EXAMINING THE RELATIONSHIP BETWEEN ACCESS TO HEALTHCARE AND MARGINALIZATION

IN TORONTO

by

Julia DiMartella-Orsi

Bachelor of Arts, University of Toronto, 2019

A Major Research Paper

presented to Ryerson University

in partial fulfilment of the

requirements for the degree of

Master of Spatial Analysis

in the program of

Spatial Analysis

Toronto, Ontario, Canada 2020

© Julia DiMartella-Orsi, 2020
Author’s Declaration

I hereby declare that I am the sole author of this MRP. This is a true copy of the MRP, including any required final revisions.

I authorize Ryerson University to lend this MRP to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this MRP by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my MRP may be made electronically available to the public.
Examining the Relationship Between Access to Healthcare and Marginalization in Toronto

Master of Spatial Analysis 2020

Julia DiMartella-Orsi
Spatial Analysis at Ryerson University

Abstract

This study analyzes the association between healthcare accessibility and marginalization using the city of Toronto as a case study. Access to healthcare is determined by the 2020 Statistics Canada Proximity Measures Database, and marginalization is measured using the four dimensions of marginalization in the 2016 Ontario Marginalization Index. Techniques such as choropleth mapping and Univariate global and local Moran’s I were applied, followed by an OLS and Lag-error regression. The results indicated that as material deprivation increased healthcare access decreased, and that as dependency increased access increased - however they were not statistically significant. Conversely, ethnic concentration and residential instability were statistically significant, and increased as accessibility to healthcare increased. Based on these results, it was concluded that greater marginalization led to greater healthcare access in Toronto. These results contradicted one of this study’s working hypotheses and recent research on healthcare accessibility, that find that as marginalization increases healthcare accessibility decreases.
Acknowledgements

I would like to thank my supervisors Dr. Evan Cleave and Dr. Jonathan Cinnamon for their time, effort, and guidance in helping me complete this research paper. I would also like to thank my family for their support and encouragement throughout the process.
# Table of Contents

Author’s Declaration .................................................................................................................. ii
Abstract ..................................................................................................................................... iii
Acknowledgements ..................................................................................................................... iv
List of Tables ................................................................................................................................. vii
List of Figures ................................................................................................................................. viii
Chapter 1: Introduction .................................................................................................................. 1
  1.1 Research Context .................................................................................................................. 1
  1.2 Research Objectives ............................................................................................................. 2
  1.3 Structure of the Report ......................................................................................................... 3
Chapter 2: Literature Review .......................................................................................................... 4
  2.1 Healthcare Accessibility ........................................................................................................ 4
  2.2 Health Indices ....................................................................................................................... 7
  2.3 Proximity Measures Database (PMD) .................................................................................. 9
  2.4 The Ontario Marginalization Index (ON-Marg)................................................................. 11
Chapter 3: Methodology ................................................................................................................. 16
  3.1 Methodological Context ....................................................................................................... 16
  3.2 Study Area: The City of Toronto ......................................................................................... 17
  3.3 Data and Data Sources ........................................................................................................ 19
  3.4 Analytical Approach ............................................................................................................ 20
  3.4.1 Choropleth Mapping ....................................................................................................... 21
  3.4.2 Exploratory Spatial Data Analysis .................................................................................. 21
  3.4.3 Spatial Regression Analysis ............................................................................................ 23
Chapter 4: Results .......................................................................................................................... 25
  4.1 Choropleth Mapping ............................................................................................................ 25
  4.2 Univariate Global and Local Moran’s I ............................................................................. 31
  4.3 OLS Regression .................................................................................................................... 37
  4.4 Lag-error Regression ............................................................................................................ 37
Chapter 5: Discussion & Conclusion ............................................................................................. 40
  5.1 Summary of the Research .................................................................................................. 40
  5.2 Discussion ............................................................................................................................ 41
  5.3 Implications ......................................................................................................................... 43
  5.4 Strengths and Contributions of the Research ................................................................. 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5  Limitations and Future Research</td>
<td>44</td>
</tr>
<tr>
<td>5.6  Conclusion</td>
<td>45</td>
</tr>
<tr>
<td>Appendix A – LISA Significance Maps</td>
<td>46</td>
</tr>
<tr>
<td>References</td>
<td>49</td>
</tr>
</tbody>
</table>
List of Tables

Table 2.1: PMD Network Distance Cut-off by Facility ................................................................. 10
Table 2.2: Ontario Marginalization Index Indicators ........................................................................ 12
Table 4.1: Z-value and Moran’s I for each ON-Marg dimension and the PMD ................................ 32
Table 4.2: OLS Regression Results ................................................................................................ 37
Table 4.3: Lag-error Regression Results .......................................................................................... 38
List of Figures

