Forecasting the Local Risk of Type 2 Diabetes and the Effects of Prevention Programs Using Spatial Microsimulation: Development and Use of the TropISM Model

by

Peter Driezen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Public Health and Health Systems

Waterloo, Ontario, Canada, 2017

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Simulation modelling has become an important tool in social science research, though such models are less commonly used in population health. Spatial microsimulation models provide a unique way to estimate health outcomes at the small area level and forecast the local effects of potential health interventions. Spatial microsimulation models combine geographically rich census data with information rich survey data to generate synthetic small area populations containing the wealth of information available in both types of data. Model outputs can therefore be used to inform the local delivery of health promotion programs.

This research developed and validated the Type 2 dIabetes Spatial Microsimulation ("TropISM") model of chronic disease risk factors and outcomes for the 140 neighbourhoods of metropolitan Toronto. The model was developed using Canadian census data from 2006 and the 2005 Canadian Community Health Survey. The five-year incidence of diabetes was estimated using the Diabetes Population Risk Tool (DPoRT 2.0), a population-level risk algorithm that forecasts disease incidence using risk factor information routinely collected in population health surveys, together with the synthetic TropISM population. By leveraging both models, it was possible to evaluate the effects of hypothetical weight loss interventions on potential reductions in diabetes incidence at the neighbourhood level. Synthetic, neighbourhood specific prevalence estimates of diabetes were also used to estimate potential spatial accessibility to diabetes education programs within metropolitan Toronto. Accessibility was estimated using a two-step floating catchment area model, a type of spatial interaction model used to estimate area specific provider-to-population ratios.

Results indicate that although the TropISM model accurately replicated demographic characteristics of Toronto's 140 neighbourhoods, it underestimated the true prevalence of type 2 diabetes, hypertension, and heart disease among men. In addition, TropISM captured broad spatial patterns in disease prevalence, but was unable to capture the spatial variability in known prevalence assessed from administrate health databases maintained by the Institute for Clinical Evaluative Sciences. Irrespective of these limitations, when the DPoRT model used synthetic TropISM population to forecast diabetes incidence, the overall five-year forecast

incidence was comparable to the true incidence of disease (4.9% vs. 5.8%, respectively; 95% uncertainty interval: 4.2%-5.9%).

At the neighbourhood level, the true incidence of diabetes fell within the range of forecast uncertainty in 65 neighbourhoods, while forecast incidence was underestimated in 52 neighbourhoods, most of which were located in Scarborough and Etobicoke. Again, broad spatial patterns in forecast incidence were captured by TropISM even though forecasts did not capture the spatial variability in true incidence rates.

When the synthetic TropISM population was used to assess the ex ante effects of population-level weight loss programs on the future incidence of diabetes in silico, the entire population of high-risk, overweight individuals having a body mass index ≥ 25 kg/m^2 would have to lose 17% of its baseline body weight to produce a reduction in diabetes incidence of just one percentage point across metropolitan Toronto. Greater reductions in incidence were observed in neighbourhoods comprised of larger proportions of visible minorities and immigrants, even though the baseline prevalence of overweight and obesity tended to be slightly lower in these neighbourhoods compared to the metropolitan average (39% vs. 41.2%, respectively).

Finally, the two-step floating catchment area model was used with synthetic, neighbourhood-specific counts of type 2 diabetes to conduct an exploratory analysis of spatial accessibility to diabetes education programs located throughout metropolitan Toronto. Results point to a potential mismatch between population demand for services and potential spatial access. In particular, some neighbourhoods within Scarborough had relatively higher prevalence of type 2 diabetes but lower access to diabetes education programs while neighbourhoods within central Toronto tended to have greater spatial access and lower prevalence rates of type 2 diabetes. Disparities in service provision suggest additional resources could be devoted to diabetes management in high-prevalence, low service neighbourhoods.

In light of its short-comings, TropISM model results suggest how spatial microsimulation models can be improved to produce more accurate neighbourhood-specific estimates of diabetes prevalence. Importantly, TropISM was able to capture broad spatial patterns in diabetes prevalence and incidence, providing insight

into where weight loss programs may contribute to greater reductions in diabetes incidence and identifying factors that might influence those reductions. This information can be used to customize health promotion interventions to the particular needs of specific communities.

In conclusion, the TropISM spatial microsimulation model was able to (a) predict the consequences of different weight loss programs on projected diabetes incidence and (b) identify potential mismatches between existing demand for health promotion programs and the geographic availability of those resources. This locally relevant information enables public health planners to better allocate scarce resources to communities of greatest need. This research therefore illustrates how spatial microsimulation modelling can be used as a spatial decision support tool for local public health planning.

Acknowledgements

This research represents the culmination of a long journey. Several people played instrumental roles in encouraging me to see this work through to its completion. First and foremost, I would like thank my advisor, Dr. Ian McKillop, for his continued words of encouragement and instructive feedback which helped shape the final product presented here. I would also like to thank my committee members, Dr. Rob Feick and Dr. Joon Lee, for their critical feedback and thoughtful suggestions. All of your help is greatly appreciated!

A critical piece of this research would not have been possible without the help of Dr. Walter Wodchis and Mr. Chris Bai from the Health System Performance Research Network at the University of Toronto. Dr. Wodchis graciously provided access to administrative health data from the Institute for Clinical Evaluative Sciences while Mr. Chris Bai extracted the necessary data and provided me with summary datasets needed to validate the TropISM model.

I would also like to thank my sister-in-law, Dr. Diana van de Hoef, for taking the time to proofread the penultimate draft of my thesis and catching some amusing typographical errors.

Finally, this work would not have been possible without the support of my wife, Jennifer, and our two children, Marc and Breanna. Thanks for keeping me smiling!

Dedication

To my loving wife Jennifer,
and to Marc and Breanna,
for your unending patience,
encouragement, and continued support.
I couldn't have done it without you!

Table of Contents

List of Tables

List of Figures

Chapter

1

Introduction

1.1 The Role of Local Information in Decision Making

Evidence-based decision making is a cornerstone of medicine and has come to play a critical role in public health practice. Evidence-based public health relies on the best available scientific information to guide the implementation and delivery of interventions designed to improve population health (Brownson, Fielding, & Maylahn, 2009). Health policy makers rely on evidence to allocate scarce resources. Health policy making is a rational decision making process where policy makers:

- identify the objectives of a health program or policy,
- identify the options available to meet those objectives,
- predict the consequences of each option, and
- evaluate the consequences (Kaati, Sjostrom, & Vester, 2004).

A unique feature of evidence-based public health is the local context in which decisions are made. Often, public health practitioners and policy makers must transfer evidence about the effectiveness of an intervention developed in one particular setting to a new locale. Doing so, however, changes the conditions under which an intervention was originally delivered and may result in that intervention being less effective. Thus, the local context is important for decision making and public health decision makers should incorporate that information into the decision making process.

The importance of local information in decision making might be best illustrated with an example of swarm behaviour. In *The Smart Swarm,* Miller (2010) describes how a colony of ants can make seemingly intelligent decisions even though an individual ant is by no means intelligent. When searching for a source of food, the Argentine ant secretes pheromones that guide other ants to a food source. When choosing between two different paths to a food source, a colony of Argentine ants will usually choose the shorter path. The key to this intelligent "decision" is the pheromone trail. As more ants use the shorter trail, it accumulates a greater concentration of pheromones, inducing other ants to take the shorter path. In this case, the ants are using *local* information to collectively make a decision about the shortest path to a source of food.

Likewise, local information is important for decision making in public health. Congdon (2012) argues that geographically disaggregated forecasts of future diabetes prevalence are necessary for tailoring public health interventions to the local context. People living in different areas may not respond in the same ways to the same policy or program. Ballas and Clarke (2001); Ballas, Clarke, and Wiemers (2006) and Ballas, Clarke, Dorling, and Rossiter (2007a) argue that it is important to understand and forecast the local effects of national level policies. Understanding how different localities vary in their responses to national policies allows policy makers to better respond to the needs of their local communities. Moreover, predicting the local consequences of different policy options equips decision makers with the necessary information needed to select one policy or program over another. Predicting the consequences of health programs or policies is a crucial step in the decision making process (Kaati et al., 2004).

1.2 Microsimulation Modelling in Public Health Research

Simulation modelling has become an important tool in social science research. It is a valid way of conducting scientific research, following on the heels of deductive and inductive methods (Axelrod, 1997; Garson, 2009). The goal of simulation modelling is to understand the behaviour of complex systems using simpler models. According to Harrison, Lin, Carroll, and Carley (2007), one of the goals of simulation modelling is to *predict* how a system might respond to different sets

of initial conditions or how a given set of initial conditions might result in different outcomes if different processes are modelled. These types of simulations are useful for predicting the possible effects of a policy, allowing researchers to "compress time" and experiment with the effects of different policies without having to implement them in real-world settings (Turban & Aronson, 2001). Simulation models therefore provide public health decision makers with useful ways to predict the consequences of health programs and policies.

Microsimulation models comprise one type of simulation modelling used in public health research. These models simulate a population of individuals to predict the aggregate behaviour of a larger system (Spielauer, 2011). Each individual possesses a set of traits that may change over time (Klevmarken, 2008). A set of rules and relationships govern individual-level behaviour, specifying how individuals change over time or how they respond to certain conditions. Using the simulated population of individuals, it is possible to predict the effects of a change in policy on the entire population.

Microsimulation modelling has its roots in economics. First suggested by Orcutt in 1957, early microsimulation models were developed to model the effects of different public policies on the American labour market (Birkin & Clarke, 2011; Orcutt, Caldwell, & Wertheimer II, 1976). Orcutt's original DYNASIM model used a cross-sectional database of individuals to project demographic and economic events over time, including marriage, fertility, education, employment, earned income, retirement, and death (Ross, 1991). As illustrated by the DYNASIM model, there are two key features common to all microsimulation models. The first is a micro dataset of individual-level data; these data often come from sample surveys of some target population. For example, Kopec et al. (2010a) used data from the 2001 Canadian Community Health Survey, as well as physician diagnosed incidence rates of osteoarthritis and statistical regression models, to forecast the prevalence of osteoarthritis in the Canadian population.

The second component of a microsimulation is a model that specifies how individuals in the simulated population change over time (Zucchelli, Jones, & Rice, 2012). The Population Health Model is a continuous-time microsimulation of the Canadian population that ages individuals and predicts the likelihood that each individual experiences different life events, such as diabetes, hypertension, heart

disease, or cancer (Nadeau et al., 2013). Regression models estimated from the National Population Health Survey define several models used to predict these events. Predicted events are averaged over the entire simulated population to forecast the average prevalence of each event at some future time point.

Spatial microsimulation models form one unique subclass of models that simulate populations within specific geographic areas. Tanton and Edwards (2013) state that spatial microsimulation models are primarily used for (a) small area estimation, (b) small area projection, and (c) small area policy modelling. The unique feature of spatial microsimulation is that it explicitly focuses on *where* simulated individuals live. Geographic location is important in these models because it might affect how individuals behave and respond to public policies. Spatial microsimulation models therefore attempt to simulate realistic populations for specific small areas, such as census dissemination areas, in order to forecast the geographic effects of policies applied at regional or national levels (Ballas, Clarke, Dorling, Rigby, & Wheeler, 2006; Ballas et al., 2007a; Ballas, Rossiter, Thomas, Clarke, & Dorling, 2005a; van Leeuwen, 2010a). Spatial microsimulation models are useful because detailed data are often unavailable at fine geographic levels and the effects of public policies might vary over geographic space (Ballas et al., 2005a). Collecting detailed data using survey methods is often too expensive given varying demands for detailed data and the limited resources available to collect such data.

In summary, microsimulation models assess how individual-level changes generate aggregate regularities. In other words, microsimulation is a bottom-up approach to forecast population-level changes and plausible responses to public policies (van Leeuwen, 2010a). Thus, microsimulation models can be used to assess the effects of alternative policy options *before* they are implemented (Clarke, 1996). This is important for policies whose effects may not manifest for long periods of time following implementation.

1.3 Research Objectives

This research will demonstrate that spatial microsimulation is a useful tool for local health planning. Using type 2 diabetes as an example, it will demonstrate how

spatial microsimulation can inform the delivery of health promotion programs at the local level. The specific objectives of this research are:

1. To develop the "Type 2 dIabetes Spatial Microsimulation" model ("TropISM"), a spatial microsimulation of type 2 diabetes and its risk factors, for neighbourhoods within metropolitan Toronto.
2. To assess the validity of the TropISM model by comparing small area estimates generated by the model to known estimates obtained from external data sources.
3. To use the simulated TropISM population in conjunction with the Diabetes Population Risk Tool (Manuel, Rosella, Tuna, Bennett, & Stukel, 2013; Rosella, Lebenbaum, Li, Wang, & Manuel, 2014; Rosella, Manuel, Burchill, & Stukel, 2011) to
    a) forecast the five-year incidence of diabetes at the neighbourhood level,
    b) project how weight loss in high-risk individuals affects the neighbourhood-level incidence of diabetes, and
    c) assess the uncertainty of these projections using probabilistic sensitivity analysis.
4. To use TropISM model outputs with a spatial interaction model to
    a) measure potential spatial access to diabetes education programs for neighbourhoods within metropolitan Toronto,
    b) measure potential spatial access to diabetes prevention programs, implemented in community recreation centres, for neighbourhoods within metropolitan Toronto, and
    c) assess the uncertainty surrounding potential spatial access to health promotion programs using probabilistic sensitivity analysis.

1.4 Unique Research Contributions

This research applies the current state of spatial microsimulation practice to the field of public health program planning to demonstrate that this type of simulation modeling can generate realistic small area populations composed of individuals having varying chronic disease risk factor profiles. Using these synthetic populations, it is possible to forecast the future risk of chronic disease at the small

area level. These forecasts can identify areas having the greatest need for chronic disease prevention and management programs. Specifically, this research will

1. Develop and validate the TropISM spatial microsimulation model of type 2 diabetes and its risk factors. Model outputs will be used to forecast the five-year incidence of diabetes at the neighbourhood level. By including a diverse set of chronic disease risk factors, the TropISM model can be re-used and adapted to forecast the incidence of other chronic diseases.
2. Examine how the forecast incidence of diabetes changes under different weight loss scenarios. Subjecting the TropISM population to different intervention scenarios provides insight into the intensity of intervention required to produce population-level reductions in risk. Since intervention effects may vary over geographic space, it is also possible to assess potential inequities in intervention effectiveness.
3. Assess the uncertainty around the projected risk of diabetes using probabilistic sensitivity analysis. Formal sensitivity analysis of spatial microsimulation outputs is relatively rare in the literature. Therefore, this research will demonstrate that probabilistic sensitivity analysis should become a standard component of modelling policy scenarios using spatial microsimulation models.
4. Combine TropISM model outputs with a spatial interaction model to estimate how spatial accessibility to diabetes management and prevention programs varies over geographic space. By identifying areas of low accessibility among high-risk populations, it is possible to allocate additional resources to these areas in order to improve local health outcomes.

In summary, this research demonstrates that spatial microsimulation modelling provides public health practitioners with a relevant spatial decision support tool that can be used to inform the delivery of chronic disease prevention and management programs.

1.5 Thesis Organization

This thesis is divided among seven chapters. Chapter 2 reviews the epidemiology of type 2 diabetes, from global to local perspectives. It outlines important risk

factors that contribute to the risk of developing type 2 diabetes. This chapter also examines the prevention of type 2 diabetes in high risk populations and discusses the burden of type 2 diabetes in Ontario. Currently, little information is available describing the burden of type 2 diabetes at the local level and how local resources might address this burden.

In order to inform the development of the TropISM model, Chapter 3 reviews the field of spatial microsimulation. It describes the evolution of spatial microsimulation and key assumptions underlying the technique. New developments in the field are discussed and important limitations are highlighted. Chapter 4 develops the TropISM model for metropolitan Toronto using data from the 2006 Canadian Census and the 2005 Canadian Community Health Survey. The model is validated using population health outcome data obtained from the Institute for Clinical Evaluative Sciences. Strengths and limitations of the model are discussed along with implications for future research.

Chapter 5 uses the simulated TropISM population to forecast the five-year incidence of diabetes at the neighbourhood level using the Diabetes Population Risk Tool (DPoRT). It develops several scenarios that assess how forecast incidence changes as a function of individual-level weight loss within neighbourhoods. A baseline scenario is also modelled to forecast expected incidence assuming the simulated TropISM population maintains its current body weight. The baseline forecast incidence is compared against external incidence estimates obtained from the Ontario Diabetes Database maintained by the Institute for Clinical Evaluative Sciences. A formal probabilistic sensitivity analysis is conducted to assess forecast uncertainty and to assess which components of the DPoRT model contribute most to forecast uncertainty.

Chapter 6 uses simulated neighbourhood counts of type 2 diabetes and overweight (BMI $\geq 25\,\mathrm{kg/m^2}$) to conduct an exploratory analysis of potential spatial access to health promotion programs geared toward diabetes management and prevention. Using the two-step floating catchment area model, a type of spatial interaction model, this chapter examines how potential access varies across Toronto's neighbourhoods and where disparities might exist between neighbourhood-level prevalence and access. Probabilistic sensitivity analysis is then used to quantify the uncertainty around estimated access. The implications of varying access are

discussed in light of which neighbourhoods within Toronto have the greatest need for additional diabetes management and prevention services. Strengths and limitations of the results are discussed along with implications for the local delivery of these health promotion programs.

Chapter

2

The Epidemiology of Type 2 Diabetes

2.1 Background

Diabetes mellitus is a chronic metabolic disease characterized by high levels of circulating blood glucose. There are three main types: type 1, type 2, and gestational. Type 1 diabetes is an autoimmune disease in which the body's immune system destroys the β cells of the pancreas. Without these cells, the pancreas cannot produce insulin and the body's cells cannot absorb glucose leading to hyperglycemia, a state of chronic, elevated blood glucose (Cigolle, Blaum, & Halter, 2009; Public Health Agency of Canada, 2011). In type 2 diabetes, the body's cells become increasingly resistant to the effects of insulin resulting in a relative insulin deficiency. This also results in hyperglycemia. The third type, gestational diabetes, may develop during pregnancy but typically resolves following pregnancy (Public Health Agency of Canada, 2011).

Type 2 diabetes is the most prevalent type of diabetes, accounting for 90% to 95% of all cases (Caspersen, Thomas, Boseman, Beckles, & Albright, 2012; Millar & Young, 2003; National Diabetes Information Clearinghouse (NDIC), 2011). Globally, 366 million people had type 2 diabetes in 2011 (Whiting, Guariguata, Weil, & Shaw, 2011). Demographic projections that do not account for changing inci-

dence rates forecast that 552 million people, or 9.9% of the world's population, will be living with 2 diabetes by 2030 (Whiting et al., 2011). Projected prevalence rates vary by region, from a low of 5.9% among African nations to highs of 12.6% in North America and 14.3% among Middle Eastern countries (Whiting et al., 2011).

Left unmanaged, diabetes is a risk factor for other chronic diseases. Hyperglycemia wreaks havoc on blood vessels, causing atherosclerosis, retinopathy, and vision loss. It also causes long term kidney damage which can lead to end-stage renal disease (Caspersen et al., 2012; Klein, Saaddine, & Klein, 2011; Public Health Agency of Canada, 2011). Vascular damage reduces blood flow to the nervous system, damaging nerve tissue and causing pain and numbness in the extremities, foot ulceration, and bone damage (Boulten & Bowling, 2011; Public Health Agency of Canada, 2011). To help prevent these complications, type 2 diabetes must be managed effectively.

Management relies on maintaining blood glucose concentrations within normal ranges (fasting blood glucose concentrations of 4–7 mmol/L and postprandial blood glucose concentrations of 5–10 mmol/L as recommended by the Canadian Diabetes Association, 2008). Management strategies need to be tailored to the individual patient, but typically include dietary changes, increased physical activity, weight loss, and smoking cessation (Ripsin, Kang, & Urban, 2009). These lifestyle changes can delay the progression of type 2 diabetes and prevent long-term complications (Shamseddeen, Getty, Hamdallah, & Ali, 2011).

Incorporating all of these changes into daily routines is a challenging task for type 2 diabetics. To improve glycemic control, newly diagnosed diabetics need to learn how to monitor their blood glucose, make appropriate dietary changes, plan their meals, exercise more, and monitor any self-administered medications that have been prescribed. Effectively managing type 2 diabetes is difficult and often requires support from a multi-disciplinary team of health professionals, nutritionists, and diabetes educators (Armour, Norris, Jack, Zhang, & Fisher, 2005).

2.2 Risk Factors

Globally, the number of people diagnosed with type 2 diabetes has more than doubled since 1980 (Danaei et al., 2011). Much of this increase is attributable to de-

mographic factors, such as population growth and aging. The prevalence of overweight and obesity also increased during this time period (Finucane et al., 2011). This is cause for concern, since excess body fat is one of the strongest predictors of type 2 diabetes (Shamseddeen et al., 2011). Although excess body fat is the result of energy imbalance, where energy consumed from food exceeds that expended through exercise and activities of daily living, the modern environment plays a role. In particular, it has become too easy to consume energy-dense but nutrient poor foods (Caballero, 2007; Seidell, 2000). Moreover, large segments of the population are sedentary and fail to engage in regular physical activity. Thus, modern, urban environments are thought to promote obesity through societal factors and policies that encourage weight gain (Egger & Swinburn, 1997). In short, the risk factors for type 2 diabetes are many and include modifiable risk factors such as lifestyle choices and behaviours, non-modifiable risk factors, such as age, ethnicity, and genetics. The risk of developing diabetes is further influenced by social determinants of disease including income and education. Often, the most disadvantaged segments of society face the greatest risks of developing type 2 diabetes (Raphael et al., 2003).

2.2.1 Modifiable risk factors

One of the most important factors influencing the risk of developing type 2 diabetes is excess body fat (Shamseddeen et al., 2011): 65% to 80% of all cases of type 2 diabetes can be attributed to obesity (Costacou & Mayer-Davis, 2003; Paulweber et al., 2010; Seidell, 2000). Overweight and obesity increase insulin secretion and lead to insulin resistance (Burr, Shephard, & Riddell, 2012; Kahn, 2003; Srikanth & Deedwania, 2011). Weight loss, on the other hand, reduces the risk of developing diabetes. Diet and physical activity influence body weight by controlling energy balance (Caballero, 2007; van Dam, 2003).

High fat diets, especially diets high in saturated fats, are associated with decreased insulin sensitivity, increased insulin secretion, glucose intolerance, and hyperglycemia (Costacou & Mayer-Davis, 2003; Feskens et al., 1995; Hu, van Dam, & Liu, 2001). The body's physiologic response to certain foods also affects the risk of diabetes. Some foods increase blood glucose concentrations more than others;

such foods have a higher glycemic index. High glycemic index diets chronically increase blood glucose concentrations which can lead to glucose intolerance (Barclay et al., 2008; Costacou & Mayer-Davis, 2003). For example, white rice has an especially high glycemic index. Hu, Pan, Malik, and Sun (2012) estimated that each additional serving in daily rice consumption increased the risk of diabetes by 11%.

In a similar vein, sugar-sweetened beverages (e.g., soft drinks and energy drinks such as "Red Bull") have a high sugar content and almost no nutritional value. Not only do these drinks have a high glycemic index, they readily contribute to weight gain because of their high sugar content and their inability to generate satiety. As a result, it is difficult for people to reduce their caloric intakes during subsequent meals (Malik, Popkin, Bray, Després, & Hu, 2010). One study of more than 50,000 nurses found that nurses who increased their consumption of sweetened sugar beverages over an eight year period gained, on average, 8 kg of additional body weight (Schulze et al., 2004).

Physical activity plays an important role in preventing the onset of type 2 diabetes while a sedentary lifestyle increases the risk. Physical activity maintains a balance between total energy intake and energy expenditure. A positive energy balance results in weight gain over time while a negative energy balance results in weight loss (van Dam, 2003). By preventing weight gain, physical activity (a) prevents the onset of diabetes and (b) slows its progression following diagnosis (Bassuk & Manson, 2005). Observational studies examining the relationship between physical activity and incident type 2 diabetes have found significantly lower risks of developing diabetes among physically active men and women. In particular, moderately active women have 25% to 34% lower risk of developing diabetes compared to sedentary women (Folsom, Kushi, & Hong, 2000; Hu et al., 1999; Weinstein et al., 2004). Men who are moderately active for 40 minutes per week are 56% less likely to develop diabetes compared to less active men (Lynch et al., 1996). Conversely, sedentary lifestyles increase the risk of developing diabetes. In a British study of government employees, women who engaged in less then 1.5 hours per week of moderate to vigorous physical activity had a 71% greater risk of diabetes compared to more active women. In the same study, inactive men had a 52% greater risk of diabetes compared to active men (Kumari, Head, & Marmot,

2004).

Randomized controlled trials of physical activity interventions support these findings, noting that prediabetic individuals randomized to a regular exercise intervention program were significantly less likely to convert to diabetes compared to individuals randomized to control groups (Knowler et al., 2002; Pan et al., 1997; Tuomilehto et al., 2001). One study noted that the risk of diabetes in the intervention group was significantly lower even in the *absence* of weight loss. Thus, population interventions need to encourage high-risk adults to increase their levels of physical activity in order to reduce the incidence of type 2 diabetes.

Although diet and physical activity are modifiable risk factors, these behaviours are socially constructed and are influenced by culture and public policies. For example, energy dense foods are readily accessible in many urban settings. Urbanization, along with modern technology and transportation, have increased the number of people in society who are, by and large, sedentary (Caballero, 2007; Seidell, 2000). Modern environments have become "obesogenic", readily promoting weight gain and chronic diseases caused by overweight and obesity. Obesity, therefore, could be viewed as a normal physiologic response to an abnormal environment (Egger & Swinburn, 1997).

This shift in thinking acknowledges the complex interplay of societal factors and public policies that promote weight gain or encourage weight loss. On a population level, preventing type 2 diabetes requires preventing obesity, which in turn, requires structural changes to modern urban environments. Public policies must not only encourage healthy lifestyles, but must also promote healthy environments that encourage active transportation (e.g., walking to work), discourage car use, promote access to fresh foods, and discourage access to energy dense foods (e.g., fast foods). Multi-sectoral cooperation is therefore essential to obesity prevention at the population level and requires input from public health, urban planning, economics, and tax policy (Caballero, 2007; Egger & Swinburn, 1997; James, 1997).

2.2.2 Non-modifiable risk factors

Of the different non-modifiable risk factors for type 2 diabetes, age is one of the most important (Paulweber et al., 2010). The prevalence of type 2 diabetes increases with age; in European countries, less than 10% of people younger than 60 have type 2 diabetes while more than 20% of people older than 80 have type 2 diabetes (Paulweber et al., 2010). Similar trends are seen in Canada, where the prevalence of diabetes increases steadily with age. In 2008–2009, 2.6% of Canadians aged 35–39 had diabetes compared to 16.6% of Canadians aged 60–64 and 25% of Canadians aged 70 and older (Public Health Agency of Canada, 2011).

The incidence of diabetes also increases with age. One Norwegian study followed 26,168 people from 1995 to 2005 and documented increased incidence of diabetes among older participants. While incident cases of diabetes were negligible among 25–29 year olds, incidence rates peaked among men aged 60–69 (7.5 cases/1000). Incidence peaked later in women. For each 10 year increment in age, the risk of diabetes increased by 67% in men and 36% in women (Joseph, Svartberg, Njølstad, & Schrimer, 2010). Similar increases in the incidence are seen among Canadians. Incidence rates among Canadians younger than 35 in 2008–2009 were $\leq$ 2.3/1000. Among 40–44 year olds, incidence rates were 5.1/1000; among 70–74 year olds, incidence rates increased to 20.3/1000. In all age groups, the incidence of diabetes is higher among Canadian men than it is among Canadian women (Public Health Agency of Canada, 2011).

Ethnicity is another non-modifiable risk factor associated with type 2 diabetes, although the role that it plays in the disease process remains unclear (Carulli et al., 2005; Davis, 2008; Paulweber et al., 2010). What is clear is that some ethnic groups have a greater risk of developing diabetes compared to others, likely due to genetic differences between these groups (Paulweber et al., 2010; Public Health Agency of Canada, 2011). Specifically, people of European descent are less likely to develop diabetes compared to people of African, Asian, and Hispanic American descent. Moreover, modifiable risk factors tend to cluster among some ethnic groups, which may contribute to the increased risk of disease (Public Health Agency of Canada, 2011).

2.2.3 Social determinants of type 2 diabetes

Socioeconomic factors also influence the risk of developing type 2 diabetes. Disadvantaged populations, including the poor, under-educated, and certain ethnic groups, are disproportionately affected by type 2 diabetes (Raphael et al., 2003). Diabetes is more prevalent among people having lower household incomes (Hux & Tang, 2003; James, Young, Mustard, & Blanchard, 1997; Wilkins, Berthelot, & Ng, 2002). Poverty and deprivation play an important role in the development of type 2 diabetes. These social determinants influence health behaviours by limiting

- educational opportunities which might affect one's ability to make health promoting decisions,
- access to healthy foods,
- opportunities to participate in leisure time physical activity, and
- access to health services, which might delay diagnosis (Agardh, Allebeck, Hallqvist, Moradi, & Sidorchuk, 2011; Raphael et al., 2003).

One meta-analysis of 15 cohort and 8 case-control studies estimated the risk of developing diabetes as a function of socioeconomic class, defined using education, occupation, and income (Agardh et al., 2011). Overall, the least educated groups had a 41% greater risk of developing diabetes. Low social class, measured by occupation, was associated with a 31% increased risk of diabetes while low income was associated with a 40% increased risk compared to the highest socioeconomic groups (Agardh et al., 2011). Likewise, disadvantaged groups face greater difficulties managing their disease (Raphael et al., 2003). As a result, it is important to address social inequalities to prevent diabetes among disadvantaged groups and empower them to better manage their disease, preventing future complications.

2.3 Prevention of Type 2 Diabetes

2.3.1 Randomized controlled trials of diabetes prevention programs

Effective diabetes prevention programs encourage lifestyle changes that promote weight loss. The goal of these programs is to prevent high-risk individuals from developing type 2 diabetes. To promote weight loss, high-risk individuals must

increase their levels of physical activity and reduce their intake of carbohydrates and fats (Albright & Gregg, 2013; Ramachandran & Snehalatha, 2011).

Randomized controlled trials have demonstrated that type 2 diabetes can be prevented in high-risk populations (Albright & Gregg, 2013), typically defined on the basis of impaired glucose tolerance and the presence of one or more risk factors, such as overweight/obesity. Five randomized controlled trials conducted in different populations (China, Japan, Finland, the United States, and India) have demonstrated similar effects of intensive lifestyle modification on preventing type 2 diabetes (Knowler et al., 2002; Kosaka, Noda, & Kuzuya, 2005; Pan et al., 1997; Ramachandran et al., 2006; Tuomilehto et al., 2001). Most of these trials enrolled around 500 participants, with the exception of an American trial that enrolled 3234 participants (Knowler et al., 2002).

Each of these trials shared common features, from the individuals enrolled in the trials to the lifestyle interventions they employed. First, these trials enrolled high-risk individuals, defined as individuals who showed signs of impaired glucose tolerance (blood glucose concentrations between 7.8–11.0 mmol/L). Four of the five trials also used body mass index to either stratify individuals into risk groups (Kosaka et al., 2005; Pan et al., 1997) or as an inclusion criterion, where only overweight individuals having a BMI > 24 kg/m^2 were included (Knowler et al., 2002; Tuomilehto et al., 2001). All but one trial (Kosaka et al., 2005) included both men and women while two trials enrolled participants from specific age groups. Across all trials, the average participant was between 45 to 55 years old.[1]

In each trial, diet and/or exercise formed the basis of the lifestyle intervention. Participants randomized to the control condition received advice about strategies they could use to change their diets, lose weight, and increase their physical activity in order to control their blood glucose. The lifestyle interventions shared common features, including:

1. Reduced calorie diets where only 25–30% of daily energy came from dietary fat (Knowler et al., 2002; Pan et al., 1997; Ramachandran et al., 2006; Tuomilehto et al., 2001),

[1] Tuomilehto et al. (2001) enrolled participants aged 40–65 while Ramachandran et al. (2006) enrolled participants aged 35–55.

2. Increased exercise, defined as moderate levels of physical activity lasting for at least 30 minutes per day (Knowler et al., 2002; Pan et al., 1997; Ramachandran et al., 2006; Tuomilehto et al., 2001),
3. Weight loss, defined as a loss of 5–7% of baseline body weight (Knowler et al., 2002; Tuomilehto et al., 2001) or a loss of 0.5–1.0 kg/month (Kosaka et al., 2005), *or*
4. Intensive one-on-one counselling provided by physicians or other specialists (Knowler et al., 2002; Kosaka et al., 2005; Pan et al., 1997; Tuomilehto et al., 2001). Counselling was designed to help participants achieve the intervention goals of making dietary changes, increasing exercise, and losing weight. Ramachandran et al. (2006) provided monthly telephone counselling to ensure adherence to the lifestyle intervention.

In addition to these features, two of the trials included a pharmacological intervention where metformin was used either alone (Knowler et al., 2002) or in combination with lifestyle modification (Ramachandran et al., 2006). In both trials, the metformin intervention was added to test whether it could prevent the onset of type 2 diabetes in high-risk individuals.

In all of these trials, the development of type 2 diabetes was the primary outcome measure, defined as having a blood glucose concentration $\geq$ 11.1 mmol/L two hours after an oral glucose tolerance test on two successive occasions. The follow-up period ranged from three to six years. Intervention adherence was assessed and compared across intervention and control groups. In all trials, participants in the intervention groups consumed fewer calories, lost more weight, and exercised more than participants in the control group.

Each of these trials demonstrated that intensive lifestyle intervention significantly reduced the risk of type 2 diabetes in high-risk individuals: risk reductions ranged from 31–68%. Trials that included a lifestyle intervention and a metformin intervention found that metformin was *not* more effective than intensive lifestyle modification, either alone or when it was combined with lifestyle intervention (Knowler et al., 2002; Ramachandran et al., 2006). Risk reductions were stable across different ethnic groups and weight loss was an important component of intervention success (Davies, Tringham, Troughton, & Khunti, 2004). Hamman, Wing, Edelstein, Lachin, and Delahanty (2006) demonstrated that every kilogram

of body weight lost due to lifestyle modification produced a 16% reduction in diabetes incidence.

2.3.2 Community-based prevention initiatives

Although randomized controlled trials have demonstrated the efficacy of lifestyle interventions in preventing type 2 diabetes among high risk individuals, such trials are delivered under ideal experimental conditions. For example, the US Diabetes Prevention Program used one-on-one counselling sessions to help participants in the intervention condition achieve physical activity, nutrition, and weight loss goals (Knowler et al., 2002). Such intensive interventions are difficult to deliver in real-world settings (Ali, Echouffo-Tcheugui, & Williamson, 2012; Glasgow, Lichtenstein, & Marcus, 2003). Moreover, participants in randomized controlled trials may not be representative of the broader population of high-risk individuals. For example, in the Diabetes Prevention Program (Knowler et al., 2002), a greater proportion of participants in the lifestyle intervention were women, Caucasian, and younger than the average high-risk individual. These limitations may prevent generalization of trial results. Therefore, it is important to study the effectiveness of lifestyle interventions in community settings to ensure that results can be replicated in real-world settings.

Studies that implemented lifestyle interventions in community settings have achieved encouraging results. A meta-analysis of 28 translation trials implementing interventions based on the Diabetes Prevention Program showed significant weight loss among high-risk individuals after participating in the intervention for at least one year (Ali et al., 2012). High-risk individuals included people with prediabetes or people having a body mass index ≥ 25 kg/m^2 and who had at least one other risk factor. On average, high-risk individuals participating in these studies lost four percent of their baseline body weight after 12 months (Ali et al., 2012).

Twelve of the 28 interventions were delivered in community settings, such as community centres, recreation centres, and faith-based organizations. Another eleven were conducted in health care facilities while four used electronic media to deliver lifestyle training to participants. Consistent results were observed regardless of the type of health professional delivering the program. Weight loss was

slightly greater in studies using medical or other health professionals compared to trials using lay community members to deliver the intervention, although the difference was not statistically significant (Ali et al., 2012).

The number of intervention sessions included as part of the lifestyle modification program was also important. On average, programs that delivered a greater number of sessions produced greater weight loss. Although the prevention of diabetes was not studied as an outcome, Ali et al. (2012) argue that because weight loss is the most important predictor of diabetes prevention, weight loss is a suitable outcome for studies examining the real-world effectiveness of diabetes prevention programs.

That said, one of the studies included in this meta-analysis demonstrated significant reductions in blood glucose concentrations. Katula et al. (2011) examined the effectiveness of volunteer community health workers delivering a weight loss program based on the Diabetes Prevention Program. In this study, 301 high-risk individuals[2] were randomized to lifestyle intervention or usual care. Lifestyle intervention consisted of group education sessions delivered on a weekly basis for six months followed by monthly sessions for another six months. Usual care consisted of two diet counselling sessions with a nutritionist followed by monthly newsletters discussing healthy lifestyles. After 12 months, those in the lifestyle intervention group lost an average of 5.7 pounds of body weight compared to participants in the control condition. Fasting blood glucose concentrations were also significantly lower in the intervention group compared to the control group, with the intervention group showing a 0.21 mmol/L reduction. Based on these findings, it appears that even in real-world settings it is possible for high-risk individuals to lose a clinically meaningful amount of body weight that could prevent diabetes.

Community-based interventions typically cost less than the interventions delivered as part of a randomized controlled trial (Simmons, Unwin, & Griffin, 2010). However, community-based trials often face difficulties when it comes to targeting and enrolling the most vulnerable groups who face the highest risks of developing diabetes. Of the 28 trials included in the meta-analysis conducted by Ali et al. (2012), most participants were female and non-Hispanic white, whereas type 2 di-

[2] Defined as individuals having prediabetes and a body mass index between 25 and 40 kg/m^2

abetes tends to be more prevalent among men and minority ethnic groups (Public Health Agency of Canada, 2011). Socially and economically disadvantaged individuals have a greater risk of developing type 2 diabetes and it may be more difficult to target these individuals.

In short, it is important to understand the factors that contribute to the success of community-based intervention programs (Jackson, 2009). Translation studies of the Diabetes Prevention Program that used local YMCAs to deliver lifestyle intervention programs have demonstrated promising results (Ackermann, Finch, Brizendine, Zhou, & Marrero, 2008; Ackermann & Marrero, 2007; Vojta, Koehler, Longjohn, Lever, & Caputo, 2013). As Vojta et al. (2013) argue, YMCA centres across the United States are promising avenues for delivery of diabetes prevention programs because almost 60% of the US population lives within 5 kilometres of a YMCA. Programs delivered by YMCA staff would cost less than programs delivered by health care professionals (e.g., physicians, nurses, physiotherapists, or nutritionists/dietitians). In addition, YMCAs typically have a high penetration rate in many communities. These community centres also do not turn people away because of an inability to pay for services (Ackermann & Marrero, 2007). This is important for lifestyle modification programs that target chronic diseases because disadvantaged, low-income groups are often disproportionately affected by chronic diseases (Ackermann & Marrero, 2007; Raphael et al., 2003).

Attendance is another factor that influences the effectiveness of community-based prevention programs; low attendance limits their effectiveness (Ali et al., 2012). Lifestyle change and weight loss are not possible if people do not actively participate in the intervention. Another factor that influences adherence is how well patients monitor their participation. Self-monitoring helps participants track caloric intakes and the amount of physical activity they need to engage in to lose weight.

In summary, community-based prevention programs can produce beneficial results in high-risk populations. Although these programs may not achieve the same results of the randomized controlled trials upon which they were based, extant evidence demonstrates that community-based diabetes prevention is possible if high-risk groups lose meaningful amounts of body weight.

2.4 Burden of Disease

The prevalence of diabetes increased in many regions around the world since 1980 (Danaei et al., 2011). In 2011, 366 million people had diabetes, a number that is expected to grow to 552 million by 2030 (Whiting et al., 2011). Diabetes is most prevalent in Western Pacific, North American, and Mediterranean countries, where 10.1%, 11.1%, and 12.5% of adults aged 20–79 had diabetes in 2011, respectively. By 2030, prevalence in these regions is expected to increase to 11.6%, 12.6%, and 14.3%, respectively (Whiting et al., 2011).

Canadian data reveal similar trends: since the late 1990s, prevalence increased in all age groups. The largest increase occurred among adults aged 75–79, where prevalence increased from 14.0% in 1998 to 25.5% in 2008 (Public Health Agency of Canada, 2011). In 2011, 2.7 million adults, or 10.8% of the Canadian population aged 20–79, were living with diabetes. This number is expected to increase to 3.7 million by 2030 (Whiting et al., 2011). Prevalence also varies regionally within Canada. In 2008, some of the highest rates of disease[3] were observed in Newfoundland & Labrador (6.5%), Nova Scotia (6.1%), and Ontario (6.0%), while Alberta (4.9%) and Nunavut (4.4%) experienced the lowest rates of disease (Public Health Agency of Canada, 2011). However, these estimates need to be interpreted cautiously, since 20% to 40% of all of cases of diabetes remain undiagnosed (Canadian Diabetes Association, 2015; Hux & Tang, 2003; Public Health Agency of Canada, 2011; Rosella, Lebenbaum, Fitzpatrick, Zuk, & Booth, 2015).

Excess body fat, in particular overweight (BMI $\geq$ 25 kg/m^2) and obesity (BMI $\geq$ 30 kg/m^2), is one of the strongest predictors of type 2 diabetes, accounting for as much as 80% of all cases (Costacou & Mayer-Davis, 2003; Paulweber et al., 2010; Seidell, 2000). In Canada in 2010, diabetes was three times more prevalent among overweight and obese individuals compared to normal weight individuals (Public Health Agency of Canada, 2011). Like diabetes, average body weight has increased on a global scale since 1980 (Finucane et al., 2011). Among high income countries, the United States experienced the largest increase in average body weight. Among US men, average BMI increased from 25.5 kg/m^2 in 1980 to 28.5 kg/m^2

[3] Age-standardized estimates from the Canadian Chronic Disease Surveillance System, all age groups.

in 2008; in US women, average BMI increased from 25.0 kg/m^2 in 1980 to 28.3 kg/m^2 in 2008. A similar trend was observed in Canada: by 2008, average BMI was 27.5 kg/m^2 among men and 26.7 kg/m^2 among women. Significant weight gain also occurred in Australasian, South American, North African, and Middle Eastern countries (Finucane et al., 2011).

The global increase in average BMI suggests that obesity became more common during this time. The age-standardized prevalence of obesity almost doubled globally from 1980 to 2008, increasing from 4.8% to 9.8% in men and from 7.9% to 13.8% in women (Finucane et al., 2011). Obesity was most prevalent among North American men in 2008, where just under 30% of men were obese. Obesity was least prevalent in south Asia (Finucane et al., 2011).

Because obesity accounts for a large proportion of all cases of type 2 diabetes, it is reasonable to expect that the incidence of disease has increased over time. Statistics from high income countries support this claim. In Denmark, the incidence of diabetes increased by 5.3% per year from 1995 to 2004 (Carstensen, Kristensen, Ottosen, & Borch-Johnsen, 2008). In the United Kingdom, incidence rates increased from 1.7/1000 in 1991 to 4.5/1000 in 2002. By 2010, incidence had risen to 5.2/1000 (Holden et al., 2013).

In the United States, the incidence of diabetes increased in all age groups from 1990 to 2010 (Centers for Disease Control and Prevention, 2007). The largest increases were observed in people aged 45 and older, where incidence increased from approximately 6 cases/1000 in 1990 to 13 cases/1000 in 2010 (Centers for Disease Control and Prevention, 2007). Incidence rates vary geographically as well: from 2004–2009, incidence was highest in the American South (e.g., Louisiana, Mississippi) and lowest in the American Midwest (e.g., Wisconsin, Iowa; Centers for Disease Control and Prevnention, 2013). In Canada, however, diabetes incidence remained relatively stable between 1998 and 2009, fluctuating between 5.6 cases/1000 and 6.1 cases/1000 (Public Health Agency of Canada, 2011). Small increases were observed in Ontario, Saskatchewan, British Columbia, and the Northwest Territories.

Even in the absence of increasing incidence, type 2 diabetes places a large burden on society in terms of health service utilization and health care spending. Considering that treatment is costly (Lipscombe & Hux, 2007), that the popula-

tions of many nations around the world are aging, and that type 2 diabetes is more prevalent among older age groups, it is reasonable to expect that health service utilization for diabetes and health care costs will increase in the future. Indeed, national spending on diabetes is a function of a country's age structure: societies that have a larger proportion of older people spend more on diabetes health care, since costs tend to be higher among older people (Zhang et al., 2010).

In 2010, $376 billion to $672 billion (USD) were spent on diabetes, accounting for 12% of the world's health expenditures (Zhang et al., 2010). Eighty percent of nations spend 5%–13% of their national health budgets on diabetes. High-income countries spend proportionately more on diabetes related health care than low- and middle-income countries, even though the latter group tends to have more people living with diabetes. Per capita spending on diabetes is highest in the United States and lowest in India.

A large part of diabetes related health care spending is attributable to the long-term complications that develop from poorly managed diabetes. In Canada, 5500 deaths per year are directly attributable to diabetes; this number rises to 25,000 deaths per year when complications from diabetes are included (e.g., cardiovascular disease, stroke, and end-stage renal disease; O'Brien, Patrick, & Caro, 2003). In the year 2000, 21% of Canadian diabetics had heart disease while 25% of cardiac surgeries were attributable to diabetes. Diabetes also accounted for one-third of all new cases of end-stage renal disease and half of all non-traumatic cases of lower extremity amputations (O'Brien et al., 2003). A sizable proportion of diabetes related costs are due to complications arising from the disease. This demonstrates the importance of prevention, which includes better disease management for people diagnosed with type 2 diabetes and, more importantly, *preventing* high-risk individuals from developing type 2 diabetes in the first place.

2.5 The Local Burden of Diabetes in Ontario

Provincially, Ontario had the third highest prevalence of diabetes in Canada in 2008–2009 (Public Health Agency of Canada, 2011). From 1995–2005, the overall age-sex adjusted prevalence of diabetes increased from 4.9% to 8.9%.[4] The largest

[4]Standardized to the 2001 Ontario population.

increase in prevalence occurred among men aged 50 and older, increasing by almost seven percentage points from 11.8% in 1995 to 19.2% in 2005. Prevalence among women the same age increased from 9.6% to 15.4%. From 1997–2003, incidence rates in this age group increased from 13.5/1000 to 15.9/1000 in men and from 10.7/1000 to 12.7/1000 among women (Lipscombe & Hux, 2007).

Mortality from diabetes decreased during this period. Overall, age-sex adjusted mortality rates decreased from 17.6/1000 in 1995 to 13.3/1000 in 2005. Decreases were seen in all age groups (Lipscombe & Hux, 2007). More recent data spanning a 14-year period from 1996–2010 demonstrate similar declines in mortality (Lind et al., 2013). Based on these data, it is reasonable to conclude that increases in prevalence were due to increasing incidence *and* decreasing mortality rates (Lipscombe & Hux, 2007). In other words, more people developed diabetes during this time period and those diagnosed with diabetes seemed to live longer with the disease. Lind et al. (2013) suggest that declines in mortality might be the result of more aggressive treatment, improving overall survival rates. However, they also note that in 2009 there were more diabetics living with the disease for a shorter time period than in 1995 as a result of improved screening among high-risk groups. Regardless, trend data clearly indicate that both the prevalence and incidence of diabetes have increased since the 1990s.

Substantial regional variation in diabetes burden exists within Ontario. As might be expected, areas containing a higher proportion of ethnic minorities have higher diabetes prevalence rates. In 1999, diabetes prevalence was higher than the provincial average of 6.2%[5] in most regions of Northern Ontario, Toronto, Peel Regional Municipality, and Essex, Brant, and Elgin Counties in Southwestern Ontario (Hux & Tang, 2003). The high prevalence rates in the northern areas of Ontario are cause for concern, especially because these regions tend to be more rural and have less access to health services.

Geographic variation in diabetes burden is also apparent at the neighbourhood level. Using data from the Ontario Diabetes Database, the Institute for Clinical Evaluative Sciences examined neighbourhood-level differences in the prevalence of diabetes across the City of Toronto (Booth, Creatore, Gozdyra, & Glazier,

[5] Age-sex adjusted.

2007). In this study, 140 neighbourhoods were defined using adjacent census tracts having similar socioeconomic characteristics. Each neighbourhood was comprised of 7000–20,000 residents and was contained by existing natural (e.g., rivers) and man-made (e.g., streets) boundaries (City of Toronto, 2013a; Creatore, Gozdyra, Booth, & Glazier, 2007a).

In 2001–2002, the overall age-sex adjusted prevalence of diabetes[6] in Toronto was 5.5% (Booth et al., 2007). Neighbourhoods in northwest Etobicoke and northeast Scarborough had the highest prevalence of diabetes (> 6.6%) while neighbourhoods in the central core along Yonge Street and in the southwest, between Bloor Street and the Gardiner Expressway, had the lowest prevalence of disease (< 4.0%; Booth et al., 2007). These spatial trends have persisted over time. By 2012, the city-wide average prevalence of diabetes may have reached 11.8%, varying form a low of 6% in the central core along Yonge Street to more than 15% in northwest Etobicoke and northeast Scarborough (Toronto Community Health Profiles Partnership, 2015).

The social determinants of diabetes also varied across Toronto. For example, unemployment rates were lowest in the central core and neighbourhoods in this area had average annual household incomes of more than $100,000 (Creatore, Gozdyra, Booth, Ross, & Glazier, 2007b). Proximity to parks and public and private recreation centres showed significant spatial heterogeneity as did the number of family physicians, diabetes specialists[7], and diabetes education programs.

An additional analysis was conducted to examine how diabetes prevalence covaried with these social determinants. Using local indicators of spatial association, Creatore et al. (2007b) demonstrated that neighbourhoods having high diabetes prevalence tended to have low average annual household incomes, high unemployment rates, and a high percentage of residents having less than a high school education. On the other hand, neighbourhoods having low diabetes prevalence rates tended to show the opposite relationship. These neighbourhoods had high annual household incomes, low unemployment rates, and a high percentage of residents having at least a secondary school education (Creatore et al., 2007b).

Less intuitive relationships were found between diabetes prevalence and prox-

[6] Type 1 and type 2 diabetes.
[7] Ophthalmologists, endocrinologists and optometrists

imity to parks and recreation spaces. Neighbourhoods where diabetes was higher than the city average had both a low and a high number of parks and recreation centres per 10,000 population. Neighbourhoods west of Yonge Street had low diabetes prevalence and a high number of parks and recreation centres while areas east of Yonge Street had low diabetes prevalence and a low number of parks and recreation centres per 10,000. From a health planning perspective, these relationships suggest that some neighbourhoods have less than optimal access to neighbourhood resources that might promote physical activity (Creatore et al., 2007c). Such neighbourhoods might be candidates for prevention programs designed to promote physical activity.

Access to health services was significantly associated with diabetes prevalence. Neighbourhoods having high prevalence tended to have (a) fewer family physicians per 10,000 residents, (b) fewer diabetes education programs per 10,000 residents, (c) longer travel times to the nearest family physician, either by private car or public transit, and (d) longer travel times to the nearest diabetes education program, either by private car or public transit (Glazier et al., 2007). These results suggest that geographic access to health care resources important for managing diabetes is poorest in areas where they are most needed.

2.6 The Ontario Diabetes Strategy

In 2008, the Ontario Ministry of Health and Long Term Care established the Ontario Diabetes Strategy to improve access to health services for diabetics and to help them better manage their disease. At this time, $741 million were earmarked for the Strategy and an additional $152 million was committed to the Strategy in 2012 to extend the program until 2016 (Silversides, Doig, & Sullivan, 2013). Funds were to be used to

1. educate high-risk groups about diabetes and ways to prevent it,
2. expand diabetics' access to "Diabetes Education Teams" (consisting of a registered nurse and a dietitian) that would work with family physicians to help patients better manage their disease, and
3. improve local and regional coordination of diabetes related health services through Diabetes Regional Coordination Centres. These Centres were de-

signed to identify and fill gaps in local service provision (Booth et al., 2012; Cook, 2011; Ministry of Health and Long Term Care, 2012).

Although the Strategy has preventive components, the bulk of the Strategy focuses on service provision. Indeed, one of the Strategy's objectives is to improve local coordination of health services for diabetics, suggesting geographic disparities exist in access to services. However, the 2012 Ontario Auditor's General report suggests that the geographical allocation of diabetes education teams throughout Ontario is inefficient. The Auditor's report found geographic overlap in service provision, overlap that resulted from duplication in service provision between Diabetes Education Teams, hospitals, and other health care providers. This led to an under-utilization of 90% of diabetes education programs (Office of the Auditor General of Ontario, 2012). The Auditor's General report concludes that the Ontario Diabetes Strategy has not lived up to expectations.

Furthermore, the Ministry of Health and Long-Term Care stipulates that each Diabetes Education Team must maintain an active caseload of 1000 patients per fiscal year. In 2008–2009, 90% of Diabetes Education Teams did not meet this benchmark while 33% had fewer than 500 active patients (Office of the Auditor General of Ontario, 2012). When asked why they could not meet the caseload, many Diabetes Education Teams stated that programs were located too close in proximity to other programs, leading to under-utilization and competition for patients. One under-utilized program was located in a rural area that had four other diabetes education programs covering the same catchment area (Office of the Auditor General of Ontario, 2012).

The Auditor's report also noted that only 3% of the original funding, or $19 million, was allocated to prevention. When Diabetes Education Programs were surveyed about prevention, two-thirds felt that more resources should be devoted to educating high-risk individuals. Under the mandate of the current Strategy, this type of preventive education is not possible because the funding for the strategy is intended to assist Ontarians *diagnosed* with diabetes (Office of the Auditor General of Ontario, 2012). Based on these findings, it appears that diabetes education services are less than optimally distributed at the local level in Ontario and that insufficient resources have been devoted to prevention. As Barnett, Pearce, and Howes (2006) note, problems of access to services for diabetics have largely been

ignored by geographers.

From a health promotion standpoint, if prevention initiatives reduce the incidence of type 2 diabetes in high-risk individuals, then it is important that such programs be geographically accessible. Prevention initiatives delivered in a community setting should consider the level of risk in the target population and its geographic distribution. This information should inform service delivery decisions so that prevention programs can be located in areas of greatest need. As Barnett et al. (2006) note, many community-based diabetes education programs do not reach high-risk populations. However, if prevention programs effectively reach high-risk groups *and* produce their desired effects, then it is important to assess how the geographic profile of risk might change at the small area level. This is an area of research where spatial microsimulation modelling can inform the planning and delivery of health promotion programs at local levels.

Chapter

3

Spatial Microsimulation Modelling

3.1 The Evolution of Spatial Microsimulation Modelling

As the field of microsimulation evolved, it spread from economics to other disciplines, including geography (Birkin & Clarke, 2011). The primary goal of geographic microsimulation is to simulate spatially disaggregated microdata that resemble reliable external data sources, such as census data, as closely as possible (Ballas et al., 2007a, 2005a). Spatial microsimulation modelling is driven by the need for geographically detailed microdata which can be used for ex ante policy evaluation (Ballas et al., 2007a; Ballas, Kingston, Stillwell, & Jin, 2007b; Ballas et al., 2005a).

Early spatial microsimulation models used population reconstruction methods to build synthetic populations on a case-by-case basis (Birkin & Clarke, 1988; Harland, Heppenstall, Smith, & Birkin, 2012; Williamson, 1996). Under this approach, synthetic microdata are simulated to match the known characteristics of local geographic units, such as Canadian census dissemination areas. For any given spatial unit, microdata are simulated using Monte Carlo sampling methods to randomly assign attributes to "individuals" in the synthetic population according to known distributions. When these synthetic data are aggregated over

the spatial units, the aggregate distributions resemble the published census distributions on simulated characteristics. Thus, if census data indicate that 15% of the population in a given geographic area has a post-secondary education, then a similar estimate should be obtained using the synthetic microdata. Although early spatial microsimulation models relied on Monte Carlo population reconstruction methods, the technique has been supplanted by "reweighting" methods. Essentially, reweighting methods link sample survey data to census data. Moreover, these methods are generally more robust than population reconstruction techniques (Harland et al., 2012; Huang & Williamson, 2001; Ryan, Maoh, & Kanaroglou, 2009).

3.2 Spatial Microsimulation via Reweighting

All reweighting methods have a common aim: to simulate spatially disaggregated microdata by linking geographically poor but information rich survey data to geographically rich but information poor census data (Birkin & Clarke, 2011; Chin et al., 2005). As such, reweighting methods can be thought of as a type of data imputation strategy, where information from a sample survey is added to census data (Haslett, Jones, Noble, & Ballas, 2010). This is done by calibrating the sampling weights for observations from a survey dataset so that they represent the small areas of interest instead of the original target population. Once the reweighting procedure has been completed for one small area, the survey sample is reused and calibrated to other small areas. The end result is a spatially disaggregated micro dataset containing the wealth of information available from the survey dataset for all small areas. Policy simulations can then be conducted using this dataset and results can be analyzed spatially (Tanton, Vidyattama, Nepal, & McNamara, 2011). Three types of reweighting methods are discussed in the next sections: iterative proportional fitting (IPF), generalized regression reweighting (GREGWT), and simulated annealing, a probabilistic, combinatorial optimization algorithm.

3.2.1 Iterative proportional fitting

Iterative proportional fitting, also referred to as deterministic reweighting, is one method used to generate synthetic, spatially disaggregated microdata. Iterative

proportional fitting calibrates a set of sampling weights from a micro dataset so that the sum of the new weights equals known totals over a set of demographic variables, or constraints, within a given small area. For example, if 500 people live in a small area and its age-sex distribution is known from census data, then iterative proportional fitting iteratively scales the sampling weights from the survey microdata so that sum of the calibrated weights in each age-sex category equals the census total in that same category. These adjusted weights can then be thought of as the probability that a surveyed respondent "lives" in the small area being fit (Ballas, 2004; Ballas et al., 2005a). The IPF method is deterministic in the sense that it produces identical results each time the algorithm is run using the same constraints (Ballas et al., 2007a; Haslett et al., 2010; Leyk, Buttenfield, & Nagle, 2013). Implicitly, iterative proportional fitting fits a log-linear statistical model in order to adjust the cells of an n X m table according to a set of marginal totals obtained from external data sources (Agresti, 1996; Norman, 1999; Speed, 1998).

Although variations of the method have been reported in the literature (see Anderson, 2007; Ballas et al., 2005b, 2007a; Ballas, Kingston, & Stillwell, 2004a; Ballas et al., 2005a; Edwards & Clarke, 2009; Smith, Clarke, & Harland, 2009; Smith, Pearce, & Harland, 2011), new weights (w_i^*) are computed iteratively according to

$$w_k^* = w_k \times \frac{s_{ij}}{m_{ij}} \tag{3.1}$$

where

w_k^* = the new weight for the k^{th} observation
w_k = the original sampling weight for the k^{th} observation
s_{ij} = the population total for row i, column j of the census constraint table
m_{ij} = the weighted total for row i, column j of the equivalent tabulation of survey microdata.

The algorithm begins by adjusting the sampling weights to the first constraint table and proceeds by calibrating the weights to all subsequent constraint tables. The algorithm repeats until the sum of the sampling weights equals the known totals defined by all constraint tables used in the reweighting process. Depending on the number of constraint tables used to calibrate the sampling weights, up to

20 iterations may be required until the adjusted weights converge to the constraint totals for the small area (Anderson, 2007; Ballas et al., 2005a).

Smith et al. (2009) modified this algorithm slightly by scaling the new weights in each iteration to the known small area population according to

$$aw_k^* = w_k^* \times \frac{t_o^c}{t_s^c}, \tag{3.2}$$

where

aw_k^* = the adjusted new weight for the k^{th} individual
w_k^* = the new weight for the k^{th} individual
t_o^c = the total population for constraint c in small area o
t_s^c = the sum over all k observations for constraint c from the survey microdata

Smith et al. (2009) argue that this adjustment is necessary because without it, the new weights become progressively smaller with each iteration. In order to avoid extremely small weights, the adjustment scales the new weights back to values that are consistent with the size of the real-world population for each small area. Once the algorithm converges, the final set of weights for all microdata records will sum to the small area population totals defined in the constraint tables. For any individual record, the final calibrated weights will often be decimal values instead of integer values. Ballas et al. 2005b; 2007a; 2005a and Lovelace and Ballas (2013) developed several integerization methods to convert fractional weights to whole numbers; however, there is no reason that the final calibrated weights should strictly be integers (Anderson, 2007; Lymer, Brown, Harding, & Yap, 2009; Tanton, Williamson, & Harding, 2007).

3.2.2 Generalized regression reweighting

Generalized regression reweighting (GREGWT) uses a regression model to calibrate a set of sampling weights to known benchmark totals (Singh & Mohl, 1996). The procedure ensures that calibrated weights are greater than zero. The GREGWT algorithm begins by scaling down the original microdata sampling weights to the size of the small area population. The algorithm then uses the benchmark constraints to iteratively update the sampling weights so that the sum of the calibrated

weights over the constraints for all microdata observations equals the sum of the benchmark totals to within a specified tolerance criterion (Tanton et al., 2011). When this criterion has been met or when the maximum number of iterations has been exceeded, the algorithm stops. Based on their experience, Tanton et al. (2011, 2007) report that few improvements to the calibrated weights can be made after 30 iterations. Once the microdata sampling weights have been reweighted to fit the small area benchmarks, the microdata are re-used to fit all other small areas.

Although the GREGWT algorithm has been shown to produce relatively robust small area estimates (Tanton et al., 2011, 2007), it has some limitations. First, due to its iterative nature, the algorithm can become trapped at a local minimum and therefore fails to find an optimal set of calibrated weights (Rahman, Harding, Tanton, & Liu, 2010). Second, for some spatial units, the GREGWT algorithm fails to converge. In other words, it is unable to find a set of calibrated weights that minimize the difference between the benchmark totals and those estimated from the synthetic data using the calibrated weights (Rahman et al., 2010; Tanton et al., 2007). This situation often arises because non-converging spatial units are atypical from the majority of units where convergence is achieved (e.g., the small area is largely an industrial area). Usually, non-converging units represent geographical areas of low population; Tanton et al. (2007) contend that it is often less important to obtain small area estimates for these atypical non-converging units. For example, if the purpose of the spatial microsimulation is to obtain small area estimates of low household income, then such estimates would be less important for non-converging areas where there is a high proportion of industrial activity, primarily because there would be very few people living in such areas (Tanton et al., 2007).

3.2.3 Simulated annealing

Although iterative proportional fitting and generalized regression reweighting are deterministic algorithms, probabilistic methods are available. Simulated annealing, one such probabilistic method, randomly selects observations from a micro dataset so that the final sample of selected records matches the actual population, as closely as possible, on key characteristics. The distributions of these characteristics, or constraints, are determined from univariate and/or multivariate tabula-

tions of census data for the small area. Again, if a small area contains 500 people having approximately similar proportions of men and women in all age groups, then probabilistic re-selection routines will randomly sample 500 records from a micro dataset so that the age-sex distribution of the final set of selected records matches census estimates for that area.

Williamson, Birkin, and Rees (1998) viewed this re-selection procedure as a combinatorial optimization problem. They argued that selection of an optimal set of records from a micro dataset can be conducted using a stochastic, iterative search algorithm that attempts to minimize the discrepancies between the aggregate set of microdata records and the actual population for a set of constraint variables. The optimality of the final selection can be assessed using different measures (see Section 3.2.4). In practice, the fit of the final selection is assessed using the Total Absolute Error (TAE), defined as the sum of absolute differences between the observed and expected tabulations (Voas & Williamson, 2001; Williamson et al., 1998). Specifically,

$$\text{TAE} = \sum_{ij} \left| O_{ij} - E_{ij} \right|, \tag{3.3}$$

where O_{ij} is the observed count for row i, column j of the tabulation of constraint variables from the synthetic data and E_{ij} is the expected count for row i, column j of the corresponding tabulation of constraint variables from the census data. The TAE is calculated for each constraint table used in the re-selection procedure and the overall TAE is the sum of all table specific errors.

The closer the selected sample comes to the known census distribution, the lower the TAE and the better the fit of the selected records. A perfect match between selected records and the census data on the constraint variables implies that TAE = 0 while the worst fit possible is a TAE twice the total table count. In general, a TAE of 0 may not be possible to achieve but a TAE that approaches 0 indicates better fit between selected observations and the true population (Williamson et al., 1998).

Re-selection methods begin by randomly selecting a number of observations from the micro dataset equal to the size of the population living in the area being fit. To improve the fit between selected observations and small area census data, a subset of observations can be randomly swapped with other records from the sur-

vey microdata until a reasonable fit has been achieved (Williamson et al., 1998). As a result, a given record may appear more than once in the selected sample, giving it greater "weight" in the final sample. Williamson et al. (1998) originally tried three different heuristic combinatorial optimization algorithms to select an optimal set of microdata records: hill climbing, genetic algorithms, and simulated annealing. Of these three methods, they found that simulated annealing was computationally efficient, producing the best match (i.e., lowest TAE) between the synthetic microdata and census estimates of the constraint variables (Williamson et al., 1998).

Simulated annealing is an optimization algorithm designed to mimic the physical process of annealing in metallurgy, whereby a metal is heated to a liquid state and then cooled to alter physical properties such as strength and hardness[1]. During the cooling process, the temperature of the metal is decreased so that the particles of the metal will eventually arrange themselves in a state of high density and minimum energy (Ballas et al., 2004a; Rahman et al., 2010). The goal of the simulated annealing algorithm is to find an optimal solution that minimizes the total absolute error. Simulated annealing allows for small reductions in fit during the iterative search such that the TAE may increase on the path to a global optimal solution, preventing the algorithm from becoming trapped at a local optimal solution. The choice of allowing a "worse" solution is governed by an equation derived from the laws of thermodynamics describing the probability of an increase in energy:

$$p(\Delta \mathrm{E}) = \exp(-\Delta \mathrm{E}/\mathrm{T}), \tag{3.4}$$

where

$$\begin{aligned} \mathrm{E} &= \text{energy} \\ p(\Delta \mathrm{E}) &= \text{the probability of an increase in energy} \\ \Delta \mathrm{E} &= \text{the change in energy due to selection of a poorer fitting combination} \\ \mathrm{T} &= \text{absolute temperature controlling the allowable degradation in fit.} \end{aligned}$$

For the purposes of spatial microsimulation, the total absolute error represents the energy E in Equation 3.4 while ΔE represents the magnitude of the change in TAE. Using the simulated annealing algorithm, new combinations of records

[1] See http://en.wikipedia.org/wiki/Annealing_(metallurgy).

are chosen to represent the small area if they improve the fit between the synthetic data and the census data. Worse combinations are selected only if $p(\Delta E)$ is equal or greater than a randomly generated number between 0 and 1. The random swapping of records is controlled by the temperature T in Equation 3.4. Choosing an initial value of T is arbitrary; Williamson et al. (1998) used values of 40 and 10 while Ballas et al. (2007b) used a value of half the number of households contained in a small area. More recently, Williamson (2007) recommends using the total number of cells from all constraint tables for the initial value of T.

During the swapping process, T randomly selected records from the possible solution are replaced by new randomly selected records from the micro dataset. The value of T is then reduced after a predefined number of successful swaps (usually $10 * T$) have been made so that the likelihood of a worse combination of records being selected decreases as the algorithm proceeds. In practice, it is common to reduce the T by 5% after a set number of successful swaps have been made (Williamson et al., 1998). As the error between the synthetic data and census data decreases, the number of observations made per swap is reduced to one. The algorithm continues until a maximum number of iterations has been reached or the error falls below some specified tolerance criterion. The algorithm is depicted graphically in Figure 3.1.

In an evaluation of the simulated annealing algorithm, Williamson et al. (1998) found that it produced a better fit (smaller TAE) between synthetic microdata and census data than either the hill-climbing or the genetic algorithms. Furthermore, altering the temperature T affected performance of the simulated annealing algorithm, such that smaller initial values of T caused the algorithm to accept fewer degradations in fit on the road to finding an optimal selection of microdata records. They also noted that simulated annealing produced good fitting synthetic data for small areas that were relatively typical of the overall population. In other words, if the distribution of a variable for a given small area is similar to the distribution of that variable in the sample data being matched, then the fit of the synthetic data will be relatively good. Conversely, overall fit degrades when an area is atypical and deviates from the sample used for matching (Williamson et al., 1998).

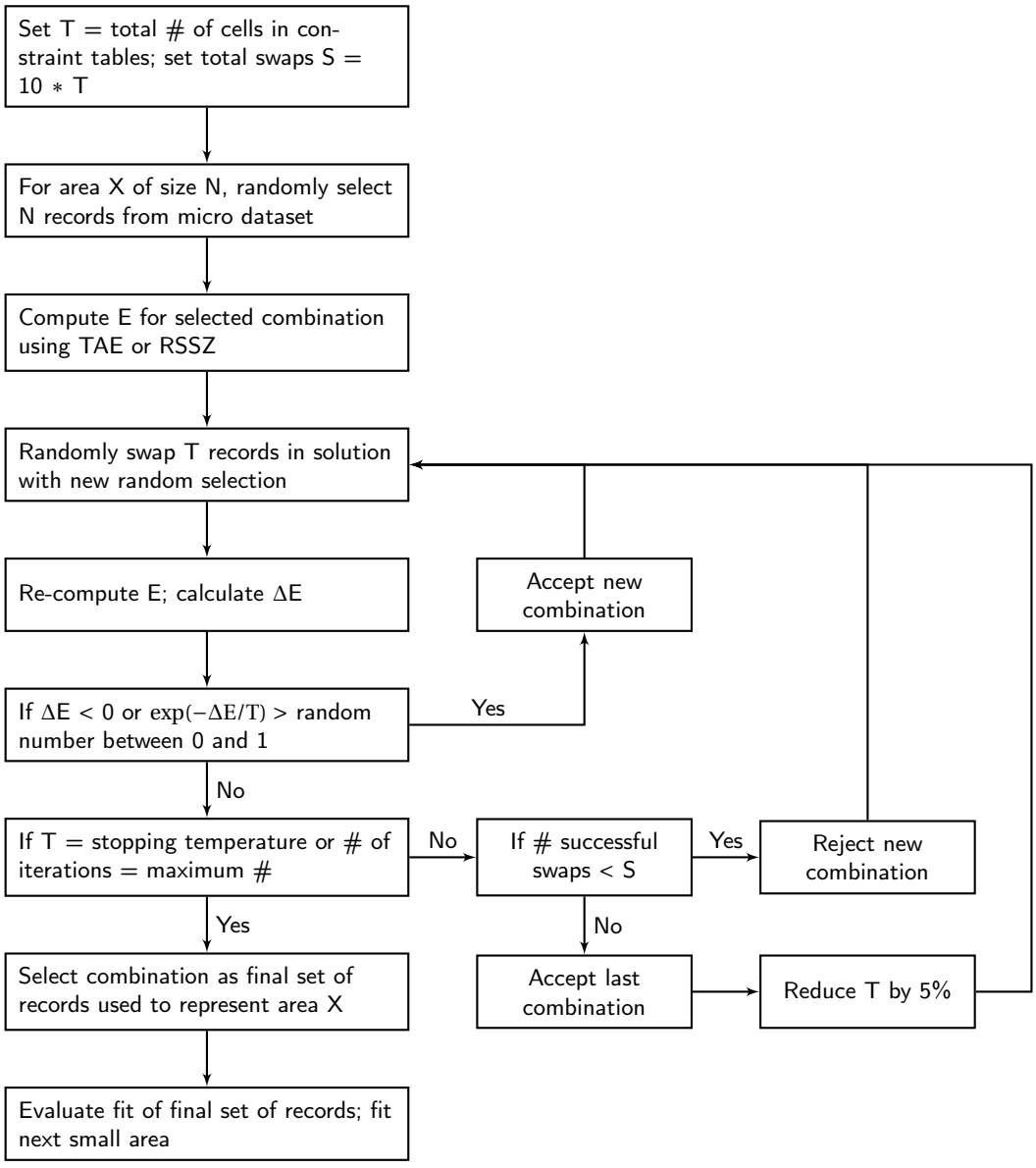


Figure 3.1. Graphical representation of the simulated annealing algorithm used in spatial microsimulation studies. Adapted from Williamson et al., 1998; Williamson, 2007; and Ballas et al., 2007b.

3.2.4 Assessing the fit of synthetic microdata

The primary goal of spatial microsimulation is to create realistic synthetic data that mimic population data as closely as possible. By fitting survey data to unique geographic areas, it is possible to impute additional information from the survey data, such as diabetes status or the presence of chronic disease risk factors, that are not available from census data. However, it is important to remember that synthetic microdata are *not* identical to actual records of households or individuals. As a result, synthetic microdata may be highly realistic for some purposes and less realistic for others.

In order to improve the fit between synthetic microdata and census data, it is important to carefully choose which variables are used as constraints in the reweighting process (Voas & Williamson, 2000). Ideally, it is important to choose a broad range of constraint variables that can be represented in as few tables as possible. Furthermore, since the goal of spatial microsimulation is to add information to geographically rich census data, the "unconstrained" variables being added must be highly correlated with constraint variables to produce valid small area estimates of unconstrained variables (Voas & Williamson, 2000). The further an area diverges from the average observation in the source micro dataset for a given *unconstrained* variable, the less likely it is that the synthetic dataset will provide reliable estimates for that variable at the small area level. Similarly, if the synthetic data cannot reproduce the distribution of constraint variables used in the reweighting procedure, there is no reason to believe they will be able to generate reliable estimates for unconstrained variables (Voas & Williamson, 2000). Thus, suitable measures should be used to assess the fit of the final synthetic data.

Such measures of fit must produce uniformly satisfactory results across all areas included in the study. As mentioned, the simplest measure of fit is the total absolute error across all tables used as constraint variables in the reweighting procedure. One limitation of the total absolute error, however, is that atypical areas will have lower fit. This reduced fit will be hidden by a global measure of fit such as total absolute error. Thus, different measures should be used to assess the overall fit of synthetic microdata. Voas and Williamson (2001) outline different measures of fit, noting that fit must be assessed during the data generation phase *and* for

the final synthetic dataset. Summary measures of fit can be used to assess fit at the cellular level (i.e., row i, column j of a constraint table) and at a tabular level.

Given the computing time required generate synthetic microdata, Voas and Williamson (2001) posit that simple statistics are preferred because they offer a fast way to evaluate successive samples. Simple measures include the total absolute error mentioned above and the standardized absolute error (SAE) which divides TAE by the expected table count N. SAE has some advantages over TAE in that it is possible to compare fit across all constraint tables because SAE gives equal weight to each table regardless of its size (Voas & Williamson, 2001). Some authors suggest that if at least 90% of small areas have less than 10% error (SAE < 0.1), then the synthetic data fit reasonably well (Riva & Smith, 2012; Smith et al., 2009; Tomintz, Clarke, & Rigby, 2009). Others have suggested that if at least 80% of small areas have less than 20% error (SAE < 0.2), then the synthetic data fit reasonably well (Clarke & Madden, 2001).

At a cellular level, Z-scores can be used to examine the fit of individual cells. Williamson et al. (1998) proposed the use of a modified Z-score to assess the fit of individual cells.[2] A unique feature of this measure is that the sum of the squared Z-scores (SSZ_m) has a χ^2 distribution with degrees of freedom equal to the number of cells in the table. Huang and Williamson (2001) further proposed the use of a relative sum of squared Z-scores ($RSSZ_m$), since the magnitude of the squared Z-scores (SSZ_m) depends on the number of table cells and the degree of error. Thus, SSZ_m increases with the size of the table. The $RSSZ_m$ divides the SSZ_m value by the table-specific 5% χ^2 critical value, where the degrees of freedom is determined by the number of cells in the table. $RSSZ_m$ is a more informative measure than SSZ_m because

1. it is a relative measure that can compare fit across several constraint tables,
2. $RSSZ_m$ statistics can be aggregated across tables to provide a measure of overall fit of the synthetic data, treating each table with equal importance,
3. the value of $RSSZ_m$ is easy to interpret: $RSSZ_m < 1$ indicates that the data fit the table and $RSSZ_m = 0$ indicates perfect fit (Ryan et al., 2009). Large

[2] $Z_m^2 = (O_{ij} - E_{ij})^2 / [E_{ij}(1 - (E_{ij}/N))]$, where O_{ij} and E_{ij} are defined as in Equation 3.3 and N is the expected table count.

values of $RSSZ_m$ indicate that the synthetic microdata are a poor fit to the constraint tables.

In summary, to evaluate the fit of synthetic microdata, measures need to be readily computed between iterations of the search algorithm. Both TAE and SAE meet this requirement, although Voas and Williamson (2001) prefer SAE because it gives equal weight to all constraint tables and can therefore be used to compare fit across tables. Once a micro dataset with reasonable fit has been selected, it is useful to compare the final observed and expected fit; Huang and Williamson (2001) recommend using $RSSZ_m$.

