically ranging from 50-80% (Kazemipur & Halli, 2000; Orfield, 2002) and municipal expenditures are tied to the types of land uses within their jurisdictions, local governments also have an incentive to engage in what is known as exclusionary zoning to maximize the number of profitable land uses and to avoid high expenditure uses
(Ihlenfeldt, 2004). For example, land use for industrial parks, offices, condominiums, and expensive single-family homes may generate more revenues than costs. On the other hand, large apartments, affordable housing, and inexpensive tract housing tend to incur negative fiscal dividends. Local governments regulate and exclude unprofitable land uses using development tools. These tools include zoning, minimum lot size, setbacks, and design standards (Duncan & Duncan, 2004). This exclusionary process can significantly impact where people are able to live (aggravating economic and racial segregation) (Rusk, 1993), the extent and quality of services and amenities municipalities are able to offer, and the presence of employment opportunities for residents. Thus, exclusionary land use planning facilitates the concentration of poverty as well as structurally disadvantaging municipalities that have a high proportion of affordable housing and other socially necessary high cost uses e.g., community centres, and should be considered as key detractors from neighbourhood well-being.

Finally, competition is pervasive among local governments throughout North America, but it is wasteful and only serves to shuffle economic activities from one place to another (Orfield, 2002). Municipalities compete with one another to attract desirable businesses and economic activities. As businesses are lured from one jurisdiction to another through financial or other incentives, one community’s loss then becomes another’s gain. Due to the level of competition, communities also have to spend valuable and scarce resources (that could be better spent elsewhere) to ensure that existing businesses do not leave for other places in the metropolitan area that offer them incentives. Thus, the lack of co-operation among municipalities fostered by a high level of fragmentation, exclusionary politics, and competition reinforces residential segregation and creates winners and losers in the highly competitive urban marketplace.

There are a number of challenges with regards to estimating the influence of broader contextual factors such as municipal fragmentation, the degree of exclusionary politics, and municipal competition on individual health status. First, in order to empirically investigate contextual effects on health in a multilevel approach at the municipal scale and beyond, we need access to individual level, neighbourhood level, and municipal level data across multiple municipalities which may be best achieved through a
collaborative approach involving intersectoral and international research partners across a 
diverse range of municipal contexts. A less ambitious approach may involve linking a 
umber of surveys that include individual-level health outcomes across diverse settings, 
and obtaining neighbourhood level and municipal level information from separate 
sources. Second, it may be difficult to assess the impact of higher scalar contextual 
factors that I identified above without further methodological developments. For 
example, if municipal fragmentation is thought to influence the spatial polarization of 
municipal income, and in turn lead to uneven distribution of social services, and finally 
driving individual health, we may first want to establish the temporal chain of these 
events through a retrospective analysis, followed by combining multilevel analysis with 
structural equation modeling to properly account for all the mediating factors that lie 
between levels of municipal fragmentation and individual health status as well as the 
hierarchal structure of the data (Goldstein, 2003). Finally, little is understood about the 
lag effects of higher scalar contextual influences, while lag effects of contextual 
influences on health is still a relatively new area of study (Blakely, Kennedy, Glass, & 
Kawachi, 2000; Gadalla & Fuller-Thomson, 2008), including both recent and historical 
contextual information (possibly by to 25 years as in the Gadalla study) with multiple lag 
times will help us understand the temporal dimension of ‘trickle-down’ in these 
contextual processes. Efforts aimed at improving the understanding of these higher scalar 
contextual influences on health may be well worth the effort since they may structure a 
number of more proximal contextual influences such as crime, local food environment, 
traffic, which were identified in this thesis. As such, addressing the health of populations 
may require that we look at not only contextual factors at the neighbourhood level, but 
also at municipal and regional levels.
Reference List