Figure 4.1: Spatial Distribution of the ON-Marg Dependency Dimension ........................................ 26
Figure 4.2: Spatial Distribution of the ON-Marg Material Deprivation Dimension ........................................ 27
Figure 4.3: Spatial Distribution of the ON-Marg Ethnic Concentration Dimension ........................................ 28
Figure 4.4: Spatial Distribution of the ON-Marg Residential Instability Dimension ........................................ 29
Figure 4.5: DA access to healthcare using the Proximity Measures Database ........................................ 30
Figure 4.6: Univariate Global Moran’s I all variable scatter plots (top row: dependency, material deprivation, ethnic concentration; bottom row: residential instability and proximity) ........................................ 31
Figure 4.7: Dependency dimension Univariate LISA cluster map ........................................ 33
Figure 4.8: Material deprivation dimension Univariate LISA cluster map ........................................ 34
Figure 4.9: Ethnic Concentration dimension Univariate LISA cluster map ........................................ 35
Figure 4.10: Residential Instability dimension Univariate LISA cluster map ........................................ 35
Figure 4.11: Proximity Measures Database (PMD) Univariate LISA cluster map ........................................ 36
List of Appendices

Appendix A – LISA Significance Maps ........................................................................................................50
Chapter 1: Introduction

1.1 Research Context

Healthcare access has been studied extensively to improve the health of people across the globe. However, many individuals continue to lack access to health services needed to live a proper healthy lifestyle. The World Health Organization (WHO) defines “health” as “a state of complete physical, mental and social wellbeing and not merely the absence of disease or infirmity” (WHO, 2020). They also argue that regardless of one’s race, religion, political belief, or economic or social condition, that they have a right to the highest attainable standard of health (WHO, 2020). Despite this claim, this right is absent from many people’s lives. Demographic factors such as income and age have become significant barriers to accessing proper healthcare and have a profound influence on one’s health. Therefore, they have become key to identifying at-risk neighbourhoods and overcoming this obstacle. As the global population continues to grow, it is especially critical that neighbourhoods with reduced access are identified, and measures aimed to improve healthcare access are organized.

According to the WHO, these differences between people that affect our health can be defined as the “social determinants of health” (WHO, 2020; Braveman, 2014). This includes the way we grow, live, work, and age in response to our environment that is moulded by aspects such as authority, wealth, and the movement of resources at multiple levels (WHO, 2020; Braveman, 2014). These conditions are responsible for the majority of the health inequalities between people and countries (WHO, 2020). When health is analyzed, it is necessary to recognize two different types of health: population health and public health. While public health often refers to the prevention of disease through organized efforts of society, such as the protection, surveillance, promotion, and assessment of health, population health is a relatively new term that has been attributed a variety of different definitions (Binns, 2015). According to Kindig (2003), population health is “the health outcomes of a group of individuals, including the distribution of such outcomes within the group... [it includes] health outcomes, patterns of health determinants, and policies and interventions that link these two” (p. 380). While public health looks primarily at the provision of healthcare services, population health
recognises that health surpasses merely spatial access to the healthcare we need. Rather, our health is also affected by uncontrollable factors (e.g., age) and rooted in our behaviors (e.g., lifestyle) and social status (e.g., income level). Many scholars have recognized the relationship between health and demographic characteristics, and seek to answer the question “why are some people, or groups of people, healthier than others?” (Evans, 1994). Those who are frequently unable to access proper healthcare due to these determinants are often marginalized populations. This may include, but is not limited to, the elderly, youth, homeless, and immigrants. Therefore, it is important that these populations are targeted when implementing healthcare interventions intended to improve access to healthcare.

1.2 Research Objectives

The purpose of this study is to analyze the association between area-level marginalization and healthcare accessibility in Toronto using the 2016 Ontario Marginalization Index (ON-Marg) and 2020 Statistics Canada Proximity Measures Database (PMD). The objectives for this study will be the following:

1) Analyze the distribution of each of the ON-Marg dimensions across Toronto’s dissemination areas (DA).

2) Examine spatial patterns of access and marginalization in Toronto and identify any clusters (high or low) within the city.

3) Explore the relationship between accessibility to healthcare calculated in the PMD and the four dimensions of marginalization in ON-Marg.

To meet these objectives, both spatial analysis and statistical analysis via Moran’s I and regression will be applied. By examining the distribution of marginalization through the four dimensions of the OM-Marg Index, as well as the proximity to healthcare in each dissemination area, at-risk neighbourhoods will be identified and highlighted as prime targets for government intervention.
1.3 Structure of the Report

This paper begins with a literature review that includes an in-depth overview of healthcare accessibility and prominent health indices, followed by an outline of the Ontario Marginalization Index and Statistics Canada’s new Proximity Measures Database (Chapter 2). The methods used in this study are then introduced, followed by an overview of the study area and data being analyzed (Chapter 3). Preliminary tables, maps, and figures are also explained in detail (Chapter 4). The results of the study are then discussed in relation to topics that appeared in the literature review, and the report concludes with a few final thoughts, observations, and recommendations (Chapter 5).
Chapter 2: Literature Review

2.1 Healthcare Accessibility

Healthcare access is a relatively broad area of research that has been studied extensively around the globe. However, there is still confusion surrounding what the term “access” truly means (Bauer, 2016). As Bauer states, “access is a multidimensional construct consisting of a variety of social, financial, geographical and personal factors”, which has resulted in a variety of different complex definitions of accessibility (Bauer, 2016, p. 1). De Mello-Sampayo specifically defines accessibility to healthcare as “the ability of a population to obtain healthcare services... it varies across space because neither health professionals nor residents are uniformly distributed” (De Mello-Sampayo, 2018, p. 739). This means that supply and demand is never evenly dispersed across space (Luo, 2003). Therefore, healthcare facilities often locate where population density is the greatest. Regardless, it