3.2.5 Data requirements

Regardless of the algorithm chosen, reweighting methods are data intensive methods that require careful consideration of the input data needed to generate synthetic small area microdata. The primary consideration is the selection of constraint variables. In order for a variable to be used as a constraint, it must be available in *both* the census dataset and the survey microdata (Smith et al., 2009). Each data source must operationalize the constraint variables in the same way (Chin et al., 2005; Smith et al., 2009; Tanton et al., 2011). Continuous variables cannot be used as constraints; such variables must be reclassified into discrete categories (Tanton et al., 2011). A more subtle requirement is that the classes of categorical variables must be equivalent (Tanton et al., 2011). For example, terms such as "unemployment" must use the same definition in both the census data and the survey data. Similarly, if the time period during which data were collected differs between data sources, then dollar amounts such as income should be adjusted for inflation to be comparable.

In addition, government statistics agencies often mask tabulated census results, especially at small area levels, to maintain confidentiality. For example, a census table of the number of people living in a given area by age and sex may have cells rounded to increments of five. As a result, constraint tables may differ slightly in the estimate of the total number of people living in that area. In order to use these different tables as constraints in the reweighting procedure, it is necessary to normalize the tables so that population totals across tables are equal (Chin

& Harding, 2006; Edwards & Clarke, 2009; Smith et al., 2009). This involves adjusting the cells of each table to some “master” table that is considered the most important constraint for the reweighting process. This adjustment may be computed as

$$(n_i / \sum_i n_i) \times t \tag{3.5}$$

where

n_i = the number of observations in cell i for a given constraint table,
$\sum_i n$ = the number of observations in each cell over all i categories, and
t = the total population from the "master" table.

A related consideration is how non-response values (e.g., “not stated” or “not applicable”) should be handled in the census data if complete census microdata are available. Chin and Harding (2006) and Tanton et al. (2011) recommend redistributing these values proportionately to the relative frequencies of the known categories. In other words, if a categorical variable such as income has four known categories and an unknown “not stated” category, the income category having the highest frequency will be assigned the greatest share of the “not stated” observations.

Finally, in order for any variable to be used as a constraint in the reweighting process, it must produce reliable estimates at the small area level (Tanton et al., 2011, 2007). If a potential constraint variable cannot produce reliable estimates of known quantities, then it is unreasonable to expect that it will provide useful information necessary for estimating unconstrained outcomes. Similarly, there must be a reasonable correlation between constraint variables and unconstrained variables (Birkin & Clarke, 2012; Chin et al., 2005; Edwards & Clarke, 2009; Haslett et al., 2010; Lymer et al., 2009; O’Donoghue et al., 2013; Procter, Clarke, & Ransley, 2008; Smith et al., 2009; Tanton et al., 2011). This is because the constraint variables are used to impute the unobserved values of the unconstrained variables for all small areas (Haslett et al., 2010). The stronger the correlation, the more reasonable the imputation (see Section 3.3).

3.2.6 Comparison of spatial microsimulation methods

All of the spatial microsimulation methods reviewed here are computationally intensive. Reweighting methods are designed to impute "missing" information at a small area level by re-calibrating the sampling weights or re-selecting observations from a survey micro dataset. The ultimate goal of reweighting methods is to create a final synthetic micro dataset that matches the actual population living in a small area as closely as possible on a set of known constraint variables. If a good fit is obtained between the true population, assessed from census data, and the synthetic population and if there is a reasonable correlation between the constraint variables and unconstrained outcomes, then reweighting methods can provide valid small area estimates of unconstrained outcomes (see Section 3.4). Each reweighting method has its own strengths and weaknesses so it is instructive to compare the performance of each method in terms of its ability to generate reliable synthetic microdata.

Harland et al. (2012) compared the performance of simulated annealing against iterative proportional fitting using data from the 2001 UK Census. They reweighted the Sample of Anonymized Records[3] for three different spatial units in the Leeds Metropolitan Area: (a) Output Areas (OA), (b) Lower Layer Super Output Areas (LLSOA) and (c) Middle Layer Super Output Areas (MLSOA). In the UK, the Output Area is the most disaggregated level of geography available from the UK Census. Each level of census geography is nested within the next higher level. Six constraint variables were used for each algorithm: gender, ethnic group, age, marital status, socioeconomic status, and highest level of education. To validate the results from each model, Harland et al. (2012) examined two-way interactions between constraint variables to determine how well each method could reproduce the number of people in each category of the cross-tabulated constraints. The validity of the methods was also assessed by comparing aggregated estimates of unconstrained variables, including (a) tenure, (b) limiting long-term illness, (c) gender crossed with hours worked and (d) economic activity crossed with car ownership. The final fit of synthetic data was assessed using the

[3] The Sample of Anonymized Records is a small, anonymized subset of individual level from the UK Census.

1. total absolute error (TAE),
2. classification error (CE), defined as the number of individuals misclassified in a cell for a particular constraint and
3. % classification error (%CE): CE/N, where N is the population of the cell of interest.

For all constraint variables, constraint interactions, and unconstrained variables, iterative proportional fitting produced greater classification error than simulated annealing (Harland et al., 2012). Classification error increased as the size of the spatial unit decreased. In other words, OAs had the highest amount of classification error while MLSOAs had the least amount. With respect to recreating observed interactions between constraint variables, iterative proportional fitting performed worse than simulated annealing, although results improved at coarser levels of census geography. This is because coarser spatial units contain larger populations and thus have a greater likelihood of containing a more representative sample of the population. In terms of unconstrained variables, there was no clear indication as to which algorithm performed better. However, for unconstrained variables that vary over geographic space (e.g., tenure), iterative proportional fitting may perform better than simulated annealing because IPF can be tailored to each spatial unit being fit by modifying the order of the constraint variables used in the model fitting process (Harland et al., 2012; Smith et al., 2009). On balance, however, simulated annealing consistently produced more accurate results for constraint variables while IPF tended to smooth results to the sample mean rather than preserve the characteristics of each spatial unit. Harland et al. (2012) also note that simulated annealing was better able to reproduce relationships (interactions) between constraint variables.

Tanton et al. (2007) compared simulated annealing against GREGWT using the 2001 Australian census to reweight the 1998–1999 Household Expenditure Survey to Statistical Local Areas (SLA) for two states: the Australian Capital Territory and New South Wales. Both univariate and multivariate constraints were used in the reweighting procedure. The objectives of this comparison were to:

1. assess which method produced better fitting synthetic data and
2. examine the validity of predictions made for unconstrained variables.

Housing stress, defined as a household spending more than 30% of its gross income on rent or a mortgage, was the primary unconstrained variable considered for the analysis of model validity.

Two variants of the simulated annealing algorithm were used to generate the synthetic data. The first used total absolute error (TAE) to select an optimal set of microdata records to fit each SLA. The second variant used the relative sum of squared Z scores (RSSZ). The fit of the final synthetic datasets was assessed for all convergent SLAs under the GREGWT algorithm and all SLAs under each simulated annealing variant. The final fit was assessed using TAE, SAE, and the overall RSSZ (RSSZ summed across all constraint tables).

With respect to overall fit, both simulated annealing variants produced better fitting synthetic data than GREGWT. In each case, the RSSZ variant produced the best fitting synthetic data followed by the TAE variant. In terms of the validity of the synthetic data, both simulated annealing and GREGWT overestimated the total number of households experiencing housing stress. However, all methods produced similar estimates of the percentage of households experiencing housing stress. When synthetic estimates were aggregated to the state level and compared against official statistics, the estimates from the GREGWT algorithm were very similar to official statistics, while those produced by the simulated annealing TAE variant slightly underestimated the percentage of households experiencing housing stress (Tanton et al., 2007). However, when synthetic estimates were compared against official estimates at the SLA level by regressing official statistics on synthetic estimates, both simulated annealing variants were deemed slightly better than the GREGWT estimates: R^2 = 0.898 for simulated annealing RSSZ, R^2 = 0.895 for simulated annealing TAE, and R^2 = 0.865 for GREGWT (Tanton et al., 2007).

The authors conclude that GREGWT produces good results which are comparable to simulated annealing. The fit of the synthetic data is slightly worse than simulated annealing but measures of predictive accuracy appear to be similar. The main limitation of GREGWT is that in some cases, the algorithm fails to converge; this is one of the primary advantages simulated annealing has over GREGWT (Tanton et al., 2007). Furthermore, the RSSZ variant of the simulated annealing algorithm tends to (a) produce better fitting synthetic data, (b) result in lower error,

and (c) produce reasonable estimates for unconstrained variables. Tanton et al. (2007) also found that increasing the number of benchmark constraints was more detrimental for GREGWT than for simulated annealing.

In summary, the results of these comparison studies suggest that simulated annealing is perhaps the most robust algorithm currently available for spatial microsimulation modelling. Iterative proportional fitting may have some advantages when the order of constraint variables is tailored to each spatial unit modelled. By modifying the order of constraint variables in the reweighting process, it is possible to produce local spatial microsimulation models which may improve overall fit (Smith et al., 2009).

3.2.7 Spatial microsimulation modelling software

Spatial microsimulation models have typically been developed using programming languages that can handle large datasets, repeated sampling, and the iterative nature of the model building process. Different research studies have used different languages to develop spatial microsimulation models, including Java, SAS, Fortran, and R. Many models have been developed in Java, although few details were provided to describe how they were developed (Ballas, Clarke, Dorling, et al., 2006; Ballas et al., 2007a; Clark, Birkin, & Heppenstall, 2014; Hynes, Farrelly, Murphy, & O'Donoghue, 2008; Hynes, Hanley, & O'Donoghue, 2010; Hynes, Morrissey, O'Donoghue, & Clarke, 2009; Kavroudakis, Ballas, & Birkin, 2013a; Kosar & Tomintz, 2014; Ma, Heppenstall, Harland, & Mitchell, 2014; Smith et al., 2011). Moreover, the source code used to develop them is not openly available. As a result, it is difficult to adapt the models to use different data sources or to replicate the original results.

Some authors have made their code openly available for others to use and adapt for their own purposes. Lovelace and Ballas (2013) developed a spatial microsimulation model of energy use for the City of Sheffield in the United Kingdom. Using survey data from the Understanding Society[4] survey and UK census data from 2001, they implemented the iterative proportional fitting algorithm in R. Similarly, Campbell and Ballas (2013) developed their spatial microsimulation

[4] http://www.understandingsociety.org.uk/

model, dubbed "SIMALBA", using iterative proportional fitting in R. Two generic packages — "ipfp" (Blocker, 2016) and "mipfp" (Barthelemy & Suesse, 2016) — can also be used to develop spatial microsimulation models in R. On the other hand, the "rakeR" package (Jones, Lovelace, & Dumont, 2016) was explicitly developed for spatial microsimulation.

Kavroudakis (2013b, 2015) implemented the simulated annealing algorithm in the "sms" package in R. This particular package seems to be in the early stages of development, since only the "Total Absolute Error" measure of fit can be used to select records for inclusion in the final model. Furthermore, it is not possible to set the initial temperate T used to prevent the algorithm from becoming trapped at a locally optimal solution.

Williamson (2007) implemented the simulated annealing algorithm in Fortran 95. Both the software, called "CO"[5], and the source code are freely available[6]. The CO package seems to provide the greatest amount of flexibility to the end user in terms of parameters that can be set to develop a spatial microsimulation model. Specifically, CO allows the user to specify

1. the initial temperature T used to prevent the algorithm from becoming stuck at a locally optimal solution,
2. the rate of decrease in temperature T, and
3. the measure used to assess the fit of selected observations, including TAE or RSSZ.

More options are available and described in greater detail in the CO user manual (Williamson, 2007).

Finally, the GREGWT algorithm has been implemented in the SAS language. The program was developed by the Australian Bureau of Statistics (Chin et al., 2005); unfortunately, it is not publicly available. Despite its seemingly widespread use in the literature, it should be noted that only researchers from Australia affiliated with the National Centre for Social and Economic Modelling at the University of Canberra have actually developed spatial microsimulation models using this algorithm and software (see, for example, Chin et al., 2005; Lymer, Brown, Yap, & Harding, 2008; McNamara, Cassells, Wicks, & Vidyattama, 2010; Rahman, Hard-

[5] CO stands for combinatorial optimization.
[6] http://pcwww.liv.ac.uk/~william/microdata/

ing, Tanton, & Liu, 2013; Tanton, Harding, & McNamara, 2010; Tanton et al., 2011). Like early models developed using Java, it is difficult for other researchers to use the GREGWT method without the software or to replicate existing results that appear in the literature.

3.3 Spatial Microsimulation Compared to Small Area Estimation

One of the strengths of spatial microsimulation modelling is its ability to forecast the effects of public policies at local geographic levels. In this sense, spatial microsimulation can be thought of as a type of small area estimation. In statistics, small area estimation models are a set of methods that attempt to produce reliable estimates for small geographic areas based on samples drawn from those areas (Pfeffermann, 2002). Typically, the sample sizes for each area are small and so statistical small area estimation models must utilize external information in the estimation process. Haslett et al. (2010) argue that spatial microsimulation and statistical small area estimation methods are fundamentally similar.

The key feature of small area estimation methods is that they use auxiliary information to make survey estimates for small areas more reliable (Ghosh & Rao, 1994; Rao, 1999). The Elbers, Lanjouw, and Lanjouw or "ELL" model is one type of small area estimation model (Elbers, Lanjouw, & Lanjouw, 2003). This model combines detailed survey data with comprehensive census data. Similar to spatial microsimulation, the ELL model assumes that some of the auxiliary information is available from both the survey data and the census data. Specifically, all explanatory variables X_i must be available in both data sources while outcome variables are only available in the survey data (Haslett et al., 2010).

The ELL model begins by estimating a mixed effects regression model from the survey data. Typically, random effects are specified for cluster level effects, where clusters are defined by the primary sampling units. Based on the fitted regression model, the outcome variable is predicted for all small areas using the auxiliary census information and the predicted regression parameters $\hat{\beta}$. This produces a set of predicted values for all census microdata records for each small area. These predicted values are then averaged within each small area. Standard errors for small area estimates are then estimated via bootstrapping techniques. Note that the ELL

method is valid only if the regression model estimated from the survey data is valid for the entire population across all small areas being estimated (Haslett et al., 2010).

Haslett et al. (2010) argue that small area estimation via ELL is similar to spatial microsimulation. Each method relies on survey data and census data to produce a pseudo-census. Both methods address the issue of incomplete census data. That is, not all census data have been observed; each method attempts to estimate the unobserved portion of the census. In other words, the complete census C can be decomposed into observed C_o and unobserved C_u portions so that $C = C_o + C_u$. In the ELL method, complete census data exist for covariates of interest but only partial information exists for the outcome variable Y. Thus, the aim of the ELL model is to replace the target

$$\varphi(C) = \varphi(C_o + C_u) \tag{3.6}$$

with an estimate of that target, defined as

$$\varphi(C^*) = \varphi(C_o + C_u^*) \tag{3.7}$$

where the missing data C_u are replaced by the surrogate C_u^*. In the ELL model, this surrogate is informed by an explicit statistical model. In other words, C_u^* is the expected value of C_u conditional on the observed data (Haslett et al., 2010):

$$\varphi(C^*) = \varphi(C_o + E[C_u \,|\, C_o]). \tag{3.8}$$

In spatial microsimulation, survey data are used to replace census observations lacking complete data. Replacements are made based on how well they match a set of constraint variables. This implies a model where the constraint variables are *assumed* to be important for predicting one or more unconstrained variables of interest. As a result, the quality of the estimates produced by a spatial microsimulation depends on the strength of the relationship (correlation) between constrained and unconstrained variables. More formally, Haslett et al. (2010) argue that producing a full census using spatial microsimulation is a function of the observed portion C_o plus the unobserved portion conditional upon the observed portion $(C_o + C_u \,|\, C_o)$, i.e.,

$$\varphi(C^*) = \varphi(C_o + C_u \,|\, C_o). \tag{3.9}$$

Haslett et al. (2010) then posit that ELL and spatial microsimulation are equivalent and demonstrate this using a statistical simulation study. Conducting eight different simulation experiments[7], Haslett et al. (2010) simulated a census and then randomly selected observations from it using a cluster sampling design. A set of covariates was generated using a multivariate normal distribution; these covariates were classified into categorical variables. An outcome variable was generated as a function of the categorical covariates. One hundred iterations were performed for each experiment. In each iteration, an ELL model was estimated as well as a spatial microsimulation model using iterative proportional fitting (Haslett et al., 2010). Results for the 100 iterations were averaged and the mean square error (Burton, Altman, Royston, & Holder, 2006) was used to evaluate the accuracy of each method. Results indicate that IPF is equivalent to the ELL model (Haslett et al., 2010).

3.4 Validity of Spatial Microsimulation

Model validation is a crucial step in encouraging policy makers to accept microsimulation results. Zucchelli et al. contend that the credibility of any microsimulation is "based on its capacity to reproduce observed data" (2012, p. 11). Models that cannot are less useful for forecasting the effects of public policies (Birkin & Clarke, 2011; Edwards & Tanton, 2013a; Hermes & Poulsen, 2012b). That said, validating spatial microsimulation models is a difficult task; in many cases, small area data do not exist which is the very reason for creating synthetic, spatially disaggregated microdata. As a result, there is often no easy way to compare model outputs to known estimates (Birkin & Clarke, 2011).

A variety of methods have been used in the spatial microsimulation literature to validate model results. Internal validation methods compare synthetic estimates of constraint variables against known values (Edwards, Clarke, Thomas, & Forman, 2011; Edwards & Tanton, 2013a). External validation methods compare synthetic estimates to external data sources, often for variables not used as constraints in the microsimulation model (Edwards et al., 2011; Edwards & Tanton, 2013a). External validation often involves aggregating small area estimates to less

[7] The eight simulation experiments were based on two census designs and four survey designs.

granular levels of geography to compare synthetic estimates to known quantities. Once there is evidence that the spatial microsimulation satisfactorily reconstructs the population *and* outcomes of interest, then and only then is it reasonable to use the synthetic microdata to conduct policy simulations.

3.4.1 Internal validation

Internal validation assesses the fit of synthetic microdata. It compares synthetic estimates against known estimates of the constraints used to reweight microdata to small areas. There are several ways to do this, but one common method involves regressing population counts from census data on synthetic estimates for all small areas included in the spatial microsimulation (Ballas et al., 2005b, 2005a; Edwards et al., 2011; Edwards & Tanton, 2013a; Gong et al., 2012; Procter et al., 2008). In this case, the objective is to determine how well synthetic estimates predict population counts for each category of all constraint variables used to develop the model. The R^2 from the regression model can be used as a goodness of fit measure: the higher the R^2, the better the fit.

Another measure used to assess the internal validity of a spatial microsimulation model is the Standard Error about Identity or SEI (Ballas et al., 2007a). This measure is similar to R^2, but instead of measuring the dispersion of observations about a line of best fit, the SEI measures the dispersion about a line having an intercept = 0 and slope = 1 (Edwards & Tanton, 2013a; Tanton et al., 2010, 2011). In other words, if known population counts of census constraints were plotted against synthetic estimates for all small areas, then a 45^o line passing through 0 would indicate perfect fit between known population counts and synthetic estimates (i.e., SEI = 1). The greater the dispersion about this line, the lower the SEI and the poorer the fit (Ballas et al., 2007a; Miranti, McNamara, Tanton, & Harding, 2011; Tanton et al., 2010, 2011). SEI is computed as

$$\text{SEI} = 1 - \frac{\sum_i (y_{ei} - y_{ci})^2}{\sum_i (y_{ci} - \overline{y_c})^2} \tag{3.10}$$

where

y_{ei} = synthetic estimate for constraint y in area i
y_{ci} = known population count from census data for constraint y in area i
$\overline{y_c}$ = average population count from census data for constraint y
$\sum_i$ = the sum over all small areas.

More recently, Edwards and Tanton (2013a) suggest that it is possible to examine the residuals of a spatial microsimulation model to validate it. Similar to statistical regression diagnostics, the residuals from a spatial microsimulation are computed as the difference between the observed and simulated values of all constraint variables used to build the model. Scatter plots of these residuals provide an indication of fit; if the residuals appear to be randomly distributed, this indicates the model fits known census distributions reasonably well (Edwards & Tanton, 2013a).

Finally, a few authors have used t-tests to examine whether the number of simulated individuals created by the microsimulation differs from known population counts for each category of all constraint variables (Edwards & Clarke, 2013b, 2009; Edwards et al., 2011). Edwards and Tanton (2013a) suggest that if the spatial microsimulation worked well, there should be no significant differences between the simulated dataset and actual values for any of the constraint variables. This suggestion, however, ignores the problem of multiple testing because many statistical tests are performed to examine whether differences exist. Using statistical significance tests alone could erroneously lead to the conclusion that the simulated data differ from the actual data. Having said that, it might be possible to account for multiple comparisons using corrections such as the Benjamini-Hochberg test to control the false-discovery rate (Benjamini & Hochberg, 1995).

3.4.2 External validation

As mentioned, external validation involves comparing synthetic estimates to external data sources, usually by aggregating small area estimates to less granular levels of geography (Chin & Harding, 2006; Edwards et al., 2011; Edwards & Tanton, 2013a). If the goal of a spatial microsimulation is to forecast the effects of a

policy or intervention on small area estimates of obesity, then it would be important to ensure that the synthetic microdata are valid before conducting policy simulations. In this case, external validation would involve estimating obesity rates at a regional level, such as Canadian health regions, based on the spatially disaggregated synthetic data. These synthetic estimates could then be compared to known estimates based on survey data, such as the Canadian Community Health Survey, which produces statistically reliable estimates at the health region level (Béland, 2002). Small differences between the two sets of estimates provide an indication that the spatial microsimulation produces valid estimates. Chin et al. (2005) suggest that relative differences of ±10% between synthetic estimates and external estimates indicate good validity while relative differences of ±15% indicate acceptable validity.

Riva and Smith (2012) suggest a slightly different approach to validity by considering construct validity and convergent validity. Construct validity is the extent to which an operationalized measure accurately measures what it intends to measure. Convergent validity is the degree of similarity between measures which should be theoretically related (Riva & Smith, 2012). Using iterative proportional fitting, Riva and Smith (2012) calibrated data from the 2004 and 2006 Health Surveys of England to Lower Super Output Areas (LSOA) in all of England using constraint variables from the 2001 Census. Constraints included age, sex, marital status, economic activity, and occupational class. Psychological distress (having signs of anxiety and/or depression) and heavy alcohol consumption (> 8 drinks in a single day for men or > 6 drinks in a single day for women) were the unconstrained outcomes estimated at the small area level using the synthetic dataset.

To assess construct validity for each of these measures, an index of mental health needs (the Mental Heath Needs Index or MINI) was constructed using the census data for all LSOAs. Deciles of this index were then correlated against deciles of the prevalence of psychological distress and heavy alcohol consumption using the Spearman rank correlation (Riva & Smith, 2012). If psychological distress and heavy alcohol consumption measure some underlying construct of mental health needs, then there should be a good correlation between these measures and the MINI.

A similar approach was used to assess convergent validity: psychological dis-

tress and heavy alcohol consumption should be related to other indicators of poor health at the small area level. Indicators of poor health included years of potential life lost; age-sex standardized rates of illness and disability; acute morbidity (age-sex standardized rates of acute health problems, assessed from emergency department admissions); and the proportion of adults younger than 60 suffering from mood and anxiety disorders (Riva & Smith, 2012). As before, each of these measures were grouped into deciles and the Spearman correlation was used to assess whether these measures were correlated at the LSOA level.

Results indicate that there is a strong correlation between synthetic estimates of psychological distress and mental health needs (as assessed by the MINI index) at the LSOA level (ρ = 0.91, Riva & Smith, 2012). There was also moderate correlation between estimates of heavy drinking and the MINI (ρ = 0.39, Riva & Smith, 2012). In other words, synthetic estimates of psychological distress seem to measure some underlying construct of mental health need at the small area level.

In addition, Riva and Smith (2012) demonstrate strong correlations between synthetic estimates of psychological distress and other indicators of poor population health. They report strong correlations between psychological distress and years of potential life lost (ρ = 0.73) and age-sex standardized rates of illness and disability (ρ = 0.81). These results suggest that areas having higher rates of psychological distress tend to have a greater number of years of life lost as well as higher rates of illness and disability. Heavy drinking shows weaker correlations between measures of poor population health. Riva and Smith (2012) conclude that synthetic estimates of psychological distress appear to demonstrate convergent validity with other measures of poor population health.

In general, when examining the correlation between synthetic estimates of an unconstrained outcome and survey based estimates of a related outcome, Edwards and Tanton (2013a) advise that estimated correlations must be reasonably strong ($\rho > 0.5$). Weaker correlations have little predictive value and may indicate that the microsimulation suffers from poor validity.

3.5 Reliability, Precision, and Uncertainty

Recent developments in the field of spatial microsimulation have explored the reliability and precision of these models. Reliability refers to the reproducibility of results under similar conditions (Porta, 2008) while precision refers to the lack of random error about a measurement. Confidence intervals are commonly used to define the imprecision of an estimate; estimates that are less precise have wider confidence intervals.

Hermes and Poulsen (2012b) examined the reliability of spatial microsimulation by estimating the prevalence of smoking in London, UK, using two different survey datasets to build separate models. They used simulated annealing to reselect observations from (a) the 2000–01 and 2001–02 General Household Survey (GHS) and (b) the 2001 British Household Panel Survey (BHPS). Observations from these surveys were selected to match small area socioeconomic and demographic constraint tables from the 2001 UK census, including age, sex, marital status, ethnicity, labour force status, occupation, and household characteristics.

Both models produced smoking prevalence estimates at the Output Area (OA) level. A Getis-Ord Gi* hotspot analysis (Ord & Getis, 1995) was conducted using these estimates to identify clusters of high (hotspot) and low (coldspot) smoking prevalence across the city. Overall, both input data sources produced similar spatial patterns of smoking prevalence, although some differences were identified. In particular, 153 OAs were identified as hotspots using the GHS source data but coldspots using the BPHS. Similarly, 211 OAs were identified as coldspots using the GHS but hotspots using the BPHS (Hermes & Poulsen, 2012b). The primary reason for the difference seems to result from the smaller sample size and less ethnic diversity of respondents contained in the BPHS compared to the GHS. These results demonstrate that while the overall spatial pattern of results is consistent (i.e., reliable) using different input data sources, some differences may exist that will affect conclusions drawn. Hermes and Poulsen (2012b) conclude that further research is needed to improve the reliability of spatial microsimulation.

In a similar vein, Rahman (2011) and Rahman et al. (2013) developed a method to estimate the precision of small area estimates produced by spatial microsimulation. They proposed that the standard error of small area estimates produced by

spatial microsimulation can be estimated as

$$\sqrt{\frac{\hat{p}_{ij}^{m}(1-\hat{p}_{ij}^{m})}{\sum_j n_{ij}}} \tag{3.11}$$

where $\hat{p}_{ij}^{m}$ = the estimated proportion of individuals (or households) in category m (e.g., number of smokers) for the i^{th} small area in a j^{th} sub-category of interest[8]. Note that n_{ij} = the estimated size of the j^{th} sub-category for the i^{th} small area. Approximate confidence intervals are then constructed using critical values of the normal distribution, e.g., Z = 1.96 for a 95% confidence interval:

$$95\%\ \text{CI} = 1.96 \pm \sqrt{\frac{\hat{p}_{ij}^{m}(1-\hat{p}_{ij}^{m})}{\sum_j n_{ij}}} \tag{3.12}$$

Initial results demonstrate that most small areas had narrow confidence intervals, although there was considerable variability in the interval estimates across all small areas studied (Rahman, 2011). However, this initial work is based on a case study of one spatial microsimulation model and not a statistical simulation study of the properties of the proposed confidence interval (Burton et al., 2006). Therefore, the coverage, or the proportion of times the interval contains the estimate of interest, of Equation 3.12 is unknown (Burton et al., 2006). Furthermore, as Wolf (2001) argues, the precision of microsimulation models is difficult to evaluate, because factors over and above sampling error contribute to model uncertainty. In the case of spatial microsimulation, which can be viewed as an imputation strategy, observations from a sample survey are re-used between different spatial units (Leyk et al., 2013). This is true of each of the algorithms discussed above. In addition, the simulated annealing algorithm selects microdata records *more* than once within a spatial unit. As a result, observations within the synthetic dataset should not be viewed as independent and identically distributed; thus, standard statistical methods for computing variance estimates and confidence intervals do not apply.

As a result, methods for evaluating the uncertainty of spatial microsimulation estimates need to be developed; possibilities include multiple imputation meth-

[8]Note that j can be ignored if there is no sub-category of interest (personal communication, A. Rahman, November 2012). Sub-categories might include age groups or sex.

ods and bootstrapping techniques, which have been suggested in the aspatial microsimulation literature (Cohen, 1991; Wolf, 2001). Whitworth, Carter, Ballas, and Moon (2016) offer a unique solution. Prior to developing a spatial microsimulation model using iterative proportional fitting, they identified relevant constraints using a mixed effects logistic regression model, using small areas within Wales[9] as the random intercept in their model. The random effects variance estimate was then used in conjunction with estimates of poor health, a binary unconstrained outcome, to generate uncertainty about small area estimates. Assuming the random error variance estimate was normally distributed with mean 0 and standard deviation from the mixed effects regression model, Whitworth et al. (2016) drew 10,000 random samples from a normal distribution and added these random errors to the log odds of each area-specific point estimate of poor health. These 10,000 samples of log odds were then back-transformed to the probability scale and a 95% "credible interval" was defined using the 2.5^{th} and 97.5^{th} percentiles. Whitworth et al. (2016) were able to compare the uncertainty in their estimates to known estimates from census data. They found 96% of the 410 small areas used in their analysis had true census values of poor health that fell within the bounds of their 95% credible intervals.

Finally, sensitivity analysis has been used in the microsimulation literature to provide a range of uncertainty around model outputs (Gilbert & Troitzsch, 2005; Kopec et al., 2010b; O'Hagan, Stevenson, & Madan, 2007; Rutter, Zaslavsky, & Feuer, 2011; Sharif et al., 2012). Conducting a sensitivity analysis of spatial microsimulation outputs is important when projecting the effects of public policies. Sensitivity analysis allows the modeller to assess the effect of model assumptions and recognizes that parameters used to project policy effects are themselves uncertain (Rutter et al., 2011). Sensitivity analysis can be conducted in deterministic ways by systematically varying one or more model inputs and assessing the resulting effects on model outcomes (Baio & Dawid, 2011; Claxton et al., 2005). Alternatively, probabilistic sensitivity analysis assumes that model parameters follow a probability distribution. As a result, policy projections can be performed using several realizations of model parameters given the selection of reasonable probability dis-

[9] Middle Layer Super Output Areas

tributions for those parameters (Claxton et al., 2005). Monte Carlo methods are then used to randomly select values for model input parameters (O'Hagan et al., 2007) and model outputs can be summarized over all iterations to describe the uncertainty around model outputs (Sharif et al., 2012). The downside of probabilistic sensitivity analysis is the computational burden required to implement Monte Carlo routines (Claxton et al., 2005; O'Hagan et al., 2007; Sharif et al., 2012). However, stochastic methods can help prevent misleading conclusions (Bryant, 2005).

To date, few authors building spatial microsimulation models report any sort of sensitivity analysis of projected policy effects. This dearth of information is unfortunate. Indeed, Kavroudakis et al. (2013a) only mention in passing that a sensitivity analysis of spatial microsimulation model outputs was performed while Lymer et al. (2008) indicate that future work will conduct a sensitivity analysis of their model outputs. Campbell and Ballas (2013) conducted a limited sensitivity analysis of the choice of the initial starting weight for the IPF algorithm, using values of one or the original survey weight. Differences were negligible but seemed to favour an initial weight of one for all microdata records. Farrell, O'Donoghue, and Morrissey (2010) examined the sensitivity of microsimulation outputs to the choice of different subsets of the survey microdata used to generate simulated populations.

Finally, using the simulated annealing algorithm, Ma et al. (2014) conducted a sensitivity analysis of a spatial microsimulation model of small area commuting behaviours in Beijing subdistricts. They found that 10 replicates of the model developed using different random number seeds produced negligible variation in small area estimates of the constraint variables across most subdistricts of Beijing. However, four areas had relatively high variation in some education categories used as constraining variables. In spite of this, they found relatively small variability in their primary unconstrained outcomes (trip frequency and travel distance) across all subdistricts. Given the simulated annealing algorithm is an optimization algorithm that attempts to match simulated totals of constraint variables against known census counts, these results are unsurprising.

In summary, little research has examined spatial microsimulation model uncertainty. Just as spatial microsimulation results need to be validated to be useful

for informing policy decisions, measures of uncertainty can help policy makers better understand the range of effects that might be produced when using spatial microsimulation models to forecast policy effects. Future research should continue to identify novel ways to assess model uncertainty.

3.6 Limitations of Spatial Microsimulation

Although reweighting methods can produce valid estimates of unconstrained variables at local geographic levels, these methods have their limitations. One of the problems that occurs with spatial microsimulation is that estimates of variables across spatial units may be too similar; in other words, estimates do not accurately reflect the true geographic variability that exists across a region (Birkin & Clarke, 2012). This homoscedasticity occurs because unconstrained variables are not matched as well as constrained variables in the reweighting process. Instead, the ability to replicate unconstrained variables depends on the strength of the correlation between constrained and unconstrained variables (Birkin & Clarke, 2012; Haslett et al., 2010). Second, certain segments of the population (e.g., visible minorities) are often poorly represented in survey microdata (Birkin & Clarke, 2012; Voas & Williamson, 2000; Williamson et al., 1998). This leads to poor representation of atypical observations in the synthetic micro dataset while more typical observations are better represented. This produces overly smooth small area estimates. Finally, there may be geographic factors that explain why unconstrained variables differ across space even though these areas may be very similar with respect to the constraint variables used in the reweighting process (Birkin & Clarke, 2012).

Beyond these issues, the different reweighting methods are not equivalent. While simulated annealing has been shown to be the most robust, iterative proportional fitting can produce valid small area estimates *provided* the order of the constraint variables used in the reweighting process is carefully considered. Because the reweighting process is conducted on a constraint-by-constraint basis, the first constraint variable fit tends to the have the closest match to known benchmarks (Harland et al., 2012; Smith et al., 2009). Several authors recommend that constraint variables be fit in order from strongest to weakest correlation with the

unconstrained outcome (Anderson, 2007, 2009, 2013; Harland et al., 2012; Smith et al., 2009).

The GREGWT algorithm also tends to be affected by the order in which constraint variables are modelled. Simulated annealing, however, is unaffected by the order of constraint variables fit during the reweighting process. This is because simulated annealing selects a random subset of observations and compares the selected sample on *all* constraint variables at the same time (Harland et al., 2012). Having said this, Smith et al. (2009) suggest using the ordering problem for the IPF algorithm advantageously by fitting local models of spatial microsimulation, using different orders of constraint variables for different spatial units in an attempt to achieve better fit.

In addition, Tanton et al. (2010) explored whether (a) the number of constraints used in the reweighting process and (b) univariate or multivariate constraints improved the fit of spatial microsimulation. Using GREGWT and Australian data from the 2001 census and the 2002–2003 and 2003–2004 Surveys of Income and Housing, they found that increasing the number of constraints in the reweighting process had little effect on the final fit and accuracy of synthetic estimates. They posit that fit and accuracy did not improve because the additional constraints added no extra information. In other words, the existing constraints provided sufficient information needed to create a realistic synthetic population. When the number of constraints was increased from 11 univariate or multivariate constraints to 14 univariate constraints only, the accuracy of the model did not improve with respect to an unconstrained outcome variable, housing stress. Tanton et al. (2010) hypothesize that the univariate constraints do not provide enough information needed to reliably estimate unconstrained variables. They conclude that multivariate benchmarks (e.g., age X sex) are better than univariate benchmarks (e.g., age only) for creating accurate synthetic microdata. Similarly, van Leeuwen (2010a) found, using iterative proportional fitting, that more constraints produce better fitting models.

Using larger sample surveys, in terms of the total number of sampled respondents, may also produce better fitting models. Several authors have explored this, with conflicting results. van Leeuwen (2010a) and Hermes and Poulsen (2012b) found that larger sample sizes produced better fitting models. However, results

from Huang and Williamson (2001) and Tanton et al. (2010) suggest otherwise. Specifically, Tanton et al. (2010) note that limiting observations from a sample survey to the area being fit produced somewhat better fitting models (in terms of the standard error about identity) compared to models using the full sample. They found that using the full sample from the Survey on Income and Housing when fitting statistical local areas from Sydney resulted in a slightly worse fitting model than when only observations from Sydney were used in the reweighting process. This was true of other capital cities modelled, including Brisbane, Adelaide, and Perth.

Huang and Williamson (2001) suggest the effects are more nuanced. They found that limiting the survey sample to observations coming from the same region being fit works well for small areas that are typical in terms of the constraint variables used. However, for atypical areas that deviate from the "average", using the full sample results in better fit (Huang & Williamson, 2001). Hermes and Poulsen (2012b) agree with this approach, noting in a reliability study of spatial microsimulation that when the British Household Panel Survey (BPHS) was used to fit the model, areas that were atypical with respect to ethnicity did not produce reliable estimates compared to a model that used the General Household Survey (GHS). They posit that the larger sample size available in the GHS resulted in a greater number of ethnically diverse observations available for fitting atypical areas. It should be noted, however, that these conflicting results may also be a result of the algorithm used to fit the models: Huang and Williamson (2001) and Hermes and Poulsen (2012b) used simulated annealing while van Leeuwen (2010a) used iterative proportional fitting and Tanton et al. (2010) used GREGWT.

Another limitation of spatial microsimulation is related to the prevalence of unconstrained variables being estimated by the model. Ballas et al. (2007a) mention that spatial microsimulation is unreliable for estimating rare outcomes. Lymer et al. (2009, 2008) corroborate this sentiment, noting that a spatial microsimulation of severe disability, a rare outcome, tended to over-estimate the true prevalence at the small area level. Further research is required to define how prevalent an unconstrained outcome should be for spatial microsimulation to produce valid small area estimates.

Finally, it is unclear how the modifiable areal unit problem (MAUP) affects

the robustness of spatial microsimulation estimates. First described by Openshaw and Taylor (1979), the modifiable areal unit problem arises because there are many, and typically arbitrary, ways to delineate geographic boundaries (Wong, 2009). For example, geographic space across an urban area may be arbitrarily divided into different zones for the purpose of collecting census data. This same area may be divided into zip codes or forward sortation areas for the purpose of delivering mail. When information is grouped into different zones and analyzed statistically, different spatial patterns may arise simply due to the arbitrary boundaries chosen for the spatial analysis.

The MAUP consists of two related problems: (a) the zoning problem and (b) the scale problem. The zoning problem refers to the idea that arbitrary boundaries may be used to sub-divide geographic space. If a given geographic region is subdivided into a *fixed number* of regions but the *boundaries* of those regions differ, then spatial analysis of these data will yield different results and potentially different spatial patterns (Stafford, Duke-Williams, & Shelton, 2008; Wong, 2009). The scale problem refers to the *number* of spatial units used to sub-divide a geographic region. If more spatial units are used to sub-divide an area, greater variation will be observed across geographic space (Stafford et al., 2008; Wong, 2009). Typically, using highly aggregated spatial units produces stronger correlations between variables (Openshaw & Taylor, 1979; Wong, 2009). Although there is no standard way of dealing with the MAUP, one possible solution that might be useful for spatial microsimulation modelling is to conduct more than one analysis using different scales (i.e., less aggregated vs. more aggregated) or different zoning systems (Wong, 2009). This type of sensitivity analysis might then provide clues as to whether or not results are robust at different spatial scales or across different zoning schemes. Indeed, some researchers in the spatial microsimulation domain recognize that the MAUP may affect results (Ballas et al., 2007a; Edwards et al., 2011; Hynes et al., 2008, 2010; Morrissey, Hynes, Clarke, & O'Donoghue, 2010; Procter et al., 2008; Tanton et al., 2010). To date, however, no one has tried to delineate the extent of the problem nor have they offered potential solutions.

3.7 Improving the Fit of Spatial Microsimulation Models

Recent developments in the field of spatial microsimulation have focused on improving the fit and validity of these models. Although forecasts produced by spatial microsimulation models have been shown to be robust and valid (Chin et al., 2005; Hermes & Poulsen, 2012b; Riva & Smith, 2012), reweighting methods may produce estimates that do not vary across geographic space as much as they should (Birkin & Clarke, 2012). To better reflect spatial heterogeneity in model forecasts, some authors used geodemographic methods to capture additional spatial effects (Birkin & Clarke, 2012; Smith et al., 2009). Geodemographic methods are a set of clustering techniques that classify spatial units according to a set of demographic attributes common between the units (Birkin & Clarke, 2009). Cluster analysis is the most common statistical technique used. Geodemographic clustering techniques capture variation in sociodemographic variables across geographic space as well as variation caused by geography itself (Birkin & Clarke, 2012).

Using the UK Sample of Anonymized Records from the 2001 Census and the 2007 National Shoppers Survey (NSS) — a commercial survey of more than 420,000 shoppers — Birkin and Clarke (2011) assessed whether geodemographic methods might improve the fit of spatial microsimulation models. Iterative proportional fitting was used to reweight the NSS to the Output Area level for the city of Leeds. Only those survey respondents arising from the same broad region containing the city of Leeds were included in the reweighting process (n = 27,000). Furthermore, each record from the NSS has a geodemographic classification code attached to it. Output Areas within the UK are classified according to these codes, which includes seven broad groups: Blue Collar Communities, City Living, Countryside, Prospering Suburbs, Constrained by Circumstances, Typical Traits, and Multicultural Blend (Birkin & Clarke, 2012). It was therefore possible to limit the reweighting process in each small area to those respondents possessing the same geodemographic code as the small area being fit. Doing so produced greater spatial variation across Output Areas (Birkin & Clarke, 2012).

Smith et al. (2009) proposed a similar procedure to improve model fit. Using data from the 2001 UK Census and the 2003-2004 Health Survey for England (HSE), Smith et al. (2009) used k-means clustering to classify Output Areas having simi-

lar sociodemographic characteristics — sex, age, ethnicity, and social class — into non-geographic groupings of Output Areas. Five unique clusters of Output Areas (OAs) were identified. Iterative proportional fitting was then used to reweight the Health Survey for England to these non-geographic clusters according to different models that altered (a) the number of constraint categories used to reweight the HSE and (b) the order in which constraint variables were entered into the modelling procedure. Their results indicate that increasing the number of constraint categories reduces the misclassification error in unconstrained outcome variables.[10] Furthermore, modifying the order in which constraints are fit affects misclassification error. This is particularly true of the IPF algorithm, where the first constraint fit should be the one most strongly correlated with the unconstrained outcome (Harland et al., 2012; Smith et al., 2009). Finally, no single model consistently produced the most accurate results in terms of misclassification error across each of the five non-geographic clusters of OAs. This gives rise to the idea of local microsimulation models, where different models are fit for different spatial units in order to produce more accurate and spatially heterogeneous synthetic populations (Smith et al., 2009). This might be especially important in areas that diverge from "average" areas which are more easily fit and tend to be more accurate (Birkin & Clarke, 2012).

Other research recognizes the importance of maintaining spatial variability in unconstrained outcomes, noting that spatial microsimulation models assume the spatial heterogeneity of unconstrained outcomes is captured by the spatial variability of the constraints used to build the model (Morrissey & O'Donoghue, 2011; Morrissey, O'Donoghue, & Farrell, 2014). This assumption is not always met, producing a mismatch between synthetic estimates of unconstrained outcomes and external estimates of those same outcomes at the small area level. As a result, Morrissey and O'Donoghue (2011) and Morrissey et al. (2014) calibrated their spatial microsimulation of the Irish local economy to known external targets of unconstrained outcomes (employment status and hospital admission). For a binary outcome, a logistic regression model can be estimated, using external data, to identify

[10] Assessed using the overall percent error between the proportion of individuals within an Output Area classified as married based on census data vs. the proportion classified as married based on the synthetic data.

important predictors of the outcome. The linear predictor from this model, along with a randomly generated stochastic error term, is applied to the synthetic data to estimate the predicted probability of each synthetic individual having the outcome of interest as a function of each individual's synthetic covariates. Within each small area, the number of individuals having the highest predicted probability of the outcome are then classified as possessing that characteristic. The number of individuals classified in this way is determined on the basis of known, external counts of the outcome in that small area. Morrissey and O'Donoghue (2011) and Morrissey et al. (2014) conclude that such calibration more faithfully reproduces the spatial variability in unconstrained outcomes. However, such calibration methods require that known, external estimates of the outcome are available. If they are, spatial microsimulation models may be calibrated in order to further explore relationships between unconstrained outcomes and other simulated characteristics that are not available from external data sources used to calibrate the model (Morrissey et al., 2014).

Burden and Steel (2015) further explore the lack of spatial variability that arises in spatial microsimulation models, noting the underlying spatial structure of the population should be retained when developing these models. Using a within-area homogeneity statistic (Steel & Tranmer, 2011),[11] they demonstrate that selection of constraint variables highly correlated with unconstrained outcomes does not necessarily replicate the spatial variation in the population. Selection of constraint variables also needs to consider whether spatial heterogeneity can be retained across the study area. The within-area homogeneity statistic can be used to identify constraints that vary meaningfully across small areas. By including these variables as constraints in the reweighting procedure, it is possible to retain the underlying spatial heterogeneity.

3.8 Summary

This chapter discussed the evolution of spatial microsimulation modelling. It reviewed the primary methods that are available to simulate small area microdata. Each simulated population is constructed to replicate known characteristics of the

[11] Analogous to the intra-class correlation coefficient.

true small area population. Once small area populations have been simulated and validated, it is possible to subject them to hypothetical policy scenarios to better understand how the effects of different policy options might vary over geographic space. Indeed, this is one of the primary uses of spatial microsimulation (Tanton & Edwards, 2013). In other words, different individuals living within a particular area might respond differently to a given policy. The aggregate behaviour of all individuals within that area describes the behaviour of the larger system. Projecting the potential effects of different policy options over geographic space provides policy makers with local information that can inform decision making. The primary objective of this research, then, is to demonstrate the utility of spatial microsimulation modelling for planning public health programs at the local level, such as neighbourhoods nested within larger urban areas or cities.

Chapter

4

TropISM: A Spatial Microsimulation Model of Type 2 Diabetes

4.1 Rationale

The overarching premise directing this research is that local geographic information can inform the delivery of health promotion programs. Using spatial microsimulation, this research will demonstrate how the potential effects of health promotion programs might vary over geographic space. Such forecasts provide decision makers with information that will help them direct scarce resources to areas of greatest need. To do this, it is necessary to demonstrate that a valid spatial microsimulation model can be constructed from geographically rich census data and information rich survey data.

As discussed in Chapter 1, spatial microsimulation is typically used for small area estimation and to forecast the potential effects of public policies at the small area level. Such local information allows decision makers to asses how policy effects might vary over geographic space. However, small area estimates and forecasts are only useful if they are valid. Therefore, if spatial microsimulation is used to forecast the possible effects of public policies, then small area estimates of rel-

evant outcomes must be validated prior to using them for policy simulations.

To date, a limited number of studies have used spatial microsimulation to estimate the prevalence of chronic diseases or their risk factors at the small area level. Different studies have used spatial microsimulation to estimate the prevalence of obesity in England (Edwards et al., 2011; Edwards & Tanton, 2013a); Rio di Janeiro (Cataife, 2014); New South Wales, Australia (Burden & Steel, 2015), and the tri-county area surrounding Detroit (Koh, Gardy, & Vojnovic, 2015). Each of these studies found significant variation in the prevalence of obesity at the small area level although limited information was provided to support the validity of their estimates. For example, Edwards and Tanton (2013a) and Cataife (2014) simply mention that simulated estimates of obesity aggregated to higher levels of geography were similar to official estimates of obesity for England and Rio de Janeiro, respectively. Burden and Steel (2015) note that simulated counts of obesity underestimated prevalent cases of obesity in New South Wales by 19 percentage points while Koh et al. (2015) found that spatial microsimulation estimates of obesity aggregated to the county level slightly overestimated prevalence when compared to official estimates from the US Centres for Disease Control and Prevention. Simulated estimates for Wayne and Macomb counties were only slightly higher than official estimates ($\leq$ 1.1 percentages points) while simulated estimates for Oakland county were 4.4 percentages points greater than official estimates (31.3% vs. 26.9%, respectively). Finally, Edwards et al. (2011) compared simulated estimates of obesity to cancer rates estimated from external data sources. For cancers associated with excess body weight, they found a strong correlation between the simulated number of cases of obesity and the actual number of cancer cases for seven types of cancer, including colorectal, breast, kidney, esophageal, prostate, and endometrial cancers at the ward level in northern England.

Other studies have used spatial microsimulation to examine the prevalence of smoking at the small area level. Tomintz, Clarke, and Rigby (2008) and Tomintz et al. (2009) estimated the prevalence of smoking at the output area level for the City of Leeds in the United Kingdom. However, their main purpose was to use the simulated data to optimally locate smoking cessation services. They briefly mention that constraint variables used in the microsimulation model were reliably reproduced at the small area level. However, they did not provide a detailed

examination of the validity of their simulated estimates. Likewise, Kosar and Tomintz (2014) and Tomintz, Kosar, and Clarke (2016) used spatial microsimulation to estimate the prevalence of smoking at the municipality level in Austria. They noted broad regional variation in simulated smoking rates but failed to examine the validity of their simulated estimates at the municipality level.

Smith et al. (2011), on the other hand, conducted a more detailed assessment of the validity of simulated small area estimates of smoking prevalence. In 2006, the New Zealand census included questions about smoking behaviours. As a result, known small area estimates of smoking prevalence were available for the entire country at the Census Area Unit level. Smith et al. (2011) therefore built a spatial microsimulation model of smoking prevalence, using data from the 2003 New Zealand Health Survey. Simulated estimates were compared against known census estimates. Almost all small areas had simulated estimates that differed by less than 20% from census estimates while three-quarters had estimates that differed by less than 10%[1].

Finally, two recent studies used spatial microsimulation to estimate the small area prevalence of type 2 diabetes. In Australia, Burden and Steel (2015) found that simulated small area of diabetes aggregated to New South Wales underestimated external counts by 24 percentage points. However, the main purpose of their study was to better identify how constraint variables should be chosen when developing spatial microsimulation models. Clark et al. (2014) developed a spatial microsimulation model of morbidity for the elderly population of England at the local authority level. When prevalence estimates of various chronic diseases were aggregated to distinct groups of local authorities, they found that simulated estimates typically differed by less than ±3 percentage points from survey-based estimates. For diabetes, estimates outside of London were most accurate, differing by less that ±2.5 percentage points. In summary, extant research demonstrates that spatial microsimulation models can produce valid small area estimates of chronic disease risk factors. As all of the studies reported here developed their simulation models under different circumstances (e.g,. different settings, different constraint variables, and different reweighting methods), the direction and mag-

[1] Relative differences.

nitude of the differences found between simulated estimates and external survey estimates might, in part, be explained by such methodological choices.

4.1.1 Study objectives

This part of the research consists of two related objectives. First, this research develops a spatial microsimulation model of type 2 diabetes and its risk factors for the City of Toronto. The model, called "TropISM" will use census data from 2006 to reweight data from the 2005 Canadian Community Health Survey to neighbourhoods within Toronto. Following Booth et al. (2007), neighbourhoods are defined as groups of census tracts having similar sociodemographic characteristics. The unconstrained outcomes of interest include: (a) type 2 diabetes, (b) body mass index, (c) hypertension, (d) heart disease, and (e) smoking. These outcomes were selected because they are used by the Diabetes Population Risk Tool (DPoRT) to forecast the population risk of developing type 2 diabetes (Manuel et al., 2013; Rosella et al., 2014, 2011). Using the simulated neighbourhood populations, DPoRT will be used to assess the potential effects of different weight loss strategies on the spatial distribution of diabetes risk across Toronto neighbourhoods.

Second, this research assesses the validity of the TropISM spatial microsimulation model using three approaches. First, as recommended in Section 3.4, an examination of the model's internal validity will be conducted by comparing simulated estimates of the constraints used to reweight the survey data to known census values of those constraints for all neighbourhoods. Then, model's external validity will be assessed by aggregating the simulated data to the health region (city) level and comparing aggregate estimates against known estimates from the 2003, 2005, and 2007 Canadian Community Heath Surveys. If the microsimulation model performs well, simulated estimates should be comparable to known estimates from these data sources.

Next, simulated estimates will be compared to external estimates at the neighbourhood level. Where possible, comparisons will be made using the same indicator. For example, simulated estimates of type 2 diabetes will be compared to external estimates of diabetes at the neighbourhood level. Otherwise, comparisons will be made using similar indicators, or, following Edwards et al. (2011), measures as-

sociated with the indicator of interest. For example, simulated obesity rates can be compared to known cancer rates for cancers associated with obesity. Section 4.2.5 discusses the validation procedures in more detail.

4.2 Methods

4.2.1 Data sources

The City of Toronto, Canada's largest city located in south central Ontario, comprised the study area for the TropISM model. Located on the northern shore of Lake Ontario, Toronto is bounded by Highway 427 in the west, Steeles Avenue in the north, and Rouge Park in the east (Figure 4.1). TropISM was developed using publicly available data sources, including data delineating the geographic boundaries of Toronto's neighbourhoods, 2006 Canadian Census data, and data from the 2005 Canadian Community Health Survey. Geographic data identifying Toronto's 140 neighbourhoods were obtained from the City's Open Data Initiative (Social Development, Finance & Administration, City of Toronto, 2009). Each neighbourhood is comprised of two or more adjacent census tracts having similar sociodemographic characteristics. No neighbourhood is comprised of fewer than 7,000 residents; most neighbourhoods contain 15,000 to 20,000 residents (Creatore et al., 2007a). Neighbourhood boundaries were designed to remain fixed over time and to respect existing natural (e.g., rivers) and man-made (e.g., roads) boundaries (City of Toronto, 2013b).

Constraint data for the TropISM model came from the 2006 Canadian Census. Summary datasets contained estimates of sociodemographic characteristics for each census tract (Statistics Canada, 2006a, 2006b). Tract-specific counts were aggregated to the neighbourhood level so that detailed constraint information was available for all neighbourhoods. Sociodemographic constraints were selected on the basis of their association with type 2 diabetes and its risk factors. Data from the 2006 Canadian Census were used because the mandatory long form census was replaced by a large voluntary survey in 2011. This survey has been shown to contain non-negligible biases (Statistics Canada, 2013).

The 2005 Canadian Community Health Survey (CCHS) provided a rich source of unconstrained outcomes for the TropISM model. The CCHS is a cross-sectional

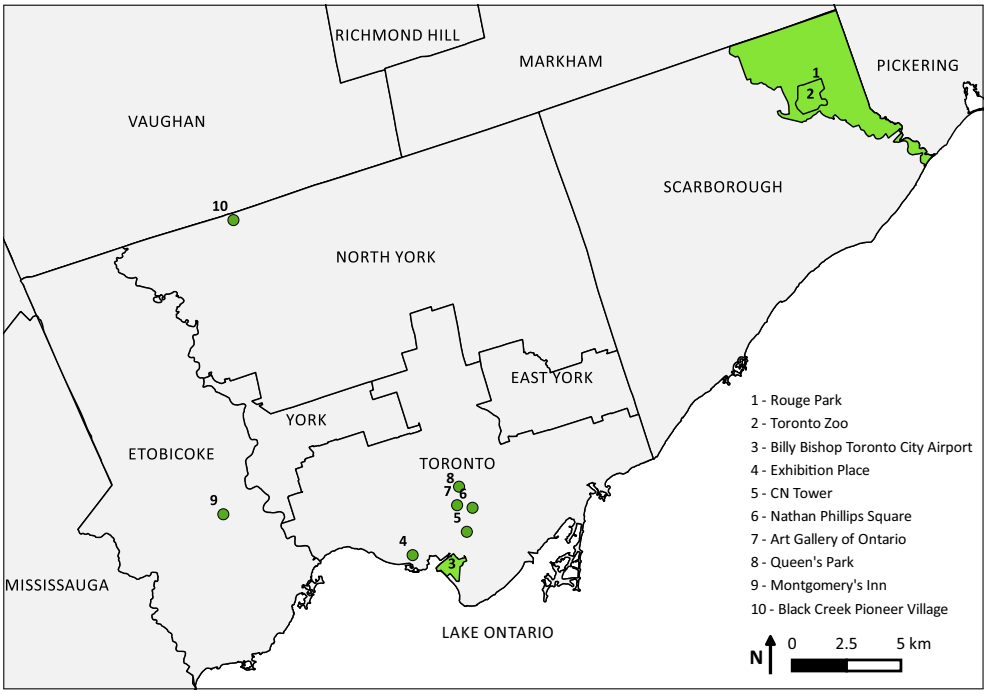


Figure 4.1. The metropolitan Toronto study area used to develop the TropISM model, which includes the boroughs of Etobicoke, North York, York, Toronto, East York, and Scarborough.

survey conducted on a biennial basis. It contains sufficient sample size to reliably estimate health outcomes and risk factors at the health region level (Béland, 2002). The CCHS uses a stratified, multi-stage sampling design to randomly select more than 120,000 respondents aged 12 and older from all Canadian provinces and territories. Respondents are interviewed in person or via computer assisted telephone interviewing methods (Thomas & Wannell, 2009).

The target population for the TropISM model included all residents of Toronto aged 15 and older. In 2005, 41,766 respondents from Ontario were interviewed for the CCHS. A subset of 38,330 Ontario respondents aged 15 and older was used to develop TropISM. Observations were reweighted using constraint variables common to both datasets. Constraint variables used the same categorical levels (e.g., sex, immigrant status) or were reclassified in the CCHS to match the constraint

categories from the census (e.g., age groups, income groups).

4.2.2 Unconstrained outcomes

Several unconstrained outcomes were simulated using these data sources, including (a) type 2 diabetes, (b) body mass index, (c) hypertension, (d) heart disease, and (e) current smoking status. These outcomes were selected because they are used by the Diabetes Population Risk Tool to predict the future risk of type 2 diabetes at the population level (Manuel et al., 2013; Rosella et al., 2014, 2011). To date, this tool has been used to project the future burden of type 2 diabetes (Rosella et al., 2011) and the potential effects of various population level interventions designed to reduce the risk of diabetes (Manuel et al., 2013). However, these studies estimated risk and risk reductions at a *provincial* level. The current study estimates risk reductions at the neighbourhood level for Toronto, an ethnically diverse Canadian city (Topping, 2012). All of these outcomes were readily identified using data routinely collected by the Canadian Community Health Survey (Rosella et al., 2011).

Type 2 diabetes status was assessed using the Ng-Dasgupta-Johnson algorithm which broadly targets all respondents who reported having diabetes in population survey microdata (Ng, Dasgupta, & Johnson, 2008). Respondents who had gestational diabetes were classified as non-diabetic. Type 1 diabetics were screened out on the basis of their age at diagnosis (younger than 30) and whether they started using insulin within six months of their diagnosis. The remainder were classified as type 2 diabetics as were respondents currently using oral anti-hyperglycemia medications.

The remaining risk factors were assessed using simpler approaches. Body mass index was computed as a continuous variable using self-reported weight (in kg) and height (in m^2) and then classified into six categories: (a) <23, (b) 23–24.99, (c) 25–29.99, (d) 30–34.99, (e) ≥ 35, and (f) unknown.[2] Hypertension and heart disease were assessed using affirmative responses to the questions “Do you have high blood pressure?” and “Do you have heart disease?”. Current smoking sta-

[2]Consistent with how BMI is used by the Diabetes Population Risk Tool.

tus was assessed on the basis of whether or not respondents currently smoke on a daily or occasional basis.

4.2.3 Selection of constraint variables

Constraint variables used to develop TropISM were selected because they

- were common to both the 2006 census dataset and the 2005 Canadian Community Health Survey,
- were used as risk factors in the Diabetes Population Risk Tool, and
- were reasonably correlated with the unconstrained outcomes.

Variables used as risk factors in the Diabetes Population Risk Tool were automatically included as constraints, including sex; age; ethnicity (white vs. non-white); immigrant status (born in Canada vs. immigrant); education (< secondary, secondary graduate/some post-secondary, and post-secondary graduate); and personal income (< $15,000, $15,000–$29,999, $30,000–$49,999, $50,000+). Although the DPoRT model only uses three broad age groups to predict future diabetes risk (< 45, 45–64, and ≥ 65), seven age groups were used to develop TropISM (15–24, 25–34, 35–44, 45–54, 55–64, 65–74, and ≥ 75).

Additional variables were considered as constraints if they were significantly associated with the five unconstrained outcomes. Candidates included housing tenure (does not own home vs. owns home) and marital status (married/common-law, widowed/separated/divorced, and single/never married). In order to assess the strength of the association between each constraint variable and the unconstrained outcomes, linear and logistic regression models were estimated using Ontario respondents aged 15 and older from the 2005 CCHS. Each constraint variable was added sequentially to the models and the increment in R^2 or pseudo-R^2 was assessed as well as the change in the Bayesian Information Criterion.[3] A sex-by-age interaction term was also included in each model. Figure 4.2 illustrates how the models were built and Figure 4.3 illustrates the improvement in model fit obtained by adding additional constraint variables to the model. A linear regression model was used to assess the amount of variation in body mass index explained by each of the constraint variables while logistic regression was used for all other

[3] Smaller values indicate better fit to the data.

Model 1: Y ~ sex
Model 2: Y ~ sex + age group
Model 3: Y ~ sex + age group + sex * age group
Model 4: Y ~ sex + age group + sex * age group + ethnicity
Model 5: Y ~ sex + age group + sex * age group + ethnicity + immigrant
Model 6: Y ~ sex + age group + sex * age group + ethnicity + immigrant + education
Model 7: Y ~ sex + age group + sex * age group + ethnicity + immigrant + education + income
Model 8: Y ~ sex + age group + sex * age group + ethnicity + immigrant + education + income + tenure
Model 9: Y ~ sex + age group + sex * age group + ethnicity + immigrant + education + income + tenure + marital status

Figure 4.2. Incremental model building strategy used to select constraint variables for TropISM model development.

outcomes. Separate logistic regression models modelled the probability of (a) being a current smoker, (b) having high blood pressure, (c) having heart disease, and (d) having type 2 diabetes.

Figure 4.3 displays the incremental fit for each model. What is immediately clear is that sex and age alone explain a large amount of variation in each of the outcomes. For the logistic regression of smoking status on the constraints, overall fit improved with each model. In other words, even when marital status was added to the model (model 9), the fit improves noticeably. A similar, but smaller effect was observed for the linear regression of body mass index on the constraints. However, for the remaining three outcomes, very little additional variance was explained when housing tenure and marital status were added to the models. Based on these results, a decision was made to develop the TropISM model using one bivariate constraint — sex crossed with age group — and the following univariate constraints: ethnicity, immigrant status, education, and personal income, yielding a total of 25 constraint categories.

Although income was chosen as a constraint variable, it should be noted that

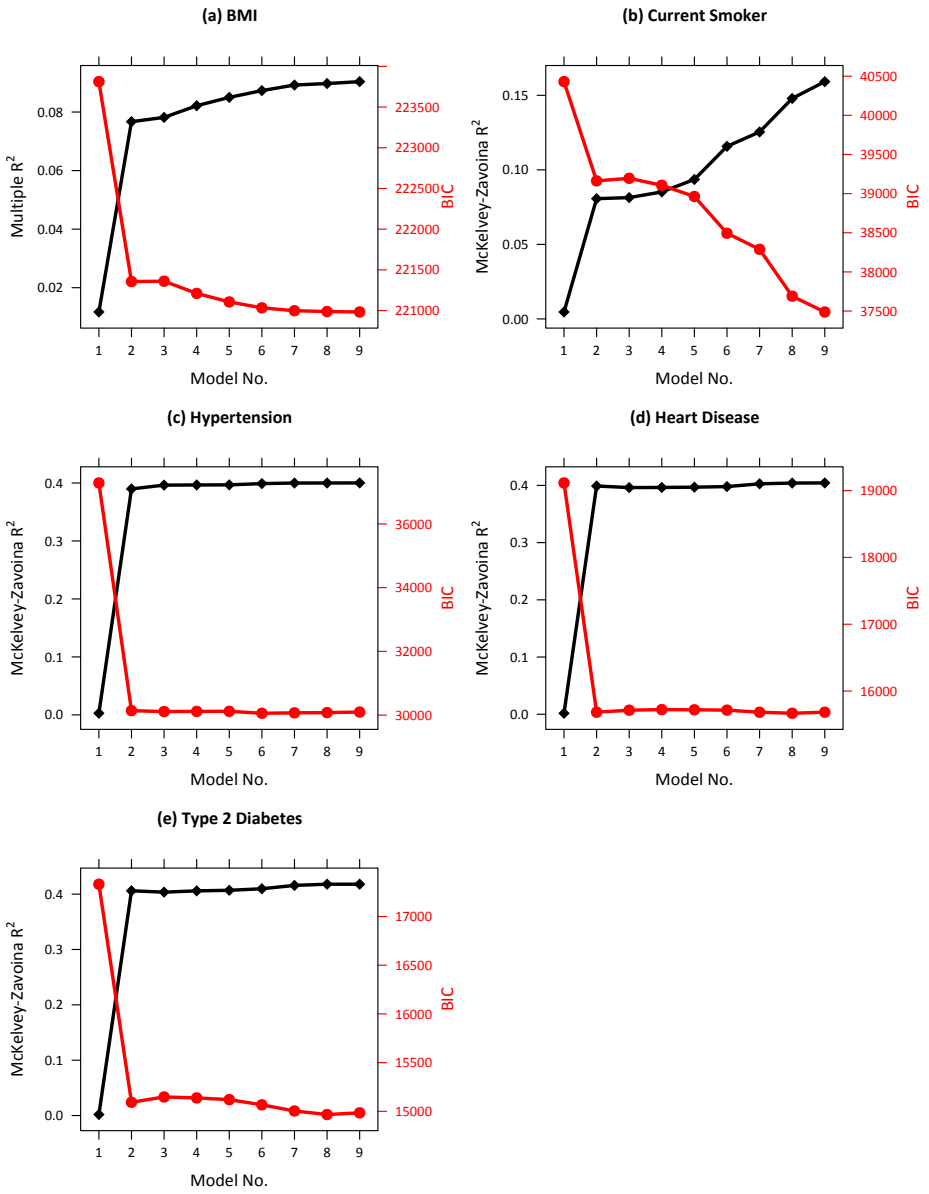


Figure 4.3. Selection of constraint variables to develop the TropISM model. Black lines depict the proportion of variance explained by each model while red lines compare model fit.

14% of respondents from the Ontario subset of the 2005 CCHS did not provide income information. Following Chin and Harding (2006) and Tanton et al. (2011), "not stated" responses were imputed. However, instead of redistributing those responses proportionately to the relative frequencies of the known categories, a regression procedure was used to impute the missing data.[4] Missing values for the other constraint variables were not imputed. Therefore, an initial microsimulation model was developed using 38,330 respondents having complete (or imputed) data for sex, age, ethnicity, immigrant status, education, and personal income.[5] Although the CCHS data were collected in 2005 and the census data were from 2006, given the short time period between data collections, reported incomes in the survey data were not adjusted for inflation.

4.2.4 Reweighting methodology

The simulated annealing algorithm, implemented in the Fortran95 "CO" software program (Williamson, 2007), was used to resample observations from the Ontario subset of the 2005 CCHS. Observations were selected so that the final distribution of observations matched known distributions of the univariate and bivariate constraints at the neighbourhood level. The 2006 Canadian census supplied constraint information at the census tract level. This information was aggregated to the neighbourhood level for each of Toronto's 140 neighbourhoods.

The simulated annealing algorithm was selected to develop the TropISM model because it is not affected by the order in which constraint variables are fit during the reweighting process, unlike iterative proportional fitting or GREGWT (Harland et al., 2012). Furthermore, there is no need to normalize constraint tables so that population totals across tables are equal (Harland et al., 2012; Williamson, 2007). In addition, there is no need to convert the final weights to integer values, since observations are resampled rather than reweighted. Moreover, existing literature suggests that the simulated annealing algorithm is the most robust algorithm currently available for spatial microsimulation. Finally, the CO software program is

[4] Data were imputed using a multinomial regression model to predict income as a function of Ontario health region, sex, age, ethnicity, immigrant status, and education. Imputation was conducted in R using the "regressionImp" function from the "VIM" (Visualization and Imputation of Missing Values) package (Templ, Alfons, Kowarik, & Prantner, 2015).

[5] Income was imputed for a total of 4,308 respondents.

freely available and has been used successfully in other settings (e.g., Burden & Steel, 2015; Hermes & Poulsen, 2012b, 2013).

Although the CO software program does not require normalization of constraint tables to ensure that small area population counts are consistent across tables, Canadian census data are masked. In other words, small area population counts have been rounded to the nearest 5 to preserve confidentiality. In addition, population counts for different sociodemographic variables do not necessarily apply to the target population of residents age 15 and older. Often, census counts apply to all age groups instead of the population age 15 and older. As a result, constraint tables were proportionately scaled to the total population age 15 and older in each neighbourhood.

For example, if a neighbourhood contained 31,000 residents, 26,000 of whom were 15 and older, and if population counts for a given constraint represented the total population, then population counts within each constraint category were scaled so that the total population for that constraint summed to 26,000 instead of 31,000. More concretely, if a constraint variable, such as ethnicity, was comprised of two categories (e.g., visible minority and non-minority), and if there were 24,000 and 7,000 people in each category, respectively, then these counts were scaled to the population age 15+ according to:

$$\frac{\text{total in constraint category } i}{\text{total population}} * \text{total population age 15+} \tag{4.1}$$

Continuing with this example, the approximate number of visible minorities age 15+ in this neighbourhood equals

$$= 24,000/31,000 * 26,000$$
$$= 20,129.$$

Following normalization of constraint tables, the CO software was used to resample observations from the CCHS microdata to match census constraints within each neighbourhood. Selection of observations was controlled by: (a) the initial temperature T, (b) the number of random swaps R made to the initial set of selected observations, and (c) the reduction in temperature T following R random swaps. The initial temperate T defines the number of observations that will be

randomly replaced in order to improve the fit between selected observations and small area constraints. Following Williamson (2007), the initial temperature T was set to the total number of cells from all constraint tables. The number of random swaps R was set to $10 * T$ and the temperature T was reduced by 5% following R successful swaps. Thus,

- the initial temperature T was set to 25,
- the number of successful random swaps R that need to be made before the temperature is reduced was set to 250, and
- following R successful swaps, the temperature was reduced by 5%. In other words, the temperature reduction parameter was set to 0.95.

Improvement in model fit was assessed using the relative sum of squared Z-scores during each step of the reweighting process, as suggested by Burden and Steel (2015); Hermes and Poulsen (2012b); Ryan et al. (2009) and Williamson (2007).

An iterative process was used to develop the TropISM model and select a final model that produced realistic estimates of unconstrained outcomes. An initial model was developed using the full Ontario subset of the 2005 Canadian Community Health Survey. Model fit statistics (total absolute error and standardized absolute error) were used to assess how well the simulated data fit the constraining tables. The simulated data were then aggregated over all 140 neighbourhoods. Simulated estimates of unconstrained outcomes were compared to external estimates of these same outcomes. Absolute and relative differences between simulated estimates and external estimates from the 2003, 2005, and 2007 Canadian Community Health Surveys[6] suggested this initial model was unable to produce realistic estimates of unconstrained outcomes. Therefore, additional models were developed using smaller subsets of the 2005 CCHS.

The next model was developed using a largely urban subset of the 2005 CCHS. In this case, the source dataset for the CO algorithm was limited to observations from Ontario health regions containing a census metropolitan area.[7] This limited the source microdata to 25,195 observations coming from 17 health regions.[8] A

[6] Appropriately weighted.

[7] A census metropolitan area, or CMA, is an urban areas having a population $\geq$ 100,000, of which at least 50,000 must live in the central core of that urban area

[8] Brant, Durham, Halton, Hamilton, Kingston-Frontenac-Lennox-Addington, Middlesex-London, Niagara, Ottawa, Peel, Peterborough, Simcoe-Muskoka, Sudbury, Thunder Bay, Waterloo,

third model used an even smaller subset of microdata observations by limiting the source microdata to those observations sampled from health regions within the Greater Toronto Area (n = 9,224).[9] A final model was developed using only those observations coming from the Toronto health region (n = 2,974). The fit of each model was assessed and simulated estimates of unconstrained outcomes were compared to known external estimates.

4.2.5 Model validation

4.2.5.1 Internal validation

Internal validation procedures assessed the fit of the simulated data against each constraint table using (a) the total absolute error (TAE) and (b) the standardized absolute error (SAE). SAE values > 0.2 for a given constraint indicate poor fit. This means that the simulated population in a particular area does not reproduce the characteristics of the known population for that constraint. Parallel coordinate plots were used to visually compare SAE values across all constraining tables and neighbourhoods.

In addition, scatterplots of known counts against simulated counts were produced for each constraint category for all model constraints (i.e., 25 constraint classes representing the five constraint tables). The standard error about identify was then used to assess how much simulated population counts deviated from known census counts for each constraint category. An SEI of 1 indicates perfect fit while values smaller than 1 indicate poorer fit.

4.2.5.2 External validation

City-level validation. The first external validation method compared simulated estimates of unconstrained outcomes aggregated over metropolitan Toronto to health region estimates from the Canadian Community Health Survey. Simulated estimates were compared to CCHS estimates from 2003, 2005, and 2007 separately by sex for (a) type 2 diabetes and (b) hypertension. Age-sex specific estimates were also compared for body mass index while estimates of personal income, current

Wellington-Dufferin-Guelph, Windsor-Essex, York and Toronto.

[9] Halton, Peel, York, Durham, and Toronto

smoking status, and heart disease were compared for men only. Absolute and relative differences between the simulated estimates and external data sources were used to assess TropISM model validity. Following Chin et al. (2005), relative differences of ±10% suggested good validity while relative differences of ±15% suggested acceptable validity.

Neighbourhood-level validation. The second external validation method compared simulated estimates of unconstrained outcomes against external estimates of the same measures or measures associated with the outcomes. External estimates were derived from administrative health data maintained by the Institute for Clinical Evaluative Sciences (Table 4.1). For diabetes, hypertension, and heart disease, simulated counts were compared against the actual number of cases in each neighbourhood using the overall concordance correlation coefficient (ρ_c) to measure the precision and accuracy of simulated counts (Barnhart, Haber, & Song, 2002; Lin, 1989, 2000).[10] Precision is defined by the Pearson correlation coefficient (ρ) and measures how much simulated counts deviate from the line of best fit. Accuracy is defined by a measure of bias (C_b) that assesses how much simulated counts deviate from the line of concordance, or the 45^o line. Accuracy decreases as C_b approaches zero. C_b is comprised of two types of deviation from the 45^o line: a scale shift and a location shift. The scale shift measures the change in slope of the line of best fit relative to the line of concordance while the location shift measures the change in height of the line of best fit. Negative values for the location parameter indicate that simulated counts under-estimate known counts for each of the disease outcomes.

For neighbourhood counts of diabetes, the simulated number of cases of type 2 diabetes were compared against known counts of diabetes ascertained from the Ontario Diabetes Database. First, simulated counts were compared against known counts. Second, simulated counts were corrected for undiagnosed cases of diabetes, assuming that 30% of cases are undiagnosed (Canadian Diabetes Association, 2015; Hux & Tang, 2003; Public Health Agency of Canada, 2011; Rosella et al., 2015). Known counts were corrected for the false positive rate in the algorithm used to identify cases of diabetes in the Ontario Diabetes Database. For diabetes,

[10] Estimated in R using the "epiR" package (Stevenson et al., 2015).

Table 4.1. Disease indicators and data sources used to externally validate unconstrained outcomes from the TropISM spatial microsimulation model.

Indicator	Years	Sex	Data Source
Diabetes	2005 (P) 2006-2010 (I)	M, W M, W	Ontario Diabetes Database
∗ Obesity related cancers: endometrial (W) + post-menopausal + breast (W) + kidney + colorectal + bladder + esophageal + pancreatic + thyroid + gall bladder + prostate (M)	2005 (I)	M, W	Ontario Cancer Registry
Lung cancer	2005 (I)	M	Ontario Cancer Registry
Chronic obstructive pulmonary disease	2005 (I)	M	Chronic Obstructive Pulmonary Disease
Hypertension	2005 (P)	M, W	Ontario Hypertension Database
Heart disease atheroslcerosis, cornary artery disease, myocardial infarction + heart attack	2005 (P)	M	Congestive Heart Failure, Ontario Myocardial Infarction Database

Notes:
P = prevalent cases, I = incident cases, M = men age 20+, W = women age 20+
∗ All cancers combined, as appropriate, for men and women.

the false positive rate is estimated to be 3% (Hux, Ivis, Flintoft, & Bica, 2002). In addition, known counts were also corrected by assuming that 95% of all cases of diabetes are type 2 diabetes (Caspersen et al., 2012; Millar & Young, 2003; National Diabetes Information Clearinghouse (NDIC), 2011).

A similar procedure was used to compare simulated counts of hypertension against known counts estimated from the Ontario Hypertension Database. Simulated counts were compared by examining the concordance between simulated counts and known counts and then by correcting simulated counts for underdiagnosis (17% of Canadians with hypertension remain undiagnosed Campbell & Ballas, 2013) and known counts for the false positive rate, estimated to be 5% (Tu, Campbell, Chen, Cauch-Dudek, & McAlister, 2007). For heart disease, no corrections were made since two different administrative data sources were used to ascertain known counts of heart disease across Toronto's neighbourhoods, and the proportion of cases coming from each source was not identified.

In addition to the concordance correlation coefficient, the average difference between simulated disease counts and known counts across all neighbourhoods was estimated. Average differences were used to identify whether TropISM under- or over-estimated known counts of these three outcomes at the neighbourhood level. A relative error measure was also considered by estimating the proportion of neighbourhoods having simulated counts lying within ± 20% of the known counts. Following Clarke and Madden (2001), if 80% of neighbourhoods had simulated counts lying within a 20% difference of known counts, then synthetic counts of diabetes, hypertension, and heart disease would be considered as reproducing known counts with reasonable accuracy.

The validity of simulated counts was further examined by classifying neighbourhoods into quintiles according to simulated and known counts of diabetes, hypertension, and heart disease. The overall percent agreement and Cohen's κ were used to compare the similarity between each classification. Higher values of each measure suggest that neighbourhoods were ranked into the same quintile, suggesting a greater degree of similarity in the relative ranking of neighbourhoods.

The validity of all unconstrained outcomes was also assessed by comparing simulated prevalence rates against known prevalence rates. For some outcomes, small area prevalence rates were not available for metropolitan Toronto. In these

cases, simulated estimates of overweight, obesity, and smoking were validated by comparing the prevalence of these outcomes to disease indicators associated with these outcomes. Similar to Edwards et al. (2011), cancers associated with excess body weight were used to assess the validity of the synthetic body weight data at the neighbourhood level. Relevant cancers included endometrial, post-menopausal breast, kidney, colorectal, bladder, esophageal, pancreatic, thyroid, gall bladder, and prostate cancer (Table 4.1). Likewise, to assess the validity of simulated smoking prevalence at the neighbourhood level, the cumulative incidence proportions of (a) lung cancer and (b) chronic obstructive pulmonary disease (COPD) were compared to simulated smoking prevalence rates. If TropISM is valid, simulated prevalence estimates for each unconstrained outcome should exhibit similar spatial patterns of (a) known prevalence rates for those outcomes or (b) incidence rates of disease indicators associated with those outcomes.

For this analysis, the Pearson correlation was used to assess the similarity between simulated prevalence rates and known prevalence rates or incidence rates of relevant disease indicators. Both raw and spatially smoothed rates[11] were compared to assess how well TropISM replicated disease prevalence at the neighbourhood level. Higher positive values of the Pearson correlation suggest that neighbourhoods having higher simulated prevalence rates tend to have higher known prevalence rates.

Neighbourhood-specific rates were also classified into quintiles to assess how well TropISM could replicate broad spatial patterns of disease prevalence. Again,

[11] Non-parametric spatially smoothed rates were estimated in R using a "rook" contiguity matrix to define contiguous neighbourhoods, following the advice of Griffith (1996) who recommends using fewer neighbours to define the underlying spatial structure, since over-specification (more neighbours) reduces statistical power. Spatially smoothed rates r were estimated as

$$r = \frac{(e_j + \overline{e}_{lag_j})}{(p_j + \overline{p}_{lag_j})} \tag{4.2}$$

where:

e_j = the number of events in neighbourhood j,
$\overline{e}_{lag_j}$ = the average number of events in all neighbourhoods bordering neighbourhood j,
p_j = the total population in neighbourhood j,
$\overline{p}_{lag_j}$ = the average population across all neighbourhoods bordering neighbourhood j

raw and spatially smoothed rates were compared. The overall percent agreement and Cohen's κ were used to assess the degree of similarity between these two classification systems. Finally, the global bivariate Moran's I statistic[12] was used to compare the spatial similarity of simulated and known rates. High positive values of the bivariate Moran's I statistic suggest that neighbourhoods having low or high simulated prevalence rates are surrounded by neighbourhoods having similarly low or high known disease rates. Bivariate LISA statistics were then computed in GeoDa[13] to identify groups of neighbourhoods having similarly high or low simulated and known prevalence rates of diabetes.

4.3 Results

4.3.1 Model development and fit

Table 4.2 compares the subsets of the 2005 CCHS used to develop the TropISM model against demographic characteristics of Toronto residents estimated from the 2006 Canadian census. In this table, CCHS estimates are unweighted so they represent the sample of respondents used to develop TropISM. Generally, the full sample of Ontario respondents differed appreciably from the characteristics of Toronto residents, especially with respect to ethnicity, birthplace, education, marital status, and housing tenure. Respondents from smaller subsets were more similar to Toronto residents.

The Toronto subset of CCHS respondents showed more similarity to the characteristics of Toronto residents compared to other subsets. Thirty-eight percent of Toronto respondents from the CCHS were visible minorities compared to 46% of Toronto residents. Similarly, 49% of respondents from Toronto were born outside Canada compared to 51% of residents. Marital status and housing tenure were closely aligned to census estimates. Respondents from Toronto having a post-secondary education were slightly over-represented compared to census estimates (59% vs. 53%, respectively), as were respondents having personal incomes of more than $50,000 per year (27% vs 24%, respectively). In spite of some differ-

[12] Estimated using PySal version 1.10.0 (Rey & Anselin, 2010) using a "rook" contiguity spatial weights matrix.

[13] P-values accounted for multiple testing using the false discovery rate adjustment in R.

ences, the characteristics of CCHS respondents from Toronto were most similar on factors associated with type 2 diabetes including ethnicity, birthplace, sex, and age.

Table 4.3 compares the fit of TropISM models developed using each subset of the CCHS. Computationally, models developed using the smallest subsets were the fastest to fit.[14] Models developed using larger subsets had smaller percentages of selected observations coming from the Toronto health region. In the model developed using the full Ontario subset, only 16% of selected observations were from the Toronto health region compared to 36% of observations in the GTA subset. This result is important because models developed using larger subsets were less able to reproduce unconstrained outcomes compared to models developed using smaller subsets (Appendix A).

Another interesting feature of the models fit using smaller subsets of microdata is that these models tended to produce larger weights. In the model developed from the full Ontario subset, 50% of observations were selected only once in any given neighbourhood, while 25% of observations were selected two to 23 times. This means that one microdata record was sampled up to 23 times to ensure that the total number of selected observations matched the total number of people in each constraint category for a given neighbourhood. In the model developed from the Toronto subset, 50% of observations were selected up to four times, while 25% of observations were selected between seven and 79 times. This result is expected because fewer observations were available in this subset, meaning more of the same observation had to be selected to ensure that the total number of selected observations matched the total number of people in each constraint category.

Table 4.3 summarizes the internal fit of each model. There were negligible differences between models in the total and standardized absolute error. Given that all models were developed using four univariate constraints and one bivariate constraint and that constraint tables were normalized prior to model development, this result is not surprising. It suggests there was sufficient variation in the

[14] Models were developed using the CO algorithm on a 64-bit Windows PC with an Intel Xeon 3.40 GHz CPU, 8 logical cores and 8 gigabytes of memory. Although the CO software was developed as a 32-bit application, it is possible to run it in parallel on a multi-core system.

Table 4.2. Demographic characteristics of the population of metropolitan Toronto in 2006 compared to the sample of respondents from different subsets of the 2005 Canadian Community Health Survey.

		CCHS*			
Characteristic	Census	ON	CMA	GTA	TOR
Sex					
Men	47.5	45.3	45.5	46.3	45.6
Women	52.5	54.7	54.5	53.7	54.4
Age, men					
20–44	51.5	44.8	48.6	52.9	52.9
45–64	32.2	33.4	31.7	29.7	27.9
65+	16.3	21.8	19.7	17.4	19.1
Age, women					
20–44	48.6	40.6	43.6	48.7	49.7
45–64	31.5	32.6	30.8	29.5	26.5
65+	19.8	26.8	25.6	21.7	23.8
Ethnicity					
Visible minority	46.4	12.4	16.3	28.3	37.7
Birthplace					
Immigrant	50.9	20.6	26.6	40.1	49.1
Education					
< Secondary	14.9	22.7	20.1	17.7	18.6
Secondary	32.5	25.3	25.3	24.8	22.0
Post-secondary	52.6	52.0	54.6	57.5	59.4
Personal income					
< $15,000	35.8	30.4	29.2	29.5	31.5
$15,000–$29,999	24.6	23.3	21.9	18.7	21.3
$30,000–$49,999	16.0	22.6	22.7	22.2	20.1
≥$50,000	23.5	23.7	26.2	29.6	27.0
Marital status					
Single	34.3	26.1	28.2	30.4	36.4
Married/common-law	50.2	53.9	52.8	53.4	46.5
Separated/widowed/divorced	15.5	20.0	18.9	16.1	17.1
Tenure					
Home is owned	56.5	73.9	72.2	72.2	56.4

*Subset of respondents from the 2005 CCHS: ON = Ontario, CMA = health regions containing a census metropolitan area, GTA = health regions from the Greater Toronto Area only, TOR = Toronto health region only.

Table 4.3. Fit of the TropISM model developed using different subsets of the 2005 Canadian Community Health Survey.

	Health Regions Included			
	Ontario	CMA	GTA	Toronto
Total sample	n = 33,380	n = 25,195	n = 9,224	n = 2,974
CO execution time	1 hr 18 min	1 hr 42 min	19 min	7 min
Selected observations*				
Toronto	15.6	18.5	35.8	100
GTA	24.0	29.3	64.2	—
Other	60.4	52.3	—	—
CO weights				
Median	1	1	2	4
Third quartile	2	2	3	7
Maximum	23	26	38	79
Assessment of fit†				
TAE‡				
Median	5	5	5	5
IQR	(3, 7)	(3, 7)	(3, 7)	(3, 7)
Range	(0, 16)	(0, 16)	(0, 16)	(0, 16)
SAE§				
Median ($\times 10^{-4}$)	3.0	3.0	3.0	3.0
IQR ($\times 10^{-4}$)	(2.0, 6.0)	(2.0, 6.0)	(2.0, 5.0)	(2.0, 5.0)
Range ($\times 10^{-3}$)	(0, 2.1)	(0, 1.2)	(0, 1.2)	(0, 1.2)

* Percentage of selected microdata observations coming from different health regions.

† Summary of fit across all neighbourhoods.

‡ Total absolute error.

§ Standardized absolute error.

constraint variables from each microdata subset for the CO algorithm to successfully replicate each neighbourhood population according to the constraints.

Figure 4.4 compares the standardized absolute error (SAE) for each constraint variable for all neighbourhoods across all models using parallel coordinate plots. There were only small differences between simulated data and census data for any constraint, regardless of the microdata subset used to develop TropISM. Models developed using the GTA and Toronto subsets seemed to fit slightly better, while some neighbourhoods had slightly higher errors when the Ontario and CMA subsets were used.

Across all subsets, the largest errors occurred in the sex by age constraint. Education and income showed fewer differences between simulated neighbourhood populations and the 2006 census, while ethnicity and immigrant populations were perfectly replicated across all models with the exception of the CMA subset. The standardized absolute error ranged from 0 to 0.0016 across all neighbourhoods for the sex by age constraint in the model developed using the full Ontario subset of microdata. This translates into differences of less than 0.2% between the simulated number of individuals in each sex-age group compared to the actual number of people in each sex-age group according to census estimates.

Model fit was also evaluated by plotting census population estimates for each neighbourhood against simulated estimates for all constraint categories. The standard error about identity (SEI) was estimated for each constraint category. Based on the Toronto subset, the SEI for all constraint categories was greater than 0.99, indicating nearly perfect fit between the simulated number of individuals in each constraint category and census counts for all neighbourhoods. In summary, the simulated annealing algorithm was able to accurately simulate the population of each neighbourhood according to the five constraints used to develop TropISM, regardless of which microdata subset was used.

4.3.2 Demographic characteristics of the simulated population

Table 4.4 summarizes the demographic characteristics of the simulated population within each of metropolitan Toronto's six boroughs. About 14% of the population in Toronto is 65 years of age or older, a somewhat smaller percentage than in

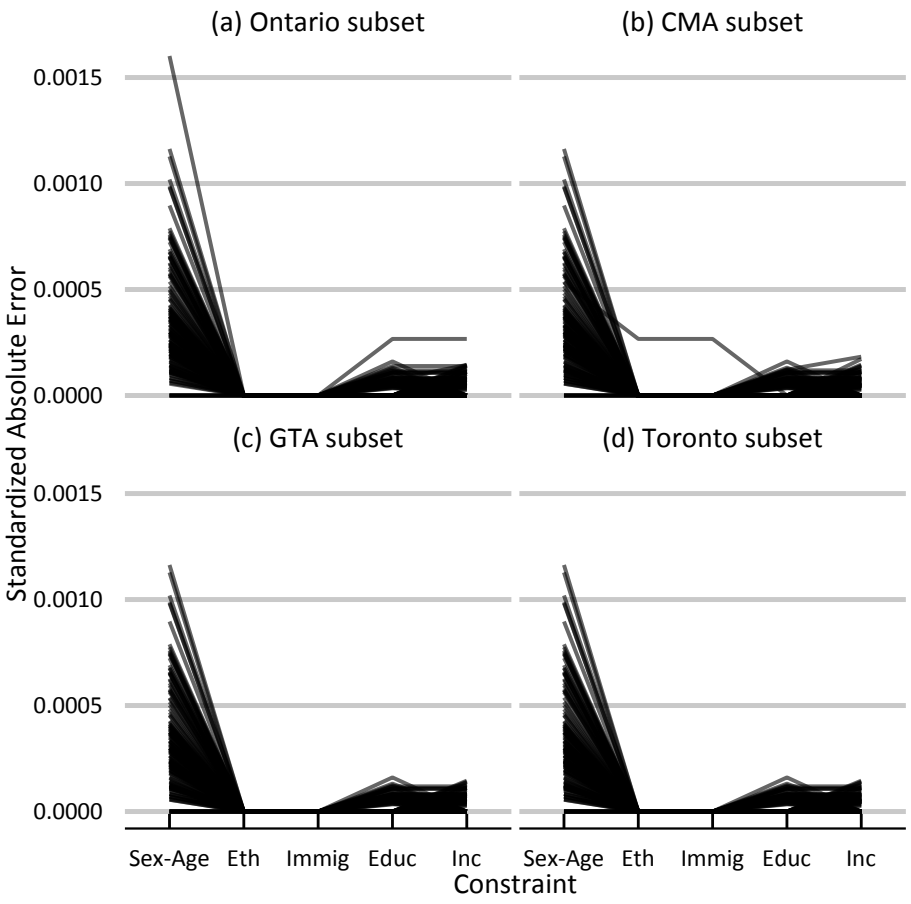


Figure 4.4. Assessment of TropISM model fit using the standardized absolute error to compare the difference between simulated neighbourhood populations and census data. Four models were developed using different subsets of the 2005 CCHS. In all models, constraints were: sex-age group ("Sex-Age"), ethnicity ("Eth"), immigrant status ("Immig"), education ("Educ"), and income ("Inc").

the other five boroughs. In Etobicoke, for example, almost 20% of the population is 65 or older.

Toronto's population is also better educated: 61% has a post-secondary education compared to only 41% of the population in York. Likewise, personal incomes tend to be higher in Toronto where 31% of residents have personal incomes greater than $50,000 per year. This is almost double the percentage in Scarborough and York. Unlike Toronto, Scarborough and North York are more ethnically diverse. In Scarborough, two-thirds of the population is comprised of visible minorities as is 52% of the population in North York. Immigrants make up more than half the population in York and almost 60% of the population in Scarborough and North York.

Table 4.4. Demographic characteristics of the simulated TropISM population by borough.

Characteristic	Etobicoke	North York	York	Toronto	East York	Scarborough
% Men age 65+	8.3	8.2	6.6	5.9	6.4	7.2
% Women age 65+	11.2	11.0	9.5	8.1	9.6	9.6
% Visible minority	37.7	51.9	42.5	30.3	37.9	67.3
% Immigrant	48.2	59.1	53.1	39.2	45.0	58.5
% Post-secondary*	49.9	53.0	40.6	60.9	54.2	46.3
% Income $50,000+†	25.0	21.7	17.2	30.7	24.8	17.3

* Percent having any post-secondary education.

† Percent having a personal income of $50,000 or more.

These borough level differences in population composition are reflected at the neighbourhood level. Noticeable spatial heterogeneity exists in the demographic profile of each neighbourhood (Appendix A, Figure A.1). More than 70% of residents in several of north Scarborough's neighbourhoods are visible minorities. In most of the remaining neighbourhoods, 50%–70% of residents are visible minorities. In Etobicoke, a north-south divide exists in the visible minority population. More than 31% of the residents in north Etobicoke are visible minorities while in south Etobicoke, fewer than 31% of the residents are visible minorities. In Toronto,

a similar divide is seen: waterfront neighbourhoods tend to have a sizeable percentage of visible minorities (31%–49%) while most of the remaining neighbourhoods are comprised of fewer than 31% visible minorities. Unsurprisingly, spatial patterns in the immigrant population are similar to spatial patterns in the visible minority population.

Two-thirds to three-quarters of residents living in central Toronto and central North York have a post secondary education. Likewise, many neighbourhoods in south Etobicoke have a well-educated citizenry where at least 45% of residents have a post-secondary education. Spatial trends in personal income tend to mimic the spatial patterns in post-secondary education. That is, neighbourhoods having a higher percentage of residents with a post-secondary education tend to have higher percentages of residents earning at least $50,000 per year.

Such differences in population composition are important to highlight from a population health perspective. As mentioned in Chapter 2, socioeconomic factors are related to a higher risk of type 2 diabetes. Furthermore, ethnic minorities have higher rates of type 2 diabetes. Thus, highlighting key differences in the demographic structure of neighbourhoods helps contextualize why some neighbourhoods have higher prevalence and forecast incidence of type 2 diabetes compared to others. However, prior to forecasting the incidence of type 2 diabetes at the neighbourhood level using the Diabetes Population Risk Tool (Chapter 5), it is necessary to validate the unconstrained outcomes of interest in the simulated TropISM population.

4.3.3 City-level validation of unconstrained outcomes

Although the CO algorithm accurately simulated neighbourhood populations according to each of the five constraints, models developed using different subsets of microdata varied in their ability to reproduce unconstrained outcomes at the city (health region) level. Generally, models developed using larger subsets of data were less able to accurately reproduce unconstrained outcomes. Appendix A compares simulated estimates aggregated to metropolitan Toronto against external estimates from the 2003, 2005, and 2007 Canadian Community Health Surveys, stratified by sex.

For men and women, the model developed using the Ontario subset of microdata showed larger differences in unconstrained outcomes between simulated estimates and CCHS estimates compared to the model developed using the Toronto subset of microdata. Based on the Ontario subset, 6.6% of simulated men had type 2 diabetes compared to an estimate of 4.9% from the 2003 CCHS (Table A.1). This represents an absolute difference of 1.7 percentage points but a relative difference of 34%. As might be expected, the TropISM estimate is closer to the 2005 estimate of 5.7% (absolute difference: 0.9 percentage points; relative difference: 15.9%). Compared to the 2007 CCHS, TropISM underestimates the prevalence of type 2 diabetes in men by 1.6 percentage points, or 19%.

Models developed using smaller subsets of microdata produced similar estimates of type 2 diabetes among men. The CMA subset produced the closest match in diabetes prevalence, where simulated prevalence was 6.4%. This represents an absolute difference of 1.5, 0.7, and -1.8 percentages points between TropISM and the 2003, 2005, and 2007 CCHS, respectively. For the model developed from the Toronto subset, the simulated prevalence of type 2 diabetes was also 6.6%.

Among women, the TropISM estimate of type 2 diabetes prevalence was closer to CCHS estimates. Based on the Ontario subset, the simulated prevalence of type 2 diabetes was 5.4%, compared to 5.2%, 4.3% and 3.6% when the CMA, GTA and Toronto subsets were used, respectively. Compared to the 2005 CCHS, TropISM overestimated the prevalence of type 2 diabetes among women, but tended to underestimate prevalence compared to estimates from the 2003 and 2007 CCHS. Based on the Toronto subset, TropISM underestimated the prevalence of type 2 diabetes by 1.3 and 3.1 percentage points, respectively. However, TropISM overestimated the prevalence of type 2 diabetes by 1.1 percentage points compared to the 2005 CCHS (a relative difference of 43%).

Although the external validation for type 2 diabetes shows mixed results for men and women, for other unconstrained outcomes, the model developed from the Toronto subset tended to produce the greatest similarity between simulated estimates and the CCHS. As might be expected, simulated estimates were most similar to CCHS estimates from 2005. Among men, all but one simulated estimate differed from known estimates by less than three percentage points in any given year. For BMI in the oldest age group, the simulated estimate differed by 3.8 per-

centage points.

In some cases, small absolute differences translated into large relative differences because the unconstrained outcomes occurred infrequently in the population. For example, the simulated prevalence of heart disease was 6.0% in men, which overestimated the known prevalence of heart disease in 2005 by 1.2 percentage points. Although this difference is small in absolute terms, in relative terms, it means the simulated estimate is 31.6% larger than the known estimate. Similar differences were seen for some of the BMI outcomes in both age groups. TropISM, in particular, seemed unable to accurately replicate the percentage of men classified as having an unknown BMI. However, these differences are less serious, because the Diabetes Population Risk Tool that was used to forecast the incidence of diabetes does not use unknown BMI status as a predictor of diabetes incidence in men.

Regardless of the microdata subset used, TropISM was unable to accurately replicate personal income when simulated estimates were compared against the 2003 CCHS. However, the 2003 CCHS used different cut-points to define income quantiles compared to the 2005 and 2007 surveys, so an inability to replicate this outcome is unsurprising.

Among women, the model developed from the Toronto microdata subset produced the most similar estimates of unconstrained outcomes at the metropolitan level. Simulated estimates of BMI prevalence differed by less than ±3 percentage points in the youngest age group, with the exception of the prevalence of underweight (BMI < 23) in 2005 (Table A.8). Among women aged 45–65, BMI prevalence was reasonably well simulated compared to CCHS estimates from 2005. Compared to 2003 estimates, TropISM overestimated the prevalence of underweight women in this age group and underestimated the prevalence of overweight women ($25 \leq$ BMI < 30). In the oldest age group, the largest differences between TropISM and CCHS estimates were found in the prevalence of healthy weight ($23 \leq$ BMI $<$ 25), obesity ($30 \leq$ BMI < 35), and unknown BMI.

4.3.4 Neighbourhood-level validation of unconstrained outcomes

At the neighbourhood level, TropISM did not accurately predict the prevalence of unconstrained outcomes. For diabetes, hypertension, and heart disease, it was possible to compare simulated counts and prevalence rates to known counts and prevalence rates ascertained from administrative health data. For diabetes and hypertension, TropISM under-predicted the total number of prevalent cases as well as prevalence rates (Tables 4.5 and 4.6, respectively). In any given neighbourhoods, TropISM predicted 926 fewer cases of type 2 diabetes, on average, compared to the known number of cases ascertained from the Ontario Diabetes Database. The extent of under-prediction was greater for women than men, such that TropISM predicted 548 fewer cases of type 2 diabetes among women compared to 378 fewer cases among men. The magnitude of under-prediction was greater for hypertension. For heart disease, however, TropISM predicted too many cases among men.

Since the TropISM model was developed from survey data, it is possible that some respondents who have diabetes or hypertension are unaware of their disease status. Therefore, crude corrections were made to the simulated prevalence of diabetes and hypertension, assuming that 30% of diabetes cases are undiagnosed and 17% of hypertension cases are undiagnosed (Campbell, McAlister, & Quan, 2013; Canadian Diabetes Association, 2015; Hux & Tang, 2003; Public Health Agency of Canada, 2011; Rosella et al., 2015). Crude corrections were also applied to the known number of cases of diabetes and hypertension, because the algorithms used to classify patients as diabetic or hypertensive have false positive rates of 3% and 5%, respectively (Hux et al., 2002; Tu et al., 2007). In addition, the Ontario Diabetes Database does not distinguish between type 1 and type 2 diabetes, so known rates were adjusted by an additional correction factor of 0.95 to reflect that 95% of all cases of diabetes are type 2 diabetes (Caspersen et al., 2012; Millar & Young, 2003; National Diabetes Information Clearinghouse (NDIC), 2011).

After these corrections were applied to simulated and known cases of diabetes and hypertension, the similarity between them improved. However, TropISM still predicted 591 fewer cases of diabetes and 1354 fewer cases of hypertension. In spite of these adjustments, only 13% of neighbourhoods fell within the 20% rel-

Table 4.5. Neighbourhood-level validation of the TropISM model: comparison of simulated counts and known counts of (a) diabetes, (b) hypertension, and (c) heart disease (men only).

	Diabetes						Hypertension						Heart disease
	Total		Men		Women		Total		Men		Women		Men
Measure	Cnt	Cor*	Cnt	Cor*	Cnt	Cor*	Cnt	Cor†	Cnt	Cor†	Cnt	Cor†	Cnt
Difference													
Mean	-926	-591	-378	-185	-548	-406	-1940	-1354	-928	-670	-1012	-684	126
(± SD)	(640)	(475)	(272)	(182)	(374)	(304)	(1136)	(894)	(564)	(459)	(585)	(455)	(73)
Relative Error													
% ±0.2	0	12.9	1.4	33.6	0	3.6	0.7	6.4	0	5.7	2.1	12.8	5.7
Concordance													
Overall ρ_c	0.360	0.598	0.515	0.804	0.225	0.373	0.464	0.649	0.403	0.573	0.516	0.706	0.678
Precision (ρ)	0.933	0.933	0.950	0.950	0.878	0.878	0.952	0.952	0.945	0.945	0.946	0.946	0.937
Accuracy	0.386	0.640	0.543	0.847	0.256	0.425	0.488	0.682	0.426	0.606	0.546	0.746	0.723
Scale shift	0.393	0.551	0.497	0.701	0.300	0.417	0.512	0.631	0.457	0.563	0.564	0.695	1.366
Location shift	-1.503	-0.877	-1.088	-0.486	-2.039	-1.382	-1.284	-0.850	-1.435	-0.982	-1.157	-0.742	0.820
Agreement‡													
Overall %	67.1		69.3		55.0		73.6		66.4		65.7		65.7
Cohen's κ	0.589		0.616		0.438		0.670		0.580		0.571		0.571

Cnt: simulated counts compared to known counts.
Cor:* simulated counts corrected for undiagnosed diabetes; known counts corrected for false positives & removal of type 1 diabetes.
Cor:† simulated counts corrected for undiagnosed hypertension; known counts corrected for false positives.
‡ Based on classifying neighbourhoods into quintiles according to simulated and known disease counts.

ative error rate for diabetes and only 6% of neighbourhoods fell within the 20% relative error rate for hypertension (Table 4.5).

The concordance correlation coefficient was then used to examine the accuracy and precision of simulated counts of unconstrained outcomes against known counts. Generally speaking, simulated counts of diabetes, hypertension, and heart disease were very precise, reflected in the high Pearson correlation between simulated and known counts (all correlations > 0.87). These high positive correlations indicate a strong increasing relationship between simulated counts and known counts. In spite of this, the overall concordance correlation between simulated and known counts was lower than the Pearson correlation, and the reason for this difference is that TropISM under-predicts true prevalence. In other words, simulated counts of diabetes and hypertension are downwardly biased. This downward bias is reflected in the negative value of the location shift parameters for diabetes and hypertension. Simulated counts of heart disease, on the other hand, are upwardly biased. Correcting simulated counts for undiagnosed cases and known counts for the false-positive rate generally improves the accuracy of simulated counts. This is especially true for simulated counts of diabetes among men, where the overall accuracy improves from 0.543 to 0.847, increasing the value of the concordance correlation from 0.515 to 0.804. Figure A.2 in Appendix A demonstrates this improvement in accuracy after these crude adjustments were applied, using prevalent cases of diabetes as an example. The left-hand column of Figure A.2 presents scatterplots of simulated counts against known counts of diabetes. The solid black line represents the concordance line, or perfect agreement. The dashed grey line represents the line of best fit between simulated and known counts while the dotted red line is the line of best fit after correcting for undiagnosed cases in the simulated counts and false positives in known counts. These plots demonstrate TropISM's high precision but low accuracy.

The right-hand column of Figure A.2 presents index plots of simulated and known counts of diabetes for each neighbourhood, ranked from lowest to highest, after applying the crude corrections. The shaded red area identifies a relative error of ±20% around known counts of diabetes and black dots represent simulated counts of diabetes for each neighbourhood. These plots demonstrate some important concepts. First, not only does TropISM under-predict the true num-

ber of prevalent cases within each neighbourhood, but the magnitude of underprediction is greater for women than for men. Second, the magnitude of underprediction increases such that neighbourhoods having a greater number of cases tend to have larger differences between simulated and known counts compared to neighbourhoods having fewer cases of diabetes. Smoothed trend lines for the simulated number of cases also suggest that the simulated number of cases across neighbourhoods tends to be too "flat". In other words, TropISM seems unable to replicate the variability in the prevalent number of cases of diabetes across neighbourhoods.

When neighbourhoods were ranked into quintiles according to simulated and known counts, a moderate amount of agreement was observed between the two classifications. Overall, 67% of neighbourhoods were classified into the same quintile using either simulated or known counts of diabetes (Table 4.5). For prevalent cases of diabetes among women, 55% of neighbourhoods were classified into the same quintile. Among men, 69% of neighbourhoods were classified into the same quintile. The overall percent agreement was similar for cases of heart disease among men: in that case, 66% of neighbourhoods were classified into the same quintile. The greatest amount of agreement in the neighbourhood classification was observed for total cases of hypertension. In that case, 74% of neighbourhoods were classified into the same quintile using either simulated or known counts of hypertension.

For each outcome, Cohen's κ statistics likewise suggest low to moderate levels of agreement in the neighbourhood classification of simulated and known counts of disease prevalence. κ values were highest for total counts of hypertension and for counts of diabetes among men (κ = 0.67 and 0.62, respectively) while the lowest agreement was observed for counts of diabetes among women. Based on these findings, neighbourhoods having low or high simulated counts of diabetes, hypertension, and heart disease tended to have low or high known counts of these outcomes (recognizing that TropISM incorrectly classifies some neighbourhoods).

Table 4.6 examines the similarity between simulated and known prevalence rates of diabetes, hypertension, and heart disease. It also examines the relationship between (a) simulated smoking prevalence and known incidence rates of lung cancer and COPD and (b) excess body weight (overweight + obesity) and can-

cers associated with obesity.[15] Relatively strong Pearson correlations were observed between simulated and known prevalence rates of diabetes (overall and for men), hypertension (overall and for women) and heart disease (men only). When prevalence rates were spatially smoothed, the strength of the Pearson correlation increased, suggesting similarity in the broad spatial patterns of simulated and known disease prevalence. The bivariate Moran's I index, a global assessment of spatial autocorrelation between simulated and known prevalence rates, also points to moderate agreement between rates. The bivariate Moran's I index was highest for diabetes prevalence rates among men (Moran's I = 0.62) and lowest for diabetes prevalence among women (Moran's I = 0.40). Not surprisingly, spatial smoothing increased the strength of the bivariate Moran's correlation coefficient, likely because the smoothing of prevalence rates induces a spatial structure that makes it easier to identify broad spatial patterns (Anselin, Lozano, & Kochinsky, 2006).

For smoking and body weight outcomes, there was no relationship between simulated prevalence rates of smoking and excess body weight and the known incidence rates of diseases that should be correlated with these risk factors (all Pearson correlations were below 0.26). These results are inconclusive: either TropISM could not accurately simulate the prevalence of these risk factors across Toronto's neighbourhoods or there is little similarity between current prevalence rates of these risk factors and current disease incidence. The latter possibility may be true given the long latent period between these risk factors and disease onset. In other words, in a given neighbourhood, the population affected by COPD and other cancers may differ substantially from the population that currently smokes or that is overweight. Thus, the only conclusion that can be reached from the lack of agreement between the simulated prevalence of risk factors and known disease incidence rates is that it cannot be determined whether TropISM accurately simulates spatial patterns of these risk factors across metropolitan Toronto.

When neighbourhoods were classified into quintiles based on simulated or known prevalence rates, the extent of agreement between the two classifications was somewhat lower than when prevalent counts were compared. Classification

[15] Kidney, colorectal, bladder, esophageal, pancreatic, and thyroid cancers, as well as endometrial and post-menopausal breast cancer in women and prostate cancer in men.

Table 4.6. Neighbourhood-level validation of risk factor prevalence against known disease prevalence (type 2 diabetes, hypertension, heart disease) or disease outcomes associated with risk factors (smoking, overweight + obesity).

			Rate Classification†				Rate Clustering	
Indicator	Correlation*		Raw		Spatial		Moran's I‡	
Simulated (Known)	ρ_r	ρ_{sp}	A_o	κ	A_o	κ	Raw	Spatial
Type 2 diabetes (diabetes§)								
Total	0.771	0.902	47.9	0.348	57.1	0.464	0.567	0.808
Men	0.817	0.947	52.1	0.402	67.9	0.589	0.620	0.859
Women	0.557	0.710	33.6	0.170	36.4	0.205	0.398	0.859
Hypertension (hypertension¶)								
Total	0.743	0.896	38.6	0.232	58.6	0.485	0.500	0.792
Men	0.631	0.824	39.3	0.241	41.4	0.268	0.411	0.727
Women	0.697	0.830	41.4	0.268	45.0	0.312	0.463	0.737
Heart disease (heart disease)								
Men	0.810	0.915	44.3	0.304	55.0	0.438	0.419	0.762
Smoking (lung cancer)								
Men	0.080	0.386	15.0	-0.062	27.0	0.089	0.140	0.351
Smoking (COPD)								
Men	0.129	0.305	22.1	0.027	22.1	0.027	0.151	0.270
Overweight + Obesity (total cancers)								
Men	0.256	0.072	15.0	-0.062	22.9	0.036	0.066	0.063
Women	0.128	0.013	20.0	0	20.0	0	0.029	0.151

* ρ_r = Pearson correlation, raw rates; ρ_{sp} = Pearson correlation, spatially smoothed rates.
† Rate classification based on classifying neighbourhoods into quintiles according to their raw and spatially smoothed disease rates. A_o = overall % agreement, κ = Cohen's kappa.
‡ Bivariate Moran's I.
§ Simulated prevalence corrected for undiagnosed cases; known prevalence corrected false positives and removal of type 1 diabetes cases.
¶ Simulated prevalence corrected for undiagnosed cases; known prevalence corrects for false positives.

of spatial rates improved overall agreement somewhat, especially for the prevalence of diabetes among men, where 68% of neighbourhoods were classified into the same quintile according to simulated or known prevalence rates (vs. 57% agreement for overall diabetes prevalence and only 36% for prevalence among women).

Figure A.3 in Appendix A compares the spatial patterns of overall diabetes prevalence in more detail. Panels (a) and (b) present raw simulated and known prevalence rates, respectively, while panels (c) and (d) present spatially smoothed rates of simulated and known prevalence. Although the absolute value of simulated prevalence in any given neighbourhood is lower than the true value, there is greater similarity in the spatial patterns of diabetes prevalence, as demonstrated in panel (e) of Figure A.3. Neighbourhoods shaded in grey are classified into the same quintile according to either simulated or known prevalence (57% of neighbourhoods). Neighbourhoods shaded in light blue were classified by TropISM into a lower quintile compared to the known rate (20% of neighbourhoods) while neighbourhoods shaded in light salmon were classified by TropISM into one higher quintile (20% of neighbourhoods).

Panel (f) depicts bivariate LISA statistics[16] identifying groups of neighbourhoods having similarly high or low simulated and known diabetes prevalence rates. Based on this map, there are several neighbourhoods in north Etobicoke, west North York, and Scarborough having high simulated and known prevalence. More specifically, neighbourhoods in these boroughs having higher simulated prevalence rates were surrounded by neighbourhoods having higher known prevalence rates. Conversely, some neighbourhoods in Toronto and south North York having lower simulated prevalence rates were surrounded by neighbourhoods having lower known prevalence rates. These results support the conclusion that TropISM was able to replicate broad spatial patterns in diabetes prevalence across metropolitan Toronto.

[16] Estimated using spatially smoothed rates.

4.3.5 Model uncertainty

Since the TropISM model was developed using the probabilistic simulated annealing algorithm, different runs of the algorithm produce different results. In order to assess aleatory uncertainty produced by different realizations of the model, 100 replicates of TropISM were developed using the CO software. Replicates were constructed from the Toronto subset of the 2005 CCHS, the run time parameters described in Section 4.2.4, and 100 unique random number seeds. For each replicate, the simulated prevalence of four chronic disease risk factors was estimated for men and women.

Figures A.4a–A.5d (Appendix A) illustrate the range of uncertainty about simulated prevalence of type 2 diabetes, hypertension, overweight ($25 \leq \text{BMI} < 30$) and obesity ($\text{BMI} \geq 30$). In each plot, black dots represent simulated prevalence estimates from the initial TropISM model for each of the 140 neighbourhoods.[17] Red dots represent the median prevalence estimate across all 100 replicates while the shaded area represents the range of simulated prevalence spanning the 2.5^{th} and 97.5^{th} percentiles (i.e., a 95% "uncertainty" interval). In each plot, neighbourhoods are ranked by the median simulated prevalence of each outcome.[18]

Based on these plots, it is clear that the range of uncertainty in simulated prevalence varies across neighbourhoods. Some neighbourhoods show a smaller range of simulated estimates than others. For example, among men, in half of the neighbourhoods, 95% of simulated estimates of type 2 diabetes differ by 0.56–1.12 percentage points. In 25% of neighbourhoods, 95% of simulated estimates differ by as much as 1.33–2.07 percentage points (corresponding to the wider uncertainty intervals depicted in Figure A.4a). In the neighbourhood having the widest uncertainty interval, simulated estimates of type 2 diabetes range from 6.62%–8.69%.

Among women, the simulated prevalence of type 2 diabetes shows a smaller range of uncertainty. In half of the neighbourhoods, 95% of simulated estimates differ by 0.46–0.71 percentage points. Even among three-quarters of all neigh-

[17] That is, one realization of 100 replicates of the TropISM model.

[18] In other words, the neighbourhood having the lowest simulated estimate for type 2 diabetes appears first in the plot. This also means that the neighbourhood appearing first in the diabetes plot for men does not necessarily correspond to the same neighbourhood for women.

bourhoods, 95% of simulated estimates differ by only 0.93 percentage points. In the neighbourhood having the largest uncertainty interval, 95% of simulated prevalence estimates range from 2.52%–4.03% (1.51 percentage points).

For the other chronic disease risk factors, there is a slightly larger range in the uncertainty of simulated prevalence. For hypertension, 95% of simulated estimates differ by 0.79–2.60 percentage points among men and women alike. In a "typical" neighbourhood, the median simulated prevalence of hypertension across 100 replicates of the model is 14.9% (95% UI: 14.0%, 15.7%). In other words, in this neighbourhood, 50% of simulated estimates are below 14.9% and 50% are above 14.9%. Ninety-five percent of simulated estimates lie between 14.0% and 15.7%.

With respect to body weight, the greatest degree of uncertainty in simulated estimates was observed for overweight men. Figure A.5a demonstrates that the median simulated prevalence of overweight ranged from a low of 30.5% to a high of 41.1% across all neighbourhoods. In half of these neighbourhoods, 95% of simulated estimates differed by 1.25–2.27 percentage points. In another 25% of neighbourhoods, 95% of simulated estimates differed by as much as 2.75 percentage points. However, in the remaining neighbourhoods, 95% of simulated estimates differed by as much as 4.1 percentage points. In the neighbourhood having the widest uncertainty interval, 95% of simulated estimates ranged from 32.4%–36.5% (median prevalence = 34.3%).

Even though the range of uncertainty in simulated prevalence varies across neighbourhoods, it should be pointed out that in most cases, the range of uncertainty is relatively narrow. Even the neighbourhood having the widest uncertainty interval for the simulated prevalence of overweight, the half-width of the interval is less than 2.2 percentage points. In survey-based research, statistical estimates are typically accurate to ± 3 percentage points, 19 times out of 20. Thus, simulated estimates for TropISM show a smaller range of uncertainty than is typically considered acceptable in other types of scientific research.

Another interesting result demonstrated by Figures A.4a–A.5d is that simulated estimates extending beyond the 95% uncertainty interval were infrequently observed. For type 2 diabetes among women (Figure A.4b), only ten neighbourhoods had a simulated estimate from the initial TropISM model that was more extreme

than 95% of all simulated estimates in those neighbourhoods.[19] In most cases, simulated estimates from any one realization of TropISM fall within the 95% uncertainty interval.

What is also evident from these results is that simulated estimates might be more or less similar to the "typical" or median estimate. For type 2 diabetes among men, a single realization of TropISM (black dots) produced estimates that were larger than the median estimate across 100 model replicates in 85 neighbourhoods (Figure A.4a). Among women, the opposite trend was seen: only 48 neighbourhoods had simulated estimates larger than the median for one particular replicate. For hypertension among women (Figure A.4d), estimates from a single realization of TropISM tended to cluster around the median estimate across all replicates in most neighbourhoods. In summary, a single realization of a spatial microsimulation model developed using the simulated annealing algorithm may infrequently produce extreme estimates of disease prevalence. Based on these results, however, a single realization seems to produce relatively "typical" results.

4.4 Discussion

These results demonstrate that it was possible, but difficult, to develop a spatial microsimulation model of several health outcomes at the neighbourhood level for metropolitan Toronto. Using a minimal set of constraint variables common to the 2006 Canadian census and the 2005 Canadian Community Health Survey (CCHS), the TropISM model simulated chronic disease risk factors (current smoking, body mass index, and hypertension) and outcomes (type 2 diabetes and heart disease) across Toronto's 140 neighbourhoods. Four separate models were developed using different subsets of the 2005 CCHS using the simulated annealing algorithm.

With respect to internal model validity or "model fit", no one subset of microdata proved better than another. Overall error was similar across all models, as indicated by the total absolute error. Standardized estimates of absolute error were only negligibly smaller for models developed using smaller subsets of microdata. Furthermore, estimates of the standard error about identity were very high for all models (> 0.999 in almost all cases). Thus, regardless of which microdata subset

[19] Identified by the black dots extending beyond the shaded area of Figure A.4b.

was used to develop TropISM, there was sufficient variation in the sample of observations present in each subset to replicate neighbourhood populations with a high degree of accuracy, at least according to the constraints used for model development.

This high degree of accuracy was likely produced by standardizing the constraint tables prior to model development. In other words, in all neighbourhoods, constraint tables were standardized so the total population count across all constraint categories was the same. Constraints were standardized to limit the variability in population counts across constraint tables that arises from different age groups used in the reporting of Canadian census data at the census tract level. In other words, some variables were already based on the population aged 15 and older (e.g., the sex by age constraint and the education constraint), while others were based on the entire population (e.g., visible minority status). Standardization to the total population aged 15+ reconciled these differences.

One advantage of developing the TropISM model with smaller subsets of microdata is that these models were computationally faster to fit, unsurprising given the simulated annealing algorithm has fewer observations to resample when fitting the model. Related to this is the finding that the maximum number of microdata data observations resampled for the simulated population increased dramatically from 23 in any one neighbourhood when the full Ontario subset of microdata was used to 79 when the Toronto subset was used. This observation is consistent with earlier work demonstrating that larger microdata samples provide a greater pool of diverse observations needed to fit small areas that differ from the overall average or "typical" small area (Hermes & Poulsen, 2012a; Huang & Williamson, 2001; van Leeuwen, 2010a). For the TropISM model, larger subsets of microdata required a smaller number of the same observations to simulate neighbourhood populations according to census estimates.

Indeed, for the model developed using the Toronto health region subset, neighbourhoods having higher weights (n = 27) differed systematically from neighbourhoods having smaller weights (n = 113).[20] Neighbourhoods having higher weights

[20] Neighbourhoods having high weights were defined as those neighbourhoods whose median weight was > 10. This meant that 50% of microdata observations resampled by the CO algorithm were selected at least 11 times to simulate the population in that neighbourhood.

tended to have a greater concentration of ethnic minorities (63% vs 38%) and immigrants (60% vs 46%) than neighbourhoods having smaller weights. Neighbourhoods having higher weights also had a greater concentration of low income.[21] In other words, a greater number of the same microdata observations were required to simulate the demographic structure of neighbourhoods having atypical populations; in this case, a greater proportion of visible minorities, immigrants, and low income residents.

Although previous spatial microsimulation research has found that larger subsets of data tend to produce better fitting models with respect to constraint variables (e.g., Hermes & Poulsen, 2012a; Huang & Williamson, 2001; van Leeuwen, 2010a), Tanton et al. (2010) note that limiting microdata observations to the area being fit may produce better fitting models. This idea is consistent with the approach used by Koh et al. (2015) who developed a spatial microsimulation model of obesity in Detroit using the subset of respondents from the 2010 Behaviour Risk Factor Surveillance System residing in the Detroit Tri-County Metropolitan Area. Their results showed that simulated obesity estimates aggregated to the county level differed by 4.4 percentage points or less.

The development of TropISM used a similar approach. Progressively smaller subsets of CCHS microdata produced better fitting simulated estimates of chronic disease risk factors and outcomes aggregated to the metropolitan level. Compared to direct survey-based estimates, simulated estimates from the model developed using the full Ontario subset of microdata differed from −7.1 percentage points to +4.9 percentage points among men. A similar range was seen among women, where simulated estimates of type 2 diabetes, hypertension, and body mass index differed by −8.4 percentage points to +4.2 percentage points.

When TropISM was developed using the smaller Toronto subset of microdata, the absolute magnitude of differences between simulated estimates and direct survey-based estimates decreased. Among men aged 45 or older, simulated estimates of overweight ($25 \leq \text{BMI} < 30$) underestimated the survey-based estimate by 4.4 percentage points. In this same group, simulated estimates of underweight ($\text{BMI} < 23$) overestimated the survey-based estimate by 3.8 percentage points.

[21] Personal incomes < $15,000/year.

Slightly smaller differences were seen among women: simulated estimates of underweight among women aged 20–44 underestimated the survey-based estimate by 3.2 percentage points. Among women aged 45–64, the simulated estimate of underweight overestimated the survey-based estimate by 2.9 percentage points.

This improvement in fit may be partially explained by the sample of microdata observations selected for the simulated population. When the full Ontario subset of microdata was used to develop TropISM, only 16% of the observations were drawn from respondents residing in Toronto. The majority (60%) were drawn from respondents residing outside Toronto and the Greater Toronto Area. As a result, aggregate prevalence estimates of unconstrained health outcomes will be influenced to a greater extent by the observations coming from areas outside Toronto. The corollary of this is that when TropISM was developed using just the Toronto subset of microdata, only those observations influence aggregate prevalence estimates, producing estimates that are more similar to direct survey-based estimates.

Although there is no standard acceptable margin of error in survey-based research and public opinion polls (Ferber, Sheatsley, Turner, & Waksberg, 1980), it is common to hear poll results reported as accurate to within ±3 percentage points 19 times out of 20 (Mendelsohn & Brent, 2001). At the aggregate health region level, most simulated estimates of health outcomes differed from direct survey-based estimates by less than this margin of error. Indeed, this range of error is similar to that found in previous spatial microsimulation studies of health outcomes (Burden & Steel, 2015; Clark et al., 2014; Edwards et al., 2011; Koh et al., 2015).

Although TropISM replicated unconstrained outcomes at an aggregate level, at the neighbourhood level, TropISM was less accurate. Compared to administrative health data, TropISM under-predicted the true prevalence of diabetes and hypertension at the neighbourhood level while it over-predicted the prevalence of heart disease among men. These limitations are addressed in more detail below. However, this study used the concordance correlation coefficient to assess the similarity between simulated counts of disease outcomes and the known number of cases. This is a novel approach, differing from previous studies using measures of relative error (Smith et al., 2011), the coefficient of determination (R^2) (Edwards et al., 2011), and, more recently, Bland-Altman plots (Timmins & Edwards, 2016).

First Smith et al. (2011) assessed the accuracy of their spatial microsimulation model of smoking prevalence in New Zealand using a relative measure of error. Specifically, they compared the simulated prevalence of smoking against the known prevalence estimated from the census using simple relative differences. They found that the simulated prevalence differed from known prevalence by 10% or less in 75% of small areas.

Edwards et al. (2011) used a different approach, comparing the precision of their spatial microsimulation model of obesity. Specifically, they correlated simulated counts of obesity against known counts of cancers associated with obesity at the small area level using linear regression. The coefficient of determination (R^2) from these models suggested their results were very precise, so that areas having low simulated counts of obesity had low counts of cancers while areas having high simulated counts had high known counts.

Finally, Timmins and Edwards (2016) proposed using Bland-Altman plots to validate spatial microsimulation models. Bland-Altman plots are commonly used in clinical studies to assess agreement, via mean differences, between two measurements Bland and Altman (1999). When a new measurement is compared against a "gold standard", these plots help identify whether the new measurement agrees with the gold standard or whether it may be systematically biased. With respect to the validation of spatial microsimulation models, Timmins and Edwards (2016) argue that Bland-Altman plots can identify small areas that were both more and less accurately simulated.

Unlike those studies, this study used the concordance correlation coefficient to assess the accuracy *and* precision of spatial microsimulation results. These results suggest that while simulated counts of diabetes, hypertension, and heart disease were precise (Pearson correlation > 0.87), simulated counts were less accurate, such that the accuracy component of the concordance correlation often fell below 0.6. Crude adjustments to simulated counts that accounted for the potential number of undiagnosed cases, as well as adjustments to known cases that accounted for false positives that accrue in the administrative source data, improved overall accuracy. When possible, the concordance correlation coefficient should be used to validate unconstrained outcomes for future studies employing spatial microsimulation methods.

Finally, this research assessed aleatory uncertainty surrounding unconstrained outcomes produced by the probabilistic simulated annealing algorithm. Few previous studies have attempted to quantify this uncertainty. Indeed, Williamson (2007) argues that minor stochastic variation in simulated outcomes produced by multiple runs of the CO software can be ignored. Echoing this thought, Hermes and Poulsen (2012b) and Ma et al. (2014) found only minor differences in unconstrained outcomes between multiple runs of the CO software. The results reported here contest those findings. First, the magnitude of aleatory uncertainty surrounding simulated estimates of disease prevalence varies across neighbourhoods. In some neighbourhoods, 95% of simulated estimates differed by as little as 0.46 percentage points. In other neighbourhoods, 95% of simulated estimates differed by as much as 4.1 percentage points.

Second, the extent of aleatory uncertainty varies depending on the particular unconstrained outcome being estimated. Uncertainty intervals tended to be narrower for simulated estimates of type 2 diabetes than they were for simulated estimates of overweight.

Finally, a single run of the simulated annealing algorithm does not appear to produce consistent simulated estimates across neighbourhoods. In some neighbourhoods, a simulated estimate may lie below the median estimate from multiple runs while in other neighbourhoods, a simulated estimate may lie above the median. Thus, any one estimate from a single run of a probabilistic algorithm may lie closer to or further from the true neighbourhood estimate. These differences across neighbourhoods might potentially exaggerate spatial patterns observed when simulated estimates are plotted cartographically.

Based on these observations, and echoing Clarke (1996), spatial microsimulation models developed using any probabilistic reweighting algorithm should present the range of aleatory uncertainty associated with unconstrained outcomes. This acknowledges that simulated estimates of unconstrained outcomes are just that, estimates that can not be known with certainty. Larger uncertainty intervals in some small areas might result from how well constraint variables predict unconstrained outcomes. If the predictive ability of the constraints is poorer in some areas compared to others, it is possible that random selection of microdata observations lacks the sensitivity needed to select appropriate observations. This could

produce greater differences in the simulated population across multiple runs of the simulated annealing algorithm. Future research should identify the circumstances that produce greater uncertainty in simulated estimates of unconstrained outcomes at the small area level.

4.4.1 Limitations

The TropISM model was developed using data from the 2006 Canadian Census. Although these data are currently 10 years old, 2006 was the last official census conducted using the mandatory long form. In 2011, the long form census was replaced with the National Household Survey (NHS), a large but voluntary probability survey. The NHS has been shown to have non-neglibible biases, including non-response bias. Overall, the NHS obtained a 68.6% response rate, comparable to other probability based surveys (Statistics Canada, 2015). Non-respondents typically differ from respondents, and Statistics Canada has determined that statistical estimates of visible minority groups, immigrants, and educational attainment will be biased (Statistics Canada, 2015). Since these factors were used as constraint variables to develop TropISM, population estimates from the NHS might have biased the simulated population. In particular, Statistics Canada determined that the NHS might over-estimate the population having a post-secondary education and the total number of south Asian visible minorities while under-estimating the total immigrant population (Statistics Canada, 2015). Since each of these factors influence the risk of type 2 diabetes, simulated small area estimates could be affected by the non-negligible biases present in the 2011 NHS. Therefore, although the dated 2006 Census was used to develop TropISM, a decision was made to use the most reliable census data currently available.

Related to this is the use of the 2005 Canadian Community Health Survey as the source of detailed microdata. In particular, there is a potential mismatch between incomes reported in the 2006 Census and the 2005 CCHS. In particular, the Census asks about income for the year ending December 31, 2005, while the CCHS asks about income in the last 12 months. Depending on when CCHS respondents completed the survey, their reported incomes may be a better reflection of their income in 2004. Although CCHS incomes could have been adjusted for inflation,

given that incomes reported in the CCHS refer to the past 12 months rather than the 2004 calendar year, and given that the publicly released microdata report incomes in broad categories, such adjustments were deemed impractical to implement. These adjustments may not have produced meaningful differences the final simulated population.

Another caveat that might bias the simulated population and associated small area estimates derived from it is that income had to be imputed for a large number of microdata records (n = 4,308 in the full Ontario subset). Incomes were imputed using a regression procedure prior to developing TropISM as a function of Ontario health region, sex, age, ethnicity, immigrant status, and education. If these variables were poor predictors of income, then respondents whose incomes were imputed might systematically bias the final simulated population. However, these variables should be reasonably correlated with income. For example, incomes tend to increase with age and higher levels of education, whereas women tend to have systematically lower incomes than men.

In addition, if the CO software sampled a large number of observations whose incomes were imputed, and if the variables used to predict income had low predictive ability, the final simulated population might be systematically biased. Overall, 10.8% of the simulated population was derived from observations whose incomes were imputed. As might be expected, the percentage of simulated records whose incomes were imputed varied across neighbourhoods, from a low of 8.1% to a high of 12.4% (median = 10.7%, median absolute deviation = 4.9%).[22] Given the vast majority of simulated records were not drawn from respondents whose incomes were imputed, any bias that might result from ineffective imputation should be minimal.

Related to this, unconstrained outcomes are based on self-reported survey data. For example, type 2 diabetes status in the CCHS microdata was classified using the Ng-Dasgupta-Johnson algorithm for population health surveys. Although this algorithm may misclassify type 2 diabetes status, the extent of misclassification is likely minor (Ng et al., 2008). In spite of this, any biases present in the survey microdata could be corrected prior to developing spatial microsimulation

[22] For symmetric distributions, the interquartile range is approximately equal to the median ± the median absolute deviation (Jones, 2002).

models to improve the accuracy of simulated estimates. Given the Ng-Dasgupta-Johnson algorithm does not correct for *undiagnosed* cases of diabetes, results from the TropISM model demonstrate that simulated small area estimates of diabetes may need to be calibrated to produce more accurate estimates that possess the correct amount of geographic variability across all areas. The crude corrections employed in this study reduced differences between simulated prevalence and known prevalence of diabetes at the neighbourhood level.

Body mass index, used to define overweight and obesity in the CCHS microdata, is also based on self-reported information. Previous research has shown that prevalence estimates of overweight and obesity based on self-reported height and weight tend to be under-estimated. This is because height tends to be over-reported while weight tends to be under-reported. Under-estimates of overweight and obesity may also vary by gender, age, and ethnicity (Li et al., 2009): in one study, underestimation was greatest among older women (age 60+) and Mexican Americans, with slight underestimation in older men (age 60+) (Gillum & Sempos, 2005). Given that self-reported data were used to develop TropISM, it is likely that simulated rates of overweight and obesity reported here underestimate true prevalence rates.

In addition, if self-reported weight and height produced biased estimates of body mass index that varied by sex, age, and visible minority status, then it is possible the nature of that bias differs by neighbourhood. For example, if the proportion of visible minorities varied importantly between two neighbourhoods, then bias in simulated prevalence rates of overweight and obesity would also differ between these neighbourhoods. It is quite likely that bias in simulated prevalence rates is greater in neighbourhoods containing larger proportions of older women and visible minorities. Unfortunately, neighbourhood-specific prevalence rates of overweight and obesity could not be validated in this study using incidence rates of cancers associated with obesity. Therefore, based on these results, it cannot be determined whether the simulated prevalence of overweight and obesity across neighbourhoods were differentially biased.

At the neighbourhood level, validation of unconstrained outcomes showed that TropISM (a) under-predicts the true prevalence of type 2 diabetes and hypertension, (b) over-predicts the true prevalence of heart disease, and (c) fails to

capture the variability in these outcomes across neighbourhoods. These results demonstrate that is insufficient to rely solely on an aggregate-level validation of spatial microsimulation results. The neighbourhood-level validation conducted in this study provides insights into how simulated small area estimates differ from known estimates derived from external administrative data sources. For example, systematic differences were observed between neighbourhoods where the simulated prevalence of diabetes was three or more percentage points *lower* than the known estimate. Specifically, neighbourhoods whose simulated prevalence was three or more percentage points lower tended to be comprised of greater percentages of visible minorities (46% vs 31%, respectively) and immigrants (53% vs 41%, respectively).

These results corroborate earlier research demonstrating that small area estimates of outcomes across spatial units do not accurately reflect the true geographic variability in these outcomes (Birkin & Clarke, 2012). Part of the reason may be that the combinatorial optimization algorithm could not capture the variability in disease outcomes in some neighbourhoods because that variability may have been absent in the source microdata. In other words, in atypical neighbourhoods comprised of higher proportions of visible minorities and immigrants, the algorithm selected greater numbers of these types of observations to represent those neighbourhood populations. If the selected observations were unaware of their true disease status (e.g., undiagnosed diabetes or hypertension) then prevalence estimates of those outcomes will be downwardly biased in those neighbourhoods.

Burden and Steel (2015) emphasize that if the goal of spatial microsimulation is to represent the geographic diversity in unconstrained outcomes, then it is important to retain the spatial structure of the population during model development. Many authors recommend using a larger array of constraint variables to develop these models while Smith et al. (2009) suggest developing local microsimulation models that consider a different ordering of constraint variables for the iterative proportional fitting method. Because TropISM was developed using the combinatorial optimization algorithm and a limited number of constraints (sex crossed with age group, ethnicity, immigrant status, education, and personal income), it may not have been possible to replicate the spatial variability in uncon-

strained outcomes. However, if additional constraints were used, especially those involving visible minority status, it may have been possible to better capture the spatial variability in unconstrained outcomes. In this case, additional constraints might include visible minority status crossed with sex, age group, income, and education.

Another way of improving both the accuracy and geographic variability of spatial microsimulation results is through calibration of unconstrained outcomes at the small area level. Morrissey and O'Donoghue (2011) and Morrissey et al. (2014) used logistic regression to calibrate their spatial microsimulation of the Irish local economy to known external targets of unconstrained outcomes. This type of calibration faithfully reproduced the spatial variability in unconstrained outcomes. In the current study, however, calibrating the full set of unconstrained outcomes to known totals would be more difficult because it requires a complete breakdown of the total population possessing each single outcome *as well as* the total population possessing multiple outcomes. Alternatively, another approach that might prevent under-estimation of unconstrained outcomes would involve adjusting the source microdata prior to model development. Li et al. (2009) used this approach prior to estimating a statistical small area estimation model of obesity across 398 communities in Massachusetts.

In spite of these shortcomings, there was some evidence that TropISM was able to replicate broad spatial patterns in the prevalence of some unconstrained outcomes (e.g., type 2 diabetes among men). TropISM seems to be able to correctly identify neighbourhoods having relatively higher or lower disease prevalence. Depending on how spatial microsimulation results are used, this might be acceptable. In other words, if the goal is to produce accurate estimates of chronic disease risk factors and outcomes, then TropISM in its current form produced unacceptable results. However, if the goal is to identify neighbourhoods across metropolitan Toronto that have relatively low or high disease prevalence in order to support public health planning decisions, such as targeting health promotion programs to neighbourhoods having relatively higher disease prevalence, then TropISM provides useful insights because it was able to identify broad spatial patterns of type 2 diabetes at the neighbourhood level.

It is worth mentioning that the TropISM model simulated *neighbourhood* pop-

ulations within metropolitan Toronto. Although more granular levels of census geography could have been used, the choice to model neighbourhood populations was influenced by two factors. First, the neighbourhood classifications used for this research were developed by the City of Toronto and are used by local government and community agencies for socioeconomic planning purposes. Since neighbourhood boundaries remain fixed over time, it is possible to identify longitudinal trends at a geographic level, thus permitting comparison of demographic, socioeconomic, and health indicators over time.

Second, the City releases other indicators relevant for health promotion planning purposes at the neighbourhood level. Such measures include neighbourhood (a) walkability, (b) access to healthy food stores, and (c) green space (Social Policy and Research, 2014). These measures, used in conjunction with simulated prevalence estimates of overweight and obesity, could help identify neighbourhoods having higher than average rates of overweight and obesity but lower than average measures of health promoting urban environments. Examining the spatial coincidence of such outcomes permits additional health promotion planning at the neighbourhood level beyond the realm of diabetes prevention.

One final limitation to consider is the modifiable areal unit problem. Different spatial patterns may have been observed if different spatial units were used for the microsimulation. Small area populations could have been simulated for different spatial units to examine the robustness of spatial patterns in simulated prevalence estimates. Doing so, however, was deemed impractical for this research, as it would have required validating model outputs to at least two spatial units. Moreover, neighbourhoods were designed by the City of Toronto to remain fixed over time and are used for local planning purposes. Since the spatial unit of analysis has legitimate uses external to TropISM, using neighbourhoods to simulate small area populations for Toronto facilitates comparison with other policy relevant research for the City. Thus, although the modifiable areal unit problem might produce different patterns of results, model results are directly relevant for community health planning and policy development.

Before using the simulated TropISM population to forecast the incidence of diabetes or for conducting hypothetical “what if” policy simulations, it is important to consider how systematic differences between the simulated population and

the true population might affect such forecasts and experiments. At the neighbourhood level, TropISM underestimates the prevalence of type 2 diabetes and hypertension in men and women and overestimates the prevalence of heart disease in men. Such differences may influence forecasts of diabetes incidence using the Diabetes Population Risk Tool. Although neighbourhood-level prevalence estimates of overweight and obesity could not be validated, aggregate estimates across metropolitan Toronto appeared to underestimate true prevalence. Such differences may influence forecasts of diabetes incidence. For example, if the neighbourhood prevalence of obesity is underestimated, then forecasts of diabetes incidence may also be underestimated. Moreover, if the prevalence of risk factors was differentially biased across neighbourhoods, this might produce more accurate forecasts of diabetes incidence in some neighbourhoods compared to others. These issues are further addressed in Chapter 5.

4.4.2 Considerations for future research

The field of spatial microsimulation research has grown over the last two decades into a scientific discipline in its own right. Additional methodological studies will improve understanding of the conditions under which spatial microsimulation models can produce valid and reliable results. Since the primary purpose of spatial microsimulation is to predict unknown outcomes at the small area level from a set of known constraint distributions supplied by census data (Haslett et al., 2010), it is important to understand how well those constraints predict unknown outcomes. Another way to assess the potential predictive ability of constraint variables prior to simulating small area populations would be to estimate prediction error associated with those constraints. In this context, cross-validation methods developed in statistics might provide additional insights into the predictive ability of a set of constraint variables.

Cross-validation is often used in predictive modelling studies, where the primary purpose of modelling is to predict new observations from a statistical model developed using different data (Shmueli, 2010). Cross-validation methods typically randomly split a dataset into five or ten subsets of approximately equal size.[23]

[23] More generally known as k-fold cross-validation, where k represents the number of subsets used.

Model fitting is conducted on each of the subsets, with one of the subsets held aside for model testing. In other words, the parameter estimates for the model are estimated using $k-1$ subsets and new observations are predicted from the estimated model using the subset that was held aside. The prediction process is repeated for each subset so that each observation in the full dataset is used for prediction once. Prediction error is then the average of the error across k subsets (Sainani, 2014; Shmueli, 2010). In a spatial microsimulation setting, once a set of constraint variables have been chosen to develop the model, if sufficient data were available (e.g., a second external dataset), the predictive ability of the chosen constraints could be assessed using cross-validation methods. Estimates of prediction error, albeit at an aggregate level, might suggest how well constraint variables predict unconstrained outcomes.

In addition, results from the TropISM model illustrate that it is difficult to reproduce the geographic variability in unconstrained outcomes using a limited set of constraint variables for model development. Future methodological studies need to investigate how this variability could be incorporated into spatial microsimulation models. In this study, model constraints were selected on the basis of regression models and the proportion of variance explained in unconstrained outcomes. Selected constraints were reasonably correlated with the unconstrained outcomes of interest. However, constraint selection was based on the subset of Ontario respondents from the 2005 Canadian Community Health Survey. At an aggregate level, these constraints might predict outcomes reasonably well, but at a small area level, prediction might be better in some areas than others. Although broad spatial patterns seem to be replicated, there is no way to know for certain whether the predictive ability of the constraint variables differed across neighbourhoods within Toronto. If the predictive ability was worse in some areas compared to others this may have contributed to the insufficient geographic heterogeneity observed in these results.

4.4.3 Summary

Although TropISM under-predicted the true prevalence of risk factors at the neighbourhood level, as will be shown in the next chapter, the simulated TropISM pop-

ulation was still useful in forecasting the five-year incidence of diabetes. In addition, although TropISM underestimates the true prevalence of type 2 diabetes and does not capture the geographic variability in diabetes prevalence across neighbourhoods, the model appeared to capture broad spatial patterns in the prevalence of risk factors throughout metropolitan Toronto. In other words, neighbourhoods having low known prevalence of type 2 diabetes tended to have low simulated prevalence while neighbourhoods having high known prevalence tended to have higher simulated prevalence. Thus, TropISM results can be used for identifying relative inequalities in important disease outcomes, including type 2 diabetes, hypertension, and heart disease in men. A better understanding of relative differences in disease outcomes at the neighbourhood level can help target health promotion programs to neighbourhoods where these conditions are more prevalent. However, the development of spatial microsimulation models in future studies should aim to improve the accuracy of these models and develop ways of preserving the geographic variability in simulated outcomes, perhaps using local models of spatial microsimulation (Smith et al., 2009) or by using a more detailed set of constraints that preserves the geographic variability in simulated outcomes (Burden & Steel, 2015).

Chapter

5

Projecting The Local Effects of Diabetes Prevention

5.1 Rationale

Randomized controlled trials of lifestyle modification programs demonstrate that weight loss in high-risk individuals significantly reduces the risk of developing type 2 diabetes (Knowler et al., 2002; Kosaka et al., 2005; Tuomilehto et al., 2001). Tuomilehto et al. (2001) found that participants in a lifestyle intervention program lost, on average, 4.2kg (±5.1kg) of their original body weight while forty-three percent of intervention participants lost more than 5% of their original body weight. Knowler et al. (2002) found that 50% of intervention participants lost at least 7% of their original body weight by the end of the intervention. In many of these trials, the intended weight loss for participants in the intervention conditions was between 5–7% of their original body weight. A recent meta-analysis of translation studies demonstrated that high-risk individuals were able to lose, on average, 4% of their baseline body weight after participating in community-based interventions for at least one year (Ali et al., 2012).

Although participants in randomized controlled trials can achieve clinically meaningful reductions in body weight, in community settings, these reductions are more difficult to attain. Moreover, adherence to interventions may be difficult

to sustain over longer periods of time. Intuitively, greater attendance and compliance with lifestyle intervention programs should result in greater weight loss. Ali et al. (2012) back this up empirically in their meta-analysis of community-based interventions. However, losing meaningful amounts of body-weight is challenging in community settings. Even in rigorous randomized studies, only a minority of participants may lose sufficient weight to significantly lower their risk of developing type 2 diabetes (Tuomilehto et al., 2001).

Since diabetes and its risk factors are not evenly distributed over geographic space (Booth et al., 2007; Creatore et al., 2007b), there is a need to project the local effects of community-based interventions designed to promote weight loss and lower the population risk of type 2 diabetes. Such projections may identify areas that respond better to interventions and those areas that might need additional attention or different approaches to diabetes prevention. For example, some areas may have populations comprised of greater proportions of seniors who might need additional help in achieving meaningful weight loss. It is possible to obtain such local projections by combining TropISM model outputs with the Diabetes Population Risk Tool (DPoRT).

The Diabetes Population Risk Tool is a population-level risk algorithm that predicts future incidence of diabetes as a function of risk factors that are routinely assessed in population health surveys (Manuel et al., 2013; Rosella et al., 2014, 2011). This tool was developed by linking data from the 1996/1997 Ontario subset of the National Population Health Survey to a provincial registry of physician-diagnosed diabetes. Using a Weibull accelerated failure time model, the risk of developing diabetes was estimated separately for men and women. Non-white ethnicity, being an immigrant, older age, and higher body mass index reduced the time to developing diabetes while having a post-secondary education delayed its onset. Among men, heart disease also reduced the time to disease onset. In other words, men and women who possess a number of risk factors tend to develop diabetes sooner than men and women who do not possess those same risk factors.

Rosella et al. (2011) validated the DPoRT model using Ontario data from the 2000/2001 Canadian Community Heath Survey and the Manitoba subset of the 1996/1997 National Population Health Survey. The incidence of diabetes in each

sample was predicted using the mean-centred linear predictor η.[1] In the Ontario validation sample, the five-year predicted risk P_5 was estimated while in the Manitoba validation sample, the nine-year predicted risk P_9 was estimated, according to

$$P = 1 - \exp(-\exp^{m}) \tag{5.1}$$

where:

$$m = \frac{\log(d) - \eta}{\text{scale}} \tag{5.2}$$

and:

"d" is the duration of follow-up time in days, either five or nine-years,

η is the mean-centred linear predictor, and

"scale" is the scale parameter from the Weibull accelerated failure time model. Predicted rates were compared to observed rates from 2005 in both Ontario and Manitoba. Predicted rates differed from observed rates by 0.4 percentage points or less. The authors concluded that population rates of diabetes incidence can be accurately predicted using the DPoRT model combined with publicly available survey data.

Manuel et al. (2013) then used this model to predict how incidence of diabetes changes as a function of different prevention strategies among a cohort of Ontarians aged 20 and older. They examined two strategies: (a) a population-level strategy where body mass index was uniformly reduced in the entire population and (b) a high-risk strategy where pharmacotherapy or lifestyle counselling was used to treat high-risk individuals. Data from the Ontario subset of the 2003 CCHS were used to develop risk projections. If the status quo were maintained, the five-year risk of developing diabetes across the entire population was 4.7%. A 10% reduction in future incidence could be obtained if BMI was 3.5% lower across the entire population, if lifestyle counselling was successfully delivered to 32% of the high-risk

[1] The DPoRT model uses a binary classification for all risk factors. Mean values for each covariate are the proportion of respondents in the sample possessing each risk factor. If 10% of the sample is non-white and 18% have hypertension, then the mean-centred covariate pattern for someone who is white and has hypertension is (0 - 0.1) + (1 - 0.18) where 0 indicates white and 1 indicates hypertension. These values are multiplied by the estimated regression coefficients for each risk factor to obtain the mean-centred linear predictor.

population, or if pharmacotherapy was successfully delivered to 65% of the high-risk population (Manuel et al., 2013). They conclude that any of the approaches by themselves would have minimal effect, given that modest weight reductions are difficult to achieve on a population level and that lifestyle and pharmacotherapy approaches would have to have a large proportion of the population adhere to treatment. Combined approaches may be more feasible and realistic.

5.1.1 Study objectives

This study builds on the work by Manuel et al. (2013). It forecasts the five-year risk of diabetes at the *neighbourhood* level in the City of Toronto using the DPoRT model and the simulated TropISM population. Different intervention scenarios are developed and applied to the high-risk population of overweight and obese individuals (BMI $\geq$ 25 kg/m^2). Intervention scenarios are informed by existing research documenting the effectiveness of lifestyle interventions designed to promote weight loss (Sections 2.3.1 and 2.3.2), including:

1. A baseline risk scenario, where no changes to body weight, and therefore BMI, are made,
2. A strategy where all high-risk individuals lose: (a) 4%, (b) 7%, (c) 10%, (d) 14%, and (e) 17% of their baseline body weight (Ali et al., 2012; Barte et al., 2010; Knowler et al., 2002)
3. A strategy where the population of high-risk individuals loses, on average, 4.2kg ($\pm$5.1kg) of baseline body weight (Tuomilehto et al., 2001), and
4. A strategy where different subsets of high risk individuals lose 10% of their baseline body weight.

This last scenario is designed to mimic less than optimal reach of public health programs. Limited evidence from the United States suggests that as few as 4–8% of at-risk individuals might participate in weight loss programs (Almeida, Shetterly, Smith-Rey, & Estabrooks, 2010; Littman, Boyko, McDonell, & Fihn, 2012). For this last scenario, 10% of overweight individuals will be randomly selected to participate in a weight loss program. Everyone who participates will lose 10% of their baseline body weight. In more optimistic scenarios, 50% and 80% of individuals will be randomly selected to participate in a weight loss program. Again, everyone who participates will lose 10% of their baseline body weight.

The simulated TropISM population was subjected to each of these scenarios. The DPoRT model was used to forecast the five-year risk of developing diabetes at the neighbourhood level. Differences between scenarios were examined to identify neighbourhoods where weight loss programs exerted greater reductions in projected incidence.

5.2 Methods

5.2.1 Projecting neighbourhood incidence rates of diabetes

In order to implement the DPoRT model, the simulated TropISM population was classified on the basis of diabetes status and body weight status. Simulated individuals classified as having type 1 or type 2 diabetes were excluded from all risk scenarios (n = 98,361) as were men whose BMI was unknown (n = 7,615).[2] The remaining population was used to forecast the incidence of diabetes (n = 1,826,569). For each individual in the population at risk, the DPoRT model estimated the five-year probability of developing diabetes based on his or her set of risk factors (Table 5.1). This probability was estimated using the mean-centered linear predictor and the average characteristics of the entire simulated population (see Manuel et al., 2013; Rosella et al., 2014, 2011). The average characteristics of the entire simulated population were used to compute the mean centred linear predictor to ensure results could be compared across different neighbourhoods.

The probability of developing diabetes was estimated under each scenario. Estimates were averaged at the neighbourhood level to derive small area projections of the five-year diabetes incidence rate. In the baseline scenario, no changes to body weight were made in the simulated population. In Scenarios 2a through 2e, all overweight individuals (BMI $\geq$ 25kg/m^2) were subjected to a weight loss program so that each individual lost 4%, 7%, 10%, 14%, or 17% of his or her baseline body weight. The five-year risk of developing diabetes was re-estimated using each individual's new body weight.

In Scenarios 3 and 4, weight loss was randomly assigned. In Scenario 3, the average weight lost in the population was assumed to be normally distributed so

[2] These simulated individuals were excluded because the DPoRT model does not use men with unknown BMI to forecast diabetes incidence.

Table 5.1. Risk factors and parameters used by the Diabetes Population Risk Tool (DPoRT)* to forecast the five-year incidence of diabetes using the simulated TropISM population.

	Women				Men		
Risk Factor	Coef.†	(SE)	Prev.†	Risk Factor	Coef.†	(SE)	Prev.†
Intercept	10.578	(0.080)	—	Intercept	10.306	(0.070)	—
Non-white	-0.453	(0.085)	0.453	Non-white	-0.570	(0.063)	0.439
Immigrant	-0.148	(0.056)	0.546	Post-secondary‡	0.188	(0.045)	0.564
Post-secondary‡	0.194	(0.048)	0.582	Highest income§	0.117	(0.055)	0.178
Hypertension	-0.410	(0.055)	0.160	Current smoker	-0.059	(0.050)	0.280
Age < 45				Hypertension	-0.363	(0.055)	0.125
BMI 23 to < 25	-0.743	(0.191)	0.078	Heart disease	-0.348	(0.066)	0.043
BMI 25 to < 30	-1.152	(0.165)	0.090	Age < 45			
BMI 30 to < 35	-1.848	(0.179)	0.029	BMI 23 to < 25	-0.552	(0.232)	0.130
BMI ≥ 35	-2.056	(0.187)	0.020	BMI 25 to < 30	-0.952	(0.207)	0.180
BMI Unknown	-1.583	(0.213)	0.028	BMI 30 to < 35	-1.716	(0.206)	0.042
Age 45–65				BMI ≥ 35–25	-2.331	(0.226)	0.010
BMI <23	-0.710	(0.177)	0.126	Age ≥ 45			
BMI 23 to < 25	-1.234	(0.185)	0.050	BMI < 23	-1.360	(0.210)	0.098
BMI 25 to < 30	-1.836	(0.167)	0.091	BMI 23 to < 25	-1.654	(0.205)	0.132
BMI 30 to < 35	-2.374	(0.173)	0.027	BMI 25 to < 30	-2.056	(0.198)	0.180
BMI ≥ 35	-2.663	(0.187)	0.014	BMI 30 to < 35	-2.551	(0.211)	0.033
BMI Unknown	-2.199	(0.193)	0.011	BMI ≥ 35	-2.935	(0.234)	0.010
Age ≥ 65				Scale	0.799	(0.020)	—
BMI <23	-1.596	(0.173)	0.062				
BMI 23 to < 25	-1.614	(0.178)	0.040				
BMI 25 to < 30	-1.983	(0.165)	0.058				
BMI 30 to < 35	-2.215	(0.184)	0.012				
BMI ≥ 35	-2.645	(0.243)	0.002				
BMI Unknown	-2.421	(0.265)	0.008				
Scale	0.842	(0.021)	—				

* Model estimates and standard errors are from DPoRT 2.0 (personal communication, Dr. L. Rosella, March 27, 2014 and Rosella et al. (2014).)

† Coef. = Coefficient; Prev. = Prevalence

‡ Post-secondary education

§ Highest income quintile

that the population of overweight individuals lost, on average, 4.2kg of its baseline body weight. To account for random variability, this Scenario was replicated 250 times and the 5-year risk of diabetes was re-estimated for each replicate. The average 5-year risk of diabetes was then computed across all replicates at the neighbourhood level along with the median 5-year risk and the 2.5th and 97.5th percentiles.

A similar procedure was used in Scenario 4, where 10%, 50%, and 80% of overweight individuals were randomly assigned to receive a weight loss program, i.e., $Y_i \sim B(1,\pi)$, where:

$$Y_i = \begin{cases} 0 & \text{if individual } i \text{ is } \textbf{not} \text{ assigned to a weight loss program,} \\ 1 & \text{if individual } i \textbf{ is} \text{ assigned to a weight loss program,} \end{cases}$$

and π = 0.1, 0.5, or 0.8. Again, this Scenario was replicated 250 times and the average 5-year risk of diabetes was computed across all replicates at the neighbourhood level, along with the median risk and 2.5th and 97.5th percentiles. All scenarios were developed using the R statistical software package [3]. Results were summarized tabularly and visualized cartographically.

5.2.2 Validating neighbourhood-level risk projections

Neighbourhood-level risk projections were validated using the number of newly diagnosed cases of diabetes ascertained from the Ontario Diabetes Database from 2006 to 2010. First, the known number of incident cases was compared to the forecast number using the concordance correlation coefficient. Forecast counts were compared to the overall number of incident cases and to the adjusted number of cases correcting for the 3% false-positive rate in the algorithm used by the Ontario Diabetes Database to identify cases of diabetes (Hux et al., 2002).

Second, the average difference between forecast counts and known counts was estimated across all neighbourhoods to assess whether the DPoRT model under- or over-predicted incidence. A measure of relative error was again used to estimate the proportion of neighbourhoods having forecast incident counts lying within ±20% of the known number of incident cases.

[3] Sample code used to predict the risk of diabetes is presented in Appendix D

Next, neighbourhoods were classified into quintiles according to the forecast and known number of incident cases within each neighbourhood. The level of agreement between these two classification systems was assessed using the overall % agreement and Cohen's κ. Higher values of these statistics suggest that neighbourhoods were ranked into the same quintile using either classification and that the DPoRT model could correctly identify neighbourhoods having high or low diabetes incidence.

The validity of the forecasts was further assessed by comparing known and forecast incidence rates. The known incidence rate of diabetes was measured using the cumulative incidence proportion,[4] estimated as

$$\text{incidence \%} = \frac{i_j}{(\text{popn}_j - p_j)} * 100 \tag{5.3}$$

where:

$$\begin{aligned} i_j &= \text{the number of incident cases in neighbourhood } j \text{ from 2006–2010,} \\ \text{popn}_j &= \text{the total population in neighbourhood } j \text{ in 2006,} \\ p_j &= \text{the number of prevalent cases in neighbourhood } j \text{ in 2005} \end{aligned}$$

The cumulative incidence proportion was then compared to the forecast proportion across all neighbourhoods. Similar to the validation of diabetes prevalence rates, incidence rates were compared using both raw and spatially smoothed rates. The Pearson correlation was used to assess the similarity between known and forecast incidence rates. Higher positive values of the Pearson correlation suggest neighbourhoods having higher known incidence rates have higher forecast rates.

Neighbourhoods were then classified into quintiles on the basis of known and forecast incidence rates, comparing both raw and spatially smoothed rates. The overall % agreement and Cohen's κ were used to assess the degree of similarity between these two classification systems. The global bivariate Moran's I statistic was also used to compare the spatial similarity between known and forecast incidence rates. In all cases, higher values suggest a greater degree of similarity between known and forecast incidence rates. Finally, bivariate LISA statistics were esti-

[4] Estimated for the population age 20+ using incident and prevalent cases corrected for the 3% false-positive rate in the algorithm used by the Ontario Diabetes Database to identify cases of diabetes.

mated using GeoDa (version1.6.6.1)[5] to identify groups of neighbourhoods having similarly high or low forecast and known incidence rates of diabetes.

5.2.3 Neighbourhood characteristics associated with projected incidence

In order to understand how projected incidence rates might vary across neighbourhoods as a function of their demographic characteristics, an exploratory spatial analysis was conducted. For this analysis, neighbourhoods were classified according to the demographic composition of their population by dividing all neighbourhoods in two groups using a median split. For example, neighbourhoods having more than the median percentage of visible minorities (36%) were classified as having a "high" percentage of visible minorities while neighbourhoods below the median were classified as having a "low" percentage of visible minorities. Neighbourhoods were classified according to body mass index, the male population older than 45, the female population older than 65, visible minorities, immigrants, post-secondary education, and income. Mean forecast incidence rates were estimated by cross-classifying each demographic characteristic against a second. Differences in projected incidence rates by these demographic groupings were tested using a spatial simultaneous autoregressive error model, estimated in R using the "spdep" package (Bivand, 2015). A "rook" contiguity matrix was used to define which areas neighboured others, and thereby account for spatial dependence in the regression model. Differences in average incidence rates by neighbourhood characteristics were tested using a model containing the main effects of two separate demographic factors and the interaction effect. Statistically significant differences were visualized using conditioned choropleth maps.

5.2.4 Uncertainty and sensitivity analysis

Even though DPoRT model outputs were validated, it is important to recognize they are not known with certainty. Model outputs are only as reliable as the input parameters used to generate them (Rausand, 2011; Thompson & Graham, 1996). When several parameters are used to forecast outcomes, the uncertainty associated with all input parameters can be propagated through the model, resulting

[5] P-values were adjusted for multiple testing using the false discovery rate adjustment in R.

in total model uncertainty that is greater than the sum of its parts (Sharif et al., 2012). Quantifying model uncertainty is an essential step in the model development process, a step that aids the interpretation of results (Sharif et al., 2012). If model uncertainty is large, decision makers may choose to acquire additional evidence to support a particular course of action rather than base their decisions on uncertain outcomes (Briggs & Sculpher, 2006).

Uncertainty may be classified into two types: aleatory and epistemic. Aleatory uncertainty deals with random variation while epistemic uncertainty deals with incomplete information or knowledge. In order to assess the aleatory uncertainty surrounding projected diabetes incidence rates, a probabilistic sensitivity analysis was conducted. Quantitative uncertainty analysis assesses the imprecision of predicted outcomes attributable to using uncertain input parameters in the modelling process. It relies on Monte Carlo methods to randomly sample input values from appropriate probability distributions (Hoare, Regan, & Wilson, 2008; Sharif et al., 2012). For each iteration of the uncertainty analysis, the model outcomes are predicted from the prediction model (i.e., the DPoRT model) and uncertainty intervals can be defined using the 2.5th and 97.5th percentiles of the predicted outcome across all iterations. An appropriate probability distribution used to describe the input parameters depends on the nature of those parameters, but the normal distribution is usually a reasonable choice (Briggs & Sculpher, 2006).

Probabilistic sensitivity analysis extends uncertainty analysis by identifying which model parameters exert the greatest influence on the uncertainty of predicted outcomes (Hoare et al., 2008; Saltelli, 2002). It quantifies how changes in input parameters affect predicted outcomes using statistical methods (Hoare et al., 2008). Two measures can be used to identify influential model parameters: the partial rank correlation coefficient (PRCC) and the sensitivity index (S_i). The partial rank correlation coefficient measures the strength of association between a model parameter and the outcome after controlling for all other parameters included in the model (Hoare et al., 2008; Marino, Hogue, Ray, & Kirschner, 2008). The PRCC takes on values between -1 and 1. The sign of the PRCC indicates the direction of change in the outcome that is associated with larger values of a given model parameter. Model parameters having large values of the PRCC suggest they contribute substantially to outcome uncertainty (Hoare et al., 2008).

The sensitivity index measures the proportion of variation in the outcome variable attributable to each model parameter (Hoare et al., 2008). It is estimated as

$$S_i = \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i))}{V(Y)} \tag{5.4}$$

where:

$V(Y)$ is the total variance in outcome Y,

$E_{X_{\sim i}}(Y|X_i)$ is the expected, or average, value of outcome Y over all input parameters *except* X_i, conditional on X_i having a fixed value and

$V_{X_i}(E_{X_{\sim i}}(Y|X_i))$ is the variance of that expectation (Hoare et al., 2008; Saltelli, Tarantola, & Campolongo, 2000).

The sensitivity index takes on values between 0 and 1. Parameters that have larger sensitivity indices exert greater influence on outcome uncertainty. S_i is used to determine which parameter, once fixed to its "true" value, produces the greatest average reduction in the variance of outcome Y (Allaire & Willcox, 2012).

Hoare et al. (2008) developed the Sampling and Sensitivity Analysis Toolbox (SaSAT) for conducting probabilistic sensitivity analysis of mathematical models. Developed in Matlab, a standalone version of SaSAT is freely available. To conduct the sensitivity analysis of projected diabetes incidence, model parameters from DPoRT 2.0 (Rosella et al., 2014) were treated as random variables and assumed to be normally distributed. To randomly sample parameter values for the sensitivity analysis, DPoRT parameter estimates were used as the mean and standard errors as the standard deviation (Table 5.1).

One thousand samples of parameter values were drawn using Latin Hypercube Sampling, a type of stratified sampling that divides the probability density function into N equiprobable intervals.[6] For each model parameter, a single value was drawn from all N intervals (Hoare et al., 2008). The projected five-year incidence of diabetes was re-estimated using these randomly selected parameters. Results were aggregated to (a) the neighbourhood level and (b) metropolitan Toronto. Since the DPoRT model differs for men and women, separate parameters were randomly sampled for each model. Uncertainty intervals were computed for forecast incidence using the 2.5th and 97.5th percentiles.

[6] Where N equals the number of samples, in this case, 1000.

Using these projected results, partial rank correlations were estimated in SaSAT between each DPoRT model parameter and the five-year incidence of diabetes. The sensitivity index was also estimated to measure the proportion of variation in forecast incidence attributable to each model parameter. The sensitivity index for all DPoRT parameters was estimated for each neighbourhood within metropolitan Toronto. This made it possible to ascertain whether uncertainty in forecast incidence was attributable to different factors in different neighbourhoods.

5.3 Results

5.3.1 Baseline projections of diabetes incidence

In the baseline scenario, the TropISM population was used to forecast the five-year incidence of diabetes in disease-free adults aged 20 and older using the Diabetes Population Risk Tool. Uncertainty estimates were generated from 1000 random samples of DPoRT model parameters. With no changes in the simulated population's risk profile, 4.9% of Toronto's adults might develop diabetes over a five year period (95% UI: 4.2–5.9%).

As might be expected, the projected incidence of diabetes increased with age among men and women alike. With the exception of the youngest age group, it is not possible to compare age-sex specific forecasts because the DPoRT model uses different age profiles to forecast incidence for men and women. However, men and women in the youngest age group (ages 20–44) tended to have similar projected incidence rates (Table 5.2).

Table 5.2 also demonstrates that the projected incidence of diabetes varied across the city, from a low of 4.2% in the borough of Toronto to a high of 5.4% in Scarborough. Fewer differences in the five-year incidence rate were seen across boroughs in the youngest age group while larger differences were seen among the oldest men and women. In addition, there was a greater amount of uncertainty around projected incidence rates for women aged 65 and older compared to other subgroups.

Figure 5.1 presents raw and spatially smoothed forecast incidence of diabetes across metropolitan Toronto. In general, neighbourhoods in Toronto and East York had the lowest forecast rates, ranging from a low of 3.1% to a high of 5.1%.

Table 5.2. Five-year forecast incidence (%) of diabetes by sex and age group across metropolitan Toronto and its boroughs. Projections were developed using the simulated TropISM population and the Diabetes Population Risk Tool.

	Overall		All Men		Men 20-44		Men 45+	
Borough (n*)	%	(95% UI†)	%	(95% UI)	%	(95% UI)	%	(95% UI)
Metro Toronto (n = 140)	4.87	(4.25, 5.88)	4.78	(3.88, 6.26)	2.16	(1.59, 3.03)	7.94	(6.15, 10.66)
Etobicoke (n = 20)	5.01	(4.36, 6.07)	4.86	(3.92, 6.36)	2.08	(1.52, 2.91)	7.67	(5.94, 10.29)
North York(n = 33)	5.11	(4.45, 6.17)	4.93	(4.00, 6.46)	2.19	(1.61, 3.06)	8.03	(6.19, 10.80)
York (n = 10)	5.01	(4.38, 6.05)	4.92	(3.98, 6.42)	2.20	(1.62, 3.08)	8.18	(6.33, 10.98)
Toronto (n = 44)	4.17	(3.64, 5.05)	4.11	(3.34, 5.36)	1.92	(1.42, 2.71)	7.25	(5.62, 9.77)
East York (n = 8)	4.68	(4.08, 5.65)	4.66	(3.76, 6.08)	2.10	(1.56, 2.94)	7.75	(6.02, 10.40)
Scarborough (n = 25)	5.44	(4.78, 6.53)	5.46	(4.39, 7.13)	2.54	(1.87, 3.55)	8.77	(6.76, 11.80)
	All Women		Women 20–44		Women 45–64		Women 65+	
Borough (n)	%	(95% UI)	%	(95% UI)	%	(95% UI)	%	(95% UI)
Metro Toronto (n = 140)	4.95	(4.05, 6.24)	2.16	(1.63, 2.87)	5.39	(4.13, 7.12)	11.65	(8.46, 15.71)
Etobicoke (n = 20)	5.13	(4.16, 6.55)	2.10	(1.60, 2.78)	5.19	(3.99, 6.88)	11.44	(8.24, 15.64)
North York (n = 33)	5.26	(4.28, 6.65)	2.23	(1.69, 2.96)	5.52	(4.22, 7.31)	11.87	(8.62, 16.02)
York (n = 10)	5.10	(4.17, 6.43)	2.27	(1.72, 3.01)	5.71	(4.40, 7.58)	11.71	(8.54, 15.84)
Toronto (n = 44)	4.22	(3.43, 5.38)	1.91	(1.44, 2.54)	4.88	(3.75, 6.45)	10.98	(7.78, 15.20)
East York (n = 8)	4.70	(3.83, 5.96)	2.09	(1.59, 2.77)	5.14	(3.96, 6.81)	11.08	(7.98, 15.13)
Scarborough (n = 25)	5.42	(4.39, 6.86)	2.42	(1.84, 3.24)	5.91	(4.53, 7.91)	12.28	(8.97, 16.42)

* n = number of neighbourhoods

†95% UI = 95% uncertainty interval.

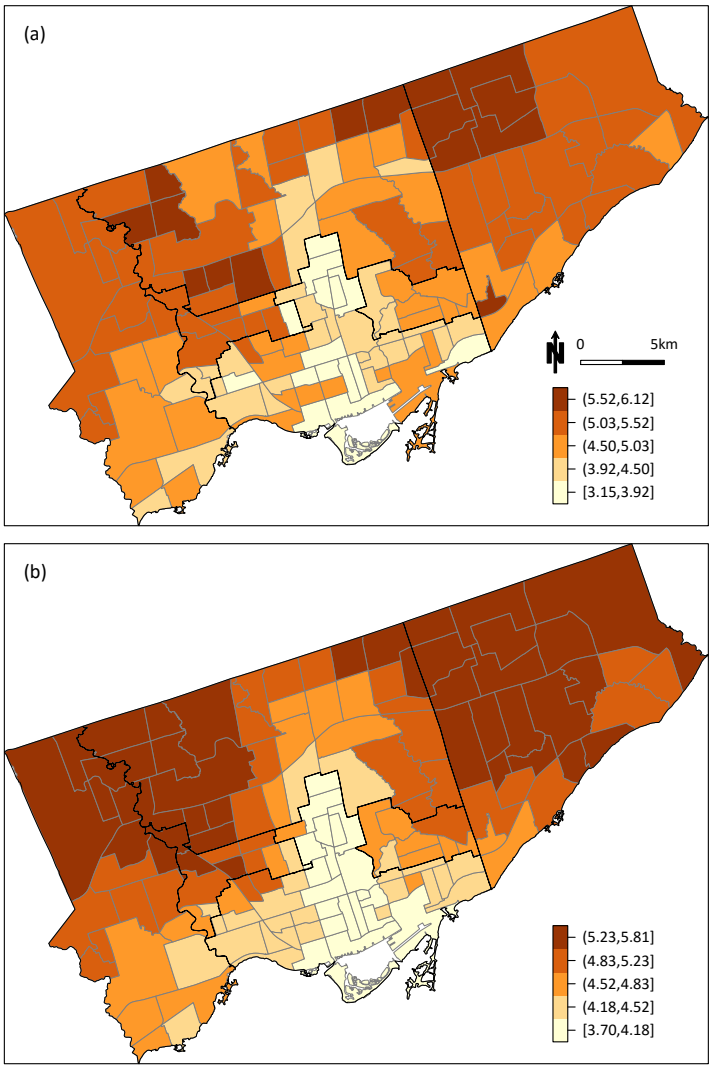


Figure 5.1. Five-year forecast incidence of diabetes across metropolitan Toronto: (a) raw incidence rate and (b) spatially smoothed incidence rate.

Forecast incidence was noticeably higher in all of Scarborough's neighbourhoods while there was a strong north-south divide in forecast incidence in Etobicoke. Forecast incidence was most variable in the boroughs of Toronto (range: 3.2–5.1%) and North York (4.2–6.0%).

5.3.1.1 Validation of projected incidence rates

Unlike the validation of simulated diabetes prevalence, the forecast incidence of diabetes was more similar to the true incidence of diabetes in Toronto from 2006 to 2010. Using cases of diabetes ascertained from the Ontario Diabetes Database, the overall five-year incidence of diabetes was 5.95%. Incidence was slightly higher among men (6.41%) than women (5.54%). After correcting for the false-positive rate, cumulative incidence proportions fell within the range of forecast uncertainty (Table 5.2). Across metropolitan Toronto, the overall corrected five-year incidence of diabetes was 5.75% whereas forecast incidence was 4.87% (95% UI: 4.25, 5.88). Similar results were obtained for men and women: corrected incidence among men was 6.20% while forecast incidence was 4.78% (95% UI: 3.88, 6.26). Corrected incidence among women was 5.35% while forecast incidence was 4.95% (95% UI: 4.05, 6.24). Thus, at an aggregate level, forecast incidence is slightly lower than actual incidence although the extent of under-prediction is greater for men than women.

On average, the DPoRT model predicted 99 fewer incident cases of diabetes in any given neighbourhood compared to the true number (Table 5.3). When the true number of incident cases was corrected for false-positives, the DPoRT model predicted 77 fewer cases using the TropISM population. Forecasts were more accurate for women than men. Overall, 62% of neighbourhoods had forecast counts lying within ±20% of known counts.

An examination of the concordance correlation coefficient illustrates that compared to the true number of incident cases, the forecast number of cases (Table 5.3) was more precise and accurate than the simulated number of prevalent cases (Table 4.5). Correcting for false positives in the true number of cases only slightly improved forecast accuracy. Moreover, scale and location shifts were less severe than they were for simulated prevalence. This is illustrated in Figure B.1 in

Table 5.3. Neighbourhood-level validation of the forecast incident counts of diabetes compared to the known number of incident cases of diabetes ascertained from the Ontario Diabetes Database (2006–2010).

	Total		Men		Women	
Measure	Cnt	Cor	Cnt	Cor	Cnt	Cor
Difference						
Mean	-99	-77	-82	-71	-17	-6
(± SD)	(201)	(189)	(109)	(103)	(97)	(92)
Relative Error						
% ±0.2	59.3	62.1	47.1	53.6	54.3	55.0
Concordance						
Overall ρ_c	0.851	0.867	0.785	0.810	0.888	0.898
Precision (ρ)	0.940	0.940	0.950	0.950	0.927	0.927
Accuracy	0.906	0.925	0.830	0.856	0.958	0.968
Scale shift	0.686	0.707	0.626	0.646	0.752	0.775
Location shift	-0.254	-0.200	-0.433	-0.380	-0.083	-0.030
Agreement						
Overall	64.0%		62.1%		65.0%	
Cohen's κ	0.554		0.527		0.562	

Cnt: forecast counts compared to known counts.
Cor: known counts corrected for false positives.

Appendix B. Overall, the line of best fit comparing the forecast number of cases to the actual number is closer to the line of concordance than it was for prevalent cases in Figure A.2. In addition, index plots of incident cases in each neighbourhood demonstrate that the trend in forecast counts lies closer to the true number of cases in most neighbourhoods.

A comparison of known and forecast incidence proportions reveals a slightly different pattern of results. Although a strong correlation was found between the true incidence proportion and forecast incidence (Table 5.4), the correlation was weaker than it was for forecast counts (Table 5.3). Stronger correlations were found between spatially smoothed incidence rates. When neighbourhoods were classified into quintiles according to known and forecast counts, the overall agreement between classifications was greater for incident counts of diabetes among women

Table 5.4. Neighbourhood-level validation of the five-year forecast incidence of diabetes (%) compared to the five-year cumulative incidence proportion (%)* estimated from the Ontario Diabetes Database (2006–2010).

	Correlation†		Rate Classification‡				Rate Clustering	
			Raw		Spatial		Moran's I§	
	ρ_r	ρ_{sp}	A_o	κ	A_o	κ	Raw	Spatial
Incidence								
Total	0.797	0.909	48.6	0.357	66.4	0.580	0.641	0.844
Men	0.820	0.933	62.1	0.527	72.9	0.661	0.617	0.858
Women	0.730	0.864	42.9	0.286	47.1	0.339	0.390	0.624

* Cumulative incidence proportions used incident cases from the Ontario Diabetes Database corrected for false positives.

† ρ_r = Pearson correlation between raw rates; ρ_{sp} = Pearson correlation between spatially smoothed rates.

‡ Rate classification based on classifying neighbourhoods into quintiles according to their raw and spatially smoothed incidence rates. A_o = overall % agreement, κ = Cohen's kappa.

§ Bivariate Moran's I.

(65% among women vs. 62% among men). However, when comparing a quintile classification of rates, the overall agreement between classifications was greater for men (43% among women vs. 62% among men). This difference may be explained by the way in which known incidence proportions were estimated. Specifically, the denominator of the incidence proportion subtracts the prevalent number of cases in each neighbourhood from the total population. Since there was a greater discrepancy between prevalent counts of diabetes at the neighbourhood level than incident counts, this discrepancy will differentially skew estimated incidence proportions across neighbourhoods. In other words, some neighbourhoods will have a greater proportion of prevalent cases removed than others, producing known incidence rates that lie further from forecast rates than others.

Figure B.2 in Appendix B compares broad spatial patterns in overall forecast

incidence against known incidence[7] in more detail. Panels (a) and (b) present raw rates (forecast and known) while panels (c) and (d) present spatially smoothed rates (forecast and known). Visual inspection of spatially smoothed rates generally reveals that areas having low forecast incidence tend to have low known incidence. Similarly, areas having high forecast incidence tend to have high known incidence. Overall, two-thirds of neighbourhoods were classified into the same quintile using either forecast or known incidence rates. Another 18% of neighbourhoods were classified into one lower quintile according to the forecast classification while 14% of neighbourhoods were classified into one higher quintile. Bivariate LISA statistics computed from spatially smoothed rates indicate that almost all neighbourhoods within Scarborough having high forecast incidence rates are surrounded by neighbourhoods having high known incidence rates. A similar cluster of neighbourhoods occurs in north Etobicoke and parts of west North York. Most neighbourhoods in Toronto having low forecast incidence rates are surrounded by neighbourhoods having low known incidence rates. Based on these results, there is a strong spatial coincidence of neighbourhoods having similarly low or high forecast and known incidence rates of diabetes.

Figure B.3 in Appendix B plots the known[8] incidence proportion against forecast incidence, ranked in ascending order of known incidence across all neighbourhoods. Forecast incidence is plotted along with the 95% uncertainty interval for each neighbourhood. Neighbourhoods having low known incidence tend to have low forecast incidence. However, as known incidence increases at the neighbourhood level, the separation between known and forecast incidence increases. Based on these results, it appears that the DPoRT model is unable to forecast the variability in known incidence across metropolitan Toronto. Thus, while aggregate forecasts of diabetes incidence are similar to known incidence proportions, forecast incidence in specific neighbourhoods may vary substantially from known incidence.

Table B.1 explores the demographic characteristics of neighbourhoods where known incidence fell outside the range of forecast uncertainty. Overall, known incidence was lower than forecast incidence in 23 neighbourhoods, 16 of which

[7] Corrected for false positives.
[8] Corrected for false positives.

were in Toronto. In these neighbourhoods, the median difference between known and forecast incidence was -0.84 percentage points (Figure B.4). In 52 neighbourhoods, known incidence was greater than forecast incidence; 15 of these neighbourhoods were in North York while 19 were in Scarborough. In this case, the median difference between known and forecast incidence was 2.30 percentage points.

Neighbourhoods where known incidence was above forecast uncertainty had, on average, a greater percentage of visible minorities (54%) than neighbourhoods whose known rates fell within or below forecast uncertainty (36% and 28%, respectively). Neighbourhoods classified above forecast uncertainty also contained a significantly greater percentage of immigrants (58%) compared to neighbourhoods lying within or below forecast uncertainty (46% and 37%, respectively). However, neighbourhoods classified below forecast uncertainty had, on average, better educated populations than neighbourhoods classified above forecast uncertainty (66% vs 53%, respectively). Interestingly, population growth between 2006 and 2011 did not differ across the three groups.

These results suggest some possible explanations for the inability of the DPoRT model to accurately forecast the variability in diabetes incidence using the simulated TropISM population. First, neighbourhoods where known incidence extended beyond forecast uncertainty differed on important demographic characteristics used by the DPoRT model to forecast incidence. Although the model predicted incidence at an aggregate level reasonably well, in neighbourhoods whose populations were atypical of the overall average population, forecast incidence was less accurately predicted. Keeping in mind that the DPoRT model was calibrated to projecting incidence in the overall Ontario population (Rosella et al., 2011), it is possible that model parameters used to forecast incidence may not be calibrated for specific subgroups within the population.

Second, TropISM under-predicted the prevalence of risk factors used by the DPoRT model to forecast incidence. In neighbourhoods where the simulated prevalence was too low, the true population at risk may not be reasonably represented, leading to less accurate forecasts. In spite of these limitations, DPoRT forecasts based on the simulated TropISM population are still useful because, in a relative sense, forecasts still identify neighbourhoods within metropolitan Toronto hav-

ing higher or lower incidence. Moreover, broad spatial patterns in incidence can also be identified. As a result, forecasts can still be used to identify neighbourhoods within metropolitan Toronto that can be targeted for planning and delivering health promotion programs. The next section therefore examines the demographic characteristics of neighbourhoods having relatively high or low forecast incidence rates.

5.3.1.2 Neighbourhood characteristics associated with projected incidence

Table 5.5 describes average projected incidence rates by demographic characteristics of Toronto neighbourhoods. Forecast incidence rates did not vary significantly in neighbourhoods whose average population BMI was greater than the metropolitan median BMI (24.9 kg/m^2) compared to neighbourhoods whose average population BMI was lower than the median. This was true regardless of other demographic factors considered. However, significant differences in forecast incidence rates were observed when neighbourhoods were cross-classified by other demographic characteristics. The largest differences in forecast incidence were observed in neighbourhoods having higher than the median percentage of immigrant women aged 65 and older versus younger neighbourhoods having relatively fewer immigrants (5.2% versus 4.4%, respectively). Figure 5.2 displays these differences cartographically using a conditioned choropleth map. Each quadrant of the figure displays neighbourhoods classified by their respective demographic profiles (i.e., immigrant and older female population structures). The upper left quadrant depicts neighbourhoods containing higher percentages of younger non-immigrant women while the lower right quadrant depicts neighbourhoods containing higher percentages of older immigrant women. Based on this plot, forecast incidence was lower in neighbourhoods located within the borough of Toronto, where there were higher percentages of younger non-immigrant women. Projected rates of diabetes were higher in Scarborough and North York, in neighbourhoods primarily composed of high percentages of immigrants.

Stronger spatial patterns were observed in forecast incidence rates when considering the visible minority and immigrant composition of neighbourhoods. The average forecast incidence of diabetes was 4.6% in neighbourhoods having rela-

Table 5.5. Average five-year forecast incidence of diabetes by demographic characteristics of metropolitan Toronto neighbourhoods.

Characteristic (median)	LL*	LH*	HL*	HH*	LR†	p
Body mass index (24.9 kg/m^2)						
% men ≥ 45	4.82	5.04	4.77	4.82	1.94	0.164
% women ≥ 65	4.81	5.09	4.57	4.94	0.77	0.381
% visible minority	4.58	5.02	4.70	5.05	0.29	0.593
% immigrant	4.63	4.97	4.68	5.21	1.86	0.173
% post-secondary education	5.02	4.74	5.02	4.63	0.83	0.362
% highest income quintile	4.99	4.60	5.20	4.64	1.48	0.223
% Men ≥ 45 (21.5%)						
% women ≥ 65	4.71	5.01	4.76	4.98	0.47	0.494
% visible minority	4.53	4.98	4.74	5.13	0.30	0.583
% immigrant	4.48	4.97	4.72	5.16	0.17	0.678
% post-secondary education	5.02	4.57	5.06	4.78	2.30	0.130
% highest income quintile	5.02^a	4.40^b	5.11^a	4.73^c	5.02	0.025
% Women ≥ 65 (9.9%)						
% visible minority	4.47	4.94	4.85	5.16	2.59	0.108
% immigrant	4.41^a	4.95^b	4.84^b	5.16^c	4.35	0.037
% post-secondary education	4.95^a	4.51^b	5.06^a	4.91^a	8.49	0.004
% highest income quintile	4.98^{ac}	4.40^b	5.15^a	4.83^c	7.80	0.005
% Visible minority (36.0%)						
% immigrant	4.56^a	5.01^b	4.93^b	5.06^b	5.12	0.024
% post-secondary education	4.89	4.53	5.09	4.95	2.99	0.084
% highest income quintile	5.02^a	4.53^b	5.07^a	4.93^a	6.01	0.014
% Immigrant (50.0%)						
% post-secondary education	4.89^a	4.47^b	5.09^a	5.01^a	6.27	0.012
% highest income quintile	5.04^a	4.54^b	5.04^a	5.05^a	8.33	0.004
% Post-secondary education (56.5%)						
% highest income quintile (14.4%)	5.10	4.80	4.93	4.55	0.28	0.599

* Demographic cross-classification: LL = low-low, LH = low-high, HL = high-low, HH = High-high. Groups having different letters are significantly different at p = 0.05, using a Bonferroni adjustment.

† LR = likelihood ratio.

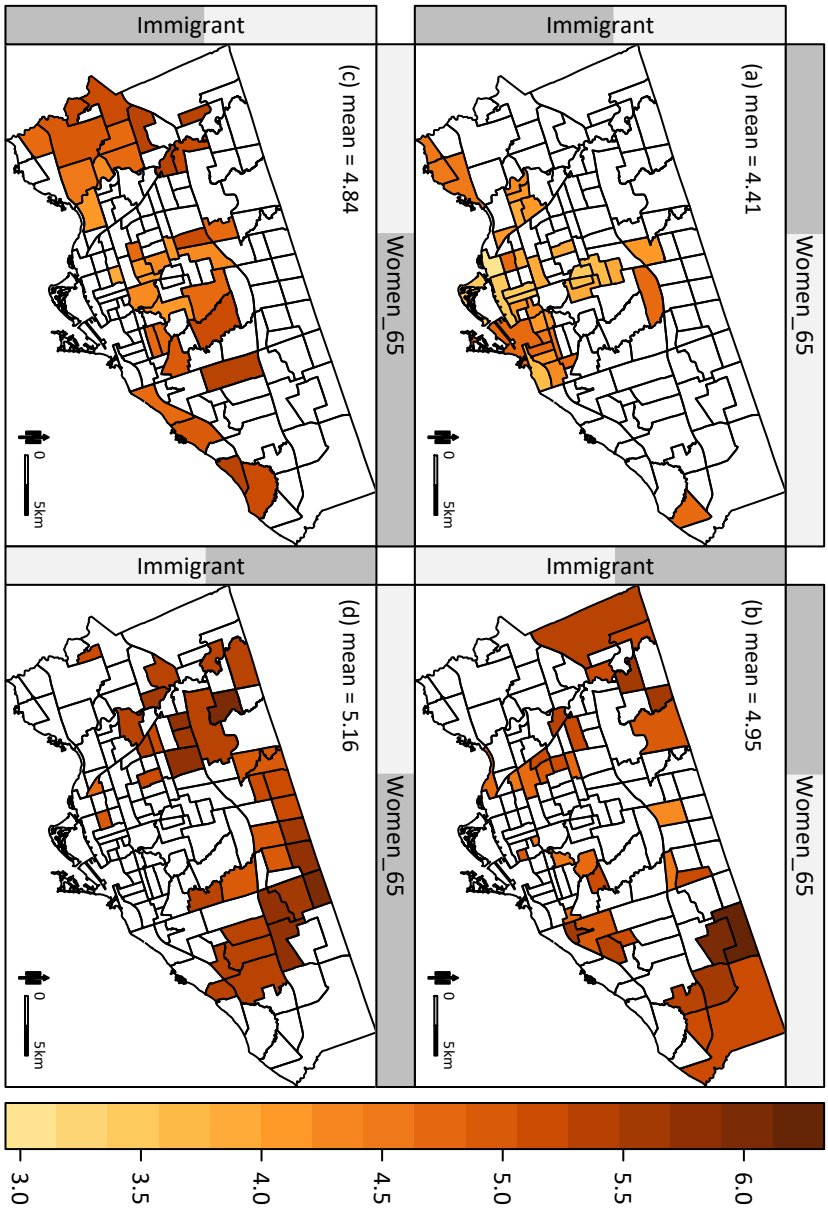


Figure 5.2. Forecast five-year incidence of diabetes (%) by neighbourhood composition: (a) low percentage of women 65+, low percentage of immigrants; (b) low percentage of women 65+, high percentage of immigrants; (c) high percentage of women 65+, low percentage of immigrants; (d) high percentage of women 65+, high percentage of immigrants.

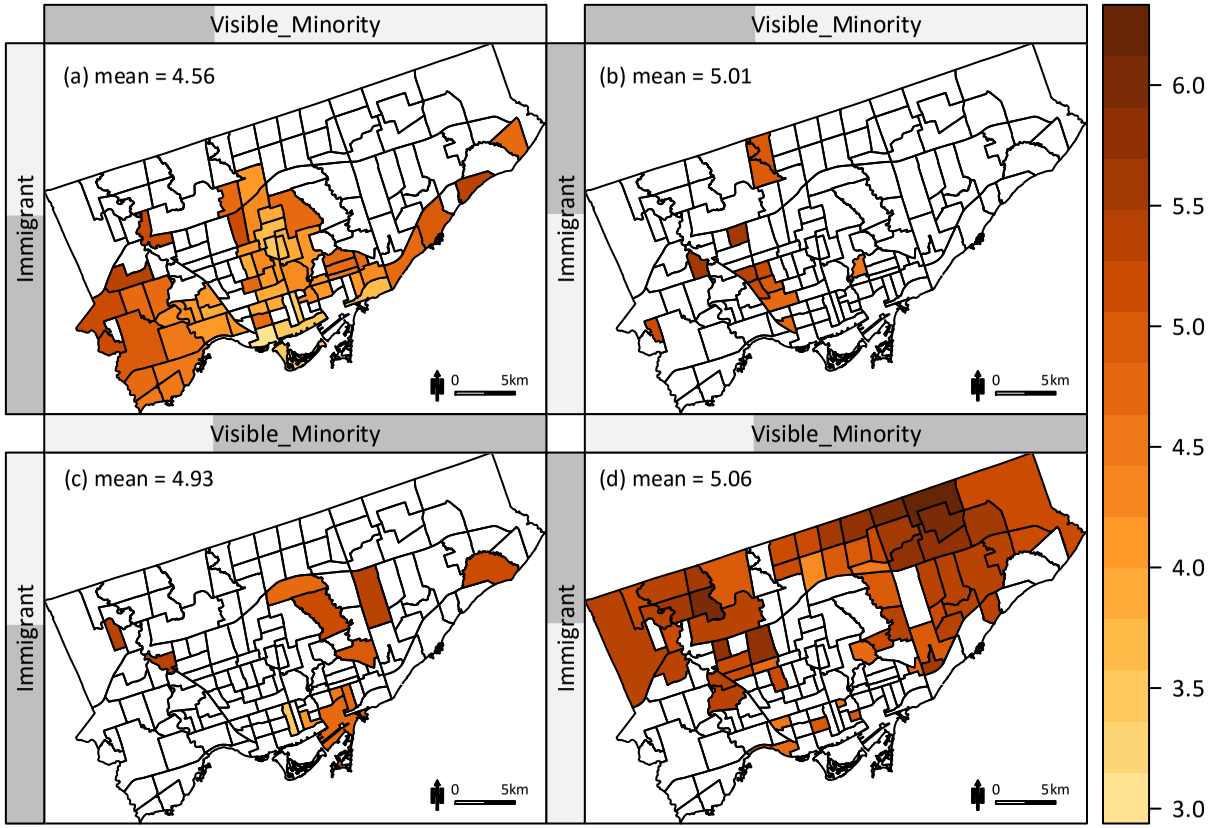
tively few visible minorities and immigrants (i.e., less than 36% of the population was from a visible minority group and less than 50% were immigrants, Figure 5.5). Conversely, in neighbourhoods having a larger percentage of visible minorities and immigrants, projected incidence was 5.1%. Figure 5.3 contrasts these differences, demonstrating a clear spatial pattern. Specifically, neighbourhoods composed of largely non-immigrant, non-visible minority populations, and low projected incidence rates tended to cluster in Toronto and south Etobicoke. Neighbourhoods having significantly higher projected incidence rates and large visible minority and immigrant populations tended to cluster in northern Etobicoke, North York, and Scarborough.

Finally, the percentage of the population having a post-secondary education influenced forecast incidence under some circumstances. Neighbourhoods composed of younger, better educated populations had lower forecast incidence rates, on average, compared to neighbourhoods composed of older, less educated populations (4.5% vs. 5.1%, respectively). Although the magnitude of the difference was similar irrespective of the male and female age profiles, average forecast incidence differed significantly only in neighbourhoods composed of higher percentages of younger, better educated women compared to neighbourhoods composed of higher percentages of older, less educated women (Table 5.5). Similar differences were seen in neighbourhoods composed of lower percentages of immigrants and higher percentages of the population having a post-secondary education.

In summary, these results indicate that the population composition of a given neighbourhood influences the forecast incidence of diabetes using the Diabetes Population Risk Tool. Based on the demographic characteristics of Toronto's neighbourhoods, significant differences in projected incidence were observed in neighbourhoods having different age, sex, visible minority, education, and income profiles (Table 5.5). Although the DPoRT model uses other risk factors to project the incidence of diabetes, these factors represent the constraints used to develop TropISM. As such, these demographic profiles should closely resemble the actual population structure of these neighbourhoods because these variables were closely replicated by the simulated annealing algorithm (see Table 4.3).

It is worth noting, however, that average forecast incidence rates presented in Table 5.5 should be interpreted cautiously. While broad spatial patterns in forecast

Figure 5.3. Forecast five-year incidence of diabetes (%) by neighbourhood composition: (a) low percentage of minorities, low percentage of immigrants; (b) low percentage of minorities, high percentage of immigrants; (c) high percentage of minorities, low percentage of immigrants; (d) high percentage of minorities, high percentage of immigrants.



incidence are reasonably correct according to the validation analysis that was conducted, the absolute magnitude of average forecast incidence rates in Table 5.5 will not accurately reflect the true incidence in some areas. For example, in neighbourhoods having greater than the median percentage of visible minorities and immigrants (Figure 5.3, panel (d)), the mean forecast incidence across these neighbourhoods is estimated to be 5.1%. However, most of these neighbourhoods were the ones in which the true incidence of diabetes was greater than forecast uncertainty. Therefore, the absolute value of diabetes incidence in Figure 5.3, panel (d) will be greater than 5.1%. In spite of this, the spatial characterization of these neighbourhoods will be approximately correct, because the DPoRT model replicated spatial patterns in forecast incidence using the TropISM population reasonably well. Therefore, the characterization of neighbourhoods presented in Table 5.5 should be approximately correct.

5.3.2 Projected incidence under weight loss scenarios

5.3.2.1 Prevalence of overweight

Until this point, only the baseline TropISM population was used to project the incidence of diabetes at the neighbourhood level. Since excess body weight is an important risk factor for diabetes, several scenarios were developed to assess how the neighbourhood incidence of diabetes might change if overweight individuals lost different amounts of body weight. One set of scenarios assumed that *all* overweight individuals in the population would lose clinically meaningful amounts of body weight. These scenarios also assumed that healthy weight and underweight individuals maintained their current body weight. A second set of scenarios assumed that different proportions of the overweight population would lose 10% of its baseline body weight, in order to mimic sub-optimal reach and effectiveness of weight loss programs.

The primary objective of these policy simulations was to assess how weight loss programs might influence the future incidence of diabetes. Thus, it is important to describe the effect of weight loss on the neighbourhood prevalence of overweight. Table 5.6 presents the simulated neighbourhood prevalence of overweight (BMI $\geq$ 25 kg/m^2) averaged across the six boroughs of metropolitan Toronto un-

der each weight loss scenario. In the baseline scenario, neighbourhoods from York have, on average, the highest simulated prevalence of overweight (43%) while neighbourhoods from Scarborough have the lowest simulated prevalence (39%). Reductions in the prevalence of overweight are approximately similar across each of the boroughs under scenarios. Compared to the baseline scenario, however, neighbourhoods in York experienced the greatest reduction in overweight prevalence from baseline to the 17% weight loss scenario. In other words, if the population of overweight individuals in York neighbourhoods lost 17% of its baseline body weight, the population prevalence of overweight could be expected to drop by 32 percentage points. This contrasts to Scarborough neighbourhoods, where the population prevalence of overweight might drop by only 29.5 percentage points.

The neighbourhood-level prevalence of overweight was higher among men than women under the baseline scenario across all neighbourhoods. Once the overweight population loses 10% of its baseline body weight, the prevalence of overweight is approximately similar among both men and women across all neighbourhoods. Under the 17% weight loss scenario, the prevalence of overweight is slightly lower among men than women.

Figure 5.4 depicts the neighbourhood-level distribution of body mass index, using kernel density estimates of the distribution under each weight loss scenario.[9] In the baseline scenario, almost half of all neighbourhoods have an average population BMI of more than 25 kg/m^2. As the population of overweight individuals loses weight, two things happen. First, average BMI across all neighbourhoods decreases from 24.88 kg/m^2 in the baseline scenario to 23.69 kg/m^2 under the 10% weight loss scenario and to 22.85 kg/m^2 under the 17% weight loss scenario. Second, neighbourhoods become more similar with respect to average BMI, as evidenced by the tighter range of density estimates under each scenario compared to the baseline scenario. This result is expected because each of the weight loss scenarios assume that the entire population of overweight individuals loses weight while body weight remains fixed in the population of under- and healthy-weight individuals.

[9] The bandwidth parameter h for the kernel density estimates was set to 0.0853 and was selected using the Sheather-Jones method.

Table 5.6. Average neighbourhood prevalence of overweight (BMI ≥ 25 kg/m^2) across metropolitan Toronto and each of its boroughs* under different weight loss scenarios.

Scenario	MT (n = 140)	ET (n = 20)	NY (n = 33)	YO (n = 10)	TO (n = 44)	EY (n = 8)	SC (n = 25)
Overall							
Baseline	41.2	42.2	40.2	42.7	42.2	42.0	39.0
4% loss	34.0	34.9	33.0	35.4	35.0	34.9	31.9
7% loss	28.1	28.8	27.2	29.2	28.9	28.8	26.5
10% loss	21.3	21.9	20.5	22.6	21.9	22.0	20.1
14% loss	13.8	14.2	13.0	14.5	14.3	14.5	13.0
17% loss	10.1	10.4	9.3	10.7	10.7	11.0	9.5
Men							
Baseline	46.1	47.3	44.7	46.5	48.3	47.2	42.7
4% loss	37.6	38.5	36.3	37.7	39.7	38.9	34.4
7% loss	30.4	31.1	29.4	30.5	31.8	31.3	28.4
10% loss	22.0	22.6	21.0	22.8	22.9	22.9	20.7
14% loss	13.6	14.0	12.8	13.9	14.3	14.4	12.9
17% loss	9.5	9.6	8.6	9.6	10.0	10.4	9.2
Women							
Baseline	36.6	37.5	36.2	39.3	36.5	37.4	35.6
4% loss	30.7	31.6	30.1	33.5	30.5	31.3	29.6
7% loss	25.9	26.7	25.3	28.2	26.0	26.6	24.7
10% loss	20.6	21.3	20.0	22.4	21.0	21.2	19.4
14% loss	13.9	14.3	13.2	15.0	14.4	14.6	13.0
17% loss	10.7	11.1	9.9	11.7	11.4	11.5	9.7

* MT = metropolitan Toronto, ET = Etobicoke, YO = York, TO = Toronto, EY = East York, SC = Scarborough.

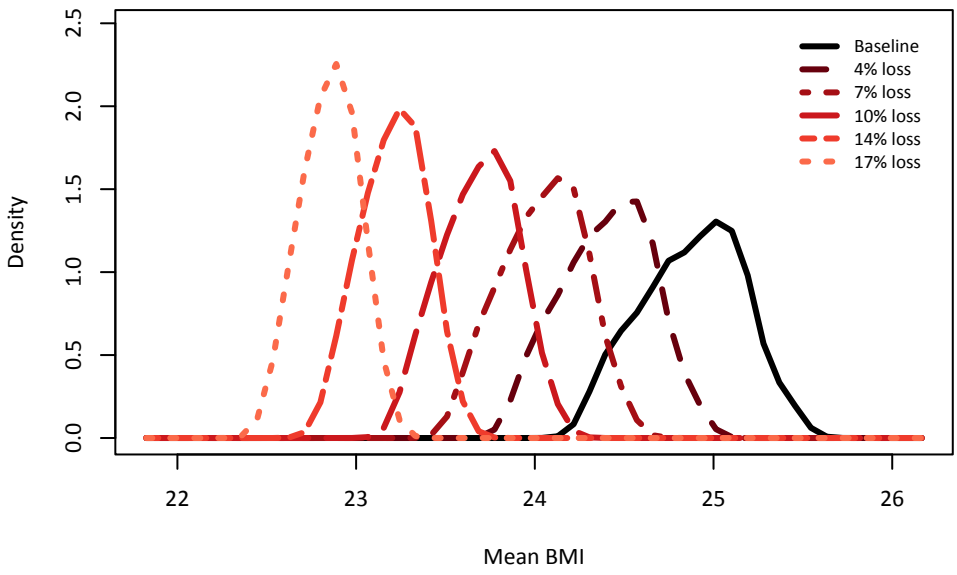


Figure 5.4. Distribution of average neighbourhood-level body mass index across metropolitan Toronto neighbourhoods under different weight loss scenarios.

5.3.2.2 Effect of population weight loss on projected diabetes incidence

Table 5.7 presents projected neighbourhood-level incidence rates averaged over all neighbourhoods under each weight loss scenario. Reductions in the incidence of diabetes are greatest when the entire population of overweight individuals loses weight, compared to only part of the population. If all overweight individuals lost 4% of their baseline body weight, the five-year incidence of diabetes might be reduced from 4.84% to 4.69%, on average, across metropolitan Toronto. Greater losses in baseline body weight are expected to produce greater reductions in the five-year incidence of diabetes. Significant changes in the neighbourhood distribution of projected incidence would not appear until the entire population of overweight individuals lost at least 10% of its baseline body weight. In this case, the average projected incidence might fall from 4.85% to 4.52% across metropolitan Toronto. A non-parametric Kruskal-Wallis test suggests a significant shift in the neighbourhood distribution of projected incidence would occur under this

scenario. In order to achieve a *one percentage point* reduction in the five-year incidence of diabetes, the entire overweight population would have to lose *17%* of its baseline body weight.

Table 5.7. Average forecast incidence of diabetes across metropolitan Toronto neighbourhoods under different weight loss scenarios: (a) complete population participation and (b) partial population participation.

			Test†		
Scenario	Mean (SD)	Median (MAD)	KW	Perm.	95% UI‡
(a) Entire population					
Baseline	4.84 (0.63)	4.87 (0.46)	—	—	4.20 – 5.86
4% loss	4.69 (0.65)	4.73 (0.47)			4.08 – 5.68
7% loss	4.65 (0.65)	4.70 (0.49)			3.94 – 5.66
10% loss	4.52 (0.65)	4.56 (0.47)	***	*	3.76 – 5.62
14% loss	4.15 (0.61)	4.19 (0.46)	***	***	3.45 – 5.22
17% loss	3.81 (0.56)	3.83 (0.42)	***	***	3.19 – 4.78
(b) Partial population participation§					
Baseline	4.84 (0.63)	4.87 (0.46)	—	—	4.20 – 5.86
10%	4.80 (0.63)	4.83 (0.45)			4.17 – 5.80
50%	4.68 (0.63)	4.72 (0.46)			3.97 – 5.66
Mean 4.2kg loss	4.64 (0.63)	4.68 (0.46)	*		3.97 – 5.66
80%	4.58 (0.64)	4.61 (0.46)	**		3.86 – 5.59
100%	4.52 (0.65)	4.56 (0.47)	***	*	3.76 – 5.62

† Tests the difference between each scenario and the baseline scenario.

KW = Kruskal-Wallis test; Perm. = permutation test of density estimates.

‡ 95% uncertainty interval

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

§ 10% weight loss in all scenarios.

Figure 5.5 presents the expected shift in average incidence at the neighbourhood level for each weight loss scenario under complete population participation. Each curve is a non-parametric kernel density estimate[10] of the neighbourhood distribution of projected incidence. Compared to the baseline scenario, significant shifts in the neighbourhood distribution of projected incidence do not oc-

[10] Bandwidth parameter $h = 0.2514$

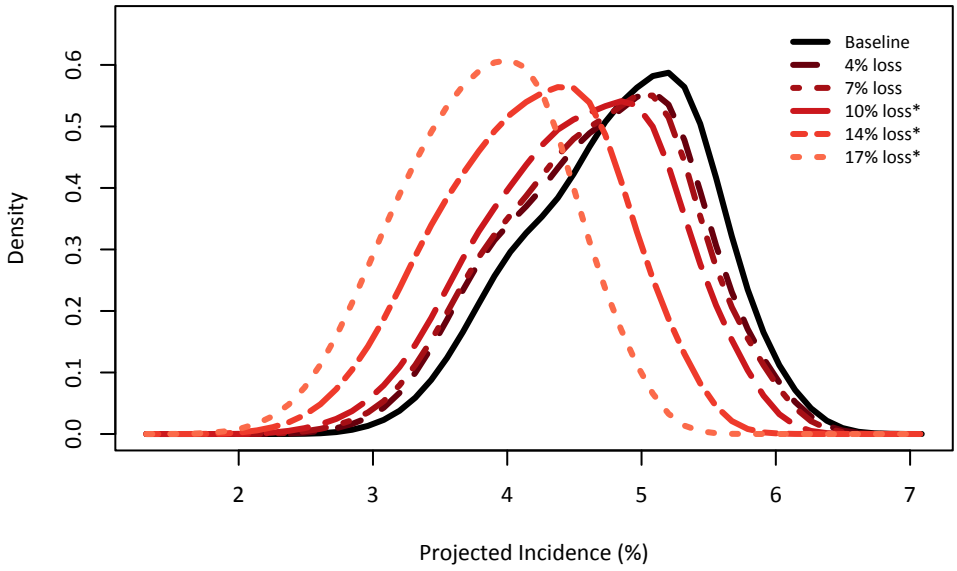


Figure 5.5. Distribution of average forecast incidence of diabetes across metropolitan Toronto neighbourhoods under different weight loss scenarios.

cur until the entire overweight population loses at least 10% of its baseline body weight. Even under this scenario, the distribution of projected incidence does not appear to deviate much from baseline. Only under the 14% and 17% weight loss scenarios do the distributions appear to be significantly different from baseline.

Another way to examine differences between scenarios is to consider the uncertainty around each set of estimates using probabilistic sensitivity analysis. The uncertainty intervals presented in Table 5.7 are based on 1000 samples of DPoRT model parameters selected using Latin hypercube sampling. They represent the range of plausible incidence projections under each scenario. In most cases, the average projected incidence lies within the uncertainty intervals for all other scenarios. The exception to this is the 17% weight loss scenario. In this case, the average projected incidence of 3.81% lies outside the 95% uncertainty interval for the baseline, 4%, and 7% weight loss scenarios. Based on these results, although the future incidence of diabetes might be reduced by a 10% reduction in baseline

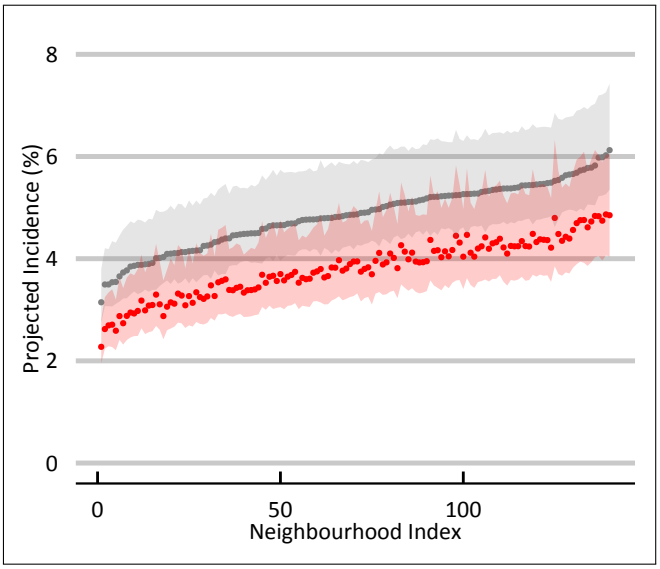


Figure 5.6. Uncertainty about forecast diabetes incidence across metropolitan Toronto neighbourhoods under the baseline (grey) and 17% weight loss (red) scenarios.

body weight, the range of uncertainty around these projections might still mean that projected incidence is not different from the baseline scenario. Indeed, important reductions in the projected incidence of diabetes might not appear until the entire population of overweight individuals loses at least 17% of its baseline body weight.

Figure 5.6 contrasts the uncertainty about these scenarios visually by ranking neighbourhoods in ascending order of forecast incidence. Baseline projections are presented in grey while projections from the 17% weight loss scenario are presented in red. The shaded ribbons surrounding each set of projections represent the 95% uncertainty interval. In almost all neighbourhoods, the projections from the 17% weight loss scenario do not lie within the uncertainty interval from the baseline scenario. This suggests there is some difference in projected incidence between these scenarios. In spite of this, there remains some overlap in the uncertainty around each scenario suggesting that a 17% weight loss might not produce reductions in incidence that differ meaningfully from the baseline scenario.

Table 5.7 also presents projected incidence averaged over all neighbourhoods for the partial population participation scenarios. In these scenarios, simulated individuals were randomly selected to lose 10% of their baseline body weight. When only 10% of the overweight population loses 10% of its baseline body weight, there is almost no perceptible shift in projected incidence. Slightly larger reductions in incidence are expected when at least 50% of the population loses 10% of its baseline body weight. The distribution of projected incidence at the neighbourhood level might only begin to shift in a meaningfully way when at least 80% of the overweight population loses at least 10% of its baseline body weight. Even then, there is enough uncertainty surrounding these estimates to conclude that projections from the 80% participation scenario might lie within the range of uncertainty surrounding the baseline scenario. In summary, small reductions in population body weight at the neighbourhood level might not produce meaningful reductions in the five-year incidence of diabetes.

Figure 5.7 summarizes projected neighbourhood-level incidence rates within each of Toronto's six boroughs. Under the baseline scenario, Scarborough neighbourhoods had the highest projected incidence (median = 5.4%) while neighbourhoods in Toronto (median = 4.1%) and East York (median = 4.7%) had the lowest projected incidence. Negligible reductions in projected incidence were produced by the 4%, 7%, and 10% weight loss scenarios. Larger reductions in body weight produced larger reductions in projected incidence across all neighbourhoods.

Figure 5.8 illustrates reductions in projected incidence between each scenario and the baseline scenario. Reductions varied between scenarios: some neighbourhoods experienced larger than average reductions in incidence (> 1 SD above the mean) while other neighbourhoods experienced smaller than average reductions (> 1 SD below the mean). Comparing the baseline scenario against the 4% weight loss scenario, most neighbourhoods in North York had lower than average reductions in projected incidence. Few neigbourhoods had larger than average reductions in projected incidence. Those that did tended to lie along Lake Ontario in southern Toronto.

In the 7% weight loss scenario, five neighbourhoods in northeast Scarborough experienced larger reductions in projected incidence compared to other neighbourhoods, where reductions tended to lie within ±1 SD of the mean reduction. In

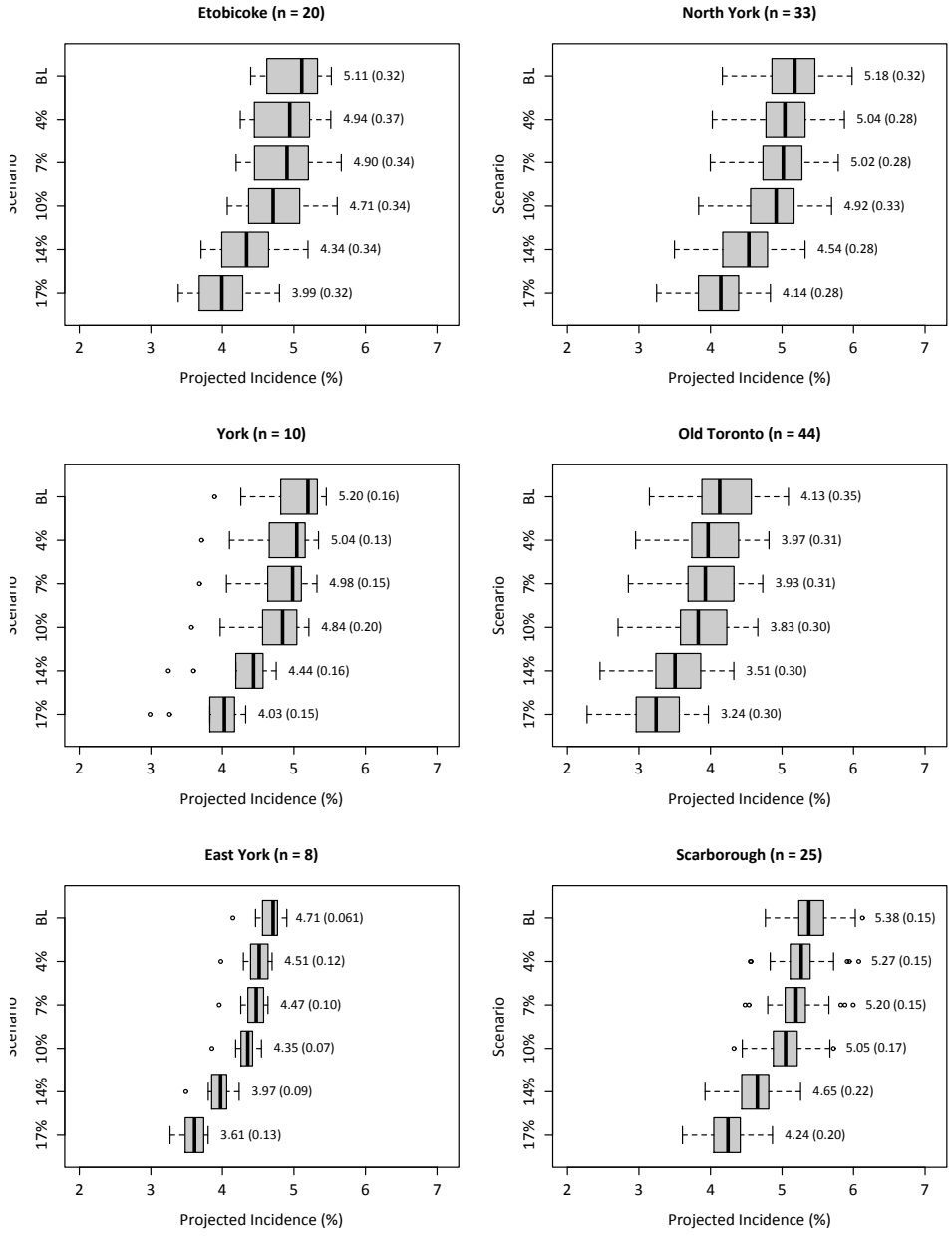


Figure 5.7. Neighbourhood distribution of forecast diabetes incidence within Toronto's six boroughs under different population weight loss scenarios. BL = baseline scenario.

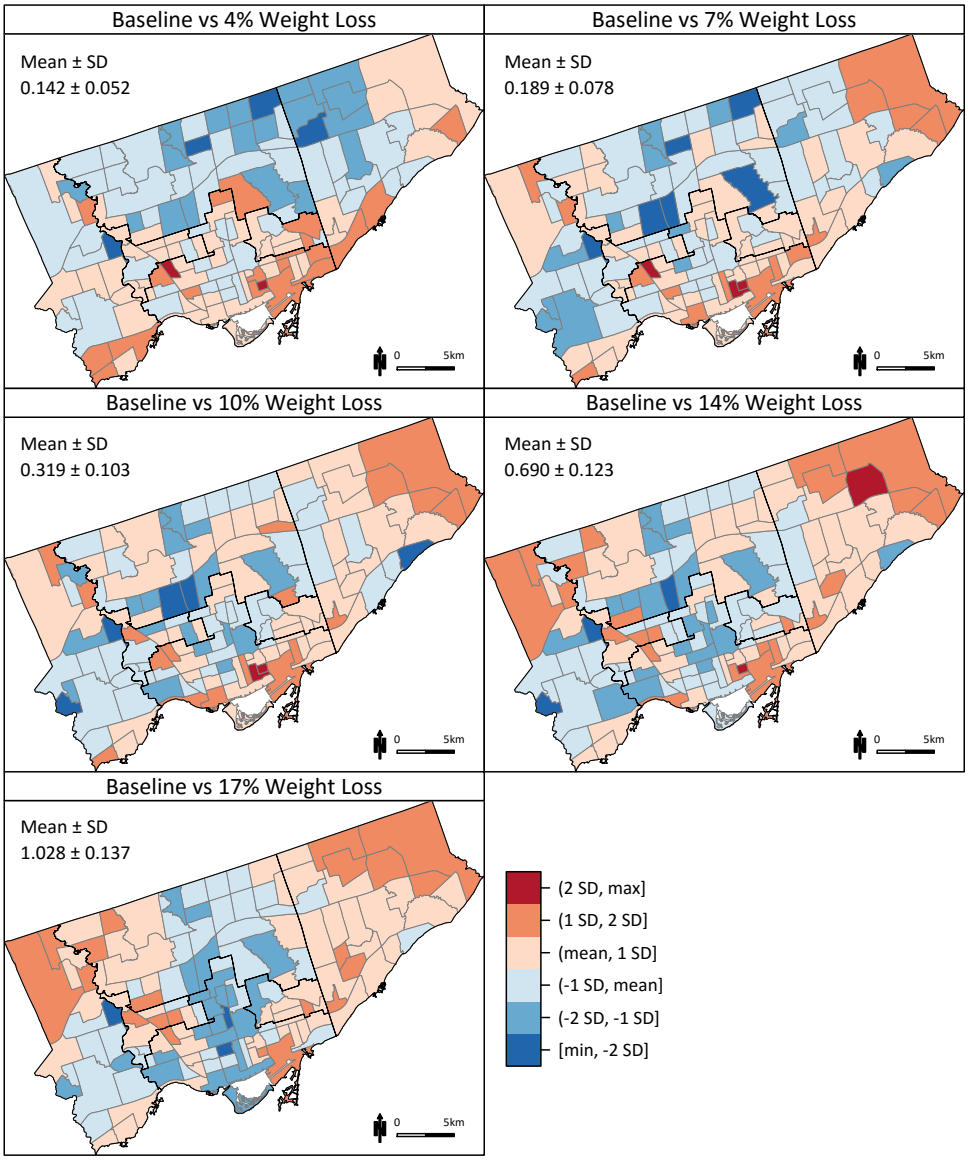


Figure 5.8. Average neighbourhood reduction in forecast diabetes incidence between each weight loss scenario and the baseline scenario.

the remaining scenarios, most of Scarborough's neighbourhoods had larger than average reductions in projected incidence. In Etobicoke, a north-south divide in projected reductions became apparent once the population of overweight indiduals lost at least 14% of its baseline body weight. Conversely, neighbourhoods in central Toronto tended to experience smaller than average reductions in projected incidence under the 14% and 17% weight loss scenarios.

These findings suggest that the populations of different neighbourhoods might not react in the same ways to health promotion programs and policies designed to promote weight loss. It might be the case that the population structure of different neighbourhoods produces such differences. For example, neighborhoods within north Etobicoke and Scarborough are composed of larger proportions of visible minorities compared to neighborhoods in Toronto. Moreover, these neighbourhoods had a lower baseline prevalence of overweight compared to the metropolitan average. Thus, it is plausible that weight loss in visible minority populations produces greater relative reductions in incidence, even though the overall projected incidence in these neighbourhoods remains above the city-wide average.

5.3.3 Probabilistic sensitivity analysis

Based on the simulated TropISM population, 4.9% of the population within any of Toronto's neighbourhoods might develop diabetes over a five-year period. An initial analysis of the uncertainty around these projections suggests that the projected five-year incidence could range from 4.2% to 5.9% (Table 5.7). Figure 5.6 demonstrates that the uncertainty about projected incidence rates is not uniform across neighbourhoods: some neighbourhoods appear to have a larger range in the uncertainty about projected rates than others. Generally, neighbourhoods having higher projected incidence seem to have larger uncertainty intervals. A formal probabilistic sensitivity analysis was conducted to identify which model parameters contributed most to uncertainty in forecast incidence.

This sensitivity analysis was conducted for each weight loss scenario in which the entire population at risk participated in the "intervention" program. Since projected rates are based on different model parameters for men and women, separate sensitivity analyses were conducted for men and women. As discussed in Sec-

tion 5.2.4, Monte Carlo methods were used to randomly sample DPoRT model parameters from a normal distribution using the estimated DPoRT parameters as the mean and their standard errors as the standard deviation (Table 5.1). One thousand sets of model parameters were used to produce 1000 projections for each neighbourhood. Results were aggregated over the entire TropISM population to assess the overall effect of uncertain inputs on the variability about projected incidence.

The partial rank correlation coefficient (PRCC) assesses the strength of the correlation between each model parameter and the uncertainty about projected diabetes incidence. Correlations closer to +1 or -1 suggest that a model parameter contributes more to forecast uncertainty than others after controlling for all other model parameters. The sensitivity index was then used to assess the proportion of variation in forecast incidence that is attributable to each model parameter.

5.3.3.1 Factors influencing forecast uncertainty

Uncertainty in projected diabetes incidence for men and women was strongly influenced by the "Intercept" and "Scale" parameters in the DPoRT model. Each parameter exerted similar effects on forecast uncertainty. Tables 5.8 and 5.9 depict the partial rank correlation coefficient (PRCC) for DPoRT model parameters for men and women, respectively. For all scenarios, each DPoRT parameter is ranked according to the absolute magnitude of the PRCC. Among men, the baseline model showed a strong negative correlation between the intercept parameter and forecast uncertainty (PRCC = -0.93) while the scale parameter had a strong positive correlation (PRCC = 0.88). More specifically, as the intercept coefficient increased in magnitude, the projected incidence of diabetes decreased. In other words, larger values of the intercept parameter produced smaller projected incidence rates. Given the DPoRT model is based on an accelerated failure time model, these results are expected. That is, a positive value of the intercept parameter means that the time to "failure" (or onset of diabetes) increases. At a population level, this reduces the projected incidence of diabetes.

On the other hand, larger values of the scale parameter increased the projected incidence of diabetes. Among men, the intercept and scale parameters ex-

Table 5.8. Influence of DPoRT model parameters on forecast uncertainty of male diabetes incidence, as measured by the partial rank correlation coefficient (PRCC) for each modelling scenario.

	Baseline	4% Loss	7% Loss	10% Loss	14% Loss	17% Loss
Parameters	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)
Intercept	-0.933 (1)	-0.935 (1)	-0.925 (1)	-0.907 (1)	-0.910 (1)	-0.895 (1)
Scale	0.884 (2)	0.895 (2)	0.888 (2)	0.874 (2)	0.883 (2)	0.862 (2)
≥ 45, $25 \leq$ BMI < 30	-0.754 (3)	-0.703 (3)	-0.577 (5)	-0.235 (7)	0.170 (8)	0.432 (6)
< 45, $25 \leq$ BMI < 30	0.566 (4)	0.578 (5)	0.589 (4)	0.597 (4)	0.612 (4)	0.643 (5)
< 45, $23 \leq$ BMI < 25	0.562 (5)	0.498 (6)	0.388 (6)	0.272 (5)	0.236 (7)	0.264 (7)
≥ 45, $30 \leq$ BMI < 35	-0.471 (6)	-0.207 (9)	-0.113 (12)	-0.190 (9)	-0.037 (15)	0.014 (16)
Hypertension	-0.320 (7)	-0.261 (7)	-0.321 (7)	-0.270 (6)	-0.265 (6)	-0.227 (8)
≥ 45, BMI ≥ 35	-0.257 (8)	-0.231 (8)	-0.180 (8)	-0.045 (16)	-0.022 (17)	0.034 (14)
≥ 45, $23 \leq$ BMI < 25	-0.245 (9)	-0.602 (4)	-0.747 (3)	-0.804 (3)	-0.821 (3)	-0.797 (3)
Non-white	-0.166 (10)	-0.109 (11)	-0.139 (10)	-0.113 (12)	-0.129 (11)	-0.140 (11)
< 45, BMI ≥ 35	-0.165 (11)	-0.075 (15)	-0.047 (15)	-0.048 (15)	0.027 (16)	-0.013 (17)
Heart disease	-0.130 (12)	-0.149 (10)	-0.116 (11)	-0.077 (13)	-0.139 (10)	-0.115 (12)
Post-secondary	0.122 (13)	0.099 (13)	0.151 (9)	0.152 (11)	0.096 (12)	0.148 (10)
< 45, $30 \leq$ BMI < 35	-0.102 (14)	-0.028 (16)	0.061 (14)	0.182 (10)	0.166 (9)	0.174 (9)
Smoker	0.039 (15)	0.107 (12)	0.014 (16)	0.069 (14)	0.043 (14)	0.025 (15)
≥ 45, BMI < 23	0.030 (16)	-0.010 (17)	-0.012 (17)	-0.196 (8)	-0.576 (5)	-0.722 (4)
Highest income	0.011 (17)	0.085 (14)	0.106 (13)	0.040 (17)	0.074 (13)	0.096 (13)

Table 5.9. Influence of DPoRT model parameters on forecast uncertainty of female diabetes incidence, as measured by the partial rank correlation coefficient (PRCC) for each modelling scenario.

	Baseline	4% Loss	7% Loss	10% Loss	14% Loss	17% Loss
Parameters	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)	PRCC (rank)
Intercept	-0.951 (1)	-0.935 (1)	-0.927 (1)	-0.911 (1)	-0.907 (1)	-0.908 (1)
Scale	0.873 (2)	0.840 (2)	0.806 (3)	0.779 (4)	0.758 (4)	0.759 (4)
≥ 65, 23 ≤ BMI < 25	-0.742 (3)	-0.811 (3)	-0.819 (2)	-0.849 (2)	-0.862 (2)	-0.874 (2)
≥ 65, BMI < 23	-0.709 (4)	-0.757 (4)	-0.782 (4)	-0.813 (3)	-0.847 (3)	-0.853 (3)
Hypertension	-0.379 (5)	-0.359 (6)	-0.286 (6)	-0.317 (6)	-0.368 (6)	-0.324 (7)
< 45, 23 ≤ BMI < 25	0.303 (6)	0.244 (7)	0.158 (10)	0.159 (8)	0.101 (14)	0.165 (10)
45–64, BMI < 23	0.297 (7)	0.037 (20)	-0.086 (13)	-0.122 (12)	-0.151 (11)	-0.164 (11)
45–64, 25 ≤ BMI < 30	-0.274 (8)	-0.200 (8)	0.008 (22)	0.088 (14)	0.172 (10)	0.283 (8)
45–64, 30 ≤ BMI < 35	-0.273 (9)	-0.096 (15)	-0.045 (17)	-0.041 (18)	-0.009 (23)	0.101 (15)
≥ 65, 25 ≤ BMI < 30	-0.273 (10)	-0.192 (10)	-0.101 (12)	-0.028 (20)	0.127 (12)	0.133 (13)
≥ 65, BMI unknown	-0.226 (11)	-0.142 (13)	-0.150 (11)	-0.128 (11)	-0.176 (9)	-0.194 (9)
< 45, 25 ≤ BMI < 30	0.194 (12)	0.192 (9)	0.245 (7)	0.264 (7)	0.299 (7)	0.342 (6)
45–64, 23 ≤ BMI < 25	-0.186 (13)	-0.366 (5)	-0.491 (5)	-0.442 (5)	-0.462 (5)	-0.494 (5)
45–64, BMI ≥ 35	-0.177 (14)	-0.065 (16)	-0.066 (15)	-0.044 (17)	0.014 (22)	-0.023 (20)
Post-secondary	0.165 (15)	0.174 (11)	0.175 (8)	0.133 (10)	0.209 (8)	0.148 (12)
Immigrant	-0.162 (16)	-0.174 (12)	-0.161 (9)	-0.113 (13)	-0.101 (15)	-0.127 (14)
≥ 65, BMI ≥ 35	-0.152 (17)	-0.103 (14)	-0.049 (16)	0.008 (22)	0.058 (17)	0.022 (21)
45–64, BMI unknown	-0.111 (18)	-0.058 (17)	-0.004 (23)	-0.141 (9)	-0.047 (19)	-0.035 (19)
Non-white	-0.099 (19)	-0.053 (19)	-0.071 (14)	-0.020 (21)	0.072 (16)	-0.013 (22)
< 45, BMI ≥ 35	-0.077 (20)	-0.022 (22)	-0.025 (19)	-0.038 (19)	0.106 (13)	0.036 (18)
≥ 65, 30 ≤ BMI < 35	-0.055 (21)	-0.034 (21)	-0.043 (18)	0.004 (23)	-0.046 (20)	0.009 (23)
< 45, 30 ≤ BMI < 35	-0.031 (22)	0.057 (18)	-0.013 (21)	0.044 (16)	0.037 (21)	0.059 (16)
< 45, BMI unknown	0.012 (23)	-0.015 (23)	-0.018 (20)	0.058 (15)	0.054 (18)	-0.048 (17)

erted the strongest influence on forecast uncertainty across all modelling scenarios. In the baseline model, body mass index parameters also influenced forecast uncertainty. Specifically, as the model coefficient for overweight among older men (aged 45+) increased, projected incidence decreased. Although this might seem counter-intuitive, it is important to remember that the average coefficient for this parameter was -2.056 (Table 5.1). Since overweight and age are both risk factors for diabetes, weaker coefficients (i.e., values that move towards zero) mean that other model parameters are influencing projected incidence. In this case, weaker values of the overweight coefficient for older men produced *lower* projected incidence rates in the baseline model.

Opposite effects were observed for the overweight ($25 \leq \text{BMI} < 30$) and healthy weight ($23 \leq \text{BMI} < 25$) parameters among men younger than 45. Across all modelling scenarios, weaker values for these coefficients (values that move towards zero) mean that these parameters exerted less influence on projected incidence rates, allowing other parameters to exert more influence. This, in turn, increased the forecast incidence of diabetes.

Under the different weight loss scenarios, most model parameters exerted a similar influence on forecast uncertainty. In other words, the sign of the PRCC was the same for a given parameter across scenarios. However, the magnitude of the PRCC differed across scenarios, meaning that the same parameter had a stronger effect on model uncertainty under some scenarios. For example, the overweight parameter ($25 \leq \text{BMI} < 30$) for older men was the third most influential parameter in the baseline scenario, but the only sixth most influential parameter in the 17% weight loss scenario. These differences indicate that forecast uncertainty was influenced by the weight loss scenarios.

Specifically, as the simulated population of high-risk individuals loses weight, there is a greater number of healthy weight individuals in that population and fewer overweight individuals. Model parameters associated with lower BMIs therefore exert greater influence on forecast uncertainty. For example, in the baseline scenario, the healthy BMI parameter for older men was ranked ninth. In the 10%, 14%, and 17% weight loss scenarios, this parameter was ranked third. Holding other parameter values constant, as the value of the healthy BMI parameter for older men increases (towards zero), the projected incidence of diabetes

decreases because (a) there are fewer overweight individuals in the population, which would increase the population incidence of diabetes, and (b) there are more healthy weight individuals in the population affected by the weaker healthy weight parameter. Each of these factors combined means that the remaining model parameters exert more influence on projected diabetes incidence, especially the intercept parameter. The overall effect, then, is to reduce the projected incidence of diabetes at the population level.

Another consistent finding to note is that some model parameters had very little influence on forecast uncertainty. Among men, the smoking status, income, heart disease, and visible minority parameters only weakly influenced forecast uncertainty. In most cases, the absolute value of the PRCC for these parameters did not exceed 0.15.

Results of the probabilistic sensitivity analysis for women were slightly different. Similar to the results for men, the intercept parameter exerted the greatest influence on forecast uncertainty across all scenarios. In all cases, larger values of the intercept coefficient were associated with lower projected incidence rates. The scale parameter was also associated with higher projected incidence rates. Unlike the results for men, the scale parameter was not consistently the second most important parameter influencing forecast uncertainty. Instead, the BMI parameters for women aged 65+ tended to be more influential than the scale parameter. Across all weight loss scenarios, the PRCC for these parameters was negative. Thus, larger values (i.e., values that move towards zero) produced lower projected incidence rates. Again, this result seems counter-intuitive. However, as the effect for these parameters become weaker, there is less influence of older healthy weight or underweight women in the model. This means that projected incidence rates are influenced by the remaining parameters in the model. Since the DPoRT model uses the youngest age group and lowest body weight group as the reference parameter (which is associated with the intercept parameter), the intercept parameter indirectly gains more influence. This is why larger, less negative values of the healthy weight and underweight parameters for older women *reduce* the forecast incidence of diabetes in the weight loss scenarios.

Response surface plots illustrate how diabetes incidence changes as a function of using different randomly sampled values for DPoRT model parameters. A

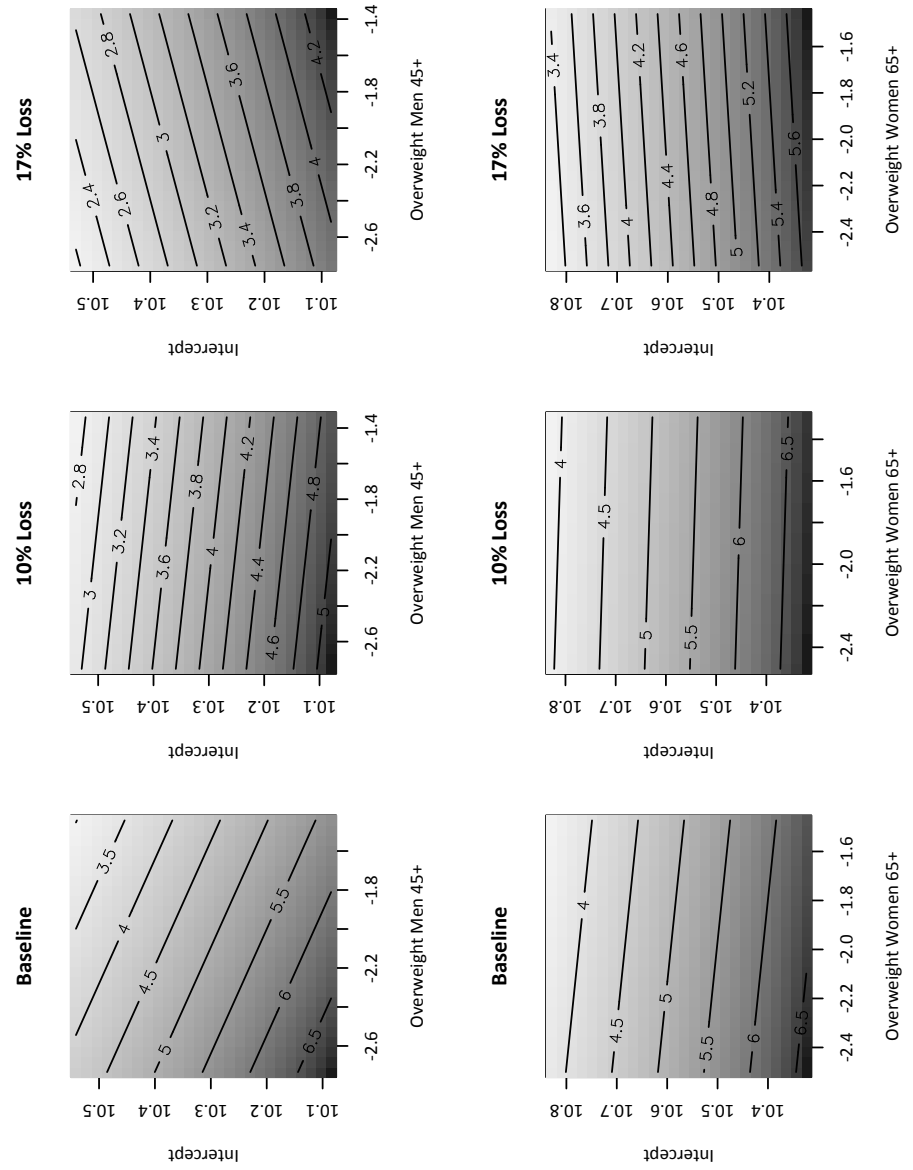


Figure 5.9. Response surface plots depicting the effect of changes in DPoRT model parameters on uncertainty surrounding the forecast incidence of diabetes for the baseline, 10% weight loss, and 17% weight loss scenarios.

response surface plot is based on a linear regression of projected incidence rates on the randomly sampled model coefficients. These plots illustrate how changes in one parameter will affect projected incidence while holding a second parameter (and all other model parameters) constant. Figure 5.9 presents six response surface plots of projected incidence under the baseline, 10% weight loss, and 17% weight loss scenarios for men and women, respectively. In these plots, the effect of the intercept parameter was plotted against the overweight parameter for the oldest men (aged 45+) and women (aged 65+). Weaker (less negative) values of the overweight parameter for older men produced lower projected incidence rates in the baseline scenario. This would be expected because weaker values for this parameter remove its influence on projected incidence, while more negative values would increase its influence on projected incidence. In the 10% weight loss scenario, weaker values of the overweight parameter have little effect on projected diabetes incidence. In the 17% weight loss scenario, weaker values for this parameter increase projected incidence. In this case, due to population weight loss, there would be fewer individuals in this category and, as a result, more in lower body weight categories. Increasing the value of this coefficient means that this parameter has less influence on projected incidence while the lower body weight parameter, if held constant, has more influence. The net effect is to increase diabetes incidence. Similar, but weaker, effects were seen for the overweight parameter for women aged 65+.

5.3.3.2 Factors explaining variability in forecast uncertainty

The sensitivity index is another metric used to identify which model parameters contribute most to forecast uncertainty. As described in Section 5.2.4, the sensitivity index ranges from 0 to 1 and can be interpreted as the proportion of variation in forecast incidence that is explained by each parameter in the DPoRT model. Tables 5.10 and 5.11 present the sensitivity index for each DPoRT parameter across all modelling scenarios for men and women, respectively. For each scenario, the five most influential parameters are ranked to easily identify which parameters contribute most to forecast uncertainty. The sensitivity index was estimated in SaSAT using regression models that regressed forecast incidence on each input

parameter. The high coefficient of determination (R^2) for each model suggests the sensitivity index was reliably estimated (Hoare et al., 2008).

Overall, similar conclusions can be drawn from this analysis as from the partial rank correlation analysis. For male- and female-specific forecasts, the intercept parameter is the most important parameter influencing uncertainty, explaining 36% to 50% of the variation in forecast uncertainty in men and as much as 58% of the variation in women. Moreover, the intercept parameter explained a greater amount of variation in the baseline scenarios than in the 17% weight loss scenario. Among men, the scale parameter was the second most influential parameter, explaining almost 30% of the variation in forecast uncertainty across all scenarios. Among women, the scale parameter was less influential, explaining only 20% of the variation in forecast uncertainty in the baseline scenario and only 11% of the variation in forecast uncertainty in the 17% weight loss scenario.

Among men, the parameter for overweight men aged 45+ explained almost 10% of the variation in forecast uncertainty in the baseline model. This parameter became less influential under greater weight loss scenarios. In the 17% weight loss scenario, this parameter explained only 2% of the variation in forecast uncertainty while the healthy weight parameter for men aged 45+ explained almost 16% of the variation in forecast uncertainty. Among women, the parameters representing the lowest BMI categories in the oldest women explained between 6% and 25% of the variation in forecast uncertainty across all scenarios.

Based on either the partial rank correlation coefficient or the sensitivity index, it is clear that the intercept and scale parameters from the DPoRT model exert the greatest influence on forecast uncertainty. Even under different weight loss scenarios, these parameters explain a substantial amount of the variation in forecast uncertainty. Having said that, certain model parameters representing different body weight categories become more influential under the weight loss scenarios. The combined results suggest that specific model parameters *as well as* the underlying structure of the population at risk influence forecast uncertainty. What is more, this analysis illustrates that a range of plausible changes in forecast incidence might result from population level interventions designed to reduce the five-year incidence of diabetes. Based on these results, substantial reductions in population body weight would be required to meaningfully reduce the five-year

Table 5.10. Proportion of forecast uncertainty in male diabetes incidence explained by DPoRT model parameters, as measured by the sensitivity index, for all modelling scenarios.

Parameter	Baseline	4% Loss	7% Loss	10% Loss	14% Loss	17% Loss
R^2	*0.9835*	*0.9828*	*0.9805*	*0.9770*	*0.9806*	*0.9775*
Intercept	$0.5021^{(1)}$	$0.4936^{(1)}$	$0.4804^{(1)}$	$0.4440^{(1)}$	$0.3957^{(1)}$	$0.3575^{(1)}$
Hypertension	0.0078	0.0069	0.0078	0.0063	0.0061	0.0051
Non-white	0.0007	0.0014	0.0007	0.0017	0.0017	0.0019
Heart disease	0.0013	0.0017	0.0023	0.0005	0.0015	0.0010
Smoker	0.0005	0.0002	0.0003	0.0002	0.0001	0.0001
Post-secondary	0.0018	0.0014	0.0016	0.0023	0.0011	0.0016
Highest income	0.0005	0.0003	0.0005	0.0006	0.0011	0.0012
$< 45, 23 \leq$ BMI < 25	$0.0393^{(4)}$	0.0258	0.0128	$0.0085^{(5)}$	0.0046	0.0057
$< 45, 25 \leq$ BMI < 30	$0.0352^{(5)}$	$0.0416^{(5)}$	$0.0455^{(4)}$	$0.0521^{(4)}$	$0.0610^{(4)}$	$0.0620^{(5)}$
$< 45, 30 \leq$ BMI < 35	0.0010	0.0000	0.0005	0.0026	0.0026	0.0031
< 45, BMI ≥ 35	0.0028	0.0010	0.0005	0.0005	0.0002	0.0002
≥ 45, BMI < 23	0.0003	0.0000	0.0001	0.0040	$0.0438^{(5)}$	$0.1056^{(4)}$
$\geq 45, 23 \leq$ BMI < 25	0.0055	$0.0436^{(4)}$	$0.0965^{(3)}$	$0.1784^{(3)}$	$0.1907^{(3)}$	$0.1580^{(3)}$
$\geq 45, 25 \leq$ BMI < 30	$0.0985^{(3)}$	$0.0709^{(3)}$	$0.0440^{(5)}$	0.0044	0.0050	0.0205
$\geq 45, 30 \leq$ BMI < 35	0.0253	0.0036	0.0013	0.0019	0.0001	0.0000
≥ 45, BMI ≥ 35	0.0071	0.0045	0.0020	0.0001	0.0000	0.0001
Scale	$0.2702^{(2)}$	$0.3034^{(2)}$	$0.3032^{(2)}$	$0.2919^{(2)}$	$0.2845^{(2)}$	$0.2764^{(2)}$

Table 5.11. Proportion of forecast uncertainty in female diabetes incidence explained by DPoRT model parameters, as measured by the sensitivity index, for all modelling scenarios.

Parameter	Baseline	4% Loss	7% Loss	10% Loss	14% Loss	17% Loss
R^2	*0.9909*	*0.9886*	*0.9876*	*0.9882*	*0.9862*	*0.9884*
Intercept	$0.5893^{(1)}$	$0.5236^{(1)}$	$0.4820^{(1)}$	$0.4163^{(1)}$	$0.3726^{(1)}$	$0.3626^{(1)}$
Hypertension	$0.0110^{(5)}$	0.0123	0.0110	0.0112	0.0104	0.0113
Non-white	0.0004	0.0005	0.0001	0.0000	0.0001	0.0000
Immigrant	0.0023	0.0025	0.0021	0.0012	0.0014	0.0013
Post-secondary	0.0026	0.0028	0.0020	0.0030	0.0024	0.0025
< 45, $23 \leq$ BMI < 25	0.0073	0.0045	0.0033	0.0015	0.0016	0.0012
< 45, $25 \leq$ BMI < 30	0.0023	0.0038	0.0042	0.0063	0.0076	0.0064
< 45, $30 \leq$ BMI < 35	0.0002	0.0000	0.0000	0.0001	0.0000	0.0003
< 45, BMI ≥ 35	0.0007	0.0000	0.0000	0.0000	0.0005	0.0001
< 45, BMI unknown	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
45–64, BMI < 23	0.0062	0.0002	0.0006	0.0019	0.0017	0.0028
45–64, $23 \leq$ BMI < 25	0.0033	$0.0125^{(5)}$	$0.0229^{(5)}$	$0.0250^{(5)}$	$0.0230^{(5)}$	$0.0273^{(5)}$
45–64, $25 \leq$ BMI < 30	0.0044	0.0019	0.0000	0.0006	0.0029	0.0040
45–64, $30 \leq$ BMI < 35	0.0048	0.0004	0.0006	0.0003	0.0000	0.0003
45–64, BMI ≥ 35	0.0018	0.0007	0.0004	0.0002	0.0000	0.0000
45–64, BMI unknown	0.0007	0.0003	0.0002	0.0003	0.0004	0.0004
≥ 65, BMI < 23	$0.0628^{(4)}$	$0.1076^{(4)}$	$0.1374^{(4)}$	$0.1834^{(3)}$	$0.2095^{(3)}$	$0.2120^{(3)}$
≥ 65, $23 \leq$ BMI < 25	$0.0834^{(3)}$	$0.1416^{(3)}$	$0.1729^{(2)}$	$0.2152^{(2)}$	$0.2492^{(2)}$	$0.2486^{(2)}$
≥ 65, $25 \leq$ BMI < 30	0.0064	0.0028	0.0010	0.0003	0.0005	0.0012
≥ 65, $30 \leq$ BMI < 35	0.0006	0.0001	0.0000	0.0000	0.0000	0.0000
≥ 65, BMI ≥ 35	0.0017	0.0006	0.0004	0.0000	0.0000	0.0000
≥ 65, BMI unknown	0.0026	0.0026	0.0024	0.0023	0.0026	0.0031
Scale	$0.2054^{(2)}$	$0.1787^{(2)}$	$0.1566^{(3)}$	$0.1308^{(4)}$	$0.1136^{(4)}$	$0.1145^{(4)}$

incidence of diabetes. The range of uncertainty about projected incidence illustrates that small changes in body weight at the population level may only produce small reductions in the future incidence of diabetes.

5.3.3.3 Geographic differences in forecast uncertainty

Results of the probabilistic sensitivity analysis presented thus far provide an overall picture of the factors that explain forecast uncertainty across the entire TropISM population for metropolitan Toronto. However, Toronto's population is heterogeneous and different neighbourhoods within the city have different demographic profiles. As a result, it is reasonable to expect that uncertainty surrounding forecast incidence might be influenced by different model parameters in different neighbourhoods. Appendix B presents results of a probabilistic sensitivity analysis conducted at the neighbourhood level. For this analysis, the sensitivity index was estimated for all DPoRT model parameters for each neighbourhood under the baseline and 10% weight loss scenarios.[11]

Similar to the overall analysis, the intercept and scale parameters explained the largest proportion of forecast uncertainty for both the baseline scenario and the 10% weight loss scenario, irrespective of gender. However, the proportion of forecast uncertainty explained by these model parameters varies across neighbourhoods. In the baseline scenario among men (Appendix B, Figure B.5), the intercept parameter accounts for as little as 40% of forecast uncertainty in some areas of Toronto. However, it explains more than 50% of forecast uncertainty in the waterfront neighbourhoods of Toronto as well as in some areas of Etobicoke and Scarborough. Similar patterns were seen among women, where the intercept parameter explained as much as 60% of forecast uncertainty (Figure B.6).

Figures B.5 (men) and B.6 (women) present neighbourhood-specific estimates of the sensitivity index for the DPoRT model parameters that contribute most to overall forecast uncertainty. These figures also present neighbourhood-specific estimates of the sensitivity index for the visible minority parameter. Among men,

[11] Table B.2 in Appendix B presents the coefficient of determination from the regression models used to estimate neighbourhood-specific sensitivity indices across Toronto's boroughs. All models had very high values for the coefficient of determination (> 0.98). Thus, neighbourhood-specific sensitivity indices were reliably estimated.

the overweight parameter for older men (age 45+) explains more than 10% of forecast uncertainty in some neighbourhoods (Figure B.5c) although it accounts for less 10% of forecast uncertainty in many neighbourhoods. The healthy weight parameter for younger men (Figure B.5d) explains only a small proportion of forecast uncertainty across all neighbourhoods. Interestingly, in some neighbourhoods, visible minority status explains at least as much of the uncertainty in forecast diabetes incidence as some of the body weight parameters among men (5%–10%) and women (9%–15%) alike. However, in many neighbourhoods, visible minority status exerts little influence on forecast uncertainty.

As in the overall analysis, the neighbourhood-specific sensitivity analysis illustrates that different model parameters exert a greater influence on forecast uncertainty in the weight loss scenarios compared to the baseline scenario. Among men and women alike, in the 10% weight loss scenario, both the intercept and scale parameters explain smaller amounts of forecast uncertainty than under the baseline scenario. Among men, the overweight BMI parameter among older men explains 3%–19% of forecast uncertainty across all neighbourhoods in the baseline scenario (Figure B.5c) while the healthy weight BMI parameter among older men explains 7%–24% of outcome uncertainty in the 10% weight loss scenario (Figure B.7c). This result might be expected, given that weight loss in this subgroup is where most of the weight loss would occur.

Another interesting finding that applies to the gender-specific forecasts under either scenario is that the visible minority parameter exerts a negligible effect on forecast uncertainty in almost all of Toronto's neighbourhoods *except* in some areas of north Scarborough, where there is a sizable visible minority population (Figure A.1 in Appendix A). Thus, although visible minority status may only account for 5% of forecast uncertainty among men and as much as 15% of forecast uncertainty among women, this parameter exerts more influence on forecast uncertainty in these neighbourhoods compared other areas of Toronto.

These results indicate that neighbourhood-specific population profiles influence DPoRT model uncertainty. Moreover, these differences are maintained under different modelling scenarios. For example, visible minority status remains an important contributor to forecast uncertainty in some areas of Scarborough, even under the 10% weight loss scenario. This suggests that in spite of popula-

tion weight loss, visible minority status still influences uncertainty, even if it only accounts for a fraction of it (5%–15%). There are clearly areas within Toronto where additional factors beyond weight loss should be considered when planning population-level health promotion programs designed to encourage weight loss.

5.4 Discussion

5.4.1 Implications

Having developed the TropISM model of type 2 diabetes and its risk factors for metropolitan Toronto, it was possible to forecast the five-year incidence of diabetes at the neighbourhood level using the Diabetes Population Risk Tool. An initial scenario was developed assuming the baseline population of individuals remained unchanged, retaining its initial risk factor profile. Under this scenario, from 2006 to 2010, 4.9% of Toronto adults aged 20 and older were predicted to develop diabetes. Previous studies of known and forecast incidence reported diabetes incidence rates of 4.1% (Lipscombe & Hux, 2007) and 4.7% (Manuel et al., 2013), respectively, for the province of Ontario. These rates are broadly consistent with the forecast incidence reported here using the simulated TropISM population.

Using newly diagnosed cases of diabetes ascertained from the Ontario Diabetes Database from 2006 to 2010, the true incidence of diabetes, after removal of false positive diagnoses, was estimated to be 5.8%. Although the DPoRT model under-predicts diabetes incidence, the true incidence was within the range of forecast uncertainty. Therefore, at an aggregate level, the DPoRT model was able to forecast diabetes incidence reasonably well using the simulated TropISM population. Sex-specific estimates of diabetes incidence also fell within the range of forecast uncertainty, although the magnitude of under-prediction was greater among men than women.

Forecast incidence varied across metropolitan Toronto; in the six boroughs, incidence ranged from a low of 4.2% in Toronto to highs of 5.1% in North York and 5.4% in Scarborough. This variation was also apparent at the neighbourhood level, although forecasts for some neighbourhoods were more accurate than others. Forecast incidence was lower than the true incidence of diabetes in several

neighbourhoods within Scarborough and some neighbourhoods of north Etobicoke and west North York. In these neighbourhoods, DPoRT under-predicted true incidence by 2.3 percentage points, on average. Conversely, DPoRT over-predicted incidence in some parts of Toronto, although the magnitude of over-prediction was much smaller (0.84 percentage points, on average).

In spite of these absolute differences between true and forecast incidence, DPoRT seemed capable of capturing broad spatial patterns in diabetes incidence at the neighbourhood level. Neighbourhoods having elevated forecast incidence rates differed from those having lower rates. In particular, neighbourhoods whose populations were comprised of larger percentages of (a) women age 65 or older, (b) immigrants, (c) visible minorities, and (d) smaller percentages of high income earners had higher forecast incidence rates in the baseline scenario. Conditioned choropleth maps suggest that these neighbourhoods clustered in particular areas of metropolitan Toronto. For instance, neighbourhoods whose population was comprised of greater than the metropolitan median percentage of women aged 65+ and immigrants tended to have higher forecast incidence rates than other neighbourhoods in Toronto. Most of these neighbourhoods were located in North York and Scarborough. Thus, the demographic composition of neighbourhoods clearly influences forecast incidence using the DPoRT model.

As discussed in Chapter 4, the TropISM model under-predicted the true prevalence of several risk factors used by the DPoRT model to forecast diabetes incidence. In spite of this, the model was able to replicate broad spatial patterns in diabetes incidence, even though forecasts in some neighbourhoods were less accurate than in others. These results suggest that the spatial patterns of some risk factors (e.g., body weight) may be approximately correct, since the DPoRT model was able to replicate spatial patterns in diabetes incidence. In addition, simulated neighbourhood populations were accurately replicated for several demographic factors used by the DPoRT model, including sex, age, education, ethnicity, and immigrant status. Because these factors were used as constraint variables to develop TropISM, their accurate replication may have contributed to more accurate forecasts of diabetes incidence.

Another factor which may have contributed to more accurate forecasts of diabetes incidence is the way in which the DPoRT model was developed. In particu-

lar, this model linked respondents from the Ontario subsets of the 2001 and 2003 Canadian Community Health Surveys to the Ontario Diabetes Database (Rosella et al., 2014). Respondents were followed to 2011 to determine whether they developed diabetes. Because survey data are subject to self-report biases, including undiagnosed cases of important risk factors (e.g., hypertension), the regression model, upon which DPoRT is based, predicted diabetes status in 2011 using survey responses subject to these biases. In other words, the DPoRT model was calibrated to predict diabetes status from data containing inherent biases. Thus, even though the simulated TropISM population under-predicts the prevalence of important risk factors, the simulated data could still be used to forecast diabetes incidence over a five-year time period.

Since forecast incidence was more accurately replicated at the neighbourhood level than diabetes prevalence, it was possible to conduct policy simulations to predict the potential impact of population weight loss on reductions in diabetes incidence. It should be recognized, however, that if the spatial distribution of excess body weight (BMI $\geq$ 25 kg/m^2) was not accurately replicated, this bias would influence the conclusions drawn from these policy simulations. Given that DPoRT was able to forecast the overall incidence of diabetes reasonably well, and given that broad spatial patterns in forecast incidence appear to be replicated, there is some reason to believe that broad spatial patterns of overweight and obesity were replicated as well. Therefore, although the magnitude of forecast reductions in incidence may not be accurate, the purpose of these policy simulations was to assess how incidence differentially changes across neighbourhoods within metropolitan Toronto as a function of population changes in body weight. These types of scenarios provide insight into how incidence rates might be reduced in different areas of the city. Public health planners can use this information to allocate appropriate health promotion programs to neighbourhoods expected to benefit most.

The weight loss scenarios developed in this research demonstrate that substantial reductions in body weight are required in the high-risk population (BMI $\geq$ 25 kg/m^2) before a noticeable reduction is observed in the five-year incidence of diabetes. If the entire high-risk population lost 4.2kg of its baseline body weight, on average, overall diabetes incidence in metropolitan Toronto might decrease by just 0.2 percentage points, from 4.84% to 4.64%. If 80% of the high-risk popula-

tion lost 10% of its baseline body weight, the five-year risk of diabetes might only decline to 4.58%. Across metropolitan Toronto, a reduction in diabetes incidence of one percentage point requires that *all* high-risk individuals lose 17% of their baseline body weight.

These scenarios demonstrate the challenges faced by health promotion programs; not only is it difficult for individuals to lose excess body weight and maintain those losses (Wu, Gao, Chen, & van Dam, 2009), on a population level, weight loss strategies in and of themselves may only produce small reductions in the future incidence of diabetes. Findings from Manuel et al. (2013) corroborate this conclusion. Those results suggest that if the entire population of adults in Ontario had a 3.5% lower BMI, 38,500 new cases of diabetes could be prevented over a five-year period. This translates into an absolute reduction in incidence of 0.47 percentage points, from 4.68% to 4.21%.

In spite of small projected reductions in incidence, it is necessary to contextualize these findings. Even though large, sustained reductions in population body weight are required to produce small reductions in diabetes incidence, body weight in Canada has risen steadily since 1978, when an estimated 14% of Canadians were obese (BMI $\geq$ 30 kg/m^2). By 2008, the prevalence of obesity increased to 25%. Estimates based on self-reported data show similar increases in obesity prevalence: in 1985, 6% of Canadians were obese according to the Canadian Health Promotion Survey. This increased to 18% in 2009 (Public Health Agency of Canada and the Canadian Institute for Health Information, 2011). Regardless of whether measured or self-reported BMI is used to define obesity, these estimates demonstrate that over a 25–30 year period, the prevalence of obesity increased by at least 11 percentage points. This being the case, sustained weight loss on a population level requires a long time horizon. Health promotion programs focusing on weight loss alone seem insufficient to effectively reduce diabetes incidence. As Manuel et al. (2013) point out, effective strategies require multiple health promotion interventions and societal changes, including changes to the built environment that promote active transportation (i.e., walking or cycling to work), expanded access to urban green space to promote physical activity, and policies that enable consumers to make healthier food choices.

Although the scenarios developed here project the potential impact of weight

loss on future diabetes incidence, their static nature assumes that weight loss is the only factor that influences incidence. Although the DPoRT model uses several risk factors to forecast incidence, these scenarios do not consider potential feedback between weight loss and concomitant reductions in hypertension or heart disease that might further influence the risk of developing diabetes. In other words, weight loss might produce better control of blood pressure which may improve cardiovascular health. All of these improvements may act together to produce greater reductions in forecast incidence than weight loss alone.

At the neighbourhood level, these scenarios demonstrate that weight loss differentially influences diabetes incidence throughout metropolitan Toronto. Reductions in incidence are expected to be larger in some neighbourhoods compared to others. This appears to be the result of the population structure within specific neighbourhoods. In other words, the initial distribution of body weight in a particular neighbourhood influences the average reduction that occurs under the different scenarios. Neighbourhoods composed of populations that are, on average, heavier will lose greater amounts of body weight under the same weight loss scenario than populations that are, on average, lighter. However, this does not necessarily translate into greater reductions in diabetes incidence. The spatial pattern of results suggests that other factors in conjunction with weight loss influence forecast incidence. Under almost every weight loss scenario, several neighbourhoods within Scarborough were predicted to experience greater than average reductions in incidence even though the baseline prevalence of overweight and obesity is lower in these neighbourhoods. What is unique to these neighbourhoods is their demographic composition: many of them have large visible minority and immigrant populations.

This result is best highlighted by comparing forecast incidence under the baseline and 17% weight loss scenarios. Compared to the baseline scenario, incidence in some Scarborough neighbourhoods is expected to decline by as much as 1.3 percentage points under the 17% weight loss scenario. In some Toronto neighbourhoods, however, incidence may only decline by one percentage point. This is in spite of the higher mean prevalence of overweight in Toronto neighbourhoods compared to Scarborough neighbourhoods (42.2% vs. 39.0%, respectively). However, Scarborough neighbourhoods have a greater proportion of visible mi-

norities compared to Toronto neighbourhoods (67% vs 30%). Results from the probabilistic sensitivity analysis support this conclusion. Although ethnicity generally exerted little influence on forecast uncertainty in most neighbourhoods, for some neighbourhoods in Scarborough, ethnicity accounted for 5% to 15% of the variation in forecast uncertainty. These findings support current clinical practice guidelines that stress the importance of targeting culturally appropriate prevention programs toward high-risk groups, including high-risk ethnic groups (Cheng, 2013).

One unique aspect of this research rarely seen in the spatial microsimulation literature is the use of probabilistic sensitivity analysis (PSA) to assess uncertainty surrounding microsimulation forecasts. Such an analysis underscores the idea that forecast results can not be known with certainty and should be treated as model-based projections. In other words, even though 4.9% of individuals living in metropolitan Toronto might develop diabetes in a five-year period, the PSA suggests five-year incidence might be as low as 4.2% or as high as 5.9%. This range of results conveys the idea that the DPoRT model merely predicts what might happen under a specific set of assumptions. In this case, it is reasonable to expect that if the baseline population does not change, in terms of its underlying demographic structure and health risk factor profile, then anywhere from 4.2% to 5.9% of the population might develop diabetes over a five-year period.

The probabilistic sensitivity analysis also provides insight into which model parameters exert the greatest influence on forecast incidence. With respect to the Diabetes Population Risk Tool, overall results and neighbourhood specific results point to the intercept and scale parameters as being the most influential. Since these parameters are specific to the type of model used to develop the Diabetes Population Risk Tool (i.e., an accelerated failure time regression model), it suggests that the form of the model used to predict diabetes incidence influences the uncertainty surrounding forecast rates. If a different model were used (e.g., a Cox proportional hazards survival model), it is possible that different factors might exert more or less influence. Thus, even though a probabilistic sensitivity analysis does not assess epistemic uncertainty, in this case, the results of the PSA suggest the form of the model used to forecast diabetes incidence influences uncertainty. Therefore, a more in-depth analysis could use different types of models to forecast

risk to assess the robustness of forecast incidence under different model forms. If different models yield similar forecasts, then decision makers can place greater confidence in forecast incidence, aiding potential decisions. For example, if forecast incidence was consistently high in specific neighbourhoods under different modelling scenarios, then it might be reasonable to target health promotion efforts toward those neighbourhoods.

Another strength of the PSA is that it allows identification of which model parameters exert greater influence on forecast uncertainty in specific locations. In other words, some parameters influenced forecast uncertainty to a greater degree in some neighbourhoods compared to others. This was demonstrated most clearly with the visible minority parameter in the DPoRT model. With the exception of some neighbourhoods in Scarborough, this parameter had negligible influence on forecast uncertainty in almost all of Toronto's neighbourhoods. In Scarborough, however, the visible minority parameter accounted for as much as 15% of the variation in forecast uncertainty, suggesting that in these neighbourhoods, visible minority status is an important factor influencing incidence. Thus, the probabilistic sensitivity analysis helped identify areas within Toronto that might require different approaches to diabetes prevention, especially in light of the finding that projected reductions in incidence were greater in these areas even though initial body weight at the population level tended to be lower compared to other neighbourhoods.

5.4.2 Limitations

Although forecast incidence was accurate for metropolitan Toronto as a whole and for some of its neighbourhoods, forecasts were less accurate in other neighbourhoods. In particular, neighbourhoods in North Etobicoke, west North York, and Scarborough had true incidence rates greater than the upper limit of forecast uncertainty. Using the simulated TropISM population, DPoRT failed to accurately capture the geographic *variability* in forecast incidence, even though it captured broad spatial patterns. In other words, forecasts in some neighbourhoods were noticeably lower than the true incidence. This may be the result TropISM's inability to capture the geographic variability in important risk factors used by the

DPoRT model to forecast incidence. That is, if some risk factors were less accurately predicted in these neighbourhoods relative to neighbourhoods in York or Toronto, then the DPoRT model may be missing crucial information needed to accurately forecast incidence.

Another factor that could have influenced DPoRT's inability to forecast the geographic heterogeneity in diabetes incidence lies in the construction and calibration of the DPoRT model itself. This model was developed using population survey data from the province of Ontario (Rosella et al., 2014), a population that is demographically different from the more ethnically diverse neighbourhoods of metropolitan Toronto. A different set of model parameters, calibrated to the demographic diversity specific to metropolitan Toronto may have produced more accurate forecasts of diabetes incidence in Toronto's more ethnically diverse neighbourhoods.

In summary, local spatial microsimulation models may have produced more representative neighbourhood populations, with respect to health risk factor profiles, which may have enabled more accurate forecasts of incidence in some neighbourhoods. In addition, DPoRT models using different parameter values for different communities may have permitted more accurate forecasts of diabetes incidence in some neighbourhoods. One way such variation might be captured is through the use of mixed-effects prediction models, that incorporate random intercepts and/or slopes to capture differences between neighbourhoods in risk prediction. Although conceptually attractive, such models might be more difficult to estimate in practice, as they would require obtaining diagnostic information about diabetes cases at more granular levels. Both of these solutions, however, should be considered in future studies that attempt to forecast diabetes incidence at increasingly refined levels of geography.

As noted in Section 4.4.1, the TropISM model was developed using data from the 2006 Canadian census. Because it is a static microsimulation model, TropISM does not account for temporal changes in population structure. As the probabilistic sensitivity analysis demonstrates, population structure influences forecast uncertainty. If the demographic structure of Toronto's neighbourhoods changed in important ways over the forecasting period, then projected diabetes incidence could be systematically biased. For example, if a neighbourhood's composition

changed due to gentrification, demographically, that neighbourhood's population might be younger and wealthier at the end of the forecast period compared to the beginning. This might result in artificially high forecast incidence in that neighbourhood as older people who are more likely to develop diabetes may have moved out of the neighbourhood.

However, population growth did not seem to influence whether incidence was more or less accurately forecast across neighbourhoods. On average, population growth across neighbourhoods did not differ according to whether or not true neighbourhood incidence fell below, within, or above forecast uncertainty (Table B.1).

Given DPoRT produced less accurate forecasts in some neighbourhoods under the baseline scenario, it is necessary to consider how such underestimation may influence predicted reductions in incidence under the different weight loss scenarios. If body mass index were underestimated at the neighbourhood level, then the weight loss scenarios developed for this research may also have downwardly biased reductions in diabetes incidence that might occur as a result of weight loss. Although this may be the case, part of the utility in examining different weight loss scenarios is to understand (a) how incidence is influenced by weight loss and (b) how changes in incidence vary over geographic space. Thus, even though population body weight may not have been accurately replicated in the baseline TropISM population, it would still be possible to assess the effect that weight loss has on changes in incidence and how those changes vary across neighbourhoods.

A more difficult problem to address is differential bias in baseline body weight. That is, if the prevalence of overweight and obesity were more accurately reproduced in some neighbourhoods compared to others, it becomes more difficult to assess how forecast incidence might change over geographic space. Unfortunately, the baseline prevalence of body weight could not be validated using the incidence of cancers associated with obesity. Therefore, potential geographic differences in forecast reductions need to be interpreted cautiously. However, forecast reductions under the 14% and 17% weight loss scenarios revealed that reductions were largest in the neighbourhoods of north Etobicoke, west North York, and Scarborough. Although forecast incidence was less accurate in these neighbour-

hoods, it is important to note that the populations of these neighbourhoods are comprised of larger proportions of visible minorities. Previous research suggests that weight gain in ethnic groups may be associated with greater risk of developing diabetes (Shai et al., 2006). Thus, it may also be the case that weight reductions in ethnic groups produce greater reductions in risk, consistent with the forecast reductions in these neighbourhoods. This finding is also consistent with the results of the probabilistic sensitivity analysis that suggest the ethnicity parameter from the DPoRT model was responsible for as much as 15% of the variation in forecast uncertainty in some of Scarborough's neighbourhoods. Therefore, although diabetes incidence was not accurately forecast in these neighbourhoods, the spatial patterns of forecast reductions in incidence as a function of weight loss observed in this study may still be valid (Figure 5.8).

Results from the probabilistic sensitivity analysis may also shed some light on whether inaccuracies in the baseline prevalence of risk factors biased forecast diabetes incidence. For example, neighbourhoods having higher values of the sensitivity index for some of the body weight parameters (Appendix B) represent locations where those parameters exerted greater influence on forecast uncertainty. If forecast incidence in these areas is less accurate (compared to known incidence rates), then forecast rates should be treated with greater caution. This type of assessment helps contextualize the results and the may suggest which model parameters might produce biased forecasts.

Another limitation that should be considered is the inability of the DPoRT model to produce stratified forecasts of diabetes incidence by comparable age-sex groups. Although the DPoRT model produces separate forecasts of diabetes risk by sex, the age groups used by the model differ for men and women. As a result, while it is possible to obtain detailed population estimates of diabetes risk for women aged 65 and older, the same is not true for men. This is because the DPoRT model uses separate body mass index parameters for two age groups in men (20–44 and 45+) but three age groups in women (20–44, 45–65 and 65+). As a result, forecasts for men 65 and older are calibrated by the BMI parameters specific to a larger age group than they are for women. Thus, risk estimates for men age 65+ are downwardly biased. Future versions of the DPoRT model may consider using the same age groups for both men and women in order to obtain comparable

estimates of diabetes in all age groups.

At a neighbourhood level, this difference may have contributed to some of the results observed here. Specifically, forecast incidence was found to be higher in neighbourhoods whose populations had higher than the median percentage of women aged 65 and older. This finding is counter-intuitive given that older men face a greater risk of developing diabetes. However, because the DPoRT model pools all men aged 45 and older to estimate population level risk, this counter-intuitive finding is, in all likelihood, a function of the DPoRT model. This result also underscores that model form contributes to epistemic uncertainty. Different model forms (i.e., a different parametrization of the DPoRT model) may produce different, and possibly conflicting, results. Results from the validation of forecast incidence support this conclusion. Specifically, forecast counts of diabetes were closer to true counts among women across all neighbourhoods than they were among men. This greater degree of accuracy might result from the sex-specific parameterization of the Diabetes Population Risk Tool.

5.4.3 Summary

In summary, this research provides preliminary evidence that it is possible to forecast the five-year incidence of diabetes at the neighbourhood level using the Diabetes Population Risk Tool. Even though incidence was not accurately forecast in all neighbourhoods of metropolitan Toronto, at an aggregate level, the true incidence of diabetes fell within the range of forecast uncertainty. Moreover, the DPoRT model captured broad spatial patterns in diabetes incidence using the simulated TropISM population. Future research should identify ways to more accurately forecast incidence at the neighbourhood level, which may involve producing more accurate simulated populations that better reflect the true health risk factor profile across all small area populations. It may also involve using different forms of the forecasting model in populations that differ substantially from the overall population. For example, it may be possible to estimate different forms of the DPoRT model using demographically unique populations. Different model forms could then be used for diverse communities to improve forecast accuracy. In addition, future research may consider more dynamic ways to forecast inci-

dence that allow greater flexibility in predicting how incidence might change as a function of changes in the population at risk. For example, a more dynamic model may consider feedback between weight loss and concomitant improvements in blood pressure and reduced risk of heart disease. All of these factors combined may result in a better ability to predict how incidence might change as a function of health promotion programs. Such refinements can improve the utility of spatial microsimulation models and risk prediction models for public health planners, allowing them to better incorporate model results into practical decisions of planning and delivering health promotion programs to neighbourhoods where they are needed most.

Chapter

6

Spatial Accessibility to Health Promotion Programs

6.1 Rationale

In order to minimize the population burden of type 2 diabetes, there needs to be a sufficient supply of resources dedicated to (a) preventing the disease in high-risk groups and (b) effectively managing it among diabetics. In high-risk groups, weight loss can significantly reduce the risk of developing type 2 diabetes. Applied to the TropISM model, the Diabetes Population Risk Tool demonstrates how the incidence of diabetes might change at the neighbourhood level under different risk reduction strategies. Given that risk factors and disease incidence are not evenly distributed over geographic space, spatial accessibility to disease prevention and management programs is an important consideration when planning their delivery.

Accessibility to health services is a multi-factorial concept that encompasses not only geographic accessibility but also availability, affordability, accommodation (e.g., wait times), and acceptability (e.g., patient perceptions of service provision, see Cromley & McLafferty, 2012; Schuurman, Bérubé, & Crooks, 2010). Geographic accessibility is concerned with the relative difficulty by which health services can be reached from different locations. It incorporates spatial proximity to

health services, where proximity is influenced by distance, travel time, and travel costs. Two main types of geographic access to health services can be defined: revealed access and potential access. Revealed access is concerned with the actual utilization of health services while potential access is concerned with probable access to services. Potential access depends on (a) the supply of physicians, practitioners, and/or services, (b) the population demand for those services, and (c) the proximity between supply and demand (Luo & Qi, 2009; Wan, Zhan, Zou, & Chow, 2012). Potential accessibility further depends on the supply of similar services in neighbouring communities. Thus, potential accessibility considers proximity to multiple service providers, all of whom are distributed throughout geographic space. Empirical studies of revealed accessibility demonstrate that an inverse relationship exists between proximity to health services and utilization of those services (Cromley & McLafferty, 2012). This frictional effect of distance, known as "distance decay," varies for different types of services. What is more, distance decay tends to be regionally and temporally specific (Wang, 2006). Sociodemographic factors, such as language or income, may exacerbate the effect of distance on utilization of health services (Cromley & McLafferty, 2012).

Different methods exist to measure geographic accessibility to health services. One of the simplest involves estimation of provider-to-population ratios within specific geographic regions (Luo & Wang, 2003). Although these ratios are easy to interpret, a crucial limitation is that they only measure the supply of health services within one specific region. They cannot incorporate cross-border accessibility, where some segments of the population might travel outside their immediate neighbourhood to access services (Joseph & Phillips, 1984; Luo & Wang, 2003; McGrail & Humphreys, 2009). Spatial interaction models have been used to overcome this limitation (Wan et al., 2012). These models measure the movement of people or products between geographic locations (Rodrigue, Comtois, & Slack, 2006).

Previous studies combined spatial interaction models with spatial microsimulation estimates of disease prevalence to estimate geographic accessibility to physicians (Morrissey, Ballas, Clarke, Hynes, & O'Donoghue, 2013b; Morrissey, Clarke, Ballas, Hynes, & O'Donoghue, 2008; Morrissey, Clarke, & O'Donoghue, 2013a). Small area counts of disease prevalence are used to estimate the demand for health care while the locations and number of health care providers estimate the supply

of services across geographic space. The models used by Morrissey et al. (2013b, 2008, 2013a) produce indices of spatial accessibility but lack an intuitive interpretation. Other types of spatial interaction models suffer this same shortcoming (e.g., Joseph & Bantock, 1982; Joseph & Phillips, 1984). In response, Luo and Wang (2003) developed the two-step floating catchment area model that measures geographic accessibility in terms of provider-to-population ratios.

As its name suggests, the two-step floating catchment area model estimates geographic accessibility in two steps. Given a set of j health service locations and k population locations, where k might represent the population weighted centroid of small area k, the provider-to-population ratio R_j for site j is estimated as:

$$R_j = \frac{S_j}{\sum_{k \in d_{kj} \leq d_0} P_k} \tag{6.1}$$

where:

S_j is the total supply of physicians or health care providers at location j,

P_k is the total population at location k, and

$d_{kj} \leq d_0$ represents all locations where the distance between health service location j and population location k is less than a predefined threshold distance d_0.

The second step sums R_j for all health service locations lying within the threshold distance d_0 for small area i (i.e., $d_{ij} \leq d_0$). In other words, the accessibility index A_i for small area i is estimated as

$$A_i = \sum_{j \in \{d_{ij} \leq d_0\}} R_j. \tag{6.2}$$

Larger values of A_i indicate greater accessibility. Measuring spatial accessibility this way assumes that all service locations beyond the threshold distance d_0 are inaccessible to the population residing in small area i. Conversely, all service locations lying within the threshold distance d_0 are equally accessible to the population residing in small area i.

The threshold distance d_0 is commonly computed in terms of travel time. Values of d_0 should be chosen according to the type of service modelled (Cromley

& McLafferty, 2012); larger values produce greater spatial smoothing of region-specific accessibility scores (Wang, 2006). Lee (1991) suggests a value of 30 minutes for measuring physician accessibility, although larger values should be used if the study region contains both urban and rural areas (Delamater, 2013). Threshold values lying between 30–60 minutes are commonly used in studies measuring spatial accessibility to health care providers (see Luo & Qi, 2009; Luo & Wang, 2003; McGrail & Humphreys, 2009; Wan et al., 2012; Wang & Luo, 2005).

Although the 2SFCA model has an intuitive interpretation and is easy to implement in standard GIS software packages (Luo & Wang, 2003), it has been criticized for its inability to meaningfully incorporate distance decay because service locations are treated as either accessible or inaccessible. Those that are accessible are given equal weight, regardless of whether they are five or thirty minutes from population location i. A more general two step model is defined by

$$A_i = \sum_{j=1}^{n} \left[\frac{S_j f(d_{ij})}{\left(\sum_{k=1}^{m} P_k f(d_{kj})\right)} \right] \tag{6.3}$$

where $f(d)$ is a generic distance decay function used to compute distance decay weights that differentially weight accessibility scores for any given location according to the distance between service locations and population centroids (Wang, 2012). The distance decay function may be modelled using a Gaussian function, $f(d_{ij}) = \exp^{-d_{ij}^2/\beta}$, where d_{ij} represents the distance between locations.[1] Note this function allows for a continuous decay rate so that populations located farther from health service locations are less likely to visit those locations. Further note that β is an empirically derived parameter that controls distance decay (Kwan, 1998; Luo & Qi, 2009). In the case of accessibility to health services, β would be estimated from the trips patients make to health service providers (Kwan, 1998; Luo & Qi, 2009; Rodrigue et al., 2006; Wang, 2006). Since this information is rarely available (Schuurman et al., 2010), Kwan (1998) suggests using values of β that approach zero at distances beyond which few people would travel to that service.

Regardless of the exact form of the 2SFCA model used to estimate spatial accessibility, any such model depends on assumptions which may or may not be rea-

[1] Other possibilities include an inverse power function and an exponential function. The Gaussian function is often preferred over the others because the rate of decay is less pronounced for locations that are closer together (Kwan, 1998).

sonable under different circumstances. Thus, as McGrail and Humphreys (2009) note, the results of these models need to be interpreted cautiously. However, as an exploratory tool, these models can provide insights into identifying areas where there may be an insufficient supply of diabetes management and prevention services. Combined with other sources of evidence (e.g., prevalence of chronic disease risk factors in the same locations), these methods provide program planners with additional information that can be used to enhance health service access and delivery.

6.1.1 Study objectives

This study demonstrates how disease and risk factor prevalence estimates from TropISM can be combined with a two-step floating catchment area model to conduct an exploratory analysis of spatial accessibility to diabetes management and prevention programs. The potential demand for these services is estimated from simulated counts of the number of cases of (a) type 2 diabetes and (b) overweight and obesity for each of Toronto's 140 neighbourhoods. Spatial accessibility to diabetes management programs will be assessed by determining the locations of diabetes education programs. Similarly, locations of community recreation centres (including YMCAs) will be used to measure potential access to diabetes prevention programs. This analysis assumes that diabetes prevention programs can be delivered by recreation centre staff. These assumptions are reasonable given that Ali et al. (2012) found that lay community members are just as effective in motivating weight loss in high-risk individuals as medical and allied health professionals.

The results of this exploratory analysis can be used to identify areas within Toronto that have relatively poor access to diabetes management and prevention programs. In this case, the overall level of accessibility in any one neighbourhood is less important than relative differences in access between neighbourhoods. Neighbourhoods having poor access to programs and a high burden of disease likely require additional health promotion resources.

6.2 Methods

6.2.1 Data sources

Several data sources were used to measure spatial accessibility to diabetes management and prevention programs. TropISM model outputs were used to estimate the demand for management and prevention services in each of Toronto's 140 neighbourhoods. The simulated number of type 2 diabetics living in each neighbourhood defined the population requiring access to diabetes management programs; in other words, the denominator of Equation 6.3. Similarly, the simulated number of overweight people (BMI ≥ 25 kg/m^2) within each neighbourhood defined the demand for prevention services. The supply of services was defined by the number of diabetes education programs within Toronto; this supply was used in the numerator of Equation 6.3. Likewise, to estimate accessibility to prevention services, the 2SFCA model used the number of community recreation centres within Toronto. Locations of each type of service were ascertained from publicly available data sources. In the case of diabetes education programs, data from the Government of Ontario were used to determine the location of each program within Toronto (http://www.ontario.ca/locations/health/). Location data for all recreation centres and YMCAs within the city were extracted from parks and recreation datasets available from the City of Toronto's Open Data Initiative (City of Toronto, 2013c). To estimate accessibility to prevention programs, it was assumed that facilities already providing community based fitness programs served as proxies for the types of facilities that could deliver diabetes prevention programs. All recreation centres within Toronto offering adult fitness programs, including cardiovascular and/or conditioning programs, were classified as potential sites for the delivery of diabetes prevention programs. In addition, potential sites had to contain at least one multi-purpose room, where group-based diabetes prevention and lifestyle modification classes could be delivered (Ackermann & Marrero, 2007; Ali et al., 2012).

All location data were geocoded using street network data for Toronto; these data were also obtained from the City's Open Data Initiative. The QGIS software package (Version 2.6) was used to geocode address data to geographic coordinates using the "MMQGIS" plugin (QGIS Development Team, 2009). Population

weighted centroids for each neighbourhood were computed using shapefiles delineating census tracts within Toronto, obtained from the University of Waterloo's Geospatial Data Centre. Population weighted centroids for each neighbourhood were computed using the population in each census tract along with its geographic centroid. Thus, if neighbourhood j contained n_j census tracts, the population weighted centroid for that neighbourhood was computed as:

$$x_j = \frac{\sum_{i=1}^{n_j} p_i x_i}{\sum_{i=1}^{n_j} p_i} \quad , \quad y_j = \frac{\sum_{i=1}^{n_j} p_i y_i}{\sum_{i=1}^{n_j} p_i}$$

where:

x_j and y_j represent the x and y coordinates of the weighted neighbourhood centroid,

x_i and y_i represent the x and y coordinates of the ith census tract comprising neighbourhood j, and

p_i represents the population of the ith census tract (Luo & Wang, 2003).

6.2.2 Model parameters

Table 6.1 summarizes the parameters used to model spatial accessibility to diabetes management and prevention programs. A Gaussian decay function was used to estimate accessibility to each type of service. Following Lee (1991) and Luo and Wang (2003), a threshold distance of 30 minutes was used to define whether diabetes education programs were accessible to the population of diabetics living in each of neighbourhood. A gradual decay function was used ($\beta = 440$) under the assumption that most people would be willing to travel up to 30 minutes for diabetes management services. Stronger decay effects were used to model spatial accessibility to diabetes prevention programs, under the assumption that people would be less willing to travel greater distances to access prevention programs, such as cardiovascular fitness programs. Therefore, a threshold distance of 20 minutes was used along with a steeper distance decay function ($\beta = 195$).

Table 6.1 also lists the total supply of services within Toronto. Overall, 64 diabetes education programs were identified; of these, 48 were determined to be generally available to type 2 diabetics (i.e., "unrestricted" programs not limited to hospital inpatients or patients belonging to a particular physician practice). Of these 48 programs, 30 were restricted to patients living within defined communities of Toronto (e.g., areas bounded by specific streets or entire boroughs). However, to simplify accessibility modelling, these restrictions were ignored.

Similarly, 148 community recreation centres (including five YMCAs) were identified within Toronto. Of these, 70 offered some type of cardiovascular fitness program and housed a multi-purpose room which could be used for classroom sessions. These 70 recreation facilities were selected as *possible* locations where diabetes prevention programs could be delivered in community settings. Spatial accessibility for each type of service was estimated using both sets of values as supply parameters in the 2SFCA model.

One limitation of the 2SFCA model is that spatial accessibility estimates are potentially downward biased near the borders of the study area (Luo & Wang, 2003; Wan et al., 2012). This bias is the result of "edge effects" wherein boundary neighbourhoods can only access services located within the City of Toronto. Although this might be a realistic assumption in some cases, it is possible that residents of these boundary neighbourhoods travel outside City limits to services located in Mississauga, Brampton, Vaughan, Markham, or Pickering. As a result, lower levels of spatial accessibility to services in boundary neighbourhoods need to be interpreted cautiously.

6.2.3 Data analysis

Wang and Luo (2005) describe how the 2SFCA model can be implemented in a geographic information system (GIS). Although they implemented the entire routine in ArcGIS, they note that a GIS is only needed to compute travel times between supply and demand locations. Once a matrix of travel times is computed, the remainder of the calculations can be performed using standard statistical software packages. Travel times between supply and demand locations were calculated using a local server version of the Open Source Routing Machine (Luxen,

Firebaugh, & Niklaus, 2014; Sabas, 2014) and a local copy of OpenStreetMap data for the City of Toronto (OpenStreetMap Wiki, 2014). OSRM uses "contraction hierarchies" to find the shortest path between two locations (Batz, Geisberger, & Sanders, 2008; Fairhurst, 2014; Sanders, Schultes, & Delling, 2008). Having computed travel times between all service locations and population weighted centroids, the 2SFCA model was estimated using R (see Appendix D). Estimates of spatial accessibility to diabetes education centres and community recreation facilities were summarized for neighbourhoods within each of Toronto's six boroughs while neighbourhood-specific estimates were displayed cartographically. Note that the accessibility index for diabetes education programs was reported as the number of programs per 1000 diabetics; for community recreation centres, the accessibility index was reported as the number of programs per 10,000 people at risk.

In addition, a relative "spatial access ratio" (SPAR) was reported (Wan et al., 2012). This ratio was computed by dividing neighbourhood-specific accessibility indices by the average accessibility index for the entire city. A ratio < 1 indicates that a given neighbourhood has relatively lower access to programs compared to the citywide average while a ratio > 1 indicates a neighbourhood has relatively higher access. A ratio equal to 1 means that access in a particular neighbourhood is the same as the overall average.

Validation of spatial accessibility. In order to validate the spatial patterns of accessibility to diabetes education programs, the known number of diabetes cases, ascertained from the Ontario Diabetes Database, was used as the demand parameter for the 2SFCA model. Again, the number of cases was corrected for false positives and removal of type 1 diabetes cases. Following model estimation, neighbourhoods were classified into quintiles based on the simulated and known demand for diabetes education programs. The overall % agreement and Cohen's κ were used to assess the similarity in the broad spatial classification of accessibility estimated using simulated and known demand. Choropleth maps of accessibility were also compared to ascertain the degree of similarity between the two sets of accessibility estimates. No validation could be conducted for accessibility to recreation centres because the known number of cases of overweight and obesity were unavailable.

6.2.4 Neighbourhood characteristics associated with accessibility

Measures of spatial accessibility to health promotion programs provide an indication of which neighbourhoods may have relatively lower access to those programs compared to others. This information can be used to direct additional resources to those neighbourhoods requiring them. Given low access neighbourhoods might be systematically different from high access neighbourhoods, it is useful to examine the demographic characteristics of neighbourhoods by relative spatial access.

Neighbourhoods were classified into three groups according to whether their accessibility index was lower, similar to, or higher than the citywide average. Using results from the sensitivity analysis discussed in Section 6.2.5, neighbourhoods whose 95% uncertainty bound contained the overall average were classified as having "average" accessibility to health promotion programs. Neighbourhoods whose upper 95% uncertainty interval was lower than the overall average were classified as having "low" access while neighbourhoods whose 95% uncertainty interval was greater than the overall average were classified as having "high" access.

Linear spatial simultaneous autoregressive error models were then estimated to test whether the average demographic characteristics of neighbourhoods differed by relative spatial access. Each demographic characteristic was regressed on relative spatial access accounting for spatial autocorrelation between contiguous neighbourhoods using a "rook" contiguity weight matrix. A likelihood ratio test was used to test whether neighbourhood characteristics differed by spatial access category. Spatial error models were estimated using the "spdep" package in R (Bivand, 2015).

Local Moran statistics were also used to identify possible clusters of neighbourhoods having relatively high or low simulated prevalence of type 2 diabetes and overweight. In addition, local Moran statistics were used to identify clusters of neighbourhoods having relatively high or low access to health promotion programs. Bivariate local Moran statistics were then used to examine the relationship between simulated prevalence and spatial access to programs. Univariate and bivariate local Moran statistics were estimated in GeoDa (version 1.6.6.1). P-values for local Moran statistics were computed in GeoDa to identify significant clusters

using a permutation test (499 permutations) and then adjusted for multiple comparisons using the false discovery rate adjustment in R (Benjamini & Hochberg, 1995).

6.2.5 Uncertainty and sensitivity analysis

A probabilistic sensitivity analysis was also conducted for the 2SFCA model. For this analysis, population demand (P_k), travel time between population demand and service locations (d_{kj}), the distance threshold (d_0), and distance decay (β) parameters were treated as uncertain inputs while the total number of services remained fixed. Although β should be empirically derived (Kwan, 1998; Luo & Qi, 2009), because this information is rarely available in practice (Schuurman et al., 2010), the distributional properties of P_k, d_{kj} and d_0 are unknown. As a result, these uncertain parameters were sampled from a triangular distribution that only requires minimum, maximum, and modal values (Table 6.1). Although there are limitations of using a triangular distribution for probabilistic sensitivity analysis (Briggs & Sculpher, 2006), doing so permits random sampling of input parameters. Modal values for each input parameter were the same as those used to conduct the main analysis while minimum and maximum values were selected according to what might be considered reasonable values, defined below.

Population demand for services was derived from the simulated number of prevalent cases of type 2 diabetes and the simulated number of overweight people in the TropISM population.[2] Using 100 replicates of the TropISM model, the difference between the 2.5$^{\text{th}}$ percentile and the median was used as the minimum value for sampling from a triangular distribution. The difference between the 97.5$^{\text{th}}$ percentile and the median was used as the maximum value. A modal value of 0 was selected for the peak of the distribution (Table 6.1).[3] Randomly sampled values were treated as *differences* to conduct the sensitivity analysis. In other words, if one neighbourhood had 800 cases of type 2 diabetes and a second had 950, then the values for P_k for the sensitivity analysis ranged from 704 to 886 in the first neighbourhood and from 854 to 1036 in the second neighbourhood, with modal values of $\approx$ 800 and 950, respectively.

[2] Neighbourhood-specific counts.
[3] These numbers were *not* neighbourhood-specific.

Table 6.1. Model parameters used to estimate spatial accessibility to health promotion programs within metropolitan Toronto using a two-step floating catchment area model and a Gaussian distance decay function.

Parameter	Diabetes Education Programs	Community Recreation Centres
Supply of services		
Total supply	64	148
Total eligible	48	70
Population demand (total cases)	Type 2 diabetes	BMI ≥ 25
Mean	692	5576
IQR	409–887	3800–7062
Range	209–2254	1898–14,450
Threshold distance d	30 minutes	20 minutes
Distance decay (β)	440	195
Probabilistic Sensitivity Analysis		
Probability distribution	Triangular	Triangular
Population demand (difference)		
Minimum	-96	-189
Maximum	86	190
Mode	0	0
Travel time (difference)		
Minimum	-0.025	-0.025
Maximum	0.25	0.25
Mode	0	0
Distance decay		
Minimum	347	136
Maximum	782	266
Mode	440	195
Threshold distance (minutes)		
Minimum	25	15
Maximum	40	25
Mode	30	20

In a similar way, travel times were assumed to be 2.5% faster to 25% slower while a modal time difference of 0 minutes was assumed. Thus, if the travel time between two locations was 20 minutes, the modal travel time for the probabilistic sensitivity analysis was approximately 20 minutes and would range from 19.5 minutes to 25 minutes. This assumes that the travel time between two locations estimated by the Open Source Routing Machine is about the optimal travel time, but might be about five minutes longer.

Finally, threshold distance parameters were sampled so that they ranged from 25 to 40 minutes for access to diabetes education programs while they ranged from 15 to 25 minutes for access to community recreation centres. Distance decay parameters were chosen so that they ranged from values of 347–782 to model access to diabetes education programs (mode = 440) and 136–266 for access to community recreation centres (mode = 195). These values assume that almost no one in the population would travel longer than 40, 45 (mode), and 60 minutes to access diabetes education programs. Likewise, no one would travel farther than 25, 30 (mode), and 35 minutes to access diabetes prevention programs offered through community recreation centres.

Partial rank correlations were estimated between each input parameter and spatial accessibility to health promotion programs using SaSAT (Hoare et al., 2008). Correlation coefficients (PRCC) were used to identify which input parameters exerted the greatest influence on outcome uncertainty. Since spatial accessibility measures were specific to each of Toronto's 140 neighbourhoods, a separate analysis was conducted for each neighbourhood. Boxplots were used to summarize the distribution of partial rank correlations across all neighbourhoods. In some cases, the effect of model parameters on outcome uncertainty varied substantially across neighbourhoods. Results were therefore stratified according to negative (PRCC < -0.3), negligible (-0.3 ≤ PRCC ≤ 0.3), and positive (PRCC > 0.3) values for the model parameter having the largest range of values.

SaSAT was also used to estimate the sensitivity index for each input parameter of the 2SFCA model. The sensitivity index estimates the proportion of uncertainty in spatial accessibility that is attributable to each input parameter. Neighbourhood-specific sensitivity indices were estimated and the distribution of these indices across all neighbourhoods was summarized using boxplots. The spatial distribu-

tions of the most influential model parameters were mapped to identify whether some model parameters exerted stronger effects on outcome uncertainty in some neighbourhoods.

6.3 Results

6.3.1 Spatial accessibility to health promotion programs

An initial set of two-step floating catchment area models were estimated to assess geographic accessibility to (a) diabetes education programs and (b) community recreation centres. Four models were estimated for each service using the following assumptions:

1. All programs,
2. Only programs meeting the inclusion criteria (Section 6.2.2),
3. A relatively "steep" distance decay parameter ($\beta_d = 440$ and $\beta_r = 195$)[4],
4. A more gradual distance decay parameter ($\beta_d = 782$ and $\beta_r = 266$), and
5. A maximum distance threshold $d_{0,d} = 30$ minutes and $d_{0,r} = 20$ minutes.[5]

Estimates of spatial accessibility are described in Table 6.2 for diabetes education programs and Table 6.3 for community recreation centres.

6.3.1.1 Spatial accessibility to diabetes education programs

As might be expected, the total number of diabetes education programs available to diabetics influences spatial accessibility (A_i) while the distance decay parameter used to model A_i exerts less influence. Median accessibility considering all diabetes education programs located throughout Toronto (n = 64, $\beta = 440$) is 0.72 per 1000, where higher values indicate greater accessibility. This translates into one program for every 1,391 diabetics living in Toronto. If accessibility is estimated using only unrestricted access programs, median A_i is reduced to 0.54 per 1000 meaning these programs have to provide services to an additional 457 type 2 diabetics, on average, throughout the city.

[4] β_d = decay parameter for diabetes education programs; β_r = decay parameter for community recreation centres.

[5] $d_{0,d}$: distance threshold for diabetes education programs; $d_{0,r}$: distance threshold for community recreation centres.

Using a distance decay parameter of 782 to estimate A_i has little effect on accessibility. In this case, the median A_i across all of Toronto's neighbourhoods is 0.53 per 1000, or one program for every 1,876 type 2 diabetics (compared to one program for every 1,848 diabetics if $\beta = 440$, Table 6.2). Similar reductions in accessibility are observed for neighbourhoods located within metropolitan Toronto's six boroughs.

Table 6.2. Neighbourhood accessibility to diabetes education programs estimated using a two-step floating catchment area model under different modelling assumptions.

	All (n = 64)		Subset (n = 48)	
	$\beta = 440$	$\beta = 782$	$\beta = 440$	$\beta = 782$
Borough	Mdn (MAD)	Mdn (MAD)	Mdn (MAD)	Mdn (MAD)
Access per 1000				
Metro Toronto	0.72 (0.12)	0.72 (0.09)	0.54 (0.08)	0.53 (0.07)
Etobicoke	0.56 (0.08)	0.57 (0.06)	0.46 (0.06)	0.45 (0.04)
North York	0.72 (0.14)	0.74 (0.08)	0.54 (0.10)	0.55 (0.08)
York	0.80 (0.05)	0.80 (0.04)	0.62 (0.03)	0.61 (0.03)
Toronto	0.85 (0.04)	0.80 (0.05)	0.62 (0.03)	0.59 (0.04)
East York	0.73 (0.09)	0.71 (0.06)	0.52 (0.06)	0.52 (0.04)
Scarborough	0.52 (0.12)	0.55 (0.10)	0.38 (0.08)	0.40 (0.08)
Spatial access ratio				
Metro Toronto		*not applicable*		
Etobicoke	0.81 (0.12)	0.84 (0.09)	0.88 (0.12)	0.88 (0.08)
North York	1.04 (0.20)	1.08 (0.12)	1.04 (0.19)	1.08 (0.15)
York	1.16 (0.07)	1.16 (0.06)	1.19 (0.06)	1.18 (0.06)
Toronto	1.22 (0.06)	1.17 (0.07)	1.19 (0.06)	1.14 (0.08)
East York	1.05 (0.12)	1.03 (0.09)	1.00 (0.11)	1.00 (0.09)
Scarborough	0.75 (0.17)	0.81 (0.15)	0.73 (0.16)	0.77 (0.15)

Mdn: median neighbourhood accessibility within boroughs

MAD: median absolute deviation (constant = 1); median ± 1 MAD ≈ interquartile range

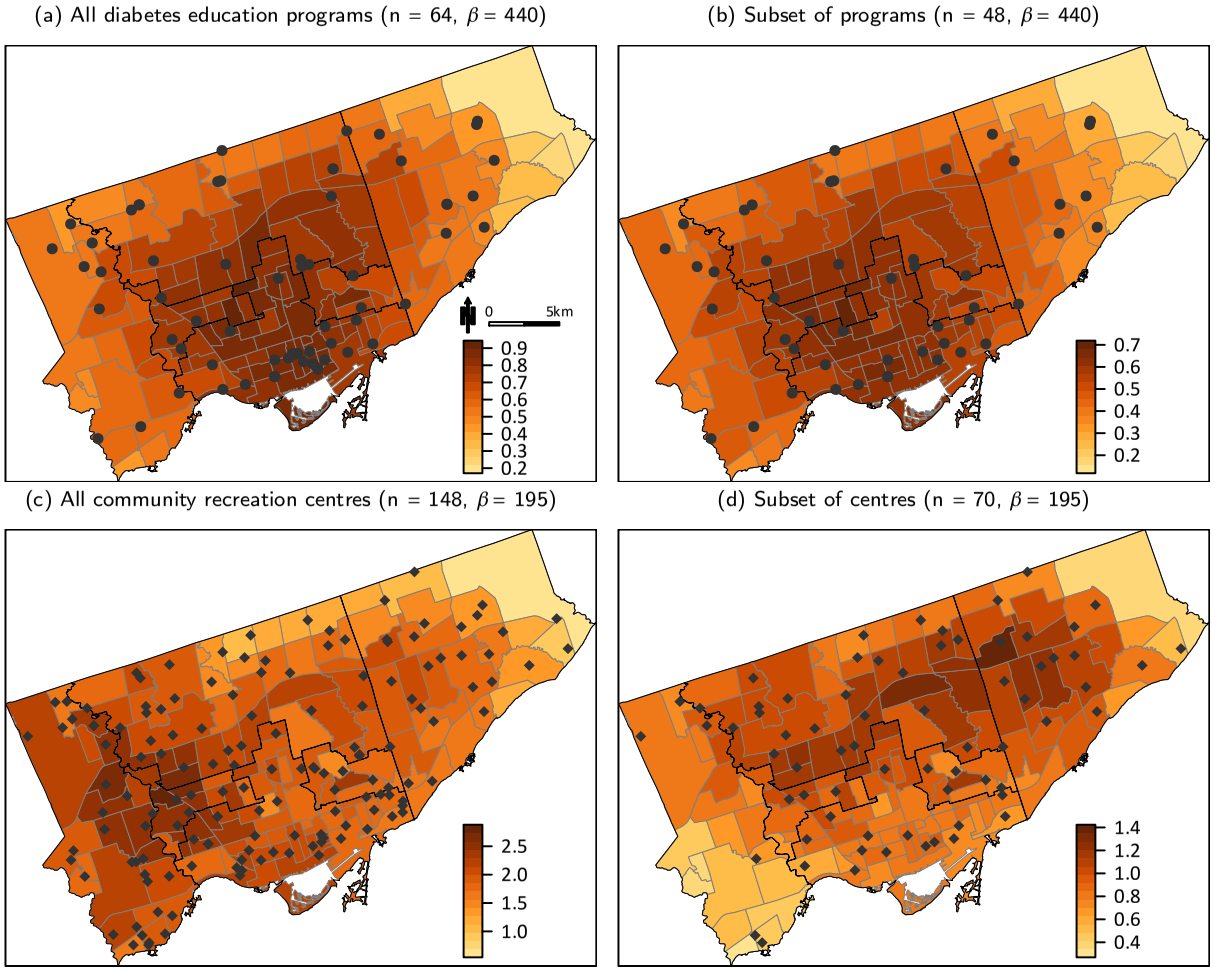
Neighbourhood accessibility to diabetes education programs was highest in York and Toronto and lowest in Etobicoke and Scarborough. The pattern of results varied slightly depending on the number of programs considered and the

distance decay parameter used to model A_i. Across all models, neighbourhoods in Scarborough had the lowest access to diabetes education programs followed by neighbourhoods in Etobicoke. When all programs were used to model accessibility, neighbourhoods in Toronto had the highest accessibility, irrespective of the distance decay parameter. However, when a subset of programs was used (n = 48), neighbourhoods in York had the highest accessibility (Table 6.2).

The spatial access ratio (SPAR), a relative measure of spatial accessibility that divides the neighbourhood specific A_i by the average A_i for metropolitan Toronto, makes it easier to identify neighbourhoods having lower access to health promotion programs and those having greater access. Using this metric, neighbourhoods in York and Toronto have the highest levels of access, irrespective of the number of programs and choice of distance decay parameter used. Access in North York and East York neighbourhoods is similar to the metropolitan average; in these boroughs, the spatial access ratio ranges from 1.00 to 1.08 across all models.

Neighbourhoods within Etobicoke and Scarborough have the lowest levels of relative access throughout metropolitan Toronto. In Scarborough, the SPAR ranges from 0.75 to 0.81; relative access is somewhat higher in Etobicoke, where it ranges from 0.81 to 0.88. A cartographic presentation of these results suggests why these neighbourhoods might have lower access: compared to neighbourhoods within Toronto, there is a relative paucity of programs in these boroughs (Figures 6.1a and 6.1b). Furthermore, programs within these boroughs are spread over larger distances than those in Toronto. Thus, diabetics living in these neighbourhoods have to travel greater distances to reach these programs. Another aspect influencing accessibility is the distance threshold (d_0). In particular, neighbourhoods in Etobicoke and Scarborough typically lie beyond the maximum travel distance of 30 minutes, meaning that many programs accessible within Toronto are not accessible to diabetics living in Etobicoke or Scarborough. Put another way, diabetics living in Toronto have more options because more programs lie within the 30-minute threshold. This, in turn, increases overall accessibility for these neighbourhoods.

Figure 6.1. Neighbourhood accessibility to health promotion programs estimated using a two-step floating catchment area model. Accessibility is measured as the number of diabetes education programs for every 1000 type 2 diabetics and the number of recreation centres for every 10,000 overweight persons.



6.3.1.2 Spatial accessibility to community recreation centres

A different pattern of results emerges when considering accessibility to community recreation centres for the population of overweight individuals (BMI ≥ 25 kg/m^2). If access is estimated using all recreation centres (n = 148), neighbourhoods in York and Etobicoke have the greatest access ($A_{i\ \text{York}}$ = 2.35 per 10,000 and $A_{i\ \text{Etobicoke}}$ = 2.20 per 10,000) while neighbourhoods in Scarborough have the poorest access. This is true irrespective of the distance decay parameter used to model A_i. However, results are sensitive to both the number and location of available services. When programs are restricted to those that currently offer cardiovascular fitness classes, accessibility is dramatically reduced across all neighbourhoods. Under this scenario, neighbourhoods within Etobicoke have the *lowest* accessibility to community recreation centres (Table 6.3). Instead, neighbourhoods in North York have the highest access to community recreation centres, followed by neighbourhoods in Scarborough and York.

Figures 6.1c and 6.1d demonstrate how accessibility changes as a function of the number and location of programs used to model A_i. What is most noticeable is the change in the geographic distribution of programs throughout the city and the neighbourhoods having the greatest access to community recreation centres. In the first scenario (Figure 6.1c), there is a high concentration of community recreation centres within Etobicoke, producing greater accessibility. However, when A_i is modelled using only those centres currently offering cardiovascular fitness classes, the number of available programs is reduced throughout metropolitan Toronto (Figure 6.1d). The reduction in Etobicoke is greater than in any other borough (78% reduction from 32 centres to seven). Although the number of centres within York is reduced by 50% (from six to three), neighbourhoods within this borough still have relatively better access to these services (SPAR = 1.05), likely because services within other boroughs lie within the 20 minute distance threshold (d_0) used to model accessibility.

Another factor that influences accessibility to community recreation centres is the overall demand for these services. Since the prevalence of overweight (BMI ≥ 25 kg/m^2) is relatively high throughout the city (mean simulated prevalence = 40.3%), the potential demand for health promotion programs designed to encour-

Table 6.3. Neighbourhood accessibility to community recreation centres estimated using a 2SFCA model with Gaussian distance decay weights and different modelling assumptions.

	All (n = 148)		Subset (n = 70)	
	$\beta = 195$	$\beta = 266$	$\beta = 195$	$\beta = 266$
Borough	Mdn (MAD)	Mdn (MAD)	Mdn (MAD)	Mdn (MAD)
Access per 10,000				
Metro Toronto	1.94 (0.21)	1.95 (0.23)	0.89 (0.11)	0.89 (0.11)
Etobicoke	2.20 (0.39)	2.11 (0.41)	0.69 (0.25)	0.69 (0.22)
North York	1.94 (0.33)	1.95 (0.32)	1.06 (0.15)	1.05 (0.17)
York	2.35 (0.44)	2.37 (0.39)	0.93 (0.11)	0.94 (0.08)
Old Toronto	1.96 (0.23)	2.00 (0.27)	0.84 (0.08)	0.84 (0.09)
East York	1.78 (0.22)	1.72 (0.23)	0.85 (0.12)	0.85 (0.10)
Scarborough	1.65 (0.33)	1.64 (0.35)	0.96 (0.24)	0.94 (0.19)
Spatial access ratio				
Metro Toronto		*not applicable*		
Etobicoke	1.14 (0.20)	1.10 (0.21)	0.78 (0.28)	0.78 (0.25)
North York	1.01 (0.17)	1.01 (0.16)	1.18 (0.17)	1.17 (0.19)
York	1.22 (0.23)	1.23 (0.20)	1.05 (0.13)	1.05 (0.09)
Old Toronto	1.02 (0.12)	1.04 (0.14)	0.94 (0.09)	0.94 (0.10)
East York	0.92 (0.11)	0.89 (0.12)	0.95 (0.14)	0.95 (0.11)
Scarborough	0.86 (0.17)	0.85 (0.18)	1.08 (0.27)	1.06 (0.22)

Mdn: median neighbourhood accessibility within boroughs

MAD: median absolute deviation (constant = 1); median ± 1 MAD ≈ interquartile range

age weight loss is greater than the potential demand for diabetes education programs. This means that limiting the number of available services redistributes total demand over fewer services. This drastically reduces accessibility across all neighbourhoods. Expressing A_i as its reciprocal demonstrates this phenomenon: when all 148 recreation centres are used to model A_i, each centre within metropolitan Toronto must provide services for ≈ 5,100 overweight individuals. When A_i is modelled using only those centres offering cardiovascular fitness classes, each centre must provide services for more than twice as many high-risk individuals (≈ 11,200 overweight individuals). Clearly, if the goal of diabetes prevention pro-

grams is to target the population at risk, then the total number *and* geographic distribution of these programs must be considered to ensure that they are readily accessible to the greatest number of people across each of Toronto's neighbourhoods.

Validation of spatial access to diabetes education programs Since the known number of cases of diabetes occurring in 2005 could be ascertained from the Ontario Diabetes Database, it was possible to re-run the 2SFCA model using known demand. Accessibility results were compared to those estimated using the simulated number of type 2 diabetes cases as the demand parameter. After neighbourhoods were classified into quintiles using both sets of results, it was found that all neighbourhoods were ranked into the same quintile using either set. That is, there was 100% agreement in the relative ranking of neighbourhoods. However, the absolute value of accessibility across all neighbourhoods was about two times greater when accessibility was estimated using simulated demand (Figure C.1). This is because the TropISM model tended to under-estimate the total number of type 2 diabetes cases across metropolitan Toronto. These lower demand estimates, in turn, lead to higher estimates of accessibility. In summary, although the absolute value of accessibility differs substantially using simulated demand, the relative spatial pattern of results is largely unaffected. Therefore, even though simulated demand over-estimates accessibility to diabetes education programs, model results can still be used to identify areas of metropolitan Toronto that have relatively lower access to diabetes education programs.

6.3.2 Uncertainty surrounding spatial accessibility

A probabilistic sensitivity analysis was conducted to assess the magnitude of uncertainty surrounding neighbourhood accessibility to health promotion programs. For this analysis, demand for services, travel time, distance decay, and the threshold distance parameters in the 2SFCA model were treated as uncertain (Table 6.1). Figure 6.2 plots the range of uncertainty about A_i for both diabetes education programs and community recreation centres. Each plot depicts average A_i for each neighbourhood, from neighbourhoods having the highest accessibility to

the lowest. Accessibility estimates for diabetes education programs presented in Figure 6.2a correspond to estimates presented in Figure 6.1b while estimates for recreation centres presented in Figure 6.2b correspond to estimates presented in Figure 6.1d.

These figures illustrate that uncertainty surrounding modelled accessibility varies across neighbourhoods. Neighbourhoods having greater access to diabetes education programs tend to have a wider range of uncertainty compared to neighbourhoods having lower access. For example, in neighbourhoods where $A_i > 0.60$ per 1000 (n = 47), the 95% uncertainty interval ranges from 0.56 to 0.73 per 1000. In neighbourhoods where $A_i < 0.40$ per 1000 (n = 18), the 95% uncertainty interval ranges from 0.27 to 0.37 per 1000.

A much different pattern of results is observed for neighbourhood access to community recreation centres. In some neighbourhoods, the range of uncertainty is much greater than it is in others. This is true even in neighbourhoods having similar A_i values. For example, A_i = 0.98 per 10,000 in both the Thistletown-Beaumond Heights neighbourhood northern Etobicoke and the Clairlea-Birchmount neighbourhood of southwest Scarborough. However, the uncertainty surrounding A_i for Thistletown-Beaumond Heights (95% UI: 0.91–1.26) is greater than it is for Clairlea-Birchmount (95% UI:0.89-0.99). What is more, uncertainty appears to be attributable to different model parameters in different neighbourhoods throughout the city. In these neighbourhoods, the distance threshold parameter d_0 accounts for 68% and 76% of outcome uncertainty, respectively. However, in Thistletown-Beaumond Heights, the travel time parameter accounts for another 28% of outcome uncertainty, while in the Clairlea-Birchmount neighbourhood, the distance decay parameter accounts for 19% of the outcome uncertainty.

These findings indicate that neighbourhood accessibility is not only influenced by the number and geographic location of available services, but also by a particular neighbourhood's location in relation to those services. Since neighbourhoods lying beyond the maximum distance threshold to service locations cannot, by definition, access those services according to the 2SFCA model, it makes sense that the distance threshold parameter strongly influences outcome uncertainty. However, this conclusion does not appear to generalize to different types of services. In the case of diabetes education programs, accessibility is strongly influenced by

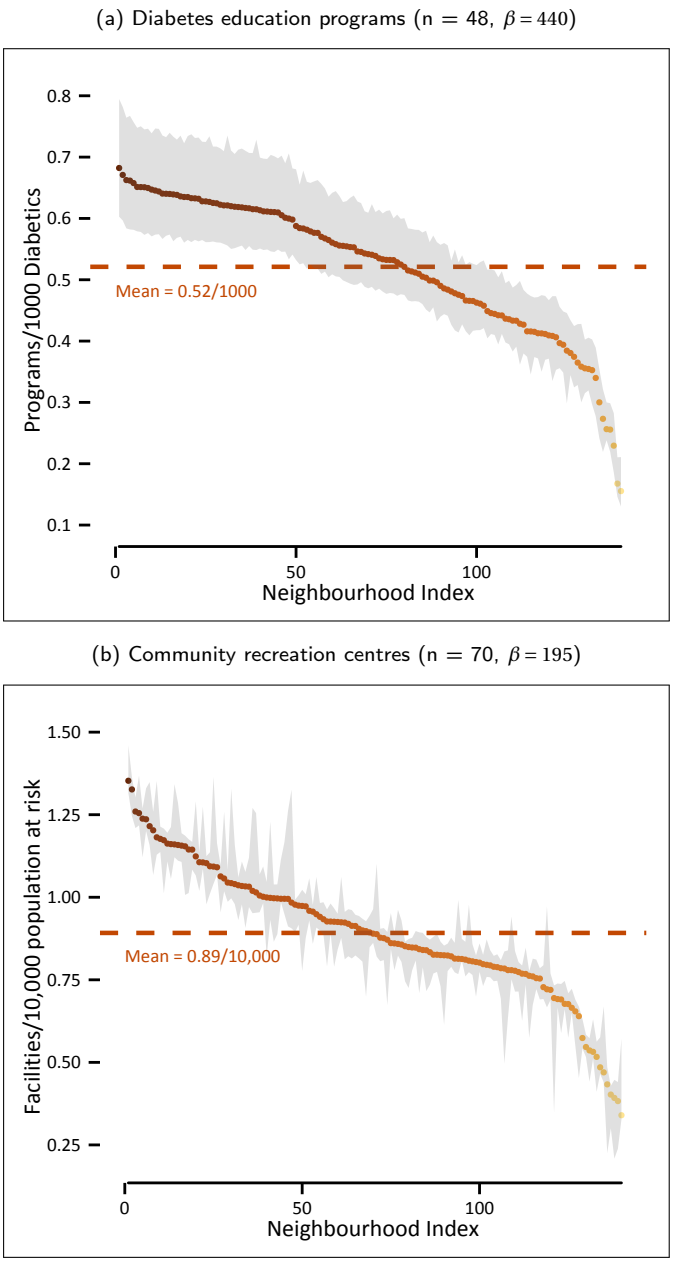


Figure 6.2. Estimated accessibility to health promotion programs and 95% uncertainty intervals.

the demand for services. For example, in almost 90% of neighbourhoods, demand for services accounts for 50% of more of the uncertainty in neighbourhood accessibility. Thus, the nature of the services offered, the population requiring access to those services, and the location of that population in relation to service locations should all be considered as important factors influencing neighbourhood accessibility. Given these results, the planning and delivery of health promotion services on a small area scale needs to consider both the location of services and the geographic distribution of service demand in order to optimally allocate programs.

6.3.3 Neighbourhood characteristics associated with accessibility

Results of the accessibility modelling conducted thus far indicate which neighbourhoods have higher or lower access to health promotion programs. From a health planning standpoint, this information may help direct additional resources to those neighbourhoods requiring them. Given low access neighbourhoods might be systematically different from high access neighbourhoods, it is beneficial to examine the demographic characteristics of neighbourhoods by relative spatial access. An additional analysis was conducted to examine whether neighbourhoods having relatively lower access to diabetes education programs and community recreation centres differed from neighbourhoods having greater access.

Neighbourhoods were classified into three groups according to whether their accessibility index was lower, similar to, or higher than the metropolitan average (see Section 6.2.4). Linear spatial error models were estimated to test whether the average demographic characteristics of neighbourhoods differed by relative spatial access. After accounting for spatial autocorrelation between contiguous neighbourhoods, no differences were seen in the demographic composition of low, average, or high access neighbourhoods. This was true for access to both types of health promotion program. Although neighbourhoods having low access to diabetes education programs seemed to have a somewhat higher prevalence of type 2 diabetes (5.1% in low vs. 4.9% in average and high access neighbourhoods), the difference was not statistically significant. Low access neighbourhoods also had a somewhat higher percentage of visible minorities (45.5%) than average (38.4%), or high access (41.2%) neighbourhoods. Again, these dif-

ferences were not statistically significant after accounting for spatial autocorrelation. A global Moran test points to strong spatial autocorrelation in estimated accessibility to both diabetes education programs (Moran's I = 0.87, $p < 0.001$) and community recreation centres (Moran's I = 0.71, $p < 0.001$). The presence of strong spatial dependence is unsurprising given that the 2SFCA model incorporates distance decay effects to estimate neighbourhood accessibility.

Table 6.4. Demographic characteristics of neighbourhoods having low, average, and high spatial access to diabetes education programs and community recreation centres.

	Relative Spatial Access			Test	
Characteristic	Low	Average	High	LR*	p-value
Diabetes education programs†	(n = 38)	(n = 49)	(n = 53)		
% men 45+	24.52	23.65	21.92	11.67	0.003
% women 65+	10.94	10.94	10.27	0.91	0.634
% visible minority	45.50	38.43	41.25	2.97	0.226
% immigrant	51.80	47.89	50.76	2.88	0.237
% post-secondary education	55.30	57.38	54.60	2.82	0.245
% highest income quintile	13.93	15.53	14.04	3.01	0.222
% type 2 diabetes	5.10	4.89	4.86	0.91	0.633
% overweight	40.58	41.10	40.55	1.82	0.403
% obese	11.08	11.18	10.96	0.89	0.641
Community recreation centres†	(n = 56)	(n = 34)	(n = 50)		
% men 45+	23.60	23.50	22.44	3.68	0.159
% women 65+	10.93	10.72	10.36	0.53	0.766
% visible minority	38.40	42.84	45.41	2.81	0.245
% immigrant	48.57	50.93	51.93	1.53	0.464
% post-secondary education	57.34	53.34	54.10	5.48	0.065
% highest income quintile	15.57	13.84	13.44	3.53	0.171
% type 2 diabetes	4.89	5.03	4.96	0.73	0.693
% overweight	41.07	40.76	40.29	2.11	0.348
% obese	11.13	11.06	10.97	0.38	0.828

* LR = likelihood ratio.

† Number of neighbourhoods falling in each spatial access category.

Given the presence of strong spatial dependence in estimated accessibility across neighbourhoods, an additional exploratory analysis was conducted to assess whether clusters of neighbourhoods having low access to programs but high prevalences of type 2 diabetes and overweight might be identified. Figure 6.3 presents the results of univariate and bivariate local Moran's I for the (a) prevalence of type 2 diabetes and spatial access to diabetes education programs and (b) prevalence of overweight and spatial access to community recreation centres. This analysis was conducted in GeoDa using a permutation test (499 replicates); p-values were adjusted for for multiple comparisons using the false discovery rate adjustment in R.

Univariate local Moran statistics suggest the presence of (a) two clusters of high diabetes prevalence among neighbourhoods in west North York and Scarborough and (b) a single cluster of low diabetes prevalence among neighbourhoods in Toronto (Figure 6.3, panels a, b & c). Conversely, local Moran statistics suggest the presence of a cluster of neighbourhoods in Scarborough having low access to diabetes education programs and a second cluster located in Toronto having high access to diabetes education programs. Bivariate local Moran statistics suggest some spatial coincidence between the two. In other words, neighbourhoods in Scarborough having a high prevalence of type 2 diabetes tend to be surrounded by neighbourhoods having low access to diabetes education programs (global bivariate Moran's I = -0.454). Conversely, there appears to be a significant cluster of low prevalence neighbourhoods in Toronto that have high access to diabetes education programs. Thus, there appears to be a mismatch between the neighbourhood prevalence of type 2 diabetes and access to diabetes education programs in some areas of metropolitan Toronto.

A similar analysis was conducted to assess the spatial coincidence of the prevalence of overweight and access to community recreation centres (Figure 6.3, panels d, e & f). In this case, univariate local Moran statistics point to a cluster of neighbourhoods in north Scarborough having a low prevalence of overweight while some neighbourhoods in south Etobicoke have a relatively high prevalence of overweight. Some of these same neighbourhoods in Etobicoke have relatively low access to community recreation centres. Bivariate local Moran statistics suggest there is some mismatch between the prevalence of overweight and access to recre-

ation centres, particularly in south Etobicoke, where some neighbourhoods having a higher prevalence of overweight are surrounded by neighbourhoods having low access. Conversely, the opposite effect is seen within some neighbourhoods of North York and Scarborough (lower prevalence and higher access). Finally, there appears to be a cluster of neighbourhoods in west North York having higher prevalence of overweight and higher access to community recreation centres. Thus, there appears to be a mismatch between potential demand for recreation programs and supply of those programs in some neighbourhoods of metropolitan Toronto. However, there may be a reasonable match may between demand for and supply of these programs in other areas (e.g., North York).

6.3.4 Probabilistic sensitivity analysis of estimated accessibility

Results from the probabilistic sensitivity analysis indicate that different model parameters contribute to the uncertainty surrounding modelled accessibility. Unlike the sensitivity analysis of forecast diabetes incidence, results for accessibility to health promotion programs have a simpler interpretation because they do not represent parameters from a statistical model but instead represent the direct effects of each model parameter on accessibility. Thus, the sign of the partial rank correlation coefficient can be interpreted in a straight-forward fashion.

Figure 6.4 demonstrates that model parameters exert similar effects on accessibility to both types of health promotion programs. This figure presents the partial rank correlation coefficient (PRCC) from a probabilistic sensitivity analysis of accessibility for all of Toronto's neighbourhoods. Since the range of the PRCC for the travel time parameter was large, neighbourhoods were stratified by values of the travel time PRCC to illustrate the relationship between each parameter and accessibility. Thus, neighbourhoods were split into three groups according to whether the PRCC for the travel time parameter took on negative (PRCC $<$ -0.3), negligible (-0.3 $\leq$ PRCC $\leq$ 0.3), or positive (PRCC $>$ 0.3) values.

Based on this analysis, simulated demand for health promotion programs exerts a similar effect on accessibility. Moreover, the value of the PRCC for the demand parameter for diabetes education programs was similar across all neighbourhoods (median PRCC = -0.9729, IQR: -0.9576 to -0.9802). There was greater

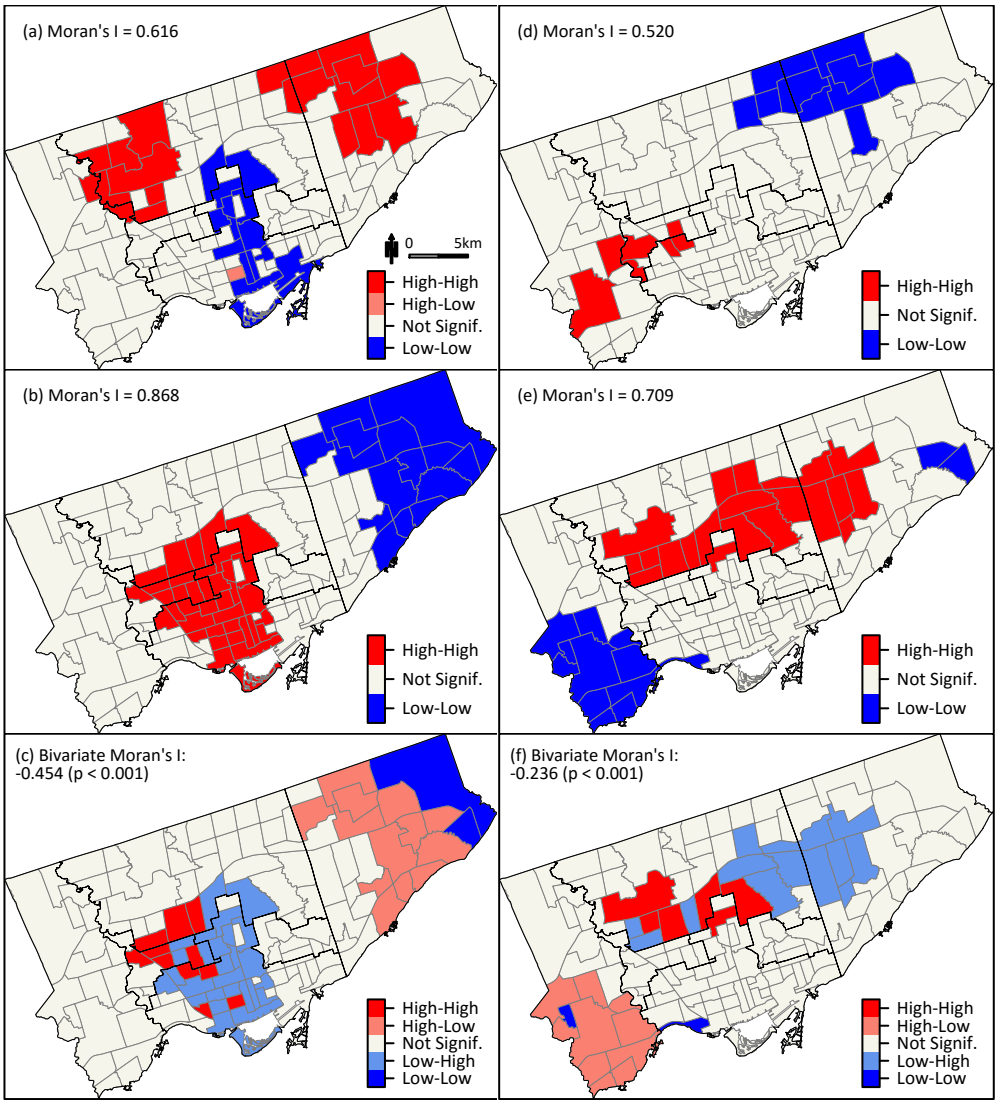
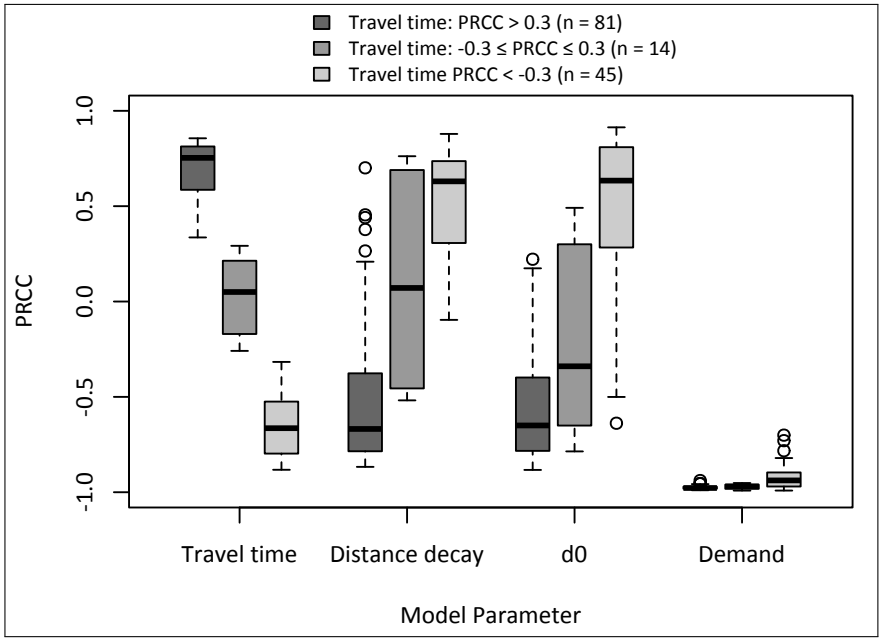


Figure 6.3. Univariate LISA statistics depicting spatial clustering of neighbourhoods by (a) simulated prevalence of type 2 diabetes, (b) access to diabetes education programs, (d) simulated prevalence of overweight and (e) access to community recreation programs. Bivariate LISA statistics in (c) represent the relationship between simulated diabetes prevalence and access to diabetes education programs. Bivariate LISA statistics in (f) represent the relationship between the simulated prevalence of overweight and access to community recreation centres.

(a) Diabetes education programs



(b) Community recreation centres

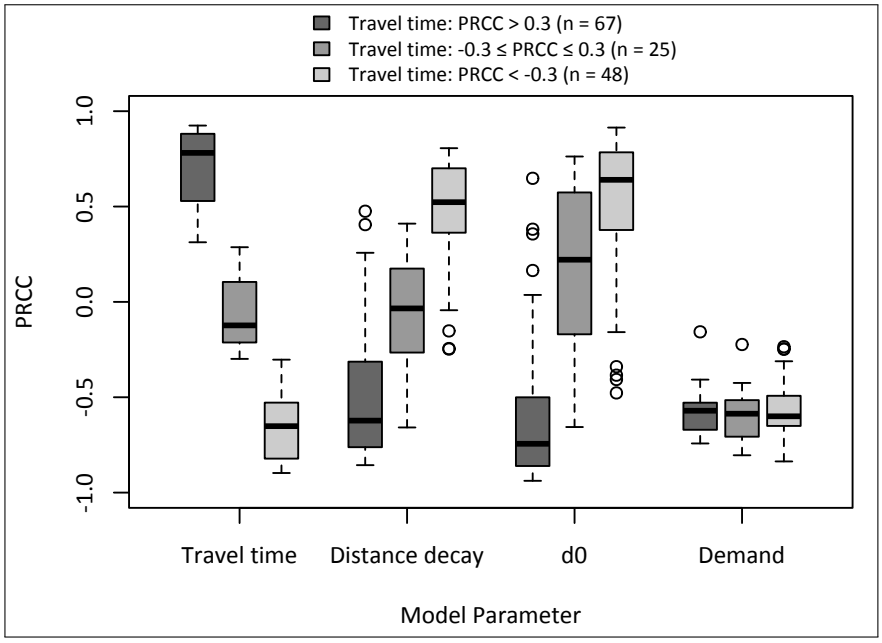


Figure 6.4. Partial rank correlation (PRCC) between parameters of the two-step floating catchment area model and estimated accessibility to (a) diabetes education programs and (b) community recreation centres.

variation in the PRCC for access to community recreation centres (median PRCC = -0.5816,IQR: -0.5142 to -0.6694). The negative sign on the coefficient indicates that as simulated demand increases, accessibility to either type of program decreases.

Different relationships were observed between the distance decay and distance threshold (d_0) parameters and modelled accessibility. Neighbourhoods having positive partial rank correlations for the travel time parameter had negative partial rank correlations for the distance decay and distance threshold parameters. Conversely, neighbourhoods having negative PRCC values for the travel time parameter had positive values for the distance decay and distance threshold parameters. These results suggest that the effect of parameter uncertainty on accessibility varies by geographic location.

In some neighbourhoods, as the value of the travel time parameter increases (i.e., longer commute times between neighbourhoods and health promotion programs), accessibility decreases. In these same neighbourhoods, as the distance threshold parameter increases, accessibility increases. In other words, when the number of programs that lie within the threshold increases, there is a concomitant increase in the total population that can access those programs. Finally, larger values of the distance decay parameter increase accessibility because larger values translate into greater values of the distance decay weights used to differentially weight accessibility scores (Equation 6.3). This, in turn, results in a greater number of services that can be accessed by the population demanding those services.

Thus far, the relationships between model parameters and estimated accessibility make intuitive sense. However, counter-intuitive results are observed in neighbourhoods having positive values of the PRCC for the travel time parameter. In these neighbourhoods, as travel time increases, accessibility increases. Furthermore, as the distance threshold and distance decay increase, accessibility decreases. Although these findings seem counter-intuitive, it must be kept in mind that the total number of services available remains constant throughout the city. Thus, if accessibility increases in some neighbourhoods, it must decrease in other neighbourhoods. As proof, consider that the sum of the population weighted, neighbourhood-specific accessibility scores must equal the total number of services available within the study area (Luo & Wang, 2003).[6]

[6]Put another way, the weighted mean of accessibility is equal to the provider-population ratio for

To further illustrate this point, Figure 6.5 demonstrates the effect d_0 exerts on accessibility, ignoring distance decay effects. It depicts the locations of two different health promotion programs (S1 and S2) along with five different population centres (P1 to P5). Two distance thresholds buffer each program location; the first depicts a 15-minute travel time threshold while the second depicts a 30-minute threshold. Population centre P1 lies within the 15-minute threshold for program S1 while population centre P2 lies within the 30-minute threshold for programs S1 and S2. If d_0 = 15 minutes, then the program-population ratio R_1 for S1 = $\frac{S1}{P1}$. Accessibility for P1 is the sum of all R_j for all programs located within 15 minutes of P1, in this case, R_1. However, if d_0 is increased to 30 minutes, then $R_1 = \frac{S1}{P1+P2}$. In other words, there is greater population demand for program S1 meaning that R_1 *decreases*. However, accessibility for P1 remains the sum of all R_j for all programs within 30 minutes of P1. Since only program S1 lies within 30 minutes of P1, accessibility for P1 also decreases. On the other hand, accessibility for P2 increases, because it lies within 30 minutes of *both* program S1 and program S2. In summary, although the results of the probabilistic sensitivity analysis seem counter-intuitive, given that the total supply of programs is fixed, if accessibility increases in some neighbourhoods, it must decrease in others due to the properties of the 2SFCA model.

Figure 6.6 further illustrates the spatial variation in the partial rank correlation for the travel time parameter. For access to diabetes education programs, most neighbourhoods in Scarborough and a few in North York have negative values of the PRCC. For access to community recreation centres, several neighbourhoods in south Etobicoke, York, Toronto, and East York have negative values of the PRCC. In these areas, access to health promotion programs decreases as travel time to these programs increases. What is interesting about the pattern of results, however, is that there seems to be a lower concentration of programs in these neighbourhoods, while there seems to be a higher concentration of programs in neighbourhoods having positive values of the PRCC. This suggests that where the supply of health promotion programs is limited, travel time is an important parameter influencing access to these programs.

Although the partial rank correlation coefficient can be used to identify model

the entire study area.

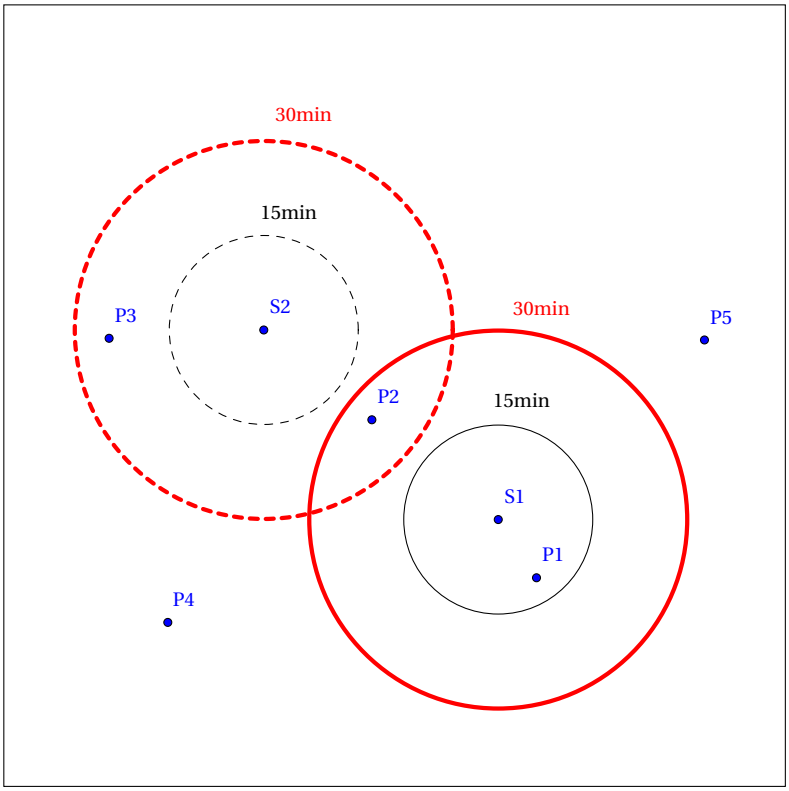
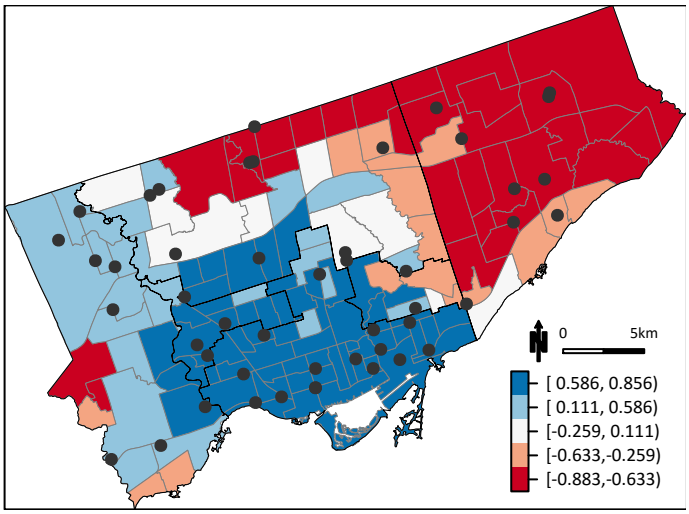


Figure 6.5. Effect of distance threshold parameter, d0, on spatial accessibility.

parameters that exert the greatest influence on outcome uncertainty, due to the large variation in PRCC across neighbourhoods, the sensitivity index (S_i) was used instead. Regression models were used to estimate the sensitivity index across all neighbourhoods. For diabetes education programs, the coefficient of determination from these models was high (median R^2 = 0.959, IQR: 0.939–0.970), suggesting that the sensitivity index was reliably estimated. However, the probabilistic sensitivity analysis of access to community recreation centres was less reliable; in this case, the median R^2 was 0.705 (IQR: 0.569–0.849). In 40 neighbourhoods, R^2 was less than 0.6, suggesting S_i was not reliably estimated. As a result, the sensitivity index was examined separately for neighbourhoods where it was reliably estimated ($R^2 \geq 0.6$) or less reliably estimated ($R^2 < 0.6$).

For access to diabetes education programs, the demand parameter exerted

(a) Diabetes education programs



(b) Community recreation centres

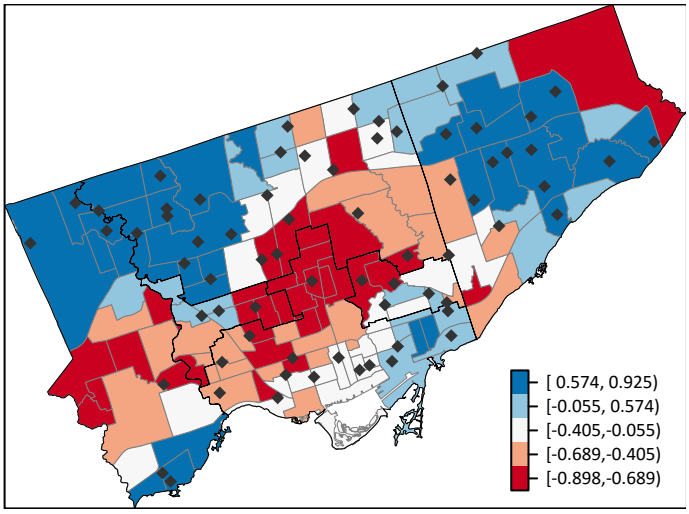


Figure 6.6. Effect of travel time on access to (a) diabetes education programs and (b) community recreation centres as measured by the partial rank correlation coefficient. Neighbourhoods shaded red have negative values of the PRCC indicating that as travel times to program locations increase, accessibility decreases.

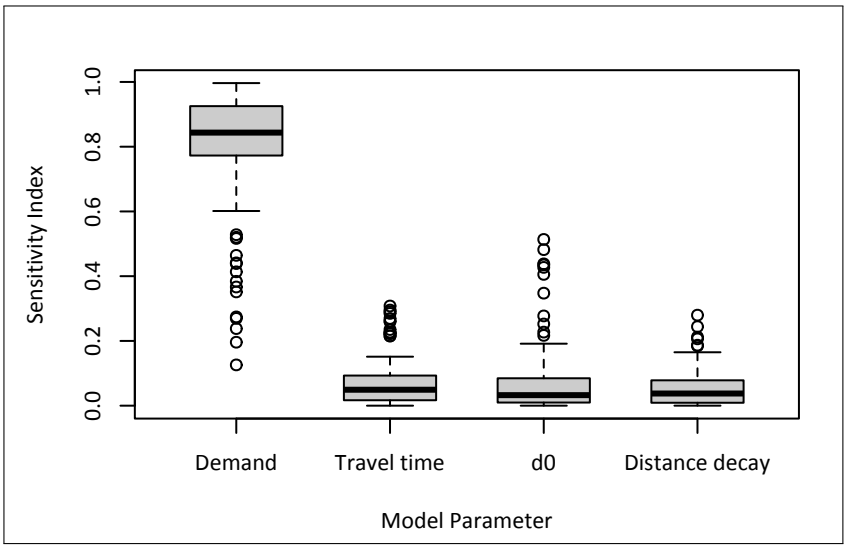
the greatest influence on outcome uncertainty, accounting for 77% or more of the variation in accessibility uncertainty for three-quarters of all neighbourhoods (median S_i = 0.844, IQR = 0.773–0.925). The remaining model parameters accounted for only small amounts of the variability in accessibility uncertainty (Figure 6.7a).

A different picture emerges for access to community recreation centres (Figure 6.7b). There was much greater variation in the sensitivity index across neighbourhoods for all model parameters. In neighbourhoods where S_i was reliably estimated (model $R^2 \geq 0.6$), the distance threshold and the travel time parameters had the most influence on outcome uncertainty. The demand and distance decay parameters were less influential. In neighbourhoods where the sensitivity index was less reliably estimated, the demand parameter might have had more influence on the uncertainty surrounding accessibility. However, the effect of this parameter varied considerably across neighbourhoods. Ultimately, the probabilistic sensitivity analysis of accessibility to community recreation centres was equivocal, in that no model parameter consistently influenced outcome uncertainty. It is possible that in this case, values used to model accessibility along with the range of values generated to conduct the probabilistic sensitivity analysis were less than optimal.

Figure 6.8 illustrates geographic differences in the sensitivity index for the demand and distance threshold parameters for accessibility to diabetes education programs. Across most neighbourhoods, the demand for diabetes education programs accounts for 53% or more of the variability in estimated accessibility. One exception occurs in neighbourhoods located in east Scarborough. Here, the distance threshold (d_0) accounts for a larger share of outcome uncertainty. Moreover, a global Moran test indicates that values of the sensitivity index for these parameters are spatially dependent, so that neighbourhoods within close proximity of one another have similar values of the sensitivity index.

Based on these results, it is clear that the simulated population demand for diabetes education programs explains most of the variability in accessibility uncertainty. However, in many Scarborough neighbourhoods, uncertainty is more strongly influenced by the distance threshold parameter. Given there are fewer diabetes education programs within Scarborough compared to the rest of the city,

(a) Diabetes education programs



(b) Community recreation centres

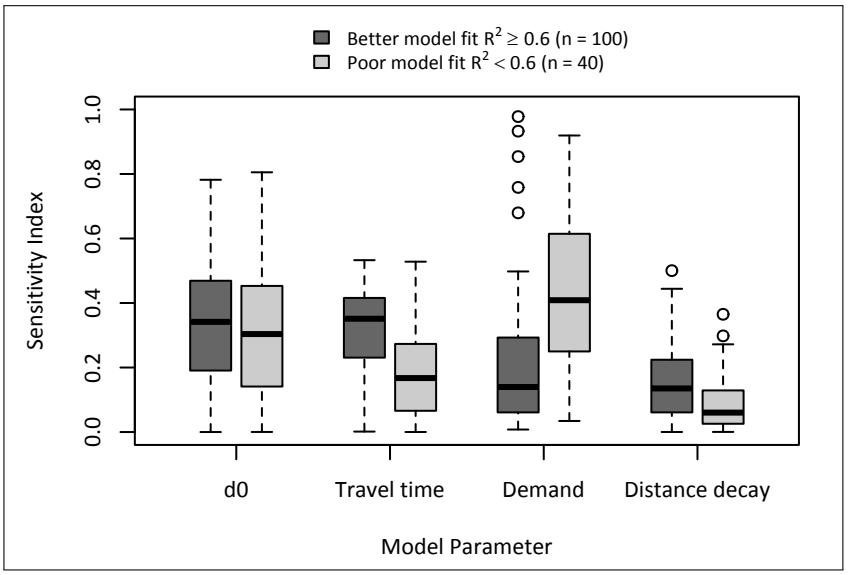
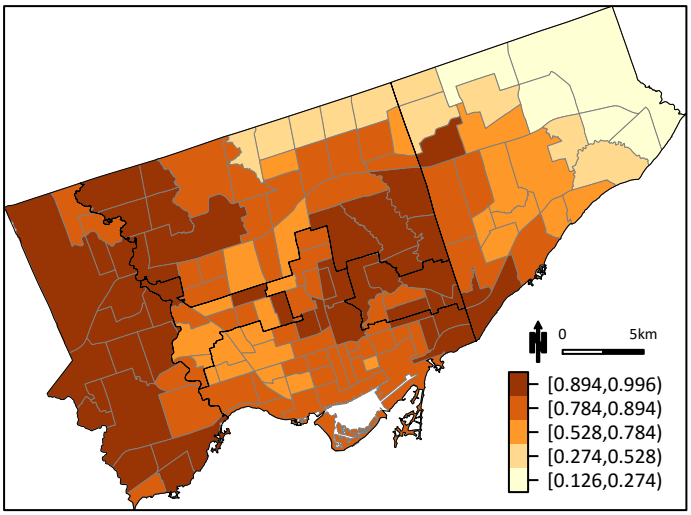


Figure 6.7. Proportion of uncertainty (sensitivity index) in spatial accessibility explained by two-step floating catchment area model parameters for (a) diabetes education programs and (b) community recreation centres.

(a) Demand parameter (Moran's I = 0.780)



(b) Distance threshold parameter (Moran's I = 0.807)

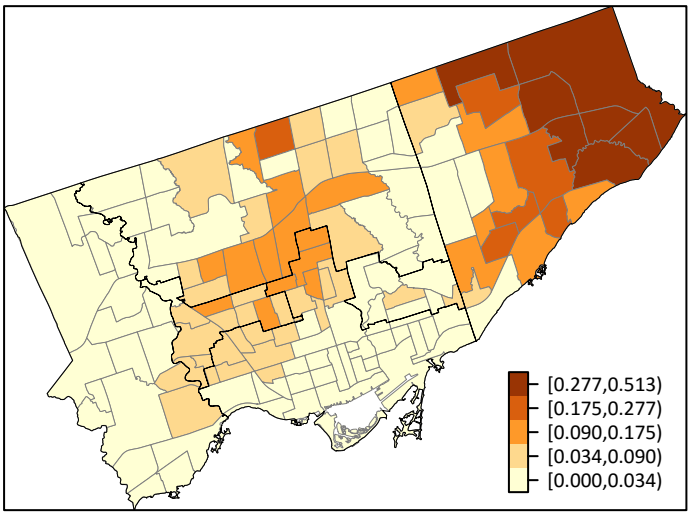
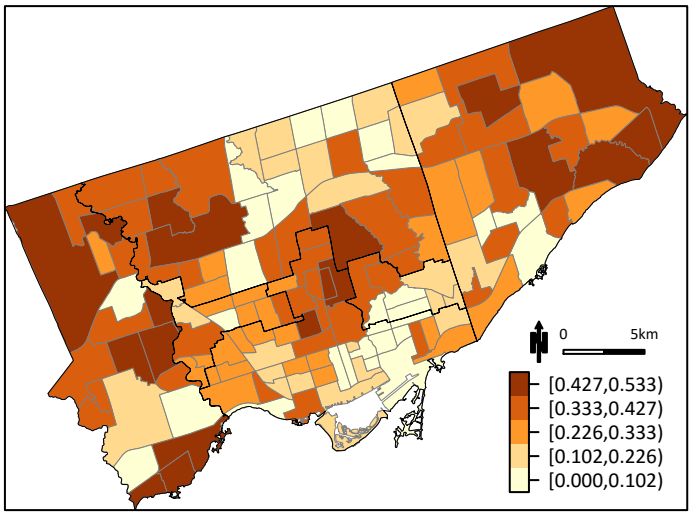


Figure 6.8. Probabilistic sensitivity analysis of accessibility to diabetes education programs: geographic variation in the sensitivity index for the (a) demand and (b) distance threshold parameters.

(a) Travel time parameter (Moran's I = 0.356)



(b) Distance threshold parameter (Moran's I = 0.251)

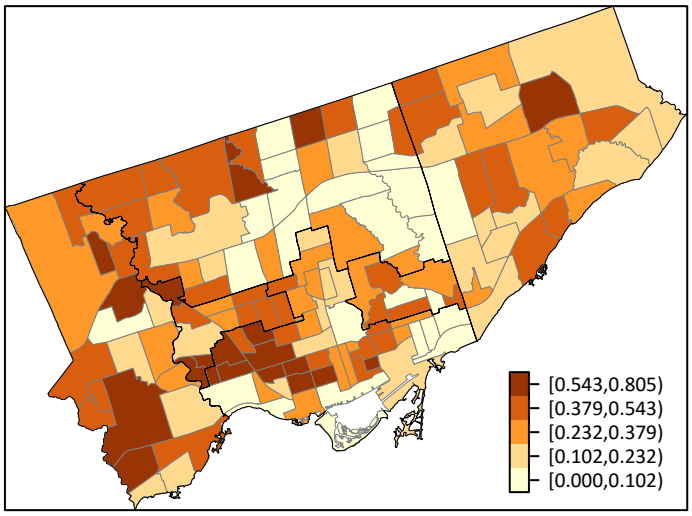


Figure 6.9. Probabilistic sensitivity analysis of accessibility to community recreation centres: geographic variation in the sensitivity index for the (a) travel time and (b) distance threshold parameters.

this result suggests that a greater number of programs might be needed in Scarborough, especially considering the simulated prevalence of type 2 diabetes is higher in these neighbourhoods (see panel (a) of Figure A.3). However, since edge effects influence the results of the 2SFCA model, this conclusion needs to be interpreted cautiously. Indeed, edge effects are likely present for these neighbourhoods because the distance threshold parameter accounts for a greater proportion of outcome uncertainty. In other words, neighbourhoods that border east Scarborough only have access to programs to the west. As a result, increasing the distance threshold means that these populations within these neighbourhoods will be able to access programs (according to the model) when they previously could not.

With respect to community recreation centres, no clear geographic patterns result from a cartographic visualization of the sensitivity index (Figure 6.9). Indeed, neighbourhoods having similar values of the sensitivity index for the travel time and distance threshold parameters are more geographically dispersed. A global Moran test supports this idea: for the travel time parameter, Moran's I = 0.356 ($p < 0.01$). For the distance threshold parameter, Moran's I = 0.251 ($p < 0.01$). Thus, there appears to be less similarity in the values of the sensitivity index for neighbourhoods within close geographic proximity of one another.

6.4 Discussion

6.4.1 Implications

Using simulated, neighbourhood-specific counts of the number of cases of type 2 diabetes and overweight, this exploratory analysis of geographic accessibility to health promotion programs suggests potential access to diabetes management and prevention programs may be limited. On average, there may only be 0.54 diabetes education programs available in Toronto for every 1000 type 2 diabetics and even fewer community recreation centres (0.89 per 10,000 overweight residents).

Although it is difficult to know whether estimates of spatial access are sufficient to meet demand, the Ontario Ministry of Health and Long-Term Care stipulates that each Diabetes Education Team have an active caseload of 1000 patients per year. If diabetes education programs were meeting this target, then the citywide average spatial access should be similar to the required caseload, or 1

program for every 1000 diabetics. Given the citywide average is only 0.54 programs/1000 type 2 diabetics, this suggests that if all diabetics in the city were using these programs, then programs might be operating *above* capacity (i.e., there is an insufficient number of programs to serve all diabetics within metropolitan Toronto). However, this conclusion assumes that all diabetics (type 1 *and* type 2) are actively using diabetes education programs, an unlikely scenario, especially given the 2012 Ontario Auditor's General report concluded that 90% of Diabetes Education Teams are under-utilized (Office of the Auditor General of Ontario, 2012). If programs are under-utilized, it is possible that diabetics either are not aware of these services or they are not inclined to use them.

Potential access is also affected by the service capacity of each program. Accessibility measures reported here assume that the number of staff in each program is constant across all programs. In other words, accessibility measures reflect access to *programs* and not total services that could be provided by each program. If the number of staff differed by program location, then estimated accessibility would change. However, all diabetes education programs are staffed by a diabetes nurse educator and a registered dietitian (Ministry of Health and Long Term Care, 2012). Therefore, the base level of service provision should be the same across all programs within Toronto. Likewise, the base level of service provision for recreation centres would be the same if each centre delivered the same number of diabetes prevention programs throughout the year (e.g., one prevention program offered biannually). Under these assumptions, it is possible to use the accessibility measures reported here to identify relative differences between neighbourhoods and where potential access may be higher or lower compared to the citywide average. In other words, the results of this exploratory analysis should be interpreted as *relative* measures of spatial access.

Access to diabetes education programs was highest in central Toronto. Neighbourhoods located in the boroughs of Toronto and York had relatively higher access compared to the metropolitan average, while neighbourhoods in Etobicoke and Scarborough had the lowest access. These findings are unsurprising given there is a greater concentration of programs in Toronto compared to other areas of the city. These general patterns held regardless of the number of programs used to model accessibility.

Different spatial patterns emerged for potential access to community recreation centres. Considering only recreation centres capable of delivering diabetes prevention programs (n = 70), neighbourhoods in North York had the greatest access (1.1 programs per 10,000 overweight population) while neighbourhoods in Etobicoke had the lowest access (0.7 programs per 10,000 overweight population). Neighbourhoods in Scarborough (0.96 programs per 10,000) and York (0.93 programs per 10,000) also had relatively higher access to community recreation centres compared to the metropolitan average (0.89 programs per 10,000).

A markedly different spatial pattern emerged when all 148 community recreation centres were used to model spatial accessibility. Under this scenario, neighbourhoods in Etobicoke had the highest access while neighbourhoods in Scarborough had the lowest access. This dramatic shift in accessibility is attributable to the both the number of recreation centres used in modelling accessibility and the location of centres capable[7] of delivering diabetes prevention programs. This finding is important for public health planning purposes. If diabetes prevention programs were delivered more locally throughout the city, then the choice of suitable locations needs to consider where programs could be located and the types of facilities in which they should be delivered. Both of these considerations influence the size of the high-risk population that could attend such programs. Geographic location-allocation models could be used as a next step in optimizing program locations and thereby the size of the target population that would attend these programs throughout the city.

This conclusion is also relevant for resource allocation decisions related to the provision of diabetes education programs. As depicted in Figures 6.1a and 6.1b, there appears to be a greater concentration of diabetes education programs in downtown Toronto. Given that diabetes education programs are often located in physician practices (Ministry of Health and Long Term Care, 2012), it may be difficult to redistribute programs in an optimal manner to ensure more equitable access throughout Toronto. However, it might be the case that additional public health resources could be directed to those areas of the city having lower access to diabetes education programs. Such efforts may help type 2 diabetics living in

[7] Defined for this analysis as a centre that already offers fitness programs *and* contains a multi-purpose room where nutrition classes could be provided.

lower access neighbourhoods better manage their diabetes.

The results reported here also demonstrate that a mismatch may exist between population demand for services and potential access. Univariate and bivariate local Moran statistics suggest that some areas of Toronto have a higher prevalence of type 2 diabetes or overweight but lower spatial access to health promotion programs (Figure 6.3). Regarding access to diabetes education programs, there are several neighbourhoods in Scarborough that have relatively higher prevalence rates of type 2 diabetes but relatively lower spatial access to diabetes education programs. Conversely, there are areas of downtown Toronto that have relatively lower prevalence rates of type 2 diabetes but greater spatial access to diabetes education programs. These findings are consistent with those reported by Glazier et al. (2007) who found higher rates of diabetes in Scarborough but low concentrations of diabetes education programs. However, the pattern of results reported for downtown Toronto differed. Generally, they found lower prevalence rates of diabetes (consistent with these findings) and lower concentrations of diabetes education programs (inconsistent with these findings). Differences might be explained by the different analytic methods used; specifically, Glazier et al. (2007) only used provider-population ratios as a measure of service provision. As mentioned in Section 6.1, these ratios do not consider availability of services in neighbouring areas.

Such disparities in service provision suggest that additional resources could be devoted to diabetes management in neighbourhoods having higher prevalence rates of diabetes but relatively poor spatial access to health promotion programs. Additional information should gathered in these areas to identify the types of resources that would help diabetics living there better manage their disease. For example, given that some neighbourhoods within Scarborough seem to have poor spatial access to diabetes education programs, additional data could be gathered to identify:

- how well diabetics are currently managing their disease,
- whether they are aware of local services to help them manage their disease, and
- whether and under what conditions they would use these types of services to help them manage their disease.

Given the complications and concomitant medical costs associated with poorly managed diabetes, preventing such complications is an important health promotion objective.

Another issue to consider is how to address excess supply of diabetes education programs in areas having lower prevalence rates of diabetes but relatively greater spatial access. First, the finding that some areas of Toronto have relatively high spatial access to diabetes education programs but relatively low prevalence rates supports the idea that some programs may be under-utilized (Office of the Auditor General of Ontario, 2012). Additional data should be gathered to identify whether diabetes education programs in these areas are operating at near- or full-capacity. If this is not the case for programs in low-prevalence, high spatial access areas, additional information might identify whether patients are drawn from a limited catchment area or could come from other areas of the city. If services are limited to diabetics residing in specific catchment areas, it might be useful to assess whether diabetics from low access neighbourhoods could attend programs in higher access neighbourhoods. In summary, the findings from this research should be used to inform future program evaluation studies that address the disease management needs of diabetics who are unable to access diabetes education programs.

Turning to diabetes prevention programs, univariate and bivariate local Moran statistics suggest that some metropolitan Toronto neighbourhoods have relatively poor access to community recreation centres in spite of having higher relative prevalence rates of overweight. A group of neighbourhoods in south Etobicoke fall under this classification. It is interesting that these neighbourhoods are the same neighbourhoods whose accessibility index was dramatically reduced if the accessibility modelling was conducted using the restricted set of recreation centres instead of the full set. This suggests that community recreation centres in south Etobicoke might not be suitable for delivering diabetes prevention programs, either because they do not currently offer cardiovascular fitness programs or because they lack a multi-purpose room suitable for classroom instruction. In this case, additional resources might be devoted to providing additional capabilities for these recreation centres, such as providing additional staff (or volunteers) to teach cardiovascular fitness classes. If recreation centres lack suitable rooms for

delivering classroom style instruction on diabetes prevention and nutrition, then alternate locations for the delivery of prevention programs should be considered.

From a methodological standpoint, the probabilistic sensitivity analysis of spatial accessibility is a relatively new and unused technique in accessibility modelling, especially with respect to the 2SFCA model. As presented here, the probabilistic sensitivity analysis demonstrates that input parameters in the 2SFCA model influence uncertainty about potential access in similar ways. First, negative values of the partial correlation coefficient for the demand parameter indicate an inverse association between demand and spatial accessibility. That is, as demand for health promotion programs increases, accessibility decreases. Although this might be intuitive, the empirical evidence presented here suggests it may be a general feature of these models. In spite of that, the strength of that association may vary for different types of services modelled. Based on these results, demand for diabetes education programs exerted a similar effect on spatial accessibility across almost all neighbourhoods. Moreover, there was little variation in the effect of demand on estimated accessibility. However, greater variation in spatial accessibility was produced by the demand parameter when modelling access to community recreation centres (Figure 6.4).

Less consistent effects were seen for the other parameters used to model spatial accessibility as there was a wide range in neighbourhood-specific PRCC values. However, stratifying partial rank correlations by the value of the correlation for the travel time parameter reveals some interesting points. First, in some neighbourhoods, there is a positive association between the travel time parameter and spatial accessibility, meaning that as travel time increases, access increases. In other neighbourhoods, as travel time increases, access decreases. Although the former seems counter-intuitive, the result is attributable to the idea that the total supply of health promotion services remains constant for accessibility modelling.

Second, high values of the partial rank correlation for the travel time parameter are associated with low PRCC values for the distance decay and distance threshold parameters. Conversely, low values of PRCC for the travel time parameter are associated with high values of PRCC for the distance decay and distance threshold parameters. This pattern of results suggests these parameters work together to influence estimates of spatial accessibility. In other words, in some neighbourhoods,

faster travel times to service locations and larger values of the distance decay and distance threshold parameters all increase spatial accessibility to health promotion programs. Given, however, that the total supply of services remains constant, this also means that in other neighbourhoods, the opposite phenomenon is observed. Finally, there are some neighbourhoods where these parameters exert negligible or weak effects on spatial accessibility to health promotion programs, as suggested by partial rank correlations lying between -0.3–0.3 for the travel time parameter.

The sensitivity index provides a way to quantify the amount of uncertainty in spatial accessibility explained by each of the 2SFCA model parameters. For access to diabetes education programs, neighbourhood-specific estimates of the sensitivity index unequivocally demonstrate that the demand parameter accounts for the majority of uncertainty in accessibility. Specifically, in 75% of all neighbourhoods, the demand parameter explains almost 80% of the uncertainty in accessibility. Cartographic presentation of these results further demonstrates this point. However, in several east Scarborough neighbourhoods, the distance threshold parameter d_0 explains a greater proportion of uncertainty in modelled accessibility than the demand parameter.

Results are less consistent with respect to spatial access to community recreation centres. First, the sensitivity index may not have been reliably estimated in 40 neighbourhoods. However, the distance threshold and travel time parameters account for 46% and 38% of the variability in spatial accessibility, respectively, in 50 of the 100 neighbourhoods where the sensitivity index was reliably estimated. Thus, uncertainty in spatial accessibility may result from the type of service that is modelled and from the specific values of parameters used to model accessibility.

For access to diabetes education programs, the distance threshold and distance decay parameters were selected on the basis of prior research modelling spatial accessibility to health care providers (Luo & Qi, 2009; Luo & Wang, 2003; McGrail & Humphreys, 2009; Wan et al., 2012; Wang & Luo, 2005). Since diabetes education programs are staffed by health care providers (nurses and registered dietitians), on a population level, decisions to attend these programs may be influenced by similar factors as those that influence the choice to attend particular physician practices or health care providers more generally.

Less empirical evidence was available to inform the choice of parameters used to model accessibility to community recreation centres. Given the nature of the demand for these services, it was assumed that the population of overweight individuals might be less inclined to travel longer distances to attend diabetes prevention programs. Thus, values for the distance decay and distance threshold parameters were chosen assuming the population at risk would not travel further than 20 minutes from their homes to attend these programs. The choice of these parameters may have made it more difficult to reliably model spatial accessibility to community recreation centres. Future research should study travel patterns to fitness programs in order to better estimate appropriate thresholds for maximum travel distance and average distance decay effects.

In summary, the results of the probabilistic sensitivity analysis demonstrate how randomly varying 2SFCA model parameters influences estimated accessibility. However, the methods used to conduct the probabilistic sensitivity analysis might have influenced the results observed. Specifically, model parameters were randomly sampled from a triangular distribution. As Briggs and Sculpher (2006) point out, a triangular probability distribution may have limited utility for conducting a probabilistic sensitivity analysis. In particular, the mode may not be the central point of the distribution (i.e., mode $\neq$ mean) and using minimum and maximum values suggests that more extreme values cannot occur. Different probability distributions, such as the normal distribution, permit the rare sampling of extreme values.

Thus, it might have been more appropriate to use a normal distribution to randomly sample parameter values for the sensitivity analysis. In this case, appropriate estimates of the mean effect of each parameter, along with its standard deviation, would be needed to randomly sample parameter values. Since such estimates were unavailable (in particular, the standard deviation), selecting a triangular distribution to sample from made it possible to conduct a probabilistic sensitivity analysis to quantify the uncertainty in spatial accessibility associated with 2SFCA model parameters. Therefore, in spite of its limitations, the probabilistic sensitivity analysis provides useful insights into how model parameters influence accessibility to health promotion programs.

6.4.2 Limitations

In this chapter, potential spatial accessibility to health promotion programs was modelled using the two-step floating catchment area model in conjunction with simulated, neighbourhood-specific counts of the demand for those services. Simulated counts of the number of prevalent cases of type 2 diabetes served as the demand parameter for modelling accessibility to diabetes education programs while simulated counts of overweight were used as the demand parameter for modelling accessibility to community recreation centres. If these simulated counts are differentially biased by geographic location, then spatial patterns in accessibility will also be biased. However, when the known number of cases of diabetes, ascertained from the Ontario Diabetes Database, was used as the demand parameter, spatial patterns of accessibility matched those using simulated demand. It must be recognized, however, that accessibility is overestimated using simulated demand by a factor of two. Thus, if accurate estimates of accessibility are required then measures of demand must not be biased. In this case, the simulated number of diabetes cases underestimates true demand producing accessibility estimates that are too optimistic. Given this research used the 2SFCA model in an exploratory way to examine *relative* differences in accessibility to health promotion programs, the spatial patterns in access to diabetes education programs reported here hold irrespective of whether accessibility is estimated using the simulated or true number of diabetes cases.

The spatial patterns of access to community recreation centres need to be interpreted more cautiously. This is because it was not possible to validate spatial patterns in the prevalence of overweight and obesity using the incidence of cancers associated with obesity. In spite of this, broad spatial patterns in forecast diabetes incidence were reasonably similar to the true incidence of diabetes at the neighbourhood level. Since incidence was forecast from a model that relies on the prevalence of overweight and obesity, there is some reason to believe that the simulated counts of overweight and obesity at the neighbourhood level may have been captured to some degree of accuracy. Thus, spatial patterns of accessibility to community recreation centres may be also be broadly accurate. Therefore, it should still be possible to identify neighbourhoods having relatively poor access

to community recreation centres using the results reported here. Future research should explore these patterns in more detail using accurate demand estimates.

It must also be recognized that accessibility estimates along the borders of metropolitan Toronto are likely downward biased due to the presence of edge effects. That is, the 2SFCA model assumes that people living in neighbourhoods lying along city borders can only access programs within the city. For modelling purposes, this means that the westernmost neighbourhoods of Etobicoke and easternmost neighbourhoods of Scarborough can only access programs to the east or west, respectively. However, neighbourhoods in central Toronto can access programs in all directions. The pattern of results depicted in Figure 6.1b and Figure 6.1d is consistent with the presence of edge effects.

Such edge effects are usually mitigated by including additional small areas outside the main study area when modelling accessibility. In this case, neighbourhoods in Mississauga, Vaughan, Markham, and Pickering could have been included. This would have meant (a) devising comparable ways of defining neighbourhoods in these cities and (b) developing the TropISM spatial microsimulation for these additional neighbourhoods. Since comparable definitions of neighbourhoods for these cities do not currently exist, it was deemed impractical to develop such definitions for the current study.

In spite of the presence of edge effects, if spatial accessibility to health promotion programs is lower in Etobicoke and Scarborough, this information could still be used for public health planning purposes. In particular, Toronto Public Health is responsible for delivering public health services to all city residents, even if some residents in Etobicoke and Scarborough might be more inclined to use programs outside metropolitan Toronto. Therefore, indicators of low accessibility to health promotion programs in these neighbourhoods warrant additional investigation. Further study should assess whether residents (a) lack sufficient access to health promotion health programs, (b) want access to these programs, and (c) would use these programs. In other words, this preliminary study of potential access to health promotion programs should be followed up by more detailed evaluation of revealed access to these programs.

Furthermore, current urban development trends in Toronto favour mixed-use development (Kane, 2014). Mixed-use development makes it possible for people

to work, eat, and play in their local neighbourhoods by providing access to commercial and civic opportunities within walking distances of their homes (Grant, 2002). Although there are different types of mixed-use development, these urban forms have the potential to improve population health (Barton, 2009). Thus, even if estimates of accessibility to health promotion programs are artificially lower in Etobicoke and Scarborough, mixed-use development encourages provision of these resources in low access neighbourhoods.

Another methodological limitation to consider is that distances between neighbourhoods and program locations were measured from population weighted neighbourhood centroids. As such, travel distances between neighbourhoods and program locations represent an average or typical measure of distance. In a neighbourhood covering a small area of land, the population weighted centroid might be a reasonable estimate of distance between any location within the neighbourhood to program locations. However, in a neighbourhood covering a large area of land, the population weighted centroid might not be a reasonable estimate of distance for any specific location within that neighbourhood. Thus, estimated spatial accessibility in these types of neighbourhoods might be less representative of average accessibility than in neighbourhoods covering smaller land areas.

It is also important to note that the travel distances used in the model represent car travel times and not other modes of transportation, such as public transit or walking. Estimates of spatial accessibility using public transit travel times would be more relevant for population subgroups, such as the elderly and immigrant populations, because these groups may not own their own vehicles. In this case, accessibility modelling would need to consider estimates of demand in these population subgroups, but also estimates of public transit travel times along a more restricted street network. Depending on the availability of public transit in specific neighbourhoods, the spatial patterns of neighbourhood accessibility may differ greatly from those presented here. For example, downtown areas of Toronto would have considerably better access to public transit than outlying areas of the city, due to greater availability of bus, subway, and streetcar transit modes, compared to outlying neighbourhoods that may only have access to buses.

Other modelling decisions may have influenced the results. First, little previous research was available to inform the choice of the distance threshold and

distance decay parameters used by the 2SFCA model. The choice for modelling accessibility to diabetes education programs was influenced by early research using a 30-minute threshold for modelling accessibility to physicians in the United States (Lee, 1991; Luo & Wang, 2003) along with a relatively gradual distance decay function. Parameters for modelling accessibility to community recreation centres were chosen under the assumption that people would *not* be willing to travel greater distances to access prevention programs such as cardiovascular fitness programs. To better model urban accessibility to health promotion programs, future research should conduct more detailed travel surveys to estimate how willing residents wanting to use these types of programs are willing to travel. Future research should also try to empirically estimate appropriate distance thresholds and distance decay functions.

In addition, there was a temporal mismatch between simulated demand for health promotion programs the the locations of diabetes education programs and community recreation centres. Program locations were based on more recent data (2014) compared to simulated demand estimates (2005). If the number of locations increased dramatically, or differentially throughout metropolitan Toronto, the relative spatial patterns in access presented here may be biased.

Finally, accessibility to diabetes education programs ignored the existing catchment areas of these programs. That is, some programs throughout Toronto are restricted to people living within a predefined area, typically bounded by specific streets within the city. Ignoring these catchment areas simplified accessibility modelling. However, it also means that accessibility estimates assume the population of any neighbourhood could access any program within the 30-minute distance threshold specified by the model. Accounting for such restrictions would alter neighbourhood-specific estimates of accessibility, especially if diabetes education programs in certain ares of Toronto were limited to specific catchment areas while others were not. Future accessibility modelling should consider ways to incorporate existing catchment areas into the modelling process.

6.4.3 Summary

This exploratory assessment of potential spatial accessibility to diabetes health promotion programs in Toronto demonstrates disparities in access may exist for some neighbourhoods throughout the city. Some communities with higher prevalence of type 2 diabetes have lower spatial access to diabetes education programs. Different comminutes with higher prevalence of overweight (BMI $\geq$ 25 kg/m^2) have lower access to community recreation centres. These results focus on specific types of diabetes prevention that might help high-risk groups better manage their disease or prevent the onset of type 2 diabetes. However, the total urban environment should be considered as well. Other community resources provide opportunities for engaging in physical activity, including community parks, recreation trails, and walkable urban neighbourhoods. Furthermore, healthy food choices can be made easier on a population level by ensuring sufficient local access to grocery stores. Future research should consider whether access to total opportunities for diabetes prevention differs at the small area level. This moves diabetes prevention from the realm of public health into the sphere of urban planning. In other words, there needs to be a movement away from obesogenic environments that promote weight gain increasing the risk of type 2 diabetes towards supportive urban environments and the totality of resources within them, resources that support healthy food choices and promote physical activity and active living. Future studies should therefore assess whether total access to all health promoting resources at the neighbourhood level is associated with lower prevalence of overweight and type 2 diabetes.

Chapter

7

Conclusions

The key premise undergirding this research is that locally relevant information is needed for effective decision making in public health practice. Program planners must decide which health promotion programs and polices will best improve the health of local populations and how the effects of those programs and policies might vary over geographic space. At the community level, differential program effects might be driven by different sociodemographic and health risk profiles of specific communities. Such differences should inform resource allocation decisions. Predicting the possible consequences of different policy options provides public health decision makers with a practical way to choose one course of action over another.

Maglio, Sepulveda, and Mabry (2014) argue that although simulation models are currently underutilized in population health, these models should complement traditional quantitative methods to better understand how different decisions influence population health outcomes. The value of simulation models is derived from their ability to examine the effects of different scenarios, thereby forecasting a range of plausible outcomes that may result from a particular course of action. Simulation models provide insight into which factors are most responsible for producing a particular outcome. Model outputs should not be viewed as exact solutions, but rather as additional evidence that informs the decision making process (Alper & Geller, 2016; Maglio et al., 2014).

In spite of its limitations, the Type 2 dIabetes Spatial Microsimulation model, or TropISM, developed in this research meets these goals. As a model of local population health in metropolitan Toronto, it demonstrates how the prevalence of type 2 diabetes and its risk factors varies across the 140 neighbourhoods of metropolitan Toronto. Although TropISM underestimates the true prevalence of type 2 diabetes and hypertension and overestimates the prevalence of heart disease among men, the neighbourhood level validation demonstrates that spatial patterns of these outcomes are predicted reasonably well. Neighbourhoods having low or high simulated prevalence tend to have low or high true prevalence. These synthetic small area estimates can be used to identify neighbourhoods experiencing a greater burden of disease.

Based on these findings, neighbourhoods within north Etobicoke, west North York, and most of Scarborough face a higher burden of type 2 diabetes compared to other neighbourhoods within metropolitan Toronto. Relative to lower risk neighbourhoods, these higher risk communities are comprised of sizeable visible minority populations. These findings point to an unequal burden of type 2 diabetes across metropolitan Toronto, suggesting that diabetes prevention and management programs should be customized to suit the particular needs of specific communities.

Although TropISM underestimates the prevalence of type 2 diabetes and its risk factors, when model outputs were used to forecast diabetes incidence using the Diabetes Population Risk Tool, TropISM produces reasonably accurate projections. Although forecast incidence was less accurate in neighbourhoods within north Etobicoke, west North York, and Scarborough, both forecast incidence and the true incidence of diabetes were higher in these communities than those from the boroughs of south Etobicoke, York, Toronto, and East York. Different weight loss scenarios suggest that when the population of high risk individuals having a body mass index greater than 25 kg/m^2 loses 10% or more of its body weight, neighbourhoods in north Etobicoke and Scarborough will experience greater reductions in diabetes incidence compared to other neighbourhoods. Forecast reductions were larger in neighbourhoods comprised of greater percentages of visible minorities and immigrants. These findings support the idea that a particular health promotion program will produce different effects in different communi-

ties. They also point to the necessity of tailoring health promotion activities to the specific needs of the communities in which they are delivered.

The evaluation of spatial accessibility to health promotion programs was another way in which TropISM was able to identify potential policy options. In particular, when the simulated prevalence of type 2 diabetes and overweight (BMI $\geq$ 25 kg/m^2) were used as the "demand" for health promotion programs, it was possible to identify neighbourhoods within metropolitan Toronto having relatively higher disease or risk factor prevalence but lower spatial access to health promotion resources. The potential mismatch between disease burden and accessibility to health promotion resources further highlights how spatial microsimulation models can inform the delivery of health promotion programs to local communities. In this case, one possible course of action would be to redistribute programs from low disease burden neighbourhoods having relatively high access to lower access neighbourhoods having greater disease burden.

Based on these findings, the TropISM spatial microsimulation model was able to (a) predict the consequences of different courses of action and (b) identify potential mismatches between existing demand for health promotion programs and the geographic availability of those resources. This locally relevant information enables public health planners to better allocate scarce resources to areas of greatest need. In summary, this research illustrates the utility of spatial microsimulation modelling as a spatial decision support tool for local public health practice.

One unique output from this research is the use of probabilistic sensitivity analysis to ascertain the magnitude of uncertainty surrounding model results. Unlike other types of sensitivity analysis, probabilistic sensitivity analysis assumes that model parameters follow a defined probability distribution. Randomly varying model parameters according to that distribution makes it possible to assess which parameters exert more influence on the uncertainty about model results. A probabilistic sensitivity analysis of forecast diabetes incidence suggests that forecast uncertainty was strongly influenced by (a) the specific parametrization of the Diabetes Population Risk Tool and (b) the underlying demographic and health risk factor profile of neighbourhood populations. In addition, some parameters, such as visible minority status, exerted greater influence on forecast incidence in some neighbourhoods compared to others. A probabilistic sensitivity analysis of spatial

access to health promotion programs suggests that the overall prevalence, or "demand", for these programs is the most important factor influencing uncertainty. Accessibility tends to decrease as demand increases.

These findings illustrate that it is possible to quantify the uncertainty surrounding small area policy modeling scenarios developed using spatial microsimulation outputs. Clarke (1996) and Whitworth et al. (2016) underscore the necessity of quantifying this uncertainty and this research offers a unique way to do so using probabilistic sensitivity analysis. Quantification of uncertainty allows public health planners to assess which factors exert greater influence on the possible outcomes of prevention programs in specific neighbourhoods. This additional information suggests how programs may need to be tailored to specific communities. For example, in some neighbourhoods it might be more important to focus on visible minority groups to achieve greater reductions in diabetes incidence. In other neighbourhoods, it might be more important to focus on excess body weight in older people. Thus, quantification of model uncertainty provides public health planners with additional information guiding resource allocation decisions.

These results also suggest how public health modellers might develop better spatial microsimulation models in future research studies. First, it is necessary to ensure that the geographic heterogeneity in unconstrained outcomes is captured by these models. Burden and Steel (2015) suggest using the within-area homogeneity statistic (Steel & Tranmer, 2011) to identify variables eliciting sufficient spatial correlation with unconstrained outcomes. Doing so should help capture sufficient spatial variation in unconstrained outcomes. Additional bivariate constraints should also be used during model development to better replicate the geographic heterogeneity in unconstrained outcomes. Given that TropISM underpredicted both diabetes prevalence and diabetes incidence in specific neighbourhoods, using additional bivariate constraints for model development may have been useful. For example, bivariate constraints crossing visible minority status with sex, age group, and education may have produced prevalence estimates that were more reflective of the true geographic variation in diabetes prevalence across metropolitan Toronto.

Related to constraint selection is the availability of suitable variables in the census data used to develop spatial microsimulation models. Currently, the Cana-

dian long-form census does not contain any information about self-rated health. Inclusion of a simple measure such as this would be useful for health planning in general, but also for development of future spatial microsimulation models using Canadian data. Given that self-rated health is correlated with both subjective health outcomes and objective health status (Goldberg, Guegen, Schmaus, Nakache, & Goldberg, 2001; Latham & Peek, 2013; Wu et al., 2013), including this measure in future censuses may ensure spatial microsimulation models better capture the geographic heterogeneity in important health outcomes at small area levels.

Finally, there needs to be greater communication between developers of spatial microsimulation models and public health practitioners who will ultimately use model outputs to inform local health promotion planning decisions. Modellers often have difficulty communicating results to decision makers (Alper & Geller, 2016; Maglio et al., 2014). Developers of spatial microsimulation models in particular need to collaborate with public health practitioners to better understand their needs for relevant small area information and how that information informs the decision making process. Additional research into all these aspects of spatial microsimulation modelling will enhance the utility of these models for public health practice. In so doing, spatial microsimulation models can provide public health practitioners with relevant spatial decision support tools that can be used to inform the delivery of chronic disease health promotion programs.

References

Ackermann, R. T., Finch, E. A., Brizendine, E., Zhou, H., & Marrero, D. G. (2008). Translating the Diabetes Prevention Program into the Community: The DEPLOY Pilot Study. *American Journal of Preventive Medicine*, *35*, 357–363. doi: 10.1016/j.amepre.2008.06.035

Ackermann, R. T., & Marrero, D. G. (2007). Adapting the Diabetes Prevention Program lifestyle intervention for delivery in the community. *Diabetes Educator*, *33*, 69–78. doi: 10.1177/0145721706297743

Agardh, E., Allebeck, P., Hallqvist, J., Moradi, T., & Sidorchuk, A. (2011). Type 2 diabetes incidence and socio-economic position: a systematic review and meta-analysis. *International Journal of Epidemiology*, *40*, 804–818. doi: 10.1093/ije/dyr029

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. New York: John Wiley and Sons.

Albright, A. L., & Gregg, E. W. (2013). Preventing type 2 diabetes in communities across the US. The National Diabetes Prevention Program. *American Journal of Preventive Medicine*, *44*, S346–S351. doi: 10.1016/j.amepre.2012.12.009

Ali, M. K., Echouffo-Tcheugui, J., & Williamson, D. F. (2012). How effective were lifestyle interventions in real-world settings that were modeled on the diabetes prevention program? *Health Affairs (Millwood)*, *31*(1), 67–75. doi: 10.1377/hlthaff.2011.1009

Allaire, D. L., & Willcox, K. E. (2012). A variance-based sensitivity index function for factor prioritization. *Reliability Engineering and System Safety*, *107*, 107–114. doi: 10.1016/j.ress.2011.08.007

Almeida, F. A., Shetterly, S., Smith-Rey, R. L., & Estabrooks, P. A. (2010). Reach and effectiveness of a weight loss intervention in patients with prediabetes in Colorado. *Preventing Chronic Disease*, *7*(5), A103. Retrieved from http://

www.cdc.gov/pcd/issues/2010/sep/09_0204.htm

Alper, J., & Geller, A. (2016). *How Modeling Can Inform Strategies to Improve Population Health: Workshop Summary.* Washington, D. C.: National Academies Press. doi: 10.17226/21807

Anderson, B. (2007). *Creating small-area income estimes: Spatial micosimulation modelling* (Tech. Rep. No. 07NRAD04639(a)). London, United Kingdom: Department of Communities and Local Government. Retrieved from http://www.communities.gov.uk/documents/communities/pdf/325286.pdf

Anderson, B. (2009). *Creating small area income deprivation estimates for Northern Ireland: Spatial microsimulation modelling* (CRESI Working Paper No. 7). Colchester: Centre for Research in Economic Sociology and Innovation, University of Essex.

Anderson, B. (2013). Estimating small-area income deprivation: An iterative proportional fitting approach. In R. Tanton & K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users* (Vol. 6, pp. 49–67). New York: Springer. doi: 10.1007/978-94-007-4623-7_4

Anselin, L., Lozano, N., & Kochinsky, J. (2006). *Rate transformations and smoothing* (Tech. Rep.). Urbana, IL: Spatial Analysis Laboratory, University of Illinois. Retrieved from http://geodacenter.asu.edu/pdf/smoothing_06.pdf

Armour, T. A., Norris, S. L., Jack, L., Zhang, X., & Fisher, L. (2005). The effectiveness of family interventions in people with diabetes mellitus: a systematic review. *Diabetic Medicine*, *22*, 1295–1305.

Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration.* Princeton, New Jersey: Princeton University Press.

Baio, G., & Dawid, A. P. (2011). Probabilistic sensitivity analysis in health economics. *Statistical Methods in Medical Research.* doi: 10.1177/0962280211419832

Ballas, D. (2004). Simulating trends in poverty and income inequality on the basis of 1991 and 2001 census data: A tale of two cities. *Area*, *36*, 146–163. doi: 10.1111/j.0004-0894.2004.00211.x

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B., & Rossiter, D. (2005b). *SimBritain:* A spatial microsimulation approach to population dynamics. *Population, Space and Place*, *11*, 13–34. doi: 10.1002/psp.351

Ballas, D., Clarke, G., Dorling, D., Rigby, J., & Wheeler, B. (2006). Using geographic information systems and spatial microsimulation for the analysis of health inequalities. *Health Informatics Journal*, *12*, 65–79. doi: 10.1177/1460458206061217

Ballas, D., Clarke, G., Dorling, D., & Rossiter, D. (2007a). Using SimBritain to model the geographical impact of national government policies. *Geographical Analysis*, *39*, 44–77. doi: 10.1111/j.1538-4632.2006.00695.x

Ballas, D., & Clarke, G. P. (2001). Modelling the local impacts of national social policies: a spatial microsimulation approach. *Environment and Planning C: Government and Policy*, *19*, 587–606. doi: 10.1068/c0003

Ballas, D., Clarke, G. P., & Wiemers, E. (2006). Spatial microsimulation for rural policy analysis in ireland: the implications of cap reforms for the national spatial strategy. *Journal of Rural Studies*, *22*, 367–378. doi: 10.1016/j.jrurstud.2006.01.002

Ballas, D., Kingston, R., & Stillwell, J. (2004a). Using a spatial microsimulation decision support system for policy scenario analysis. In J. P. Van Leeuwen & H. J. P. Timmermans (Eds.), *Recent Advances in Design & Decision Support Systems in Architecture and Urban Planning* (pp. 177–191). Dordrecht, Netherlands: Kluwer Academic Publishers.

Ballas, D., Kingston, R., Stillwell, J., & Jin, J. (2007b). Building a spatial microsimulation-based planning support system for local policy making. *Environment and Planning A*, *39*, 2482–2499. doi: 10.1068/a38441

Ballas, D., Rossiter, D., Thomas, B., Clarke, G., & Dorling, D. (2005a). *Geography matters: Simulating the local impacts of national social policies* (Tech. Rep.). York, U.K.: Joseph Rowntree Foundation. Retrieved from http://www.jrf.org.uk/sites/files/jrf/1859352669.pdf

Barclay, A. W., Petocz, P., McMillan-Price, J., Flood, V. M., Prvan, T., Mitchell, P., & Brand-Miller, J. C. (2008). Glycemic index, glycemic load, and chronic disease risk — an analysis of observational studies. *American Journal of Clinical Nutrition*, *87*, 627–637.

Barnett, J. R., Pearce, J., & Howes, P. (2006). "Help, educate, encourage?": geographical variations in the provision and utilisation of diabetes education in New Zealand. *Social Science & Medicine*, *63*, 1328–1343. doi: 10.1016/

j.socscimed.2006.03.044

Barnhart, H. X., Haber, M., & Song, J. (2002). Overall concordance correlation coefficient for evaluating agreement using multiple observers. *Biometrics*, *58*, 1020–1027. Retrieved from http://www.jstor.org/stable/3068546

Barte, J. C. M., ter Bogt, N. C. W., Bogers, R. P., Teixeira, P. J., Blissmer, B., Mori, T. A., & Bemelmans, W. J. E. (2010). Maintainance of weight loss after lifestyle interventions for overweight and obesity, a systematic review. *Obesity Reviews*, *11*, 889–906. doi: 10.1111/j.1467-789X.2010.00740.x

Barthelemy, J., & Suesse, T. (2016). mipfp: Multidimensional iterative proportional fitting and alternative models [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=mipfp (R package version 3.0-1)

Barton, H. (2009). Land use planning and health and well-being. *Land Use Policy*, *26S*, S115–S123. doi: 10.1016/j.landusepol.2009.09.008

Bassuk, S. S., & Manson, J. E. (2005). Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. *Journal of Applied Physiology*, *99*, 1193–1204. doi: 10.1152/japplphysiol.00160.2005

Batz, G. V., Geisberger, R., & Sanders, P. (2008). *Time dependent contraction heirarchies—basic algorithmic ideas* (Tehnical report). Karlsruhe, Germany: Faculty of Informatics, Universität Karlsruhe. Retrieved from http://arxiv.org/abs/0804.3947

Béland, Y. (2002). Canadian Community Health Survey—methodological overview. *Health Reports*, *13*, 9–4.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*, 289–300.

Birkin, M., & Clarke, G. (2012). The enhancement of spatial microsimulation models using geodemographics. *Annals of Regional Science*, *49*, 515–532. doi: 10.1007/s00168-011-0472-2

Birkin, M., & Clarke, G. P. (2009). Geodemographics. In R. Kitchin & N. Thrift (Eds.), *International Encyclopedia of Human Geography* (Vol. 12, pp. 382–389). New York: Elsevier.

Birkin, M., & Clarke, M. (1988). SYNTHESIS—a synthetic spatial information sys-

tem for urban and regional analysis: methods and examples. *Environment and Planning A, 20*, 1645–1671.

Birkin, M., & Clarke, M. (2011). Spatial microsimulation models: A review and a glimpse into the future. In J. Stillwell & M. Clarke (Eds.), *Population Dynamics and Projection Methods* (Vol. 4, pp. 193–208). New York: Springer. doi: 10.1007/978-90-481-8930-4_9

Bivand, R. (2015). spdep: Spatial dependence: weighting schemes, statistics and models [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=spdep (R package version 0.5-88)

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*, 135–160. doi: 10.1177/096228029900800204

Blocker, A. W. (2016). ipfp: Fast Implementation of the Iterative Proportional Fitting Procedure in C [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=ipfp (R package version 1.0.1)

Booth, G. L., Creatore, M. I., Gozdyra, P., & Glazier, R. H. (2007). Patterns of Diabetes Prevalence, Complications and Risk Factors. In R. H. Glazier, G. L. Booth, P. Gozdyra, M. I. Creatore, & M. Tynan (Eds.), *Neighbourhood Environments and Resources for Healthy Living—A Focus on Diabetes in Toronto* (pp. 17–34). Toronto, ON: Institute for Clinical Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/TDA_Chp2_press.pdf

Booth, G. L., Polsky, J. Y., Gozdyra, P., Cauch-Dudek, K., Kiran, T., Shah, B. R., ... Glazier, R. H. (2012). *Regional measures of diabetes burden in ontario* (Tech. Rep.). Toronto, ON: Institute for Clinical Evaluative Sciences.

Boulten, A. J. M., & Bowling, F. L. (2011). Diabetes and lower-extremity diseases. In K. M. V. Narayan, D. Williams, E. W. Gregg, & C. C. Cowie (Eds.), *Diabetes Public Health: From Data to Policy* (pp. 161–172). New York: Oxford University Press.

Briggs, K., A. Claxton, & Sculpher, M. (2006). Making decision models probabilistic. In A. Briggs, K. Claxton, & M. Sculpher (Eds.), *Decision modelling for health economic evaluation* (Vol. 1, pp. 77–120). New York: Oxford University Press.

Brownson, R. C., Fielding, J. E., & Maylahn, C. M. (2009). Evidence-based pub-

lic health: a fundamental concept for public health practice. *Annual Review of Public Health*, *30*, 175–201. doi: 10.1146/annurev.publhealth.031308.100134

Bryant, J. (2005). What can stochastic population projections contribute to policy analysis? *New Zealand Population Review*, *31*, 1–11. Retrieved from http://bit.ly/2d9MVRS

Burden, S., & Steel, D. (2015). Constraint choice for spatial microsimulation. *Population, Space and Place*. doi: 10.1002/psp.1942

Burr, J. F., Shephard, R. J., & Riddell, M. C. (2012). Prediabetes and type 2 diabetes mellitus. *Canadian Family Physician*, *58*, 280–284.

Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, *25*, 4279–4929. doi: 10.1002/sim.2673

Caballero, B. (2007). The global epidemic of obesity: an overview. *Epidemiologic Reviews*, *29*, 1–5. doi: 10.1093/epirev/mxm012

Campbell, M., & Ballas, D. (2013). A spatial microsimulation approach to economic policy analysis in Scotland. *Regional Science Policy & Practice*, *5*, 263–288. doi: 10.1111/rsp3.12009

Campbell, N. R. C., McAlister, F. A., & Quan, H. (2013). Monitoring and evaluating efforts to control hypertension in Canada: why, how, and what it tells us needs to be done about current care gaps. *Canadian Journal of Cardiology*, *29*, 564–570. doi: 10.1016/j.cjca.2012.05.006

Canadian Diabetes Association. (2008). Clinical practice guidelines for the prevention and management of diabetes in Canada. *Canadian Journal of Diabetes*, *32*, S1–S201.

Canadian Diabetes Association. (2015). *2015 Report on Diabetes: Driving Change* (Tech. Rep.). Toronto, ON: Canadian Diabetes Association. Retrieved from http://bit.ly/29ApAWu

Carstensen, B., Kristensen, J. K., Ottosen, P., & Borch-Johnsen, K. (2008). The Danish National Diabetes Register: Trends in incidence, prevalence and mortality. *Diabetologia*, *51*(12), 2187–2196. doi: 10.1007/s00125-008-1156-z

Carulli, L., Rondinella, S., Lombardini, S., Canedi, I., Loria, P., & Carulli, N. (2005). Review article: Diabetes, genetics and ethnicity. *Alimentary Pharmacology*

& Therapeutics, *22*(Suppl 2), 16–19. doi: 10.1111/j.1365-2036.2005.02588.x

Caspersen, C. J., Thomas, D., Boseman, L. A., Beckles, G. L. A., & Albright, A. L. (2012). Aging, diabetes and the public health system in the United States. *American Journal of Public Health*, *102*, 1482–1497. doi: 10.2015/AJPH.2011.300616

Cataife, G. (2014). Small area estimation of obesity prevalence and dietary patterns: a model applied to Rio de Janeiro city, Brazil. *Heatlh & Place*, *26*, 47–52. doi: 10.1016/j.healthplace.2013.12.004

Centers for Disease Control and Prevention. (2007). *Incidence of undiagnosed diabetes per 1,000 population aged 18–79 years, by age, United States, 1980–2010.* Retrieved from http://www.cdc.gov/diabetes/statistics/incidence/fig3.htm

Centers for Disease Control and Prevnention. (2013). *Diabetes Interactive Atlas: Diagnosed diabetes incidence age-adjusted rate per 1000.* Retrieved from http://www.cdc.gov/diabetes/atlas/countrydata/atlas.html

Cheng, A. Y. Y. (2013). Canadian diabetes association 2013 clinical practice guidelines for the prevention and management of diabetes in canada: introduction. *Canadian Journal of Diabetes*, *37*, S1–S3. doi: 10.1016/j.jcjd.2013.01.009

Chin, S. F., & Harding, A. (2006). *Regional dimensions: Creating synthetic small-area microdata and spatial microsimulation models* (Technical Paper No. 33). Canberra, Australia: National Centre for Social and Economic Modeling, University of Canberra. Retrieved from http://www.natsem.canberra.edu.au/storage/tp33.pdf

Chin, S. F., Harding, A., Lloyd, R., McNamara, J., Phillips, B., & Vu, Q. N. (2005). Spatial microsimulation using synthetic small-area estimates of income, tax and social security benefits. *Australasian Journal of Regional Studies*, *11*, 303–335.

Cigolle, C. T., Blaum, C. S., & Halter, J. B. (2009). Diabetes and cardiovascular disease prevention in older adults. *Clinical in Geriatric Medicine*, *25*, 607–641. doi: 10.1016/j.cger.2009.09.001

City of Toronto. (2013a). *Toronto Demographics.* Retrieved from http://bit.ly/29WLt3F

City of Toronto. (2013b). *Neighbourhood planning areas.* Retrieved from http://bit.ly/1f6AC5g

City of Toronto. (2013c). *Open Data: Open Government Licence—Toronto.* Retrieved from http://bit.ly/29G8Hs2

Clark, S. D., Birkin, M., & Heppenstall, A. (2014). Sub regional estimates of moribities in the english elderly population. *Health & Place*, *27*, 176–185. doi: 10.1016/j.healthplace.2014.02.010

Clarke, G. P. (1996). Microsimlation: An introduction. In G. P. Clarke (Ed.), *Microsimulation for Urban and Regional Policy Analysis* (Vol. 6, pp. 1–9). London, U.K.: Pion.

Clarke, G. P., & Madden, M. (Eds.). (2001). *Regional Science in Business.* Berlin: Springer.

Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., ... O'Hagan, T. (2005). Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics*, *14*, 339–347. doi: 10.1002/hec.985

Cohen, M. (1991). Variance estimation of microsimulation models through sample reuse. In C. F. Citro & E. A. Hanushek (Eds.), *Improving Information for Social Policy Decisions —The Use of Microsimulation Modeling* (Vol. II: Technical Papers, pp. 237–244). Washington, D. C.: National Academy Press.

Congdon, P. (2012). A model for spatially disaggregated trends and forecasts of diabetes prevalence. *Journal of Data Science*, *10*, 579–595. Retrieved from http://www.jds-online.com/file_download/368/JDS-1076.pdf

Cook, L. (2011). *Diabetes excellence in primary care CLHIN-DRCC Ontario Diabetes Strategy.* Retrieved July 13, 2013, from http://www.centraldrcc.com/CentralDRCC/images/Intro to Ontario Diabetes Strategy excellence day oct 15th.pdf

Costacou, T., & Mayer-Davis, E. J. (2003). Nutrition and prevention of type 2 diabetes. *Annual Review of Nutrition*, *23*, 147–170. doi: 10.1146/annurev.nutr.23.011702.073027

Creatore, M. I., Booth, G. L., Ross, K., Gozdyra, P., Tynan, A.-M., & Glazier, R. H. (2007c). Physical Activity and Diabetes. In R. H. Glazier, G. L. Booth, P. Gozdyra, M. I. Creatore, & M. Tynan (Eds.), *Neighbourhood Environments*

and Resources for Health Living—A Focus on Diabetes in Toronto (pp. 151–184). Toronto, ON: Institute for Clincal Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/TDA_Chp7_PartA_press.pdf

Creatore, M. I., Gozdyra, P., Booth, G. L., & Glazier, R. H. (2007a). Setting the Context. In R. H. Glazier, G. L. Booth, P. Gozdyra, M. I. Creatore, & M. Tynan (Eds.), *Neighbourhood Environments and Resources for Health Living—A Focus on Diabetes in Toronto* (pp. 1–16). Toronto, ON: Institute for Clincal Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/TDA_Chp1_press.pdf

Creatore, M. I., Gozdyra, P., Booth, G. L., Ross, K., & Glazier, R. H. (2007b). Socioeconomic Status and Diabetes. In R. H. Glazier, G. L. Booth, P. Gozdyra, M. I. Creatore, & M. Tynan (Eds.), *Neighbourhood Environments and Resources for Health Living—A Focus on Diabetes in Toronto* (pp. 46–56). Toronto, ON: Institute for Clincal Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/TDA_Chp3_Part2_press.pdf

Cromley, E. K., & McLafferty, S. L. (2012). *GIS and Public Health* (2nd ed.). New York: Guilford Press.

Danaei, G., Finucane, M. M., Lu, Y., Singh, G. M., Cowan, M. J., Paciorek, C. J., … Ezzati, M. (2011). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. *Lancet*, *378*, 31–40. doi: 10.1016/S0140-6736(11)60679-X

Davies, M. J., Tringham, J. R., Troughton, J., & Khunti, K. K. (2004). Prevention of type 2 diabetes mellitus. A review of the evidence and its application in a UK setting. *Diabetic Medicine*, *21*, 403–414. doi: 10.1111/j.1464-5491.2004.01176.x

Davis, T. M. E. (2008). Ethnic diversity in type 2 diabetes. *Diabetic Medicine*, *25*(Suppl 2), 52–56. doi: 10.1111/j.1464-5491.2008.02499.x

Delamater, P. L. (2013). Spatial accessibility in suboptimally configured health care systems: a modified two-step floating catchment area (M2SFCA) metric. *Health & Place*, *24*, 30–43. doi: 10.1016/j.healthplace.2013.07.012

Edwards, K. L., & Clarke, G. (2013b). SimObesity: combinatorial optimisation

(deterministic) model. In R. Tanton & K. L. Edwards (Eds.), *Spatial Micosimulation: A Reference Guide for Users* (Vol. 6, pp. 69–85). New York: Springer. doi: 10.1007/978-94-007-4623-7_5

Edwards, K. L., & Clarke, G. P. (2009). The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. *Social Science and Medicine, 69*, 1127–1134. doi: 10.1016/j.socscimed.2009.07.037

Edwards, K. L., Clarke, G. P., Thomas, J., & Forman, D. (2011). Internal and external validation of spatial microsimulation models: Small area estimates of adult obesity. *Applied Spatial Analysis, 4*, 281–300. doi: 10.1007/s12061-010-9056-2

Edwards, K. L., & Tanton, R. (2013a). Validation of spatial microsimulation models. In R. Tanton & K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users* (Vol. 6, pp. 249–258). New York: Springer. doi: 10.1007/978-94-007-4623-7_15

Egger, G., & Swinburn, B. (1997). An "ecological" approach to the obesity pandemic. *British Medical Journal, 315*, 477–80.

Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica, 71*, 355–364.

Fairhurst, R. (2014). *How the routing OSRM algorithm works.* Retrieved from https://help.openstreetmap.org/questions/30272/how-the-routing-osrm-algorithm-works

Farrell, N., O'Donoghue, C., & Morrissey, K. (2010). *Spatial microsimulation using quota sampling* (Teagasc Rural Economy Resaerch Series Working Paper No. 7). Carlow, Ireland: Teagasc Rural Economy Research Centre.

Ferber, R., Sheatsley, P., Turner, A., & Waksberg, J. (1980). *What is a survey?* (Tech. Rep.). Washington, DC: American Statistical Association. Retrieved from https://www.whatisasurvey.info/downloads/pamphlet_1980.pdf

Feskens, E. J., Virtanen, S. M., Rasanen, L., Tuomilehto, J., Stengøard, J., Pekkanen, J., ... Kromhout, D. (1995). Dietary factors determining diabetes and impaired glucose tolerance: a 20-year follow-up of the Finnish and Dutch cohorts of the Seven Countries Study. *Diabetes Care, 18*, 1104–1112. doi: 10.2337/diacare.18.8.1104

Finucane, M. M., Stevens, G. A., Cowan, M. J., Danaei, G., Lin, J. K., Paciorek, C. J., ... Ezzati, M. (2011). National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet*, *377*(9765), 557–567. doi: 10.1016/S0140-6736(10)62037-5

Folsom, A. R., Kushi, L. H., & Hong, C. P. (2000). Physical activity and incident diabetes mellitues in postmenopausal women. *American Journal of Public Health*, *90*, 134–138.

Garson, D. G. (2009). Computerized simulation in the social sciences: A survey and evaluation. *Simulation & Gaming*, *40*, 267–280. doi: 10.1177/1046878108322225

Ghosh, M., & Rao, J. N. K. (1994). Small area estimation: An appraisal. *Statistical Sciences*, *9*, 55–93.

Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist* (2nd ed.). Berkshire, England: Open University Press.

Gillum, R. F., & Sempos, C. T. (2005). Ethnic variation in validity of classification of overweight and obesity using self-reported weight and height in American women and men: the Third National Health and Nutrition Examination Survey. *Nutrition Journal*, *4*. doi: 10.1186/1475-2891-4-27

Glasgow, R. E., Lichtenstein, E., & Marcus, A. C. (2003). Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. *American Journal of Public Health*, *93*(8), 1261–1267. doi: 10.2105/AJPH.93.8.1261

Glazier, R. H., Weyman, J., Creatore, M. I., Ross, K., Gozdyra, P., & Booth, G. L. (2007). Community-Based Health Services and Diabetes. In editors (Ed.), *Neighbourhood Environments and Resources for Health Living—A Focus on Diabetes in Toronto* (pp. 243–270). Toronto, ON: Institute for Clinical Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/TDA_Chp11_PartA_press.pdf

Goldberg, P., Guegen, A., Schmaus, A., Nakache, J. P., & Goldberg, M. (2001). Longitudinal study of associations between perceived health status and self reported diseases in the French Gazel cohort. *Journal of Epidemiology and Community Health*, *55*, 233–238. doi: 10.1136/jech.55.4.233

Gong, H. C., McNamara, J., Vidyattama, Y., Miranti, R., Tanton, R., Harding, A., & Kendig, H. (2012). Developing spatial microsimulation estimates of small area advantage and disadvantage among older Australians. *Population, Space and Place*, *18*, 551–565. doi: 10.1002/psp.692

Grant, J. (2002). Mixed use in theory and practice: Canadian experience with implementing a planning principle. *Journal of the American Planning Association*, *68*, 71–84. doi: 10.1080/01944360208977192

Griffith, D. A. (1996). Some guidelines for specifying the geographic weights matrix contained in spatial statistical models. In S. L. Arlinghaus (Ed.), *Practical Handbook of Spatial Statistics* (pp. 65–82). Boca Raton, FL: CRC Press, Taylor & Francis.

Hamman, R. F., Wing, R. R., Edelstein, S. L., Lachin, G. A., J. M. Bray, & Delahanty, L. e. a. (2006). Effect of weight loss with lifestyle intervention on risk of diabetes. *Diabetes Care*, *29*, 2102–2107.

Harland, K., Heppenstall, A., Smith, D., & Birkin, M. (2012). Creating realistic synthetic populations at varying spatial scales: Comparative critique of population synthesis techniques. *Journal of Artificial Societies and Social Simulation*, *15*, 1. Retrieved from http://jasss.soc.surrey.ac.uk/15/1/1.html

Harrison, J. R., Lin, Z., Carroll, G. R., & Carley, K. M. (2007). Simulation modeling in organizational and management research. *Academy of Management Review*, *32*, 1229–1245.

Haslett, S., Jones, G., Noble, A., & Ballas, D. (2010). *More for less? Using statistical modelling to combine existing data sources to produce sounder, more detailed and less expensive Official Statistics* (Official Statistics Research Series No. 6). Wellington, New Zealand: Statistics New Zealand. Retrieved from http://www.statisphere.govt.nz/further-resources-and-info/official-statistics-research/series/2010/~/media/statisphere/Files/official-statistics-research-series/Volume2010/More for Less/More for less.pdf

Hermes, K., & Poulsen, M. (2012a). A review of current methods to generate synthetic spatial microdata using reweighting and future directions. *Computers, Environment and Urban Systems*, *36*, 281–290. doi: 10.1016/j.compenvurbsys.2012.03.005

Hermes, K., & Poulsen, M. (2012b). Small area estimates of smoking prevalence

in London. *Health & Place*, *18*, 630–638. doi: 10.1016/j.healthplace.2011.12.010

Hermes, K., & Poulsen, M. (2013). The intraurban geography of generalised trust in sydney. *Evironment and Planning A*, *45*, 276–294. doi: 10.1068/a44663

Hoare, A., Regan, D. G., & Wilson, D. P. (2008). Sampling and sensitivity analyses tools (SaSAT) for computational modelling. *BMC Theoretical Biology and Medical Modelling*, *5*, 4. doi: 10.1186/1742-4682-5-4

Holden, S. H., Barnett, A. H., Peters, J. R., Jenkins-Jones, S., Poole, C. D., Morgan, C. L., & Currie, C. J. (2013). The incidence of type 2 diabetes in the United Kingdom from 1991 to 2010. *Diabetes, Obesity and Metabolism*. doi: 10.1111/dom.12123

Hu, E. A., Pan, A., Malik, V., & Sun, Q. (2012). White rice consumption and risk of type 2 diabetes: meta-analysis and systematic review. *British Medical Journal*, *344*, e1454. doi: 10.1136/bmj.e1454

Hu, F. B., Sigal, R. J., Rich-Edwards, J. W., Colditz, G. A., Solomon, C. G., Willet, W. C., ... Manson, J. E. (1999). Walking compared with vigorous physical activity and risk of type 2 diabetes in women: a prospective study. *Journal of the American Medical Association*, *282*, 1433–1439.

Hu, F. B., van Dam, R. M., & Liu, S. (2001). Diet and risk of type ii diabetes: the role of types of fat and carbohydrate. *Diabetologia*, *44*, 805–817.

Huang, Z., & Williamson, P. (2001). *A comparison of synthetic reconstruction and combinatorial optimization approaches to the creation of small-aea microdata* (Working Paper No. 2001/2). Liverpool, United Kingdom: Department of Geography, University of Liverpool. Retrieved from http://pcwww.liv.ac.uk/~william/microdata/workingpapers/hw_wp_2001_2.pdf

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

Hux, J. E., & Tang, M. (2003). Patterns of prevalence and incidence of diabetes. In J. E. Hux, G. L. Booth, P. M. Slaughter, & A. Laupacis (Eds.), *Diabetes In Ontario: An ICES Practice Atlas* (pp. 1–18). Toronto, ON: Institute for Clinical Evaluative Sciences. Retrieved from http://www.ices.on.ca/file/DM_Chapter1.pdf

Hynes, S., Farrelly, N., Murphy, E., & O'Donoghue, C. (2008). Modelling habitat conservation and participation in agri-environmental schemes: A spatial microsimulation approach. *Ecological Economics*, *66*, 258–269. doi: 10.1016/j.ecolecon.2008.02.006

Hynes, S., Hanley, N., & O'Donoghue, C. (2010). A combinatorial optimization approach to nonmarket environmental benefit aggregation via simulated populations. *Land Economics*, *86*, 345–362.

Hynes, S., Morrissey, K., O'Donoghue, C., & Clarke, G. (2009). A spatial microsimulation analysis of methane emissions from Irish agriculture. *Ecological Complexity*, *6*, 135–146. doi: 10.1016/j.ecocom.2008.10.014

Jackson, L. (2009). Translating the Diabetes Prevention Program into practice. *The Diabetes Educator*, *35*(2), 309–320. doi: 10.1177/0145721708330153

James, R., Young, T. K., Mustard, C., & Blanchard, J. (1997). The health of Canadians with diabetes. *Health Reports*, *9*, 47–52.

James, W. P. T. (1997). The epidemiology of obesity: the size of the problem. *Journal of Internal Medicine*, *263*, 336–342.

Jones, M. C. (2002). Student's simplest distribution. *The Statistician*, *51*, 41–49.

Jones, P. M., Lovelace, R., & Dumont, M. (2016). rakeR: Easy Spatial Microsimulation (Raking) in R [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=rakeR (R package version 0.1.1)

Joseph, A. E., & Bantock, P. R. (1982). Measuring potential accessibility to general practitioners in rural ares: a method and case study. *Social Science and Medicine*, *16*, 85–90.

Joseph, A. E., & Phillips, D. R. (1984). *Accessibility and utilization—geographical perspectives on health care delivery*. New York: Harper and Row.

Joseph, J., Svartberg, J., Njølstad, I., & Schrimer, H. (2010). Incidence of and risk factors for type-2 diabetes in a general population: The Tromsø Study. *Scandinavian Journal of Public Health*, *38*, 768–775. doi: 10.1177/1403494810380299

Kaati, G., Sjostrom, M., & Vester, M. (2004). The quality and use of knowledge in health policy-making: A case study. *Critical Public Health*, *14*(3), 225–237. doi: 10.1080/09581590400004261

Kahn, S. E. (2003). The relative contributions of insulin resistance and beta-cell

dysfunction to the pathophysiology of Type 2 diabetes. *Diabetologia*, *46*, 3–19. doi: 10.1007/s00125-002-1009-0

Kane, L. (2014). *Toronto's mixed-use boom coming to a 'hood near you.* The Toronto Star, February 3, 2014. Retrieved from www.thestar.com

Katula, J. A., Vitolins, M. Z., Rosenberger, E. L., Blackwell, C. S., Morgan, T. M., Lawlor, M. S., & Goff, D. C. (2011). One-Year Results of a community-cased translation of the Diabetes Prevention Program: Healthy-Living Partnerships to Prevent Diabetes (HELP PD) Project. *Diabetes Care*, *34*(7), 1451–1457. doi: 10.2337/dc10-2115

Kavroudakis, D. (2013b). *Package 'sms'.* http://cran.r-project.org/web/packages/sms/sms.pdf.

Kavroudakis, D. (2015). sms: An R Package for the Construction of Microdata for Geographical Analysis. *Journal of Statistical Software*, *68*(1), 1–23. doi: 10.18637/jss.v068.i02

Kavroudakis, D., Ballas, D., & Birkin, D. (2013a). Using spatial microsimulation to model social and spatial inequalities in educational attainment. *Applied Spatial Analysis and Policy*, *6*, 1–23. doi: 10.1007/s12061-012-9075-2

Klein, R., Saaddine, J. B., & Klein, B. E. K. (2011). Diabetes and vision. In K. M. V. Narayan, D. Williams, E. W. Gregg, & C. C. Cowie (Eds.), *Diabetes Public Health: From Data to Policy* (pp. 111–134). New York: Oxford University Press.

Klevmarken, A. (2008). Dynamic microsimulation for policy analysis: Problems and solutions. In A. Klevmarken & B. Lindgren (Eds.), *Simulating an Ageing Population: A Microsimulation Approach Applied to Sweden* (Vol. 285, pp. 31–53). Bingley, United Kingdom: Emerald Group Publishing. doi: 10.1016/S0573-8555(07)00002-8

Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., & Nathan, D. M. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, *346*, 393–403.

Koh, K., Gardy, S. C., & Vojnovic, I. (2015). Using simulated data to investicate the spatial patterns of obesity prevalence at the census tract level in metropolitan detroit. *Applied Geography*, *62*, 19–28. doi: 10.1016/j.apgeog.2015.03

.016

Kopec, J., Sayre, E., Flanagan, W., Fines, P., Cibere, J., Rahman, M. M., ... Badley, E. M. (2010a). Development of a population-based microsimulation model of osteoarthritis in Canada. *Osteoarthritis and Cartilage*, *18*, 303–311. doi: 10.1016/j.joca.2009.10.010

Kopec, J. A., Finès, P., Manuel, D. A., Buckeridge, D. L., Flanagan, W. M., Oderkirk, J., ... Wolfson, M. C. (2010b). Validation of population-based disease simulation models: A review of concepts and methods. *BMC Public Health*, *10*, 710. Retrieved from http://www.biomedcentral.com/1471-2458/10/710

Kosaka, K., Noda, M., & Kuzuya, T. (2005). Prevention of type 2 diabetes by lifestyle intervention: a Japanese trial in IGT males. *Diabetes Research and Clinical Practice*, *67*, 152–162. doi: 10.1016/j.diabres.2004.06.010

Kosar, B., & Tomintz, M. (2014). simSalud: A web-based spatial microsimulation to model the health status for small areas using the example of smokers in Austria. *GI Forum*, *2*, 207–216. doi: 10.1553/giscience2014s207

Kumari, M., Head, J., & Marmot, M. (2004). Prospective study of social and other risk factors for incidence of type 2 diabetes in the Whitehall II study. *Archives of Internal Medicine*, *164*, 1873–1880.

Kwan, M.-P. (1998). Space-time and integral measures of individual accessibility: A comparative analysis using a point-based framework. *Geographical Analysis*, *30*(3), 191–216.

Latham, K., & Peek, C. W. (2013). Self-rated health and morbidity onset among late midlife U.S. adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*, 107–116. doi: 10.1093/geronb/gbs104

Lee, R. C. (1991). Current approaches to shortage area designation. *Journal of Rural Health*, *7*, 437–450.

Leyk, S., Buttenfield, B. P., & Nagle, N. N. (2013). Modeling ambiguity in census microdata allocations to improve demographic small area estimates. *Transactions in GIS*, *17*, 406–424. doi: 10.1111/j.1467-9671.2012.01366.x

Li, W., Kelsey, J. L., Zhang, Z., Lemon, S. C., Mezgebu, S., Boddie-Willis, C., & Reed, G. W. (2009). Small-area estimation and prioritizing communities for obesity control in Massachusetts. *American Journal of Public Health*, *99*, 511–519. doi: 10.2105/AJPH.2008.137364

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*, 255–268. Retrieved from http://www.jstor.org/stable/2532051

Lin, L. (2000). A note on the concordance correlation coefficient. *Biometrics*, *56*, 324–325.

Lind, M., Garcia-Rodriguez, L. A., Booth, G. L., Cea-Soriano, L., Shah, B. R., Ekeroth, G., & Lipscombe, L. L. (2013). Mortality trends in patients with and without diabetes in Ontario, Canada and the UK from 1996 to 2009: A population-based study. *Diabetologia*. doi: 10.1007/s00125-013-2949-2

Lipscombe, L. L., & Hux, J. E. (2007). Trends in diabetes prevalence, incidence, and mortality in Ontario, Canada 1995-2005: A population-based study. *Lancet*, *369*(9563), 750–756. doi: 10.1016/S0140-6736(07)60361-4

Littman, A. J., Boyko, E. J., McDonell, M. B., & Fihn, S. D. (2012). Evaluation of a weight management program for veterans. *Preventing Chronic Disease*, *9*, E99. doi: 10.5888/pcd9.110267

Lovelace, R., & Ballas, D. (2013). 'Truncte, replicate, sample': a method for creating integer weights for spatial microsimulation. *Computers, Environment and Urban Systems*, *41*, 1–11. doi: 10.1016/j.compenvurbsys.2013.03.004

Luo, W., & Qi, Y. (2009). An enhanced two-step floating catchment area (e2sfca) method for measuring spatial accesibility to primary care physicians. *Health & Place*, *15*, 1100-1107.

Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design*, *30*, 865–884.

Luxen, D., Firebaugh, J., & Niklaus, P. (2014). *Open source routing machine (osrm) wiki*. Retrieved from https://github.com/Project-OSRM/osrm-backend/wiki

Lymer, S., Brown, L., Harding, A., & Yap, M. (2009). Predicting the need for aged care services at the small area level:The CAREMOD Spatial Microsimulation Model. *International Journal of Microsimulation*, *2*, 27–42. Retrieved from http://www.microsimulation.org/ijm/V2_2/IJM_2_2_3.pdf

Lymer, S., Brown, L., Yap, M., & Harding, A. (2008). 2001 regional disability estimates for New South Wales, Australia, using spatial microsimulation. *Ap-*

plied Spatial Analysis and Policy, *1*, 99–116. doi: 10.1007/s12061-008-9006-4

Lynch, J., Helmrick, S. P., Lakka, T. A., Kaplan, G. A., Cohen, R. D., Salonen, R., & Salonen, J. T. (1996). Moderately intense physical activities and high levels of cardiorespiratory fitness reduce teh risk of non-insulin-dependent diabetes mellitus in middle-aged men. *Archives of Internal Medicine*, *156*, 1307–1314.

Ma, J., Heppenstall, A., Harland, K., & Mitchell, G. (2014). Synthesising carbon emission for mega-cities: A static spatial microsimulation of transport CO_2 from urban travel in Beijing. *Computers, Environment and Urban Systems*, *45*, 78–88. doi: 10.1016/j.compenvurbsys.2014.02.006

Maglio, P. P., Sepulveda, M. J., & Mabry, P. L. (2014). Mainstreaming modeling and simulation to accelerate public health innovation. *American Journal of Public Health*, *104*, 1181–1186. doi: 10.2105/AJPH.2014.301873

Malik, V. S., Popkin, B. M., Bray, G. A., Després, J. P., & Hu, F. B. (2010). Sugar-sweetened beverages, obesity, type 2 diabetes mellitus and cardiovasclar disease risk. *Circulation*, *121*, 1356–1364. doi: 10.1161/CIRCULATIONAHA.109.876185

Manuel, D. G., Rosella, L. C., Tuna, M., Bennett, C., & Stukel, T. A. (2013). Effectiveness of community-wide and individual high-risk strategies to prevent diabetes: a modelling study. *PLoS ONE*, *8*(1), e52963. doi: 10.1371/journal.pone.0052963

Marino, S., Hogue, I. B., Ray, C. J., & Kirschner, D. E. (2008). A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, *254*, 178–196. doi: 10.1016/j.jtbi.2008.04.011

McGrail, M. R., & Humphreys, J. S. (2009). Measuring spatial accessibility to primary care in rural areas: Improving the effectiveness of the two-step floating catchment area method. *Applied Geography*, *29*, 533–541. doi: 10.1016/j.apgeog.2008.12.003

McNamara, J., Cassells, R., Wicks, P., & Vidyattama, Y. (2010). Children in housing disadvantage in Australia: development of a summary small area index. *Housing Studies*, *5*, 625–646. doi: 10.1080/02673037.2010.483583

Mendelsohn, M., & Brent, J. (2001). Understanding polling methodology. *Canadian Journal of Policy Research*, *2*, 131–136.

Millar, W. J., & Young, T. K. (2003). Tracking diabetes: prevalence, incidence and risk factors. *Health Reports*, *14*, 35–47.

Miller, P. (2010). *The Smart Swarm: how understanding flocks, schools and colonies can make us better at communicating, decision making and getting things done*. New York: The Penguin Group.

Ministry of Health and Long Term Care. (2012). *Ontario Diabetes Strategy*. Retrieved July 7, 2013, from http://news.ontario.ca/mohltc/en/2012/11/ontario-diabetes-strategy-1.html

Miranti, R., McNamara, J., Tanton, R., & Harding, A. (2011). Poverty at the local level: National and small area poverty estimates by family type for Australia in 2006. *Applied Spatial Analysis*, *4*, 145–171. doi: 10.1007/s12061-010-9049-1

Morrissey, K., Ballas, D., Clarke, G., Hynes, S., & O'Donoghue, C. (2013b). Spatial access to health services. In C. O'Donoghue, D. Ballas, G. Clarke, S. Hynes, & K. Morrissey (Eds.), *Spatial Microsimulation for Rural Policy Analysis* (edition ed., pp. 213–230). New York: Springer.

Morrissey, K., Clarke, G., Ballas, D., Hynes, S., & O'Donoghue, C. (2008). Examining access to GP services in rural Ireland using microsimulation analysis. *Area*, *40*, 354–364. doi: 10.1111/j.1475-4762.2008.00844.x

Morrissey, K., Clarke, G., & O'Donoghue, C. (2013a). Linking static spatial microsimulation modelling to meso-scale models: the relationship between access to GP services and long-term illness. In R. Tanton & K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users* (Vol. 6, pp. 127–143). New York: Springer. doi: 10.1007/978-94-007-4623-7_8

Morrissey, K., Hynes, S., Clarke, G., & O'Donoghue, C. (2010). Examining the factors associated with depression at the small area level in Ireland using spatial microsimulation techniques. *Irish Geography*, *43*, 1–22. doi: 10.1080/00750771003696489

Morrissey, K., & O'Donoghue, C. (2011). The spatial distribution of labour force participation and market earning at the sub-national level in Ireland. *Review of Economic Analysis*, *3*, 80–101.

Morrissey, K., O'Donoghue, C., & Farrell, N. (2014). The local impact of the marine sector in Ireland: a spatial microsimulation analysis. *Spatial Economic*

Analysis. doi: 10.1080/17421772.2013.835439

Nadeau, C., Wong, S. L., Flanagan, W. M., Oderkirk, J., Manuel, D., Wall, R., & Tremblay, M. S. (2013). Development of a population-based microsimulation model of physical activity in Canada. *Health Reports*, *24*, 11–19.

National Diabetes Information Clearinghouse (NDIC). (2011). *National diabetes statistics, 2011*. http://diabetes.niddk.nih.gov/dm/pubs/statistics/.

Ng, E., Dasgupta, K., & Johnson, J. A. (2008). An algorithm to differentiate diabetic respondents in the Canadian Community Health Survey. *Health Reports*, *19*, 71–79.

Norman, P. (1999). *Putting iterative proportional fitting on the researcher's desk*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.17.2073

O'Brien, J. A., Patrick, A. R., & Caro, J. J. (2003). Cost of managing complications resulting from type 2 diabetes mellitus in Canada. *BMC Health Services Research*, *3*(1), 7. doi: 10.1186/1472-6963-3-7

O'Donoghue, C., Farell, N., Morrissey, K., Lennon, J., Ballas, D., Clarke, G., & Hynes, S. (2013). The SMILE model: construction and calibration. In C. O'Donoghue, D. Ballas, G. Clarke, S. Hynes, & K. Morrissey (Eds.), *Spatial Microsimulation for Rural Policy Analysis* (pp. 55–86). New York: Springer. doi: 10.1007/978-3-642-30026-4_4

Office of the Auditor General of Ontario. (2012). *Annual report of the Auditor General of Ontario* (Tech. Rep.). Toronto, ON: Office of the Auditor General of Ontario.

O'Hagan, A., Stevenson, M., & Madan, J. (2007). Monte Carlo probabilistic sensitivity analysis for patient level simulation models: efficient estimation of mean and variance using ANOVA. *Health Economics*, *16*, 1009–1023. doi: 10.1002/hec.1199

Openshaw, S., & Taylor, P. J. (1979). A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In N. Wriggley (Ed.), *Statistical Methods in the Spatial Sciences* (pp. 127–144). London, UK: Pion.

OpenStreetMap Wiki. (2014). *Main Page — OpenStreetMap Wiki*. Retrieved from http://wiki.openstreetmap.org/w/index.php?title=Main_Page&oldid=1060762 (Online; accessed 20-December-2014)

Orcutt, G. H., Caldwell, S., & Wertheimer II, R. (1976). *Policy exploration through microanalytic simulation*. Washington, D.C.: The Urban Institute.

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis*, *27*, 286–306.

Pan, X.-R., Cao, H.-B., Li, G.-W., Liu, P.-A., Hu, Y.-H., Jiang, X.-G., ... Howard, B. V. (1997). Effects of diet and exercies in preventing NIDDM in people with impaired glucose tolerance. *Diabetes Care*, *20*, 537–544.

Paulweber, B., Valensi, P., Lindström, J., Lalic, N. M., Greaves, C. J., McKee, M., ... Roden, M. (2010). A European evidence-based guideline for the prevention of type 2 diabetes. *Hormone and Metabolic Research*, *42*, S3–S36. doi: 10 .1055/s-0029-1240928

Pfeffermann, D. (2002). Small area estimates—new developments and directions. *International Statistical Review*, *70*, 125–143.

Porta, M. (2008). *Dictionary of Epidemiology* (5th ed.). Toronto: Oxford University Press.

Procter, K., Clarke, G., & Ransley, J. (2008). Micro-level analysis of childhood obesity, diet, physical activity, residential socioeconomic and social capital variable: Where are the obesogenic environments in Leeds? *Area*, *40*, 323–340. doi: 10.1111/j.1475-4762.2008.00822.x

Public Health Agency of Canada. (2011). *Diabetes in Canada: Facts and figures from a public health perspective*. Ottawa, Canada: Public Health Agency of Canada. Retrieved from http://www.phac-aspc.gc.ca/cd-mc/diabetes -diabete/index-eng.php

Public Health Agency of Canada and the Canadian Institute for Health Information. (2011). *Obesity in Canada: A joint report from the Public Health Agency of Canada and the Canadian Institute for Health Information* (edition ed., Vol. volume). Ottawa, Canada: Public Health Agency of Canada. Retrieved from http://www.phac-aspc.gc.ca/hp-ps/hl-mvs/oic-oac/index -eng.php#toc

QGIS Development Team. (2009). QGIS Geographic Information System [Computer software manual]. Retrieved from http://qgis.osgeo.org

Rahman, A. (2011). *Small area housing stress estimation in australia: Microsimulation modelling and statistical reliability* (Unpublished doctoral disserta-

tion). University of Canberra.

Rahman, A., Harding, A., Tanton, R., & Liu, S. (2010). Methodological issues in spatial microsimulation modelling for small area estimation. *International Journal of Microsimulation*, *3*, 3–22. Retrieved from http://www.microsimulation.org/ijm/V3_2/Volume 3 Issue 2/1_IJM_47 Proof.pdf

Rahman, A., Harding, A., Tanton, R., & Liu, S. (2013). Simulating the characteristics of popluations at the small area level: new validation techniques for a spatial microsimulation model in Australia. *Computational Statistics and Data Analysis*, *57*, 149–165. doi: 10.1016/j.csda.2012.06.018

Ramachandran, A., & Snehalatha, C. (2011). Diabetes prevention programs. *Medical Clinics of North America*, *95*, 353–372. doi: 10.1016/j.mcna.2010.11.006

Ramachandran, A., Snehalatha, C., Mary, S., Mukesh, B., Bhaskar, A. D., & Vijay, V. (2006). The Indian Diabetes Prevention Programme shows that lifesyle modification and metformin prevent type 2 daibetes in Asian Indian subject with impaired glucose tolerance (IDPP-1). *Diabetologia*, *49*, 289–297. doi: 10.1007/s00125-005-0097-z

Rao, J. N. K. (1999). Some recent advances in model-based small area estimation. *Survey Methodology*, *25*, 175–186.

Raphael, D., Anstice, S., Raine, K., McGannon, K. R., Rizvi, S. K., & Yu, V. (2003). The social determinants of the incidence and management of type 2 diabetes mellitus: Are we prepared to rethink our questions and redirect our research activities? *Leadership in Health Services*, *16*, 10–20. doi: 10.1108/13660750310486730

Rausand, M. (2011). *Risk assessment: Theory, methods, and applications*. Hoboken, NJ: John Wiley & Sons.

Rey, S. J., & Anselin, L. (2010). Pysal: A python library of spatial analytical methods. In M. M. Fisher & A. Getis (Eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications* (pp. 175–193). New York: Springer.

Ripsin, C. M., Kang, H., & Urban, R. J. (2009). Management of blood glucose in type 2 diabetes mellitus. *American Family Physician*, *79*, 29–36.

Riva, M., & Smith, D. M. (2012). Generating small-area prevalence of psychological distress and alcohol consumption: Validation of a spatial microsimulation method. *Social Psychiatry and Psychiatric Epidemiology*, *47*, 745–755. doi:

10.1007/s00127-011-0376-6

Rodrigue, J.-P., Comtois, C., & Slack, B. (2006). *The goegraphy of transport system*. New York: Routledge, Taylor & Francsis Group.

Rosella, L. C., Lebenbaum, M., Fitzpatrick, T., Zuk, A., & Booth, G. L. (2015). Prevalence of prediabetes and undiagnosed diabetes in canada (2007–2011) according to fasting plama glucose and HbA1c screnning criteria. *Diabetes Care*, *38*, 1299–1305. doi: 10.2337/dc14-2474

Rosella, L. C., Lebenbaum, M., Li, Y., Wang, J., & Manuel, D. G. (2014). Risk distribution and its influence on the population targets for diabetes prevention. *Preventive Medicine*, *58*, 17–21. doi: 10.1016/j.ypmed.2013.10.007

Rosella, L. C., Manuel, D. G., Burchill, C., & Stukel, T. A. (2011). A population-based risk algorithm for the development of diabetes: development and validation of the Diabetes Population Risk Tool (DPoRT). *Journal of Epidemiology and Community Health*, *65*, 613–620. doi: 10.1136/jech.2009.102244

Ross, C. M. (1991). DYNASIM2 and PRISM: Examples of dynamic modeling. In C. F. Citro & E. A. Hanushek (Eds.), *Improving Information for Social Policy Decisions —The Uses of Microsimulation Modeling* (Vol. II, pp. 121–138). Washington, D.C.: National Academy Press. Retrieved from http://www.nap.edu/openbook.php?record_id=1853

Rutter, C. M., Zaslavsky, A. M., & Feuer, E. J. (2011). Dynamic microsimulation models for health outcomes: A review. *Medical Decision Making*, *31*, 10–18. doi: 10.1177/0272989X10369005

Ryan, J., Maoh, H., & Kanaroglou, P. (2009). Population synthesis: Comparing the major techniques using a small, complete population of firms. *Geographical Analysis*, *41*, 181–203. doi: 10.1111/j.1538-4632.2009.00750.x

Sabas, S. (2014). *OSRMdistance*. Retrieved from https://github.com/sabas/OSRMdistance

Sainani, K. L. (2014). Explanatory versus predictive modeling. *Physical Medicine & Rehabilitation (PM&R)*, *6*, 841–844. doi: 10.1016/j.pmrj.2014.08.941

Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis*, *22*, 579–590.

Saltelli, A., Tarantola, S., & Campolongo, F. (2000). Sensitivity analysis as an ingredient of modeling. *Statistical Science*, *15*, 377–395.

Sanders, P., Schultes, D., & Delling, D. (2008). Contraction hierarchies: Faster and simpler hierarchical routing in road networks. In C. C. McGeoch (Ed.), *Lecture Notes in Computer Science* (Vol. 5038, pp. 319–333). Heidelberg, Germany: Springer.

Schulze, M. B., Manson, J. E., Ludwing, D. S., Colditz, G. A., Stampfer, M. J., Willet, W. C., & Hu, F. B. (2004). Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women. *Journal of the American Medical Association*, *292*, 927–934.

Schuurman, N., Bérubé, M., & Crooks, V. A. (2010). Measuring potential spatial access to primary health care physicians using a modified gravity model. *The Canadian Geographer*, *54*, 29–45. doi: 10.1111/j.1541-0064.2009.00301.x

Seidell, J. C. (2000). Obesity, insulin resistance and diabetes — a worldwide epidemic. *British Journal of Nutrition*, *83*(Suppl 1), S5–S8.

Shai, I., Jiang, R., Manson, J. E., Stampfer, M. J., Willet, W. C., Colditz, G. A., & Hu, F. B. (2006). Ethnicity, obesity, and risk of type 2 diabetes in women: a 20-year follow-up study. *Diabetes Care*, *29*, 1585–1590. doi: 10.2337/dc06-0057

Shamseddeen, H., Getty, J. Z., Hamdallah, I. N., & Ali, M. R. (2011). Epidemiology and economic impact of obesity and type 2 diabetes. *Surgical Clinics of North America*, *91*, 1163–1172. doi: 10.1016/j.suc.2011.08.001

Sharif, B., Kopec, J. A., Wong, H., Finès, P., Sayre, E. C., Liu, R. R., & Wolfson, M. C. (2012). Uncertainty analysis in population-based disease microsimulation models. *Epidemiology Research International*. doi: 10.1155/2012/610405

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*, 289–310. doi: 10.1214/10-STS330

Silversides, A., Doig, C., & Sullivan, T. (2013). *Canadian diabetes strategies under fire as diabetes rates continue to rise*. Retrieved July 13, 2013, from http://healthydebate.ca/2013/05/topic/managing-chronic-diseases/canadian-diabetes-strategies-under-fire-as-diabetes-rates-continue-to-rise

Simmons, R. K., Unwin, N., & Griffin, S. J. (2010). International Diabetes Federation: an updated of the evidence concerning the prevention of type 2 diabetes. *Diabetes Research and Clinical Practice*, *87*, 143–149. doi: 10.1016/j.diabres.2009.10.003

Singh, A. C., & Mohl, C. A. (1996). Understanding calibration estimators in survey sampling. *Survey Methodology*, *22*, 107–115.

Smith, D. M., Clarke, G. P., & Harland, K. (2009). Improving the synthetic generation process in spatial microsimulation models. *Environment and Planning A*, *41*, 1251–1268. doi: 10.1068/a4147

Smith, D. M., Pearce, J. R., & Harland, K. (2011). Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. *Health & Place*, *17*, 618–624. doi: 10.1016/j.healthplace.2011.01.001

Social Development, Finance & Administration, City of Toronto. (2009). *Neighbourhood planning areas* (Data file). Toronto: City of Toronto. Retrieved from http://www1.toronto.ca/City_Of_Toronto/Information_Technology/Open_Data/Data_Sets/Assets/Files/neighbourhoods.zip

Social Policy and Research. (2014). *TSNS 2020 Neighbourhood Equity Index: Methodological Documentation* (Tech. Rep.). Toronto: Social Development, Finance and Administration, City of Toronto. Retrieved from http://www.torontohealthprofiles.ca/a_documents/resources/TSNS_2020_NEI_Methodological_Documentation-Final.pdf

Speed, T. P. (1998). Iterative proportional fitting. In P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics* (Vol. 3, pp. 2116–2120). Chichester, N.Y.: John Wiley & Sons.

Spielauer, M. (2011). What is social science microsimulation? *Social Science Computer Review*, *29*, 9–20. doi: 10.1177/0894439310370085

Srikanth, S., & Deedwania, P. (2011). Primary and secondary prevention strategy for cardiovascular disease in diabetes mellitus. *Cardiology Clinics*, *29*, 47–70. doi: 10.1016/j.ccl.2010.11.004

Stafford, M., Duke-Williams, O., & Shelton, N. (2008). Small area inequalities in health: Are we underestimating them? *Social Science & Medicine*, *67*, 891–899. doi: 10.1016/j.socscimed.2008.05.028

Statistics Canada. (2006a). *Census of Population, 2006, Canada: Topic-based Tabulations* [[B2020 data files]]. Ottawa, ON.

Statistics Canada. (2006b). *Statistics Canada Census Boundaries*. [[Computer file]]. Ottawa, ON.

Statistics Canada. (2013). *National household survey (nhs): Data quality.* Retrieved February 26, 2014 from http://www12.statcan.gc.ca/NHS-ENM/2011/ref/about-apropos/nhs-enm_r005-eng.cfm.

Statistics Canada. (2015). *National Household Survey non-response bias indicators.* Retrieved October 26, 2015 from http://www12.statcan.gc.ca/nhs-enm/2011/ref/reports-rapports/sw-ep/ch5-eng.cfm.

Steel, D. G., & Tranmer, M. D. (2011). Measuring and analyizing the within group homogeneity of a multi-category variable. *Journal of Statistical Theory and Practice, 5*, 649–658. doi: 10.1080/15598608.2011.10483736

Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., ... Firestone, S. (2015). epiR: Tools for the Analysis of Epidemiological Data [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=epiR (R package version 0.9-69)

Tanton, R., & Edwards, K. L. (2013). Introduction to spatial microsimulation: History, methods and applications. In R. Tanton & K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users* (Vol. 6, pp. 3–8). New York: Springer. doi: 10.1007/978-94-007-4623-7_1

Tanton, R., Harding, A., & McNamara, J. (2010). Urban and rural estimates of poverty: Recent advances in spatial microsimulation in Australia. *Geographical Research, 48*, 52–64. doi: 10.1111/j.1745-5871.2009.00615.x

Tanton, R., Vidyattama, Y., Nepal, B., & McNamara, J. (2011). Small area estimation using a reweighting algorithm. *Journal of the Royal Statistical Society Series A: Statistics in Society, 174*, 931–951. doi: 10.1111/j.1467-985X.2011.00690.x

Tanton, R., Williamson, P., & Harding, A. (2007). *Comparing two methods of reweighting a survey file to small area data—generalised regression and combinatorial optimization* (Tech. Rep.). Canberra, Australia: National Center for Social and Economic Modelling, University of Canberra. Retrieved from http://www.natsem.canberra.edu.au/storage/Tanton et al final.pdf

Templ, M., Alfons, A., Kowarik, A., & Prantner, B. (2015). VIM: Visualization and Imputation of Missing Values [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=VIM (R package version 4.4.1)

Thomas, S., & Wannell, B. (2009). Combining cycles of the Canadian Community Health Survey. *Health Reports, 20*, 1–6.

Thompson, K. M., & Graham, J. D. (1996). Going beyond the single number: using probabilistic risk assessment to improve risk management. *Human and Ecological Risk Assessment*, *2*, 1008–1034.

Timmins, K. A., & Edwards, K. L. (2016). Validation of spatial microsimulation models: a proposal to adopt the Bland-Altman method. *International Journal of Microsimulation*, *9*, 106–122.

Tomintz, M., Kosar, B., & Clarke, G. (2016). smokeSALUD: exploring the effect of demographic change on the smoking prevalence at municipality level in Austria. *International Journal of Health Geographics*, *15*, 36. doi: 10.1186/s12942-016-0066-4

Tomintz, M. N., Clarke, G. P., & Rigby, J. E. (2008). The geography of smoking in Leeds: Estimating individual smoking rates and the implications for the location of stop smoking services. *Area*, *40*, 341–353.

Tomintz, M. N., Clarke, G. P., & Rigby, J. E. (2009). Planning the location of stop smoking services at the local leve: A geographic analysis. *Journal of Smoking Cessation*, *4*, 61–73. doi: 10.1375/jsc.4.2.61

Topping, D. (2012, December 6). Toronto's most, and least, diverse neigbhourhoods. *The GRID TO*. Retrieved from http://www.thegridto.com/city/places/torontos-most-and-least-diverse-neighbourhoods/#pager

Toronto Community Health Profiles Partnership. (2015). *Age-Standardized Diabetes Prevalence Rate (%) Among Adults 20+, 2012: Toronto Neighbourhoods*. Retrieved from http://www.torontohealthprofiles.ca/a_documents/TM_allCateg_maps/TM_maps_AHD/1_AHD_Diabetes_All_MF_N_2012_20plus_RR.pdf

Tu, K., Campbell, N. R. C., Chen, Z. L., Cauch-Dudek, K. J., & McAlister, F. A. (2007). Accuracy of administrative databases in identifying patients with hypertension. *Open Medicine*, *1*, 18–26. Retrieved from http://www.openmedicine.ca/article/view/17/35

Tuomilehto, J., Lindström, J., Eriksson, J. G., Valle, T. T., Hamalainen, H., Illanne-Parikka, P., ... Uusitupa, M. (2001). Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. *New England Journal of Medicine*, *344*, 1342–1350.

Turban, E., & Aronson, J. E. (2001). *Decision support systems and intelligent systems*

(6th ed.). Upper Saddle River, NJ: Prentice Hall.

van Dam, R. M. (2003). The epidemiology of lifestyle and risk for type 2 diabetes. *European Journal of Epidemiology*, *18*, 1115–1125.

van Leeuwen, E. S. (2010a). Microsimulation of Rural Households. In E. S. van Leeuwen (Ed.), *Urban-Rural Interactions: Towns as Focus Points in Rural Development* (pp. 115–135). Berlin: Springer-Verlag. doi: 10.1007/978-3-7908-2407-0_6

Voas, D., & Williamson, P. (2000). An evaluation of the combinatorial optimization approach to the creation of synthetic microdata. *International Journal of Population Geography*, *6*, 349–366.

Voas, D., & Williamson, P. (2001). Evaluating goodness-of-fit measures for synthetic microdata. *Geographical & Environmental Modelling*, *5*, 177–200.

Vojta, D., Koehler, T. B., Longjohn, M., Lever, J. A., & Caputo, N. F. (2013). A coordinated national model for diabetes prevention: Linking health systems to an evidence-based community program. *American Journal of Preventive Medicine*, *44*(4 Suppl 4), S301–S306. doi: 10.1016/j.amepre.2012.12.018

Wan, N., Zhan, B., Zou, B., & Chow, E. (2012). A relative spatial access assessment approach for analyzing potential access to colorectal cancer services in texas. *Applied Geography*, *32*, 291–299. doi: 10.1016/j.apgeog.2011.05.001

Wang, F. (2006). *Quantitative Methods and Applications in GIS*. Boca Raton, FL: CRC/Taylor & Francis.

Wang, F. (2012). Measurement, optimization, and impact of health care accessibility: a methodological review. *Annals of the Association of American Geographers*, *102*, 1104–1112. doi: 10.1080/00045608.2012.657146

Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Health & Place*, *11*, 131–146. doi: 10.1016/j.healthplace.2004.02.003

Weinstein, A. R., Sesso, H. D., Lee, I. M., Cook, N. R., Manson, J. E., Buring, J. E., & Gaziano, J. M. (2004). Relationship of physical activity vs body mass index with type 2 diabetes in women. *Journal of the American Medical Association*, *292*, 1188–1194.

Whiting, D. R., Guariguata, L., Weil, C., & Shaw, J. (2011). IDF diabetes atlas: Global

estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Research and Clinical Practice*, *94*(3), 311–321. doi: 10.1016/j.diabres.2011.10.029

Whitworth, A., Carter, E., Ballas, D., & Moon, G. (2016). Estimating uncertainty in spatial microsimulation approaches to small area estimation: a new approach to solving an old problem. *Computers, Environment and Urban Systems*. doi: 10.1016/j.compenvurbsys.2016.06.004

Wilkins, R., Berthelot, J. M., & Ng, E. (2002). Trends in mortality by neighbourhood income in urban canada from 1971 to 1996. *Health Reports*, *12*, 1–28. Retrieved from http://www.statcan.gc.ca/pub/82-003-s/2002001/pdf/82-003-s2002007-eng.pdf

Williamson, P. (1996). Community care policies for the elderly, 1981 and 1991: A microsimulation approach. In G. P. Clarke (Ed.), *Microsimulation for Urban and Regional Policy Analysis* (Vol. 6, pp. 64–87). London, U.K.: Pion.

Williamson, P. (2007). *CO Instruction Manual* (Working Paper No. 2007/1). Liverpool, United Kingdom: University of Liverpool. Retrieved from http://bit.ly/29IhsVX

Williamson, P., Birkin, M., & Rees, P. (1998). The estimation of population microdata by using data from small area statistics and samples of anonymized records. *Environment and Planning A*, *30*, 785–816.

Wolf, D. A. (2001). The role of microsimulation in longitudinal data analysis. *Canadian Studies in Population*, *28*, 313–339.

Wong, D. (2009). The modifiable areal unit problem (maup). In A. S. Fotheringham & P. Rogerson (Eds.), *The SAGE Handbook of Spatial Analysis* (pp. 105–124). Thousand Oaks, CA: SAGE Publications.

Wu, S., Wang, R., Zhao, Y., Ma., X., Wu, M., Yan, X., & He, J. (2013). The relationship between self-rated ehalth and objective health status: a population-based study. *BMC Public Health*, *13*, 320. doi: 10.1186/1471-2458-13-320

Wu, T., Gao, X., Chen, M., & van Dam, R. M. (2009). Long-term effectiveness of diet-plus-exercise interventions vs. diet-only interventions for weight loss: a meta-analysis. *Obesity Reviews*, *10*, 313–323. doi: 10.1111/j.1467-789X.2008.00547.x

Zhang, P., Zhang, X., Brown, J., Vistisen, D., Sicree, R., Shaw, J., & Nichols, G. (2010). Global healthcare expenditure on diabetes for 2010 and 2030. *Diabetes Re-*

search and Clinical Practice, *87*(3), 293–301. doi: 10.1016/j.diabres.2010.01.026

Zucchelli, E., Jones, A. M., & Rice, N. (2012). The evaluation of health policies through dynamic microsimulation methods. *International Journal of Microsimulation*, *5*, 2–20. Retrieved from http://www.microsimulation.org/ijm/V5_1/1_IJM_5_1_spring_2012_Zuchelli_Rice_Jones.pdf

Appendix

A

TropISM: Model Validation and Uncertainty

A.1 Model Validation

Table A.1. Simulated prevalence of unconstrained outcomes among men compared to the Canadian Community Health Survey for the TropISM model developed using the Ontario subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Men, ages 20+										
Type 2 diabetes	6.6	4.9	1.7	**34.8**	5.7	0.9	*15.9*	8.2	-1.6	*-19.2*
Income (80th %ile)	18.5	36.5	-18.0	**-49.4**	17.7	0.8	4.4	14.3	4.2	**29.4**
Current smoker	26.4	25.5	1.0	3.8	24.8	1.7	6.7	25.7	0.7	2.8
Hypertension	17.8	17.0	0.7	4.4	13.1	4.7	**35.6**	15.0	2.8	*18.3*
Heart disease	6.5	5.3	1.3	**24.0**	4.8	1.7	**35.0**	5.9	0.6	10.9
Men, ages 20–44										
BMI < 23	25.5	33.1	-7.6	**-23**	32.6	-7.1	**-21.8**	27.9	-2.4	-8.5
23 ≤ BMI < 25	22.6	24.8	-2.2	-8.8	21.2	1.4	6.5	26.5	-3.9	*-14.8*
25 ≤ BMI < 30	36.8	32.6	4.1	12.6	34.4	2.4	6.9	32.9	3.8	11.7
30 ≤ BMI < 35	10.9	6.6	4.3	**64.5**	7.7	3.2	**42.2**	6.3	4.6	**73.1**
BMI ≥ 35	3.4	1.3	2.1	**162.8**	2.1	1.2	**57.2**	4.3	-1.0	**-22.3**
BMI unknown	0.8	1.5	-0.7	**-45.1**	1.9	-1.1	**-56.8**	2.1	-1.2	**-59.3**
Men, ages 45+										
BMI < 23	18.1	21.3	-3.2	-15.0	17.4	0.7	4.2	20.8	-2.7	-12.8
23 ≤ BMI < 25	21.3	23.4	-2.2	-9.3	26.8	-5.5	**-20.5**	20.1	1.2	5.7
25 ≤ BMI < 30	42.1	39.6	2.5	6.3	43.3	-1.2	-2.8	40.6	1.4	3.5
30 ≤ BMI < 35	13.6	11.8	1.8	14.9	8.6	4.9	**57.0**	12.7	0.9	7.2
BMI ≥ 35	4.1	3.2	1.0	**30.4**	2.6	1.6	**60.8**	2.2	1.9	**86.4**
BMI unknown	0.9	0.7	0.1	19.3	1.4	-0.5	**-37.4**	3.6	-2.7	**-76.1**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.2. Simulated prevalence of unconstrained outcomes among men compared to the Canadian Community Health Survey for the TropISM model developed using the CMA subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Men, ages 20+										
Type 2 diabetes	6.4	4.9	1.5	**30.3**	5.7	0.7	12.0	8.2	-1.8	**-21.9**
Income (80th %ile)	17.5	36.5	-19.1	**-52.2**	17.7	-0.2	-1.4	14.3	3.2	**22.2**
Current smoker	25.9	25.5	0.4	1.7	24.8	1.1	4.6	25.7	0.2	0.8
Hypertension	17.7	17.0	0.6	3.7	13.1	4.6	**34.7**	15.0	2.6	*17.6*
Heart disease	6.2	5.3	0.9	*17.7*	4.8	1.4	**28.2**	5.9	0.3	5.3
Men, ages 20–44										
BMI < 23	26.1	33.2	-7.1	**-21.5**	33.1	-7.0	**-21.2**	28.0	-1.9	-6.9
23 ≤ BMI < 25	22.8	25.6	-2.8	-10.8	20.7	2.1	10.1	26.4	-3.5	-13.3
25 ≤ BMI < 30	36.8	31.7	5.1	*16.0*	34.4	2.4	6.9	32.9	3.9	11.7
30 ≤ BMI < 35	10.4	6.6	3.8	**58.4**	7.7	2.7	**34.9**	6.3	4.1	**65.2**
BMI ≥ 35	3.0	1.4	1.6	**119.7**	2.1	0.9	**42.3**	4.3	-1.3	**-31.1**
BMI unknown	0.9	1.5	-0.6	**-42.5**	1.9	-1.0	**-54.7**	2.1	-1.2	**-57.4**
Men, ages 45+										
BMI < 23	18.3	21.3	-3.0	-14.2	17.5	0.8	4.2	20.9	-2.7	-12.7
23 ≤ BMI < 25	22.0	23.6	-1.6	-6.8	26.6	-4.6	*-17.3*	20.0	2.0	10.2
25 ≤ BMI < 30	41.6	39.6	2.0	5.0	43.3	-1.7	-3.8	40.6	1.0	2.5
30 ≤ BMI < 35	13.2	11.6	1.7	14.4	8.6	4.6	**53.1**	12.9	0.3	2.5
BMI ≥ 35	3.9	3.2	0.8	**24.9**	2.6	1.3	**54.0**	2.0	2.0	**101.1**
BMI unknown	0.9	0.7	0.2	**26.5**	1.4	-0.5	**-33.6**	3.6	-2.7	**-74.7**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.3. Simulated prevalence of unconstrained outcomes among men compared to the Canadian Community Health Survey for the TropISM model developed using the GTA subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Men, ages 20+										
Type 2 diabetes	6.7	4.9	1.8	**36.3**	5.7	1.0	*17.2*	8.2	-1.5	*-18.3*
Income (80th %ile)	17.5	36.5	-19.1	**-52.2**	17.7	-0.2	-1.4	14.3	3.2	**22.2**
Current smoker	24.9	25.5	-0.6	-2.2	24.8	0.1	0.5	25.7	-0.8	-3.1
Hypertension	17.7	17.0	0.7	4.0	13.1	4.6	**35.1**	15.0	2.7	*17.9*
Heart disease	6.5	5.3	1.2	**23.7**	4.8	1.7	**34.7**	5.9	0.6	10.6
Men, ages 20–44										
BMI < 23	27.6	33.1	-5.5	*-16.6*	32.6	-5.0	*-15.3*	27.9	-0.3	-0.9
23 ≤ BMI < 25	22.3	24.8	-2.4	-9.8	21.2	1.1	5.4	26.5	-4.2	*-15.7*
25 ≤ BMI < 30	36.5	32.6	3.9	11.9	34.4	2.1	6.2	32.9	3.6	11.0
30 ≤ BMI < 35	9.8	6.6	3.2	**47.9**	7.7	2.1	**27.9**	6.3	3.5	**55.6**
BMI ≥ 35	2.9	1.3	1.6	**124.7**	2.1	0.8	**34.4**	4.3	-1.4	**-33.5**
BMI unknown	0.8	1.5	-0.7	**-49.2**	1.9	-1.1	**-60.0**	2.1	-1.3	**-62.4**
Men, ages 45+										
BMI < 23	20.7	21.3	-0.6	-3.0	17.4	3.3	*18.9*	20.8	-0.1	-0.6
23 ≤ BMI < 25	23.7	23.4	0.3	1.3	26.8	-3.1	-11.3	20.1	3.6	*18.0*
25 ≤ BMI < 30	40.4	39.6	0.9	2.2	43.3	-2.9	-6.6	40.6	-0.2	-0.5
30 ≤ BMI < 35	10.6	11.8	-1.2	-10.4	8.6	2.0	**22.5**	12.7	-2.1	*-16.4*
BMI ≥ 35	3.7	3.2	0.5	*16.7*	2.6	1.1	**43.9**	2.2	1.5	**66.7**
BMI unknown	0.9	0.7	0.2	**24.4**	1.4	-0.5	**-34.7**	3.6	-2.7	**-75.1**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.4. Simulated prevalence of unconstrained outcomes among men compared to the Canadian Community Health Survey for the TropISM model developed using the Toronto health region subset of CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Men, ages 20+										
Type 2 diabetes	6.6	4.9	1.7	**33.9**	5.7	0.9	*15.1*	8.2	-1.6	*-19.8*
Income (80th %ile)	16.8	36.5	-19.7	**-54.1**	17.7	-0.9	-5.3	14.3	2.5	*17.4*
Current smoker	26.8	25.5	1.3	5.1	24.8	2.0	8.0	25.7	1.1	4.1
Hypertension	14.6	17.0	-2.4	-13.9	13.1	1.5	11.8	15.0	-0.4	-2.5
Heart disease	6.0	5.3	0.7	13.7	4.8	1.2	**23.9**	5.9	0.1	1.8
Men, ages 20–44										
BMI < 23	33.3	33.1	0.2	0.4	32.6	0.7	1.9	27.9	5.4	*19.3*
23 ≤ BMI < 25	23.4	24.8	-1.4	-5.6	21.2	2.2	10.3	26.5	-3.1	-11.8
25 ≤ BMI < 30	32.4	32.6	-0.2	-0.6	34.4	-2.0	-5.7	32.9	-0.5	-1.4
30 ≤ BMI < 35	7.8	6.6	1.2	*18.0*	7.7	0.1	2.0	6.3	1.5	**24.1**
BMI ≥ 35	1.8	1.3	0.5	**37.9**	2.1	-0.3	*-17.6*	4.3	-2.5	**-59.2**
BMI unknown	1.3	1.5	-0.2	-13.1	1.9	-0.6	**-31.6**	2.1	-0.8	**-35.7**
Men, ages 45+										
BMI < 23	21.2	21.3	-0.1	-0.5	17.4	3.8	**21.9**	20.8	0.4	2.0
23 ≤ BMI < 25	28.1	23.4	4.7	*20.0*	26.8	1.3	5.1	20.1	8.0	**39.8**
25 ≤ BMI < 30	38.9	39.6	-0.7	-1.7	43.3	-4.4	-10.2	40.6	-1.7	-4.3
30 ≤ BMI < 35	8.4	11.8	-3.4	**-28.9**	8.6	-0.2	-2.9	12.7	-4.3	**-33.7**
BMI ≥ 35	2.9	3.2	-0.3	-7.9	2.6	0.3	13.6	2.2	0.7	**31.6**
BMI unknown	0.5	0.7	-0.2	**-31.7**	1.4	-0.9	**-64.1**	3.6	-3.1	**-86.3**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.5. Simulated prevalence of unconstrained outcomes among women compared to the Canadian Community Health Survey for the TropISM model developed using the Ontario subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Women, ages 20+										
Type 2 diabetes	5.4	5.2	0.2	5.2	2.5	2.9	**116.0**	6.7	-1.2	*-18.3*
Hypertension	19.7	16.6	3.1	*18.9*	15.5	4.2	**26.9**	18.5	1.1	6.4
Women, ages 20–44										
BMI < 23	45.5	51.6	-6.1	-11.8	54.0	-8.4	*-15.6*	52.5	-6.9	-13.2
23 ≤ BMI < 25	15.6	15.0	0.6	3.5	14.3	1.3	8.8	14.9	0.7	4.6
25 ≤ BMI < 30	19.8	19.2	0.6	3.4	17.4	2.5	14.1	18.1	1.7	9.8
30 ≤ BMI < 35	7.3	6.1	1.2	*19.8*	5.6	1.7	**30.4**	4.6	2.6	**56.4**
BMI ≥ 35	4.4	2.9	1.5	**50.1**	2.4	1.9	**80.1**	2.9	1.5	**51.7**
BMI unknown	7.4	5.2	2.2	**43.3**	6.4	1.1	*17.1*	7.1	0.3	5.4
Women, ages 45–64										
BMI < 23	29.7	29.8	-0.04	-0.1	35.7	-5.9	*-16.6*	32.3	-2.6	-7.9
23 ≤ BMI < 25	18.4	17.3	1.1	6.6	17.3	1.1	6.6	18.7	-0.3	-1.4
25 ≤ BMI < 30	29.9	32.7	-2.8	-8.7	28.6	1.2	4.4	30.4	-0.5	-1.6
30 ≤ BMI < 35	11.8	10.4	1.4	13.2	8.5	3.2	**37.6**	8.2	3.6	**43.3**
BMI ≥ 35	6.6	5.7	0.9	*16.6*	4.8	1.9	**39.5**	6.2	0.4	7.3
BMI unknown	3.6	4.2	-0.6	-14.1	5.1	-1.6	**-30.3**	4.3	-0.7	*-16.1*
Women, ages 65+										
BMI < 23	30.2	27.6	2.6	9.6	30.9	-0.7	-2.1	27.7	2.5	9.0
23 ≤ BMI < 25	18.6	22.7	-4.1	*-18.3*	21.3	-2.8	-12.9	17.0	1.6	9.1
25 ≤ BMI < 30	33.0	29.7	3.3	11.2	32.1	0.9	2.7	29.4	3.6	12.3
30 ≤ BMI < 35	11.4	10.5	0.9	9.3	7.7	3.7	**48.8**	12.4	-1.0	-8.0
BMI ≥ 35	3.2	4.3	-1.1	**-25.4**	3.0	0.3	9.5	3.6	-0.4	-9.9
BMI unknown	3.5	5.2	-1.7	**-32.4**	5.0	-1.5	**-29.7**	9.8	-6.3	**-64.3**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.6. Simulated prevalence of unconstrained outcomes among women compared to the Canadian Community Health Survey for the TropISM model developed using the CMA subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Women, ages 20+										
Type 2 diabetes	5.2	5.2	-0.01	-0.1	2.5	2.6	**105.0**	6.7	-1.5	**-22.4**
Hypertension	19.3	16.6	2.7	*16.4*	15.5	3.8	**24.2**	18.5	0.8	4.2
Women, ages 20–44										
BMI < 23	46.5	52.1	-5.6	-10.7	54.4	-7.9	-14.6	53.2	-6.7	-12.6
23 ≤ BMI < 25	15.3	14.4	0.9	6	13.8	1.4	10.4	14.2	1.1	7.8
25 ≤ BMI < 30	20.0	19.4	0.6	3	17.4	2.6	14.9	18.1	1.9	10.6
30 ≤ BMI < 35	7.0	6.0	1.0	*15.3*	5.6	1.4	**24.7**	4.6	2.3	**49.6**
BMI ≥ 35	4.1	2.9	1.2	**42.1**	2.4	1.7	**70.5**	2.9	1.2	**43.6**
BMI unknown	7.2	5.2	2.0	**38.4**	6.4	0.8	13.1	7.1	0.1	1.7
Women, ages 45–64										
BMI < 23	31.2	29.4	1.8	6.3	36.1	-4.9	-13.7	33.0	-1.8	-5.5
23 ≤ BMI < 25	17.7	17.4	0.3	2.0	17.0	0.7	4.3	18.0	-0.3	-1.3
25 ≤ BMI < 30	29.7	31.2	-1.5	-4.9	28.4	1.2	4.4	30.4	-0.7	-2.2
30 ≤ BMI < 35	11.7	12.2	-0.5	-3.9	8.5	3.1	**36.9**	8.2	3.5	**42.5**
BMI ≥ 35	6.1	5.7	0.4	7.4	4.8	1.4	**28.5**	6.2	-0.1	-1.2
BMI unknown	3.6	4.2	-0.6	-13.8	5.1	-1.5	**-30.1**	4.3	-0.7	*-15.9*
Women, ages 65+										
BMI < 23	31.9	29.5	2.4	8.1	31.7	0.2	0.6	28.3	3.6	12.6
23 ≤ BMI < 25	17.1	21.1	-4.0	*-18.9*	20.7	-3.5	*-17.1*	16.5	0.6	3.9
25 ≤ BMI < 30	33.0	29.3	3.7	12.6	32.0	0.9	2.9	29.4	3.6	12.2
30 ≤ BMI < 35	11.3	10.6	0.7	6.4	7.7	3.6	**47.3**	12.4	-1.1	-9.0
BMI ≥ 35	3.3	4.3	-1.0	**-23.5**	3.0	0.4	12.4	3.6	-0.3	-7.5
BMI unknown	3.4	5.2	-1.8	**-34.2**	5.0	-1.6	**-31.6**	9.8	-6.4	**-65.2**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.7. Simulated prevalence of unconstrained outcomes among women compared to the Canadian Community Health Survey for the TropISM model developed using the GTA subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Women, ages 20+										
Type 2 diabetes	4.3	5.2	-0.9	*-17.6*	2.5	1.8	**69.1**	6.7	-2.4	**-36.0**
Hypertension	18.4	16.6	1.8	11.1	15.5	2.9	*18.6*	18.5	-0.1	-0.5
Women, ages 20–44										
BMI < 23	49.3	51.6	-2.3	-4.4	54	-4.6	-8.6	52.5	-3.2	-6.0
23 ≤ BMI < 25	15.2	15.0	0.2	1.1	14.3	0.9	6.3	14.9	0.3	2.1
25 ≤ BMI < 30	19.0	19.2	-0.2	-0.8	17.4	1.6	9.5	18.1	0.9	5.4
30 ≤ BMI < 35	6.1	6.1	0.05	0.8	5.6	0.5	9.7	4.6	1.5	**31.6**
BMI ≥ 35	3.3	2.9	0.4	14.8	2.4	0.9	**37.8**	2.9	0.4	*16.0*
BMI unknown	7.0	5.2	1.8	**34.5**	6.4	0.6	9.9	7.1	-0.1	-1.1
Women, ages 45–64										
BMI < 23	34.1	29.8	4.3	14.3	35.7	-1.6	-4.5	32.3	1.8	5.4
23 ≤ BMI < 25	17.0	17.3	-0.3	-1.8	17.3	-0.3	-1.9	18.7	-1.7	-9.2
25 ≤ BMI < 30	29.7	32.7	-3.0	-9.1	28.6	1.1	3.9	30.4	-0.7	-2.1
30 ≤ BMI < 35	10.5	10.4	0.1	0.7	8.5	2.0	**22.4**	8.2	2.3	**27.4**
BMI ≥ 35	5.1	5.7	-0.6	-10.1	4.8	0.3	7.5	6.2	-1.1	*-17.3*
BMI unknown	3.7	4.2	-0.5	-11.2	5.1	-1.4	**-28.0**	4.3	-0.6	-13.4
Women, ages 65+										
BMI < 23	33.0	27.6	5.4	*19.4*	30.9	2.1	6.6	27.7	5.2	18.8
23 ≤ BMI < 25	18.8	22.7	-3.9	*-17.4*	21.3	-2.5	-12.0	17.0	1.8	10.3
25 ≤ BMI < 30	33.2	29.7	3.5	11.8	32.1	1.1	3.2	29.4	3.8	12.8
30 ≤ BMI < 35	9.8	10.5	-0.7	-5.8	7.7	2.1	**28.3**	12.4	-2.6	**-20.7**
BMI ≥ 35	2.4	4.3	-1.9	**-44.1**	3.0	-0.6	*-17.9*	3.6	-1.2	**-32.4**
BMI unknown	2.8	5.2	-2.4	**-45.7**	5.0	-2.2	**-43.6**	9.8	-7.0	**-71.3**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

Table A.8. Simulated prevalence of unconstrained outcomes among women compared to the Canadian Community Health Survey for the TropISM model developed using the Toronto health region subset of the CCHS.

		2003			2005			2007		
	TropISM	CCHS	Abs	Rel	CCHS	Abs	Rel	CCHS	Abs	Rel
Women, ages 20+										
Type 2 diabetes	3.6	5.2	-1.6	**-30.5**	2.5	1.1	**42.7**	6.7	-3.1	**-46.0**
Hypertension	17.4	16.6	0.8	5.0	15.5	1.9	12.1	18.5	-1.1	-6.0
Women, ages 20–44										
BMI < 23	50.7	51.6	-0.9	-1.7	54	-3.2	-6.0	52.5	-1.7	-3.3
23 ≤ BMI < 25	15.8	15	0.7	4.7	14.3	1.5	10.0	14.9	0.9	5.8
25 ≤ BMI < 30	18.2	19.2	-1.0	-5.3	17.4	0.8	4.5	18.1	0.1	0.6
30 ≤ BMI < 35	5.8	6.1	-0.3	-5.1	5.6	0.8	3.2	4.6	1.1	**23.9**
BMI ≥ 35	4.0	2.9	1.1	**36.1**	2.4	1.6	**63.3**	2.9	1.1	**37.6**
BMI unknown	5.6	5.2	0.4	8.4	6.4	-0.8	-11.4	7.1	-1.4	**-20.3**
Women, ages 45–64										
BMI < 23	38.5	29.8	8.7	**29.4**	35.7	2.9	8.0	32.3	6.2	*19.3*
23 ≤ BMI < 25	16.4	17.3	-0.9	-4.8	17.3	-0.9	-4.8	18.7	-2.2	-12.0
25 ≤ BMI < 30	27.9	32.7	-4.8	-14.7	28.6	-0.7	-2.5	30.4	-2.5	-8.1
30 ≤ BMI < 35	8.4	10.4	-2.0	*-19.5*	8.5	-0.1	-2.0	8.2	0.2	2.0
BMI ≥ 35	5.4	5.7	-0.3	-5.7	4.8	0.6	12.8	6.2	-0.8	-13.3
BMI unknown	3.4	4.2	-0.8	*-18.6*	5.1	-1.7	**-34.0**	4.3	-0.9	**-20.6**
Women, ages 65+										
BMI < 23	30.9	27.6	3.3	11.9	30.9	-0.02	-0.1	27.7	3.1	11.3
23 ≤ BMI < 25	21.6	22.7	-1.1	-5.0	21.3	0.3	1.2	17.0	4.6	**26.8**
25 ≤ BMI < 30	32.1	29.7	2.4	8.1	32.1	-0.05	-0.2	29.4	2.7	9.1
30 ≤ BMI < 35	7.8	10.5	-2.7	**-25.7**	7.7	0.1	1.2	12.4	-4.7	**-37.4**
BMI ≥ 35	3.1	4.3	-1.2	**-27.5**	3.0	0.1	6.5	3.6	-0.4	-12.4
BMI unknown	4.5	5.2	-0.7	-12.7	5.0	-0.5	-9.3	9.8	-5.3	**-53.9**

Abs: Absolute difference in percentage points. Rel: Relative percent difference.

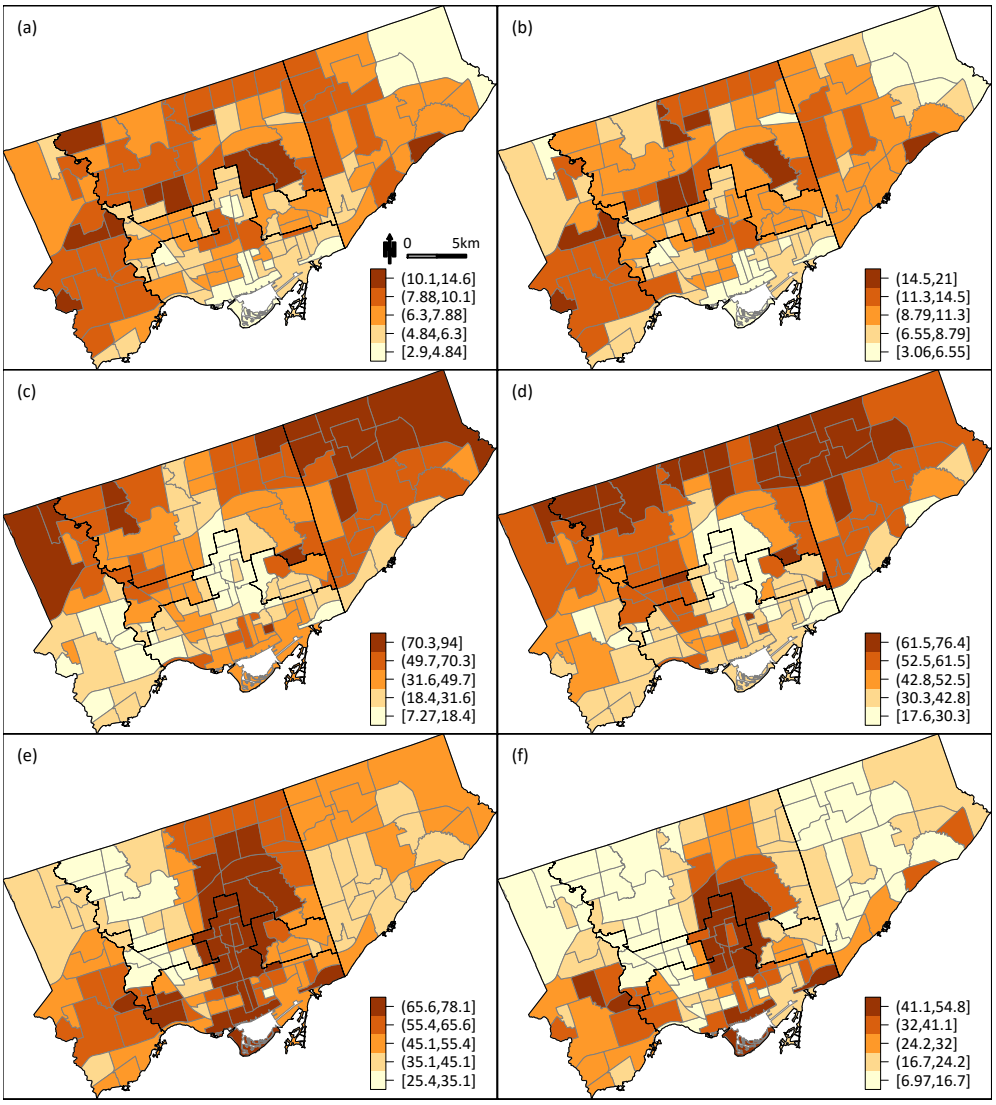


Figure A.1. Demographic characteristics of the simulated TropISM population. Neighbourhood distribution of the population of (a) men aged 65 and older, (b) women aged 65 and older, (c) visible minorities, (d) immigrants, (e) educational attainment (post-secondary), and (f) personal incomes greater than $50,000/year.

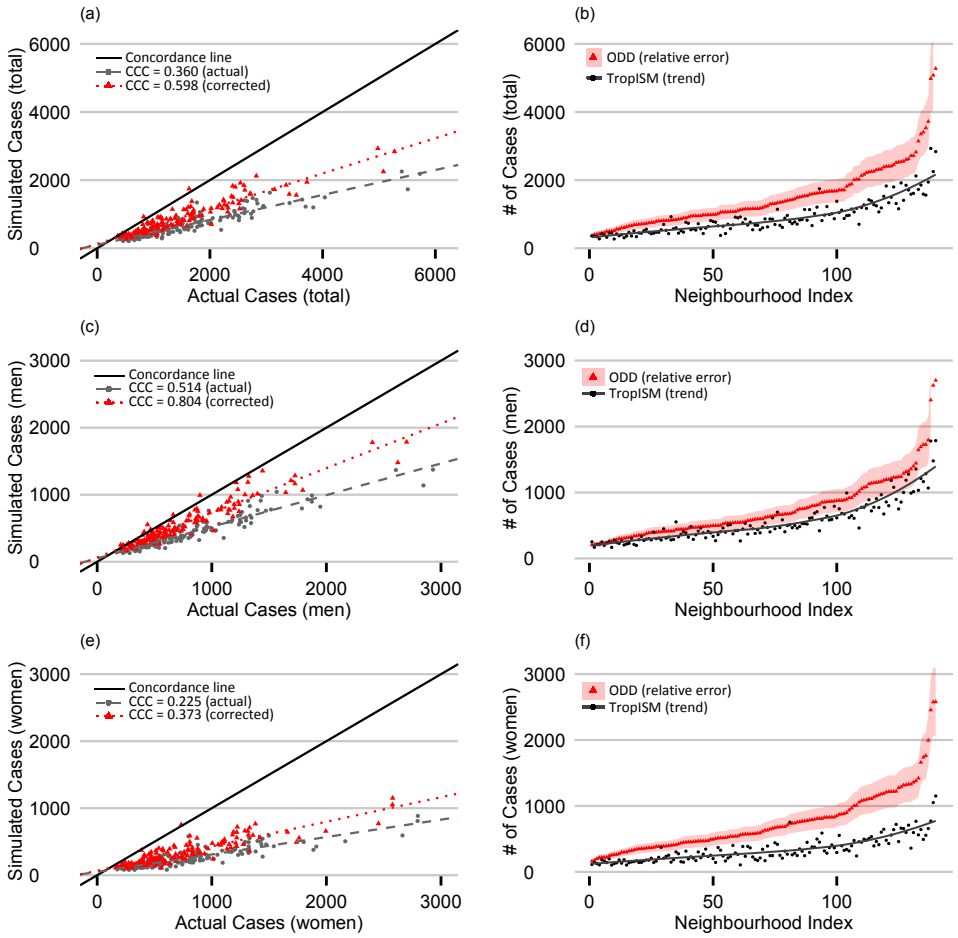


Figure A.2. Neighbourhood-level validation of the simulated number of prevalent cases of diabetes compared to the actual number of cases in 2005, ascertained from the Ontario Diabetes Database (ODD). Panels (a), (c) & (e): accuracy and precision of the simulated number of cases. Panels (b), (d) & (f): difference between actual cases and TropISM, ranked in ascending order by the actual number of cases in each neighbourhood.

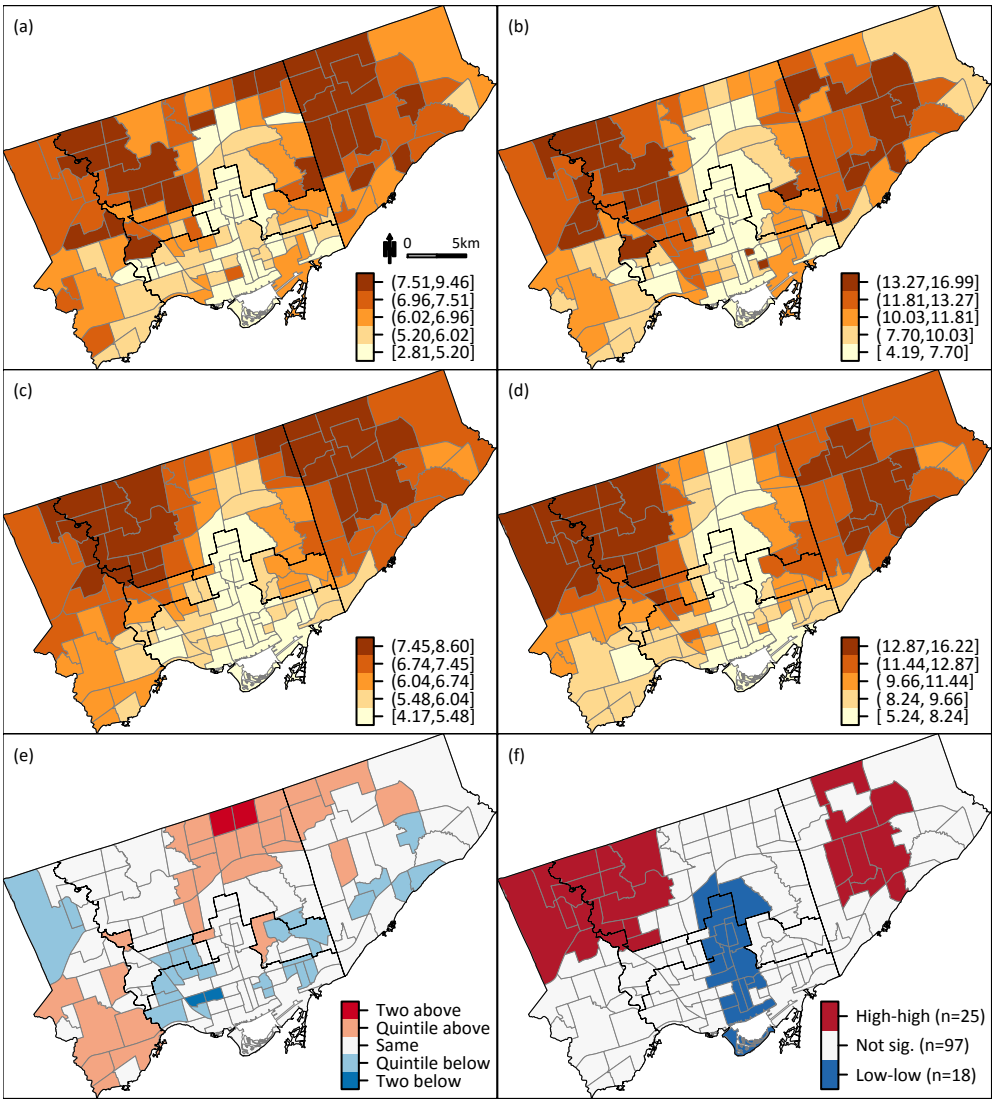


Figure A.3. Simulated prevalence (%) of type 2 diabetes (corrected for undiagnosed cases) compared to prevalence estimates from the Ontario Diabetes Database (ODD, corrected for false positives and removal of type 1 diabetes). (a) TropISM, raw prevalence rate; (b) ODD, raw prevalence rate; (c) TropISM, spatially smoothed prevalence rate; (d) ODD, spatially smoothed prevalence rate; (e) difference in the quintile rankings of spatially smoothed rates (TropISM vs. ODD); (f) bivariate LISA map depicting areas having similarly high (or low) simulated and known prevalence of type 2 diabetes.

A.2 Model Uncertainty

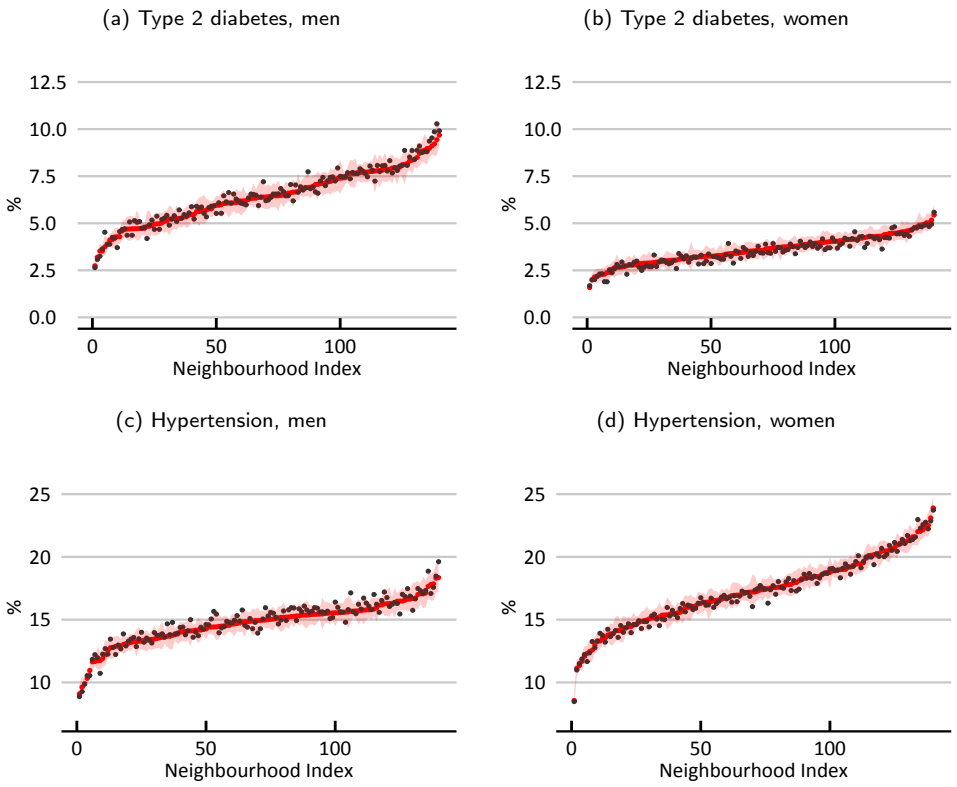


Figure A.4. TropISM model uncertainty in the simulated prevalence of type 2 diabetes and hypertension among men and women.

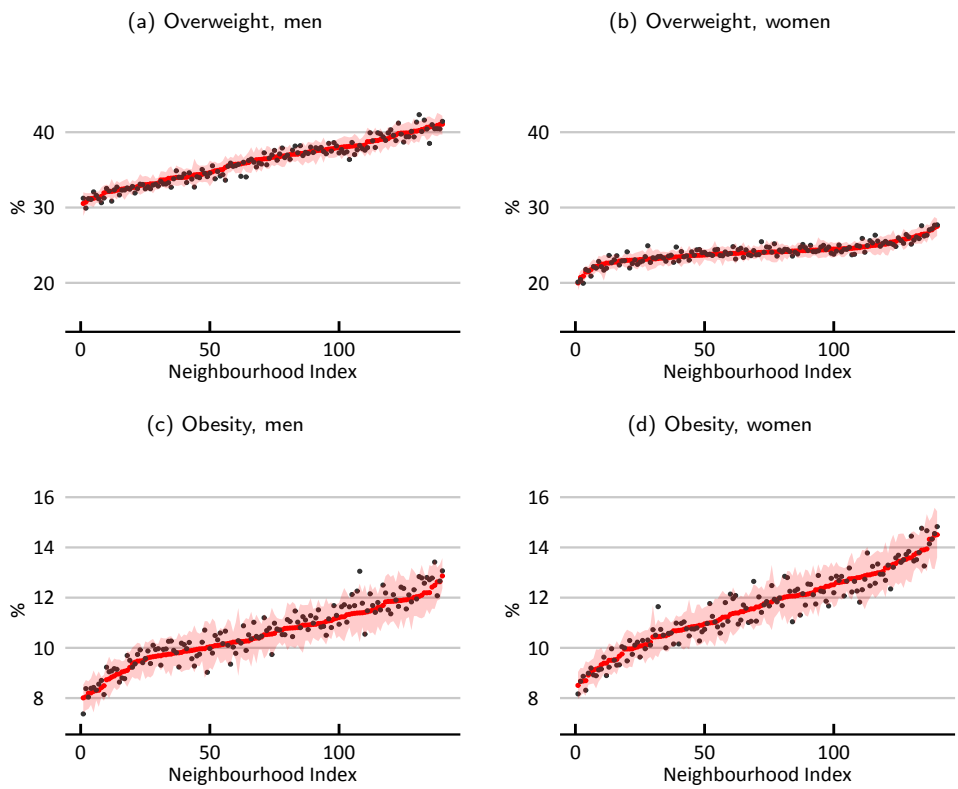


Figure A.5. TropISM model uncertainty in the simulated prevalence of overweight and obesity among men and women.

Appendix

B

Validation and Probabilistic Sensitivity Analysis of Forecast Incidence

B.1 Validation

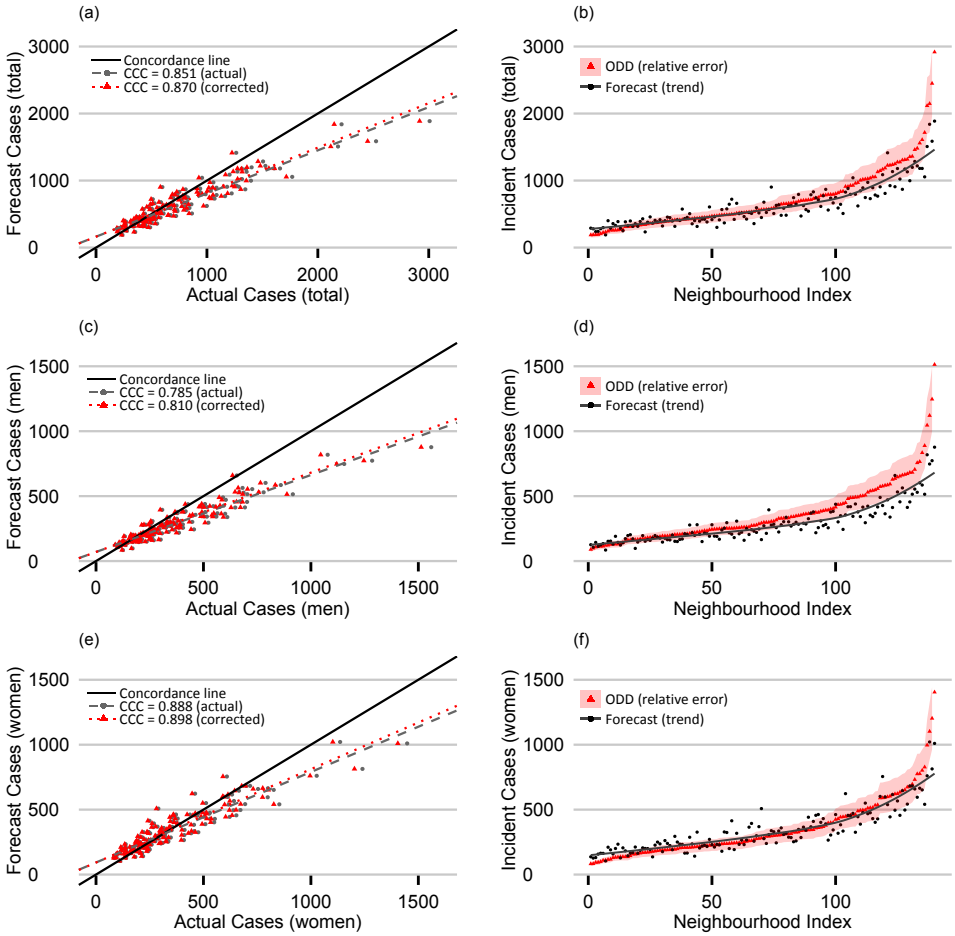


Figure B.1. Neighbourhood-level validation of the forecast number of incident cases of diabetes compared to the actual number of incident cases from 2006–2010, ascertained from the Ontario Diabetes Database (ODD). Panels (a), (c) & (e): accuracy and precision of the forecast number of cases. Panels (b), (d) & (f): difference between actual incident cases and the forecast number, ranked in ascending order by the actual number of cases in each neighbourhood.

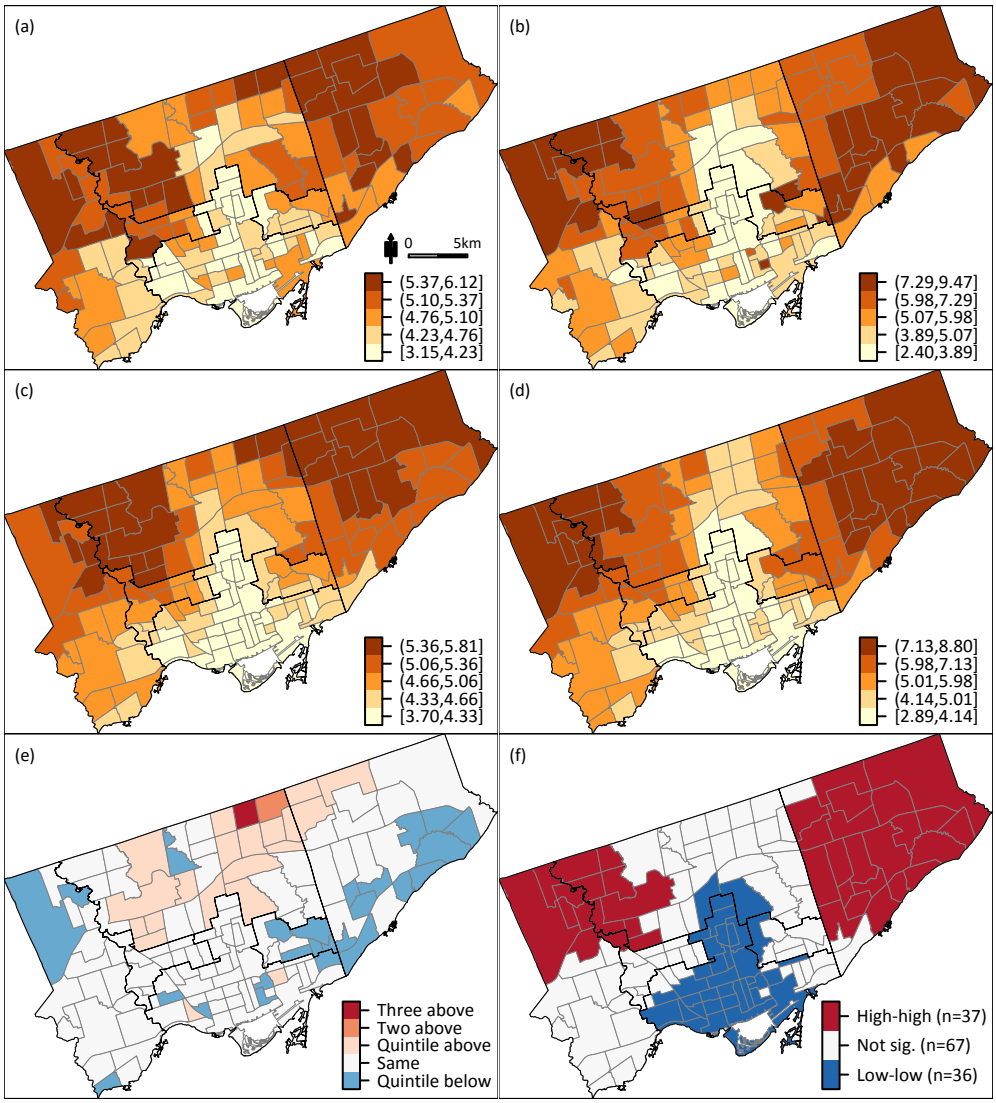


Figure B.2. Forecast five-year incidence (%) of diabetes at the neighbourhood level compared to the cumulative five-year incidence proportion (%) of diabetes from 2006–2010 (ascertained from the Ontario Diabetes Database (ODD) with removal of false positives). (a) Forecast incidence, raw (%); (b) ODD cumulative incidence, raw (%); (c) forecast incidence, spatially smoothed (%); (d) ODD cumulative incidence, spatially smoothed (%); (e) difference in the quintile rankings of spatially smoothed rates (forecast vs. ODD); (f) bivariate LISA map depicting areas having similarly high (or low) forecast and known incidence of diabetes.

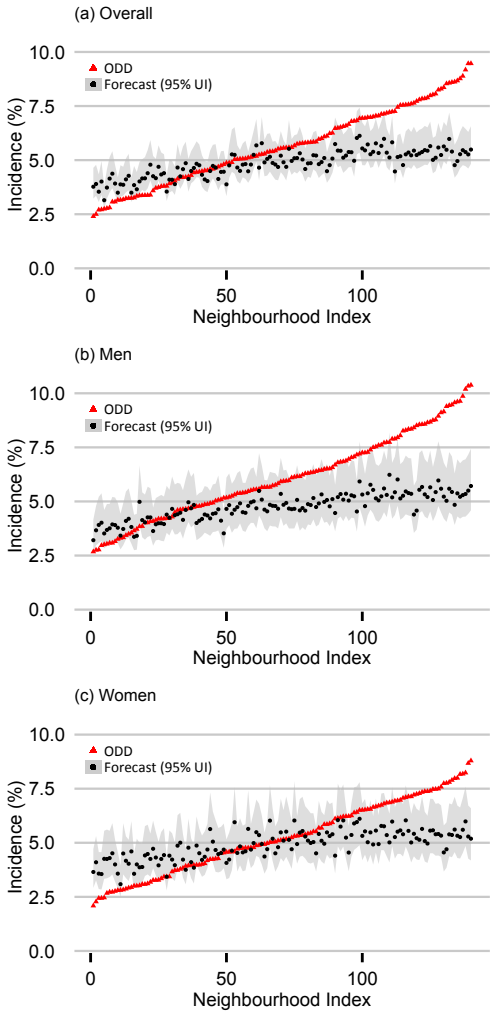


Figure B.3. Cumulative incidence proportion (%) of diabetes (2006–2010) ascertained from the Ontario Diabetes Database (ODD) compared to the five-year forecast incidence and uncertainty of diabetes across Toronto neighbourhoods.

Table B.1. Demographic characteristics of neighbourhoods where the known five-year incidence of diabetes fell below, within, or above the 95% uncertainty interval of forecast five-year incidence.

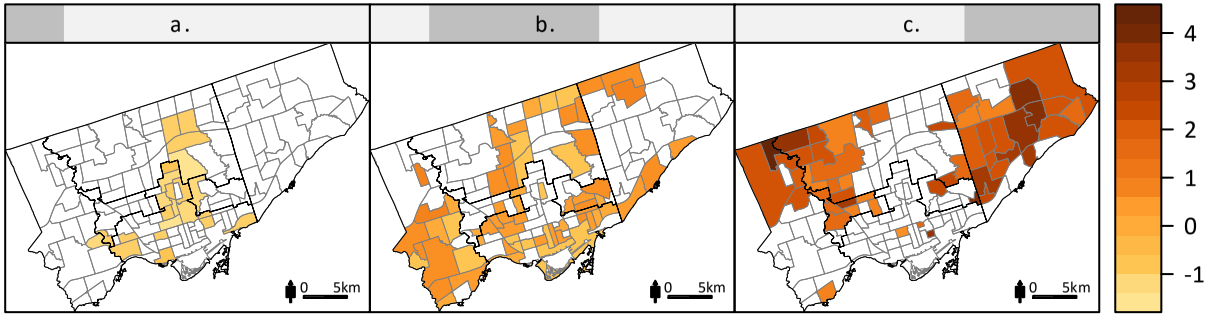
Characteristic	Below	Within	Above	LR*	p
Borough (n†, %)					
Etobicoke	1 (5.0)	13 (65.0)	6 (30.0)	—	
North York	4 (12.1)	14 (42.4)	15 (45.5)	—	
York	1 (10.0)	2 (20.0)	7 (70.0)	—	
Toronto	16 (36.4)	25 (56.8)	3 (6.8)	—	
East York	1 (12.5)	5 (62.5)	2 (25.0)	—	
Scarborough	0 (0)	6 (24.0)	19 (76.0)	—	
Overall	23 (16.4)	65 (46.4)	52 (37.1)	—	
% men ≥ 45	21.98^{a}	22.10^{a}	20.10^{b}	13.21	0.001
% women ≥ 65	9.67^{a,b}	10.95^{b}	9.20^{a}	10.94	0.004
% visible minority	28.15^{a}	35.51^{a}	53.82^{b}	42.39	< 0.001
% immigrant	37.37^{a}	46.53^{b}	57.56^{c}	49.95	< 0.001
% post-secondary education	65.76^{a}	57.02^{b}	53.15^{c}	31.11	< 0.001
% highest income quintile	22.79^{a}	16.30^{b}	11.24^{c}	67.82	< 0.001
% population growth‡	5.35	4.21	5.64	0.79	0.674

* Differences in average characteristics of neighbourhoods where the known cumulative incidence proportion was below, within, or above the 95% uncertainty interval of forecast incidence. Differences were tested using a linear spatial error regression model. Groups having different letters are significantly different at p = 0.05 using a Bonferroni adjustment. LR = likelihood ratio.

† Number of neighbourhoods within each borough.

‡ Average population growth between 2006 and 2011, in percentage points.

Figure B.4. Absolute difference (in percentage points) between the known cumulative incidence of diabetes ascertained from the Ontario Diabetes Database (ODD) and forecast incidence from the Diabetes Population Risk Tool (DPoRT) stratified by whether known incidence fell (a) below, (b) within, or (c) above forecast uncertainty.



Distribution			
Min:	-1.40	-0.55	0.99
Q1:	-1.17	-0.14	1.46
Median:	-0.84	0.31	2.30
Q3:	-0.67	0.62	2.78
Max:	-0.40	1.15	4.20

B.2 Probabilistic Sensitivity Analysis

Table B.2. Coefficient of determination (R^2) from regression models used to estimate neighbourhood-specific sensitivity indices for a probabilistic sensitivity analysis of DPoRT model parameters contributing to uncertainty about forecast type 2 diabetes incidence under different modelling scenarios.

	Baseline		10% Loss	
Borough	Median	(MAD)	Median	(MAD)
Men				
Etobicoke	0.9831	(0.0004)	0.9738	(0.0004)
North York	0.9833	(0.0003)	0.9769	(0.0005)
York	0.9836	(0.0003)	0.9775	(0.0003)
Toronto	0.9831	(0.0004)	0.9778	(0.0003)
East York	0.9837	(0.0001)	0.9777	(0.0006)
Scarborough	0.9838	(0.0002)	0.9768	(0.0003)
Women				
Etobicoke	0.9901	(0.0008)	0.9869	(0.0008)
North York	0.9906	(0.0005)	0.9880	(0.0010)
York	0.9911	(0.0003)	0.9886	(0.0003)
Old Toronto	0.9906	(0.0006)	0.9872	(0.0010)
East York	0.9906	(0.0004)	0.9872	(0.0004)
Scarborough	0.9914	(0.0002)	0.9889	(0.0008)

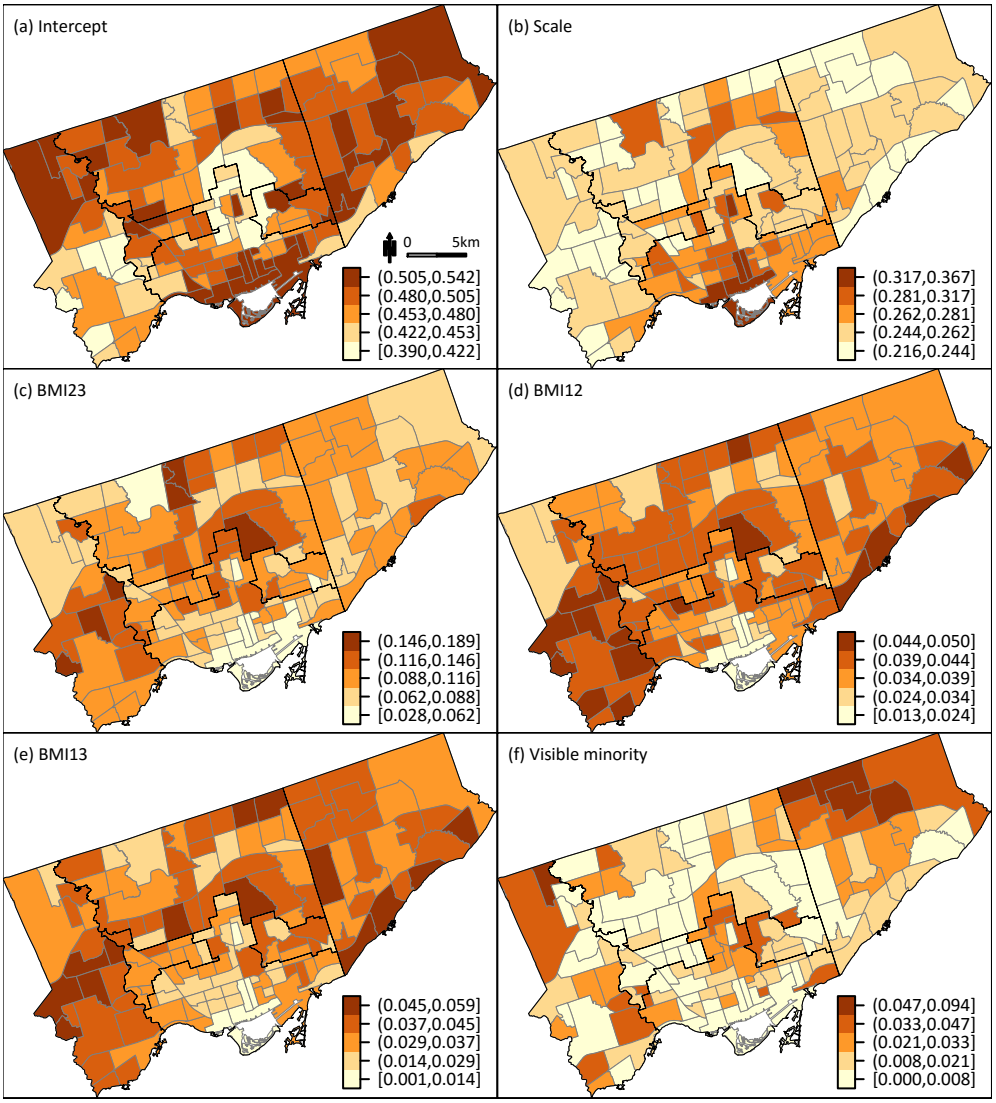


Figure B.5. Geographic distribution of the sensitivity index associated with DPoRT model parameters used to forecast diabetes incidence among men under the baseline scenario: (a) intercept parameter, (b) scale parameter, (c) overweight men aged 45+, (d) healthy weight men aged 20–44, (e) overweight men aged 20–44, and (f) visible minorities.

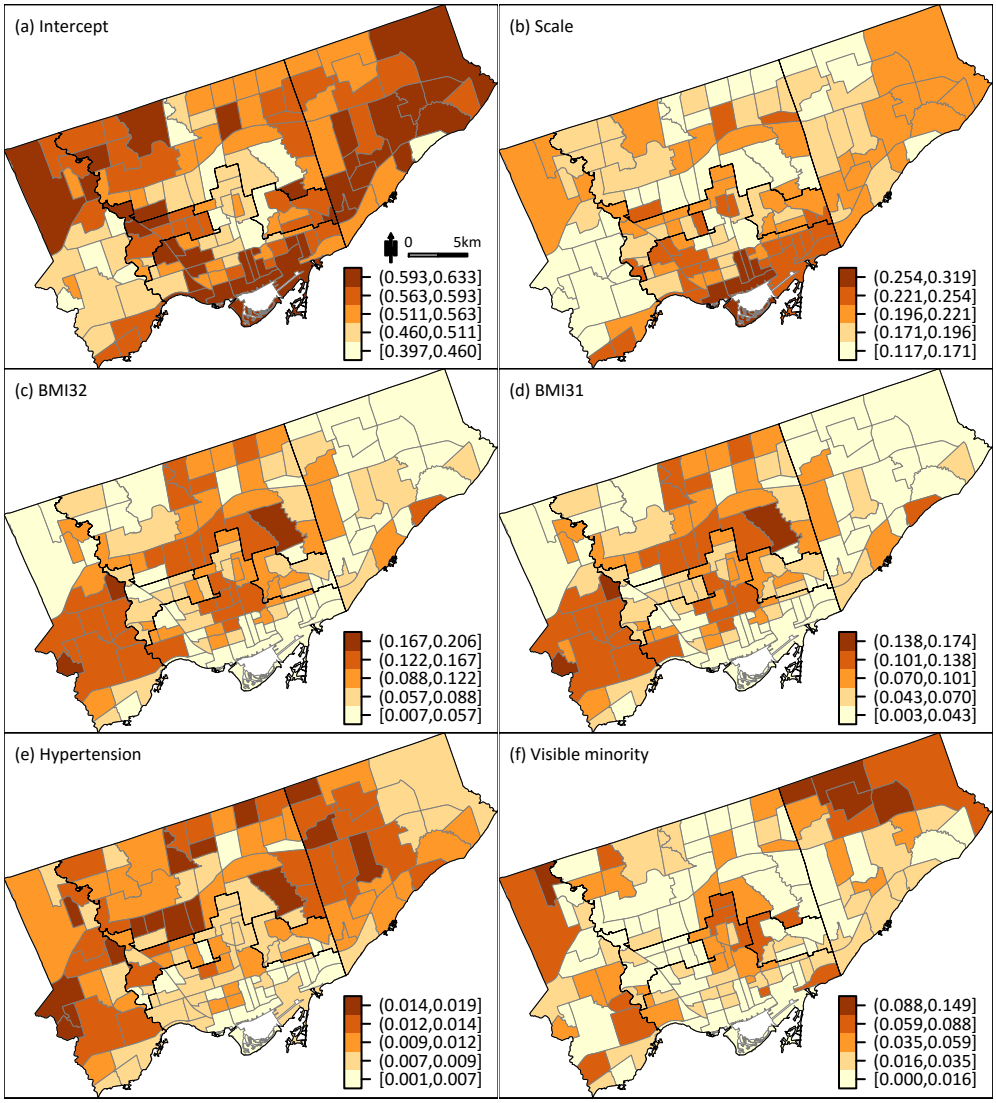


Figure B.6. Geographic distribution of the sensitivity index associated with DPoRT model parameters used to forecast diabetes incidence among women under the baseline scenario: (a) intercept parameter, (b) scale parameter, (c) normal weight women aged 65+, (d) underweight women aged 65+, (e) hypertension, and (f) visible minorities.

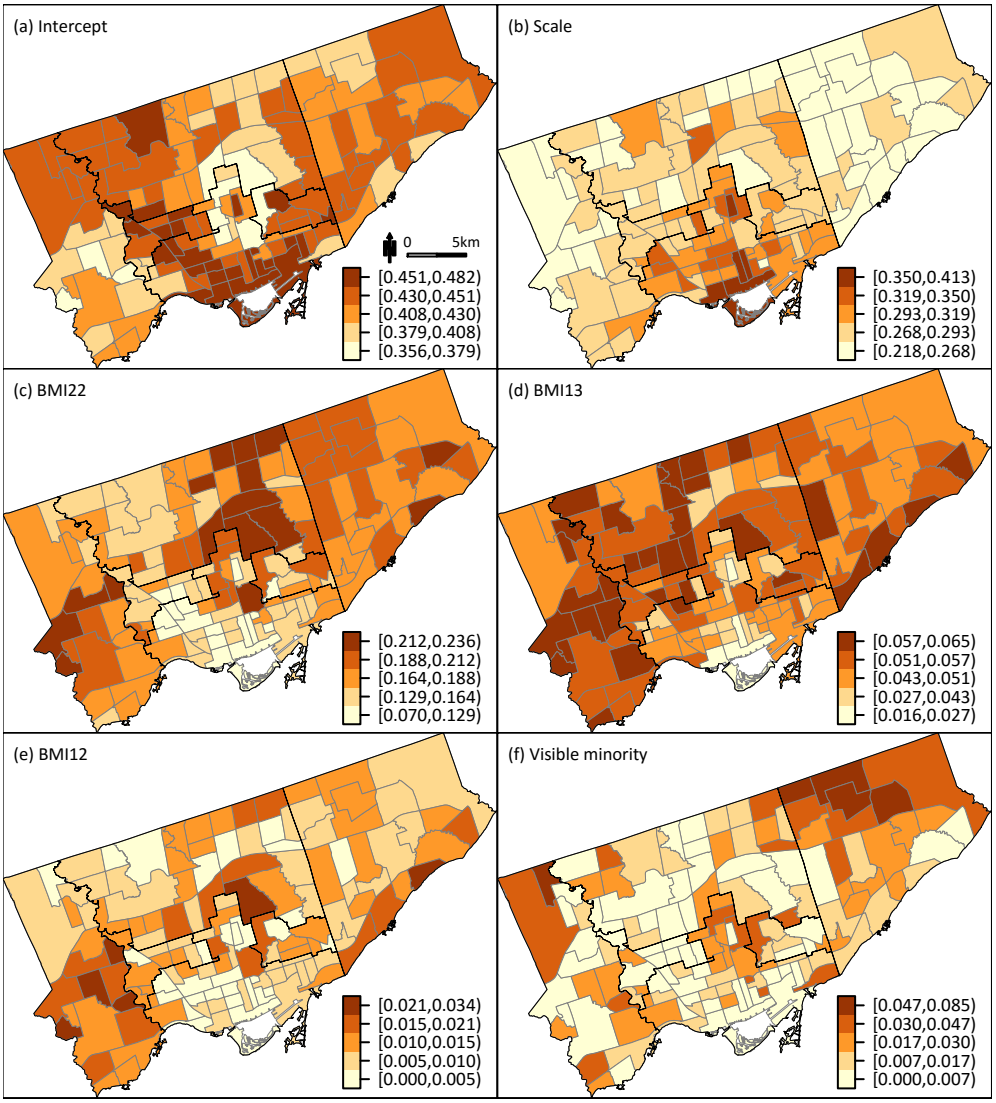


Figure B.7. Geographic distribution of the sensitivity index associated with DPoRT model parameters used to forecast diabetes incidence among men under the 10% weight loss scenario: (a) intercept parameter, (b) scale parameter, (c) healthy weight men aged 45+, (d) overweight men aged 20–44, (e) healthy weight men aged 20–44, and (f) visible minorities.

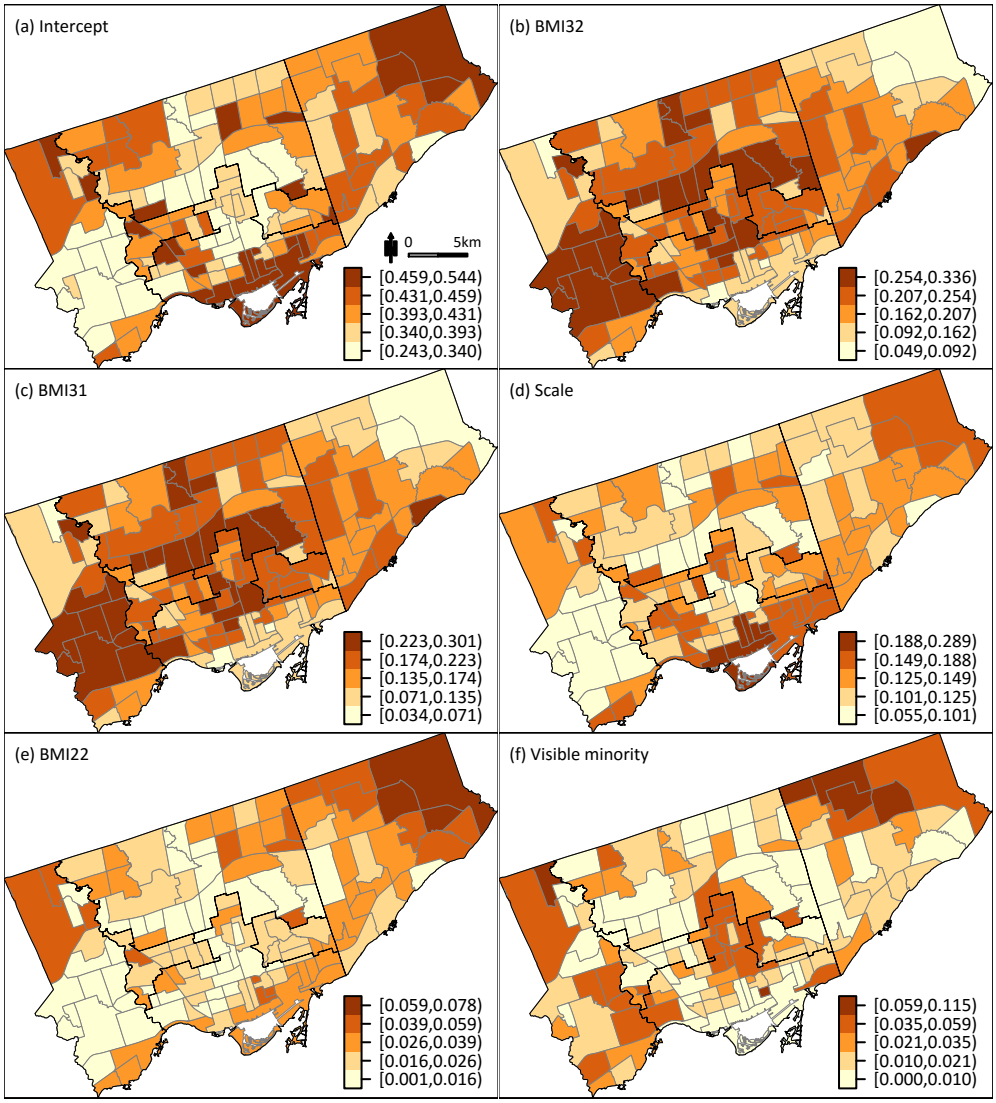


Figure B.8. Geographic distribution of the sensitivity index associated with DPoRT model parameters used to forecast diabetes incidence among women under the 10% weight loss scenario: (a) intercept parameter, (b) healthy weight women aged 65+, (c) underweight women aged 65+, (d) scale parameter, (e) healthy weight women aged 45–64, and (f) visible minorities.

Appendix

C

Spatial Accessibility using Simulated vs. Known Demand for Diabetes Education Programs

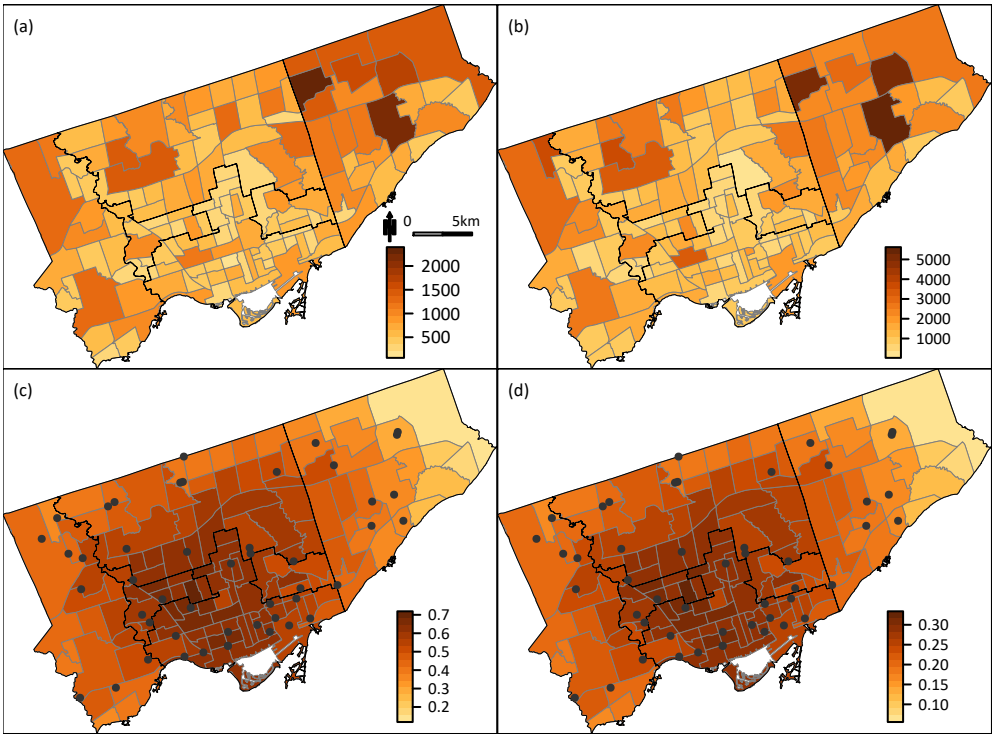


Figure C.1. Comparison of accessibility to diabetes education programs using simulated demand from TropISM vs. known demand ascertained from the Ontario Diabetes Database (ODD) in 2005. (a) Total simulated cases of type 2 diabetes; (b) known number of cases (ODD); (c) accessibility using simulated demand; and (d) accessibility using known demand.

Appendix

D

R Syntax

D.1 Forecasting Diabetes Incidence using DPoRT

Listing D.1. Functions to forecast diabetes incidence using the Diabetes Population Risk Tool. Functions were packaged into a personal R library for convenient loading.

```
## --- Required libraries ----------
require(dplyr)

## --- Functions --------------------
dport.centredCovar <- function (df, means, join.by)
{
  ## This function uses the data frame produced by dport.updateBMI as
  ## input to compute the mean centred versions of the covariates used
  ## to forecast the future risk of type 2 diabetes using the DPoRT
  ## model.
  .tmp <- left_join(df, means, by = join.by)
  return(
    mutate(
      .tmp, minority.c = MINORITY - minority,
      immigrant.c = IMMIGRANT - immigrant,
      postsecedu.c = POSTSECEDU - postsecedu,
      incomeq5.c = INCOMEQ5 - incomeq5, smoker.c = SMOKER - smoker,
      hypertension.c = HYPERTENSION - hypertension,
```

```
      heartdisease.c = HEARTDISEASE - heartdisease,
      mbmi12.c = ifelse(MALE == 1, MBMI12 - mbmi12, 0),
      mbmi13.c = ifelse(MALE == 1, MBMI13 - mbmi13, 0),
      mbmi14.c = ifelse(MALE == 1, MBMI14 - mbmi14, 0),
      mbmi15.c = ifelse(MALE == 1, MBMI15 - mbmi15, 0),
      mbmi21.c = ifelse(MALE == 1, MBMI21 - mbmi21, 0),
      mbmi22.c = ifelse(MALE == 1, MBMI22 - mbmi22, 0),
      mbmi23.c = ifelse(MALE == 1, MBMI23 - mbmi23, 0),
      mbmi24.c = ifelse(MALE == 1, MBMI24 - mbmi24, 0),
      mbmi25.c = ifelse(MALE == 1, MBMI25 - mbmi25, 0),
      fbmi12.c = ifelse(MALE == 0, FBMI12 - fbmi12, 0),
      fbmi13.c = ifelse(MALE == 0, FBMI13 - fbmi13, 0),
      fbmi14.c = ifelse(MALE == 0, FBMI14 - fbmi14, 0),
      fbmi15.c = ifelse(MALE == 0, FBMI15 - fbmi15, 0),
      fbmi16.c = ifelse(MALE == 0, FBMI16 - fbmi16, 0),
      fbmi21.c = ifelse(MALE == 0, FBMI21 - fbmi21, 0),
      fbmi22.c = ifelse(MALE == 0, FBMI22 - fbmi22, 0),
      fbmi23.c = ifelse(MALE == 0, FBMI23 - fbmi23, 0),
      fbmi24.c = ifelse(MALE == 0, FBMI24 - fbmi24, 0),
      fbmi25.c = ifelse(MALE == 0, FBMI25 - fbmi25, 0),
      fbmi26.c = ifelse(MALE == 0, FBMI26 - fbmi26, 0),
      fbmi31.c = ifelse(MALE == 0, FBMI31 - fbmi31, 0),
      fbmi32.c = ifelse(MALE == 0, FBMI32 - fbmi32, 0),
      fbmi33.c = ifelse(MALE == 0, FBMI33 - fbmi33, 0),
      fbmi34.c = ifelse(MALE == 0, FBMI34 - fbmi34, 0),
      fbmi35.c = ifelse(MALE == 0, FBMI35 - fbmi35, 0),
      fbmi36.c = ifelse(MALE == 0, FBMI36 - fbmi36, 0)))
}

dport.predictRisk <- function (centredDf, dport.parm, horizon = 5)
{
  ## Forecast the risk of type 2 diabetes using the DPoRT 2.0 model.
  .base <- select(centredDf, PID:MALE, AGEDPORTM, AGEDPORTF,
                  minority.c:fbmi36.c)
  .mp <- dport.parm$male
  .fp <- dport.parm$female
  return(
    mutate(
      .base, mu = ifelse(MALE == 1, .mp$B0 +
      .mp$HBP * hypertension.c + .mp$VISMIN * minority.c +
```

```
      .mp$CVD * heartdisease.c + .mp$SMOKER * smoker.c +
      .mp$POSTSEC * postsecedu.c + .mp$INCOME * incomeq5.c +
      .mp$BMI12 * mbmi12.c + .mp$BMI13 * mbmi13.c +
      .mp$BMI14 * mbmi14.c + .mp$BMI15 * mbmi15.c +
      .mp$BMI21 * mbmi21.c + .mp$BMI22 * mbmi22.c +
      .mp$BMI23 * mbmi23.c + .mp$BMI24 * mbmi24.c +
      .mp$BMI25 * mbmi25.c,
      .fp$B0 + .fp$HBP * hypertension.c + .fp$VISMIN * minority.c +
      .fp$IMMIGR * immigrant.c + .fp$POSTSEC * postsecedu.c +
      .fp$BMI12 * fbmi12.c + .fp$BMI13 * fbmi13.c +
      .fp$BMI14 * fbmi14.c + .fp$BMI15 * fbmi15.c +
      .fp$BMI16 * fbmi16.c + .fp$BMI21 * fbmi21.c +
      .fp$BMI22 * fbmi22.c + .fp$BMI23 * fbmi23.c +
      .fp$BMI24 * fbmi24.c + .fp$BMI25 * fbmi25.c +
      .fp$BMI26 * fbmi26.c + .fp$BMI31 * fbmi31.c +
      .fp$BMI32 * fbmi32.c + .fp$BMI33 * fbmi33.c +
      .fp$BMI34 * fbmi34.c + .fp$BMI35 * fbmi35.c +
      .fp$BMI36 * fbmi36.c),
     m = ifelse(MALE == 1, (log(365.242 * horizon) - mu) /
      .mp$SCALE, (log(365.242 * horizon) - mu)/.fp$SCALE),
      p = 1 - exp(-exp(m))))
}

dport.wtloss <- function (df, type, amount, stddev, subset = FALSE, subset.prop)
{
  ## This function is used to develop weight loss scenarios for
  ## projecting the future risk of diabetes using the Diabetes
  ## Population Risk Tool (DPoRT). Risk is projected at the
  ## neighbourhood level for metroplitan Toronto using a simulated
  ## population. The simulated population (TropISM) was developed using
  ## spatial microsimulation (combinatorial optimization via simulated
  ## annealing).

  if (!type %in% c("prop", "mean")) {
    stop("Only␣proportionate␣or␣average␣(mean)␣weight␣losses␣are␣allowed.")
  }
  if (type == "prop") {
    if (!(amount > 0 & amount < 1)) {
      stop("Proportionate␣weight␣losses␣must␣lie␣between␣0␣and␣1")
    }
```

```
  if (!subset) {
    LOSS <- 1
  }
  else {
    .riskn <- sum(group_size(group_by(df, BMICLS))[-c(1, 2)])
    df$LOSS <- with(df, ifelse(!BMICLS %in% c("< 23", "23 to < 25"),
                    rbinom(.riskn, 1, subset.prop), 0))
  }
  .tmp <- mutate(
    df,
    BODYWT2 = ifelse(!BMICLS %in% c("< 23", "23 to < 25"),
                     BODYWT - (BODYWT * amount * LOSS), BODYWT),
    BMI2 = BODYWT2/(BODYHT^2),
    BMICLS2 = ifelse(BMI2 >= 0 & BMI2 < 23, 1, 0),
    BMICLS2 = ifelse(BMI2 >= 23 & BMI2 < 25, 2, BMICLS2),
    BMICLS2 = ifelse(BMI2 >= 25 & BMI2 < 30, 3, BMICLS2),
    BMICLS2 = ifelse(BMI2 >= 30 & BMI2 < 35, 4, BMICLS2),
    BMICLS2 = ifelse(BMI2 >= 35 & BMI2 < 60, 5, BMICLS2),
    BMICLS2 = ifelse(is.na(BMI2), 6, BMICLS2),
    BMICLS2 = factor(BMICLS2, labels = levels(BMICLS)))
  return(.tmp)
  }
  else {
    .loss <- rnorm(
      nrow(subset(df, !BMICLS %in% c("< 23", "23 to < 25"))),
                 mean = amount, sd = stddev)
    .tmp <- mutate(
      df,
      WTLOSS = ifelse(!BMICLS %in% c("< 23", "23 to < 25"), .loss, 0),
      BODYWT2 = BODYWT - WTLOSS, BMI2 = BODYWT2/(BODYHT^2),
      BMICLS2 = ifelse(BMI2 >= 0 & BMI2 < 23, 1, 0),
      BMICLS2 = ifelse(BMI2 >= 23 & BMI2 < 25, 2, BMICLS2),
      BMICLS2 = ifelse(BMI2 >= 25 & BMI2 < 30, 3, BMICLS2),
      BMICLS2 = ifelse(BMI2 >= 30 & BMI2 < 35, 4, BMICLS2),
      BMICLS2 = ifelse(BMI2 >= 35 & BMI2 < 60, 5, BMICLS2),
      BMICLS2 = ifelse(is.na(BMI2), 6, BMICLS2),
      BMICLS2 = factor(BMICLS2, labels = levels(BMICLS)))
    return(.tmp)
  }
}
```

```
dport.updateBMI <- function (wtLossDf)
{
  ## Uses data frame output from 'dport.wtloss' as input to define a
  ## series of indicator variables that define age-sex specific BMI
  ## categories. The output from this function is required to compute
  ## mean centred covariates needed to forecast risk of diabetes.
  .tmp <- select(wtLossDf, PID:BODYHT, BODYWT = BODYWT2, BMI = BMI2,
                 BMICLS = BMICLS2, MINORITY:HEARTDISEASE)
  .tmp <- mutate(
    .tmp,
    MBMI12 = ifelse(MALE == 1 & AGEDPORTM == "20-44" & BMICLS == "23 to < 25", 1, 0),
    MBMI13 = ifelse(MALE == 1 & AGEDPORTM == "20-44" & BMICLS == "25 to < 30", 1, 0),
    MBMI14 = ifelse(MALE == 1 & AGEDPORTM == "20-44" & BMICLS == "30 to < 35", 1, 0),
    MBMI15 = ifelse(MALE == 1 & AGEDPORTM == "20-44" & BMICLS == ">= 35", 1, 0),
    MBMI21 = ifelse(MALE == 1 & AGEDPORTM == "45+" & BMICLS == "< 23", 1, 0),
    MBMI22 = ifelse(MALE == 1 & AGEDPORTM == "45+" & BMICLS == "23 to < 25", 1, 0),
    MBMI23 = ifelse(MALE == 1 & AGEDPORTM == "45+" & BMICLS == "25 to < 30", 1, 0),
    MBMI24 = ifelse(MALE == 1 & AGEDPORTM == "45+" & BMICLS == "30 to < 35", 1, 0),
    MBMI25 = ifelse(MALE == 1 & AGEDPORTM == "45+" & BMICLS == ">= 35", 1, 0),
    FBMI12 = ifelse(MALE == 0 & AGEDPORTF == "20-44" & BMICLS == "23 to < 25", 1, 0),
    FBMI13 = ifelse(MALE == 0 & AGEDPORTF == "20-44" & BMICLS == "25 to < 30", 1, 0),
    FBMI14 = ifelse(MALE == 0 & AGEDPORTF == "20-44" & BMICLS == "30 to < 35", 1, 0),
    FBMI15 = ifelse(MALE == 0 & AGEDPORTF == "20-44" & BMICLS == ">= 35", 1, 0),
    FBMI16 = ifelse(MALE == 0 & AGEDPORTF == "20-44" & BMICLS == "Unknown", 1, 0),
    FBMI21 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == "23 to < 25", 1, 0),
    FBMI22 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == "23 to < 25", 1, 0),
    FBMI23 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == "25 to < 30", 1, 0),
    FBMI24 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == "30 to < 35", 1, 0),
    FBMI25 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == ">= 35", 1, 0),
    FBMI26 = ifelse(MALE == 0 & AGEDPORTF == "45-64" & BMICLS == "Unknown", 1, 0),
    FBMI31 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == "23 to < 25", 1, 0),
    FBMI32 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == "23 to < 25", 1, 0),
    FBMI33 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == "25 to < 30", 1, 0),
    FBMI34 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == "30 to < 35", 1, 0),
    FBMI35 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == ">= 35", 1, 0),
    FBMI36 = ifelse(MALE == 0 & AGEDPORTF == "65+" & BMICLS == "Unknown", 1, 0))
    return(.tmp)
}
```

```
dport.summary <- function (forecastDf)
{
  ## Summarizes predicted risks
  .o <- select(summarize(group_by(forecastDf, HOODID), mean(p) * 100),
               HOODID, Overall = 2)
  .s <- select(summarize(group_by(forecastDf, HOODID, MALE),
               mean(p) * 100), HOODID, MALE, f = 3)
  .sa <- select(summarize(group_by(forecastDf, HOODID, MALE,
                AGEDPORTF), mean(p) * 100), HOODID, MALE, AGEDPORTF,
                f = 4)
  .m <- select(summarize(group_by(forecastDf, HOODID, MALE,
               AGEDPORTM), mean(p) * 100), HOODID, MALE, AGEDPORTM,
               f = 4)
  .tmp <- select(inner_join(.o, subset(.s, MALE == 1), by = "HOODID"),
                 HOODID, Overall, Male = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 1 & AGEDPORTF == "20-44"),
                 by = "HOODID"), HOODID:Male, Male.20to44 = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 1 & AGEDPORTF == "45-64"),
                 by = "HOODID"), HOODID:Male.20to44, Male.45to64 = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 1 & AGEDPORTF == "65+"),
                 by = "HOODID"), HOODID:Male.45to64, Male.65plus = f)
  .tmp <- select(inner_join(.tmp, subset(.m, MALE == 1 & AGEDPORTM == "45+"),
                 by = "HOODID"), HOODID:Male.65plus, Male.45plus = f)
  .tmp <- select(inner_join(.tmp, subset(.s, MALE == 0), by = "HOODID"),
                 HOODID:Male.45plus, Female = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 0 & AGEDPORTF == "20-44"),
                 by = "HOODID"), HOODID:Female, Female.20to44 = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 0 & AGEDPORTF == "45-64"),
                 by = "HOODID"), HOODID:Female.20to44, Female.45to64 = f)
  .tmp <- select(inner_join(.tmp, subset(.sa, MALE == 0 & AGEDPORTF == "65+"),
                 by = "HOODID"), HOODID:Female.45to64, Female.65plus = f)
  return(as.data.frame(.tmp))
}

dport.uncertainty <- function(x)
{
  ## Summarizies multiple runs of DPoRT model from a
  ## probabilistic sensitivity analysis
  mu <- mean(x)
  quants <- quantile(x, c(0.025, 0.25, 0.5, 0.75, 0.975))
```

```
  result <- data.frame(mu, rbind(quants))
  names(result) <- c("Mean", "p2.5", "p25", "p50", "p75", "p97.5")
  return(result)
}

## Required data, packaged into DPoRT library
data(riskpop)              ## baseline population at risk
data(tropism.prevalence)   ## average prevalence of risk factors by sex
data(dport.parm)           ## DPoRT 2.0 model parameters
data(tropism.hoodpop)      ## neighbourhood population counts, metropolitan Toronto
```

Listing D.2. Sample code used to forecast diabetes incidence under the baseline scenario (no weight loss) and the 10% weight loss scenario.

```
## --- Libraries ----------------------
require(dplyr)
require(DPoRT)

## --- Data ---------------------------
data(riskpop)
data(tropism.prevalence)
data(dport.parm)
data(tropism.hoodpop)

## --- Scenarios ----------------------
## Baseline
baseline <- riskpop
baseline.centred <- dport.centredCovar(
  baseline, tropism.prevalence, join.by = "MALE")
baseline.risk <- dport.predictRisk(baseline.centred, dport.parm)
forecast.baseline <- dport.summary(baseline.risk)

## 10% weight loss
wtloss.10pct <- riskpop
wtloss.10pct <- dport.wtloss(wtloss.10pct, type = "prop", amount = 0.1)
wtloss.10pct.bmi <- dport.updateBMI(wtloss.10pct)
wtloss.10pct.centred <- dport.centredCovar(
  wtloss.10pct.bmi, tropism.prevalence, join.by = "MALE")
wtloss.10pct.risk <- dport.predictRisk(wtloss.10pct.centred, dport.parm)
forecast.10pct <- dport.summary(wtloss.10pct.risk)
```

Listing D.3. R code used to conduct 1000 runs of the DPoRT model for a probabilistic sensitivity analysis of the baseline forecast of diabetes incidence.

```
## --- Libraries ----------------------
require(data.table)
require(dplyr)
require(DPoRT)
require(foreach)
require(doParallel)

## --- Data ---------------------------
## simulated TropISM data
data(riskpop)
data(tropism.prevalence)
## data(dport.parm)

## contains same info as dport.parm, only 1000 samples (normally distributed)
## from DPoRT 2.0, treating the arameter estimate as the mean and the StdErr
## as the StdDev for random sampling (Latin Hypercube Sampling in SaSAT)
load("dport_baseline_sa_samples_combined.rda")

## setup for parallel computing, 1000 replicates in data frame
replicates <- seq(1:nrow(sab$male))

## Sensitivity analysis, baseline scenario
baseline <- riskpop
baseline.centred <- dport.centredCovar(
  baseline, tropism.prevalence, join.by = "MALE")

## use 8 available cores on CPU
cl <- makeCluster(8)
registerDoParallel(cl)
loop.start <- Sys.time()

forecast <- foreach(i = replicates,.packages = c("DPoRT")) %dopar% {
  dport.parm <- lapply(sab, function(x) x[i, , drop = FALSE])
  baseline.risk <- dport.predictRisk(baseline.centred, dport.parm)
  forecast <- dport.summary(baseline.risk)
}
```

```
loop.end <- Sys.time()
print(loop.end - loop.start)
stopCluster(cl)
rm(cl)

baselinePSA <- data.table::rbindlist(forecast)
baselinePSA$replicate <- sort(rep(seq(1, 1000), 140))
save(baselinePSA, file = "DPoRT-Projected-Risk-Baseline-PSA.rda")
```

D.2 Spatial Accessibility to Health Promotion Programs

Listing D.4. Sample code used to estimate spatial accessibility to diabetes education programs using a two-step floating catchment area model.

```
## --- Libraries ----------------------
require(dplyr)

## --- Functions ----------------------
gaussian.beta <- function(distance)
{
  # 'distance' is maximum distance beyond which few interaction
  # between locations i and j occur
  d <- distance**2
  b <- abs(d/log(0.01))
  return(b)
}

gaussian.decay <- function(distance, beta)
{
  w <- exp(-distance^2/beta)
  return(w)
}

distance.decay <- function(bounds, limit, type = "gaussian")
{
  n.zones <- max(dim(bounds))
  midpts <- bounds[, 1] + (bounds[, 2] - bounds[, 1]) / 2
```

```
  if(type == "gaussian") {
    b <- round(gaussian.beta(limit))
    w <- round(gaussian.decay(midpts, b), 3)
  }
  zones <- c(c(1:n.zones), 0)
  w <- c(w, 0)
  dcy <- data.frame(zone = zones, decay = w)
  return(dcy)
}

g2sfca.summary <- function(AIdx)
{
  .mdn <- median(AIdx)
  .mad <- mad(AIdx, constant = 1)
  result <- data.frame(Median = .mdn, MAD = .mad)
  return(result)
}

## --- Data ------------------------
data(dep2sfca)
data(tropism_hood_demand)

## restrict analysis to subset of programs (n = 48)
programs <- subset(dep2sfca$diabetes.programs, SubPopnOnly == 0 & PatientsOnly == 0)
distances <- subset(dep2sfca$diabetes.distances, depid %in% programs$DEPID)
t2d.demand <- select(demand[["Type2Diabetes"]], hoodid, demand = t2d)
services <- data.frame(depid = programs$DEPID, services = rep(1, nrow(programs)))

## --- G2SFCA Model ----------------
## subset of programs (n = 48), beta = 440
gaussian.Rj <- select(distances, hoodid, depid, minutes) %>%
    mutate(decay = round(gaussian.decay(distance = minutes, beta = 440), 3)) %>%
    mutate(decay = ifelse(minutes > 30, 0, decay)) %>%
    merge(t2d.demand, by = "hoodid") %>%
    mutate(scaled.demand = demand * decay) %>%
    group_by(depid) %>%
    summarise(demand = sum(scaled.demand)) %>%
    left_join(services, by = "depid") %>%
    mutate(R = services/demand)
```

```
gaussian.Ai <- select(gaussian.Rj, depid, R) %>%
    left_join(distances, by = "depid") %>%
    mutate(decay = round(gaussian.decay(distance = minutes, beta = 440), 3)) %>%
    select(hoodid, depid, minutes, decay, R) %>%
    mutate(decay = ifelse(minutes > 30, 0, decay)) %>%
    mutate(scaled.R = R * decay) %>%
    group_by(hoodid) %>%
    summarise(A = sum(scaled.R)) %>%
    mutate(SPAR = A/mean(A), popPerPgm = 1/A)

access.subset.g2sfca.440 <- data.frame(gaussian.Ai) %>% mutate(A1000 = A * 1000)

## check
a <- data.frame(gaussian.Ai)
sum(t2d.demand$demand * a$A)
```

Listing D.5. R code used to conduct 1000 runs of the two-step floating catchment area model for a probabilistic sensitivity analysis of spatial accessibility to diabetes education programs.

```
## --- Libraries ----------------------
require(dplyr)
require(foreach)
require(doParallel)

## --- Functions ----------------------
ai.lo <- function(x) {quantile(x, probs = 0.025)}
ai.hi <- function(x) {quantile(x, probs = 0.975)}

## --- Data ---------------------------
data(dep2sfca)
data(tropism_hood_demand)
trop.demand <- demand[[5]]
rm(demand)

dtmp <- dep2sfca$diabetes.distances
ptmp <- dep2sfca$diabetes.programs

## limit programs to those broadly available
programs <- subset(ptmp, SubPopnOnly == 0 & PatientsOnly == 0)
distances <- subset(dtmp, depid %in% programs$DEPID)
services <- data.frame(depid = programs$DEPID, services = rep(1, nrow(programs)))
t2d.demand <- select(trop.demand, hoodid, demand = t2d)
saparm <- dep2sfca$diabetes.uncertainty

## --- Uncertainty --------------------
replicates <- seq(1, nrow(saparm))
cl <- makeCluster(8)

## setup || computing
registerDoParallel(cl)
loop.start <- Sys.time()

access <- foreach(i = replicates, .packages = c("dplyr"), .combine = rbind) %dopar% {
  sa <- saparm[i, ]
  .trvltime <- sa$trvltime
```

```
  .demand <- sa$demand
  .decay <- sa$decay
  .d0 <- sa$d0

  .distances <- select(distances, hoodid, depid, minutes) %>%
    mutate(minutes = round(minutes + (.trvltime * minutes), 2))
  .t2ddemand <- mutate(t2d.demand, demand = demand + .demand)

  Rj <- select(.distances, hoodid, depid, minutes) %>%
    mutate(decay = round(gaussian.decay(distance = minutes, beta = .decay), 3)) %>%
    mutate(decay = ifelse(minutes > .d0, 0, decay)) %>%
    merge(.t2ddemand, by = "hoodid") %>%
    mutate(scaled.demand = demand * decay) %>%
    group_by(depid) %>%
    summarise(demand = sum(scaled.demand)) %>%
    left_join(services, by = "depid") %>%
    mutate(R = services/demand)

  Ai <- select(Rj, depid, R) %>%
    left_join(.distances, by = "depid") %>%
    mutate(decay = round(gaussian.decay(distance = minutes, beta = .decay), 3)) %>%
    select(hoodid, depid, minutes, decay, R) %>%
    mutate(scaled.R = R * decay) %>%
    group_by(hoodid) %>%
    summarise(A = sum(scaled.R)) %>%
    mutate(SPAR = A/mean(A), popPerDep = 1/A)

  access <- Ai %>% mutate(replicate = i)
}

loop.end <- Sys.time()
print(loop.end - loop.start)
stopCluster(cl)
rm(cl)

access <- data.frame(access)

access.ui <- group_by(access, hoodid) %>%
  summarise(lo.Ai = ai.lo(A),
            mdn.Ai = median(A),
```

```
            hi.Ai = ai.hi(A)) %>%
  data.frame()

popPer.ui <- group_by(access, hoodid) %>%
  summarise(lo.popPer = ai.lo(popPerDep),
            mdn.popPer = median(popPerDep),
            hi.popPer = ai.hi(popPerDep)) %>%
  data.frame()

## --- Output ----------------------
save(access, file = "deppsa.rda")
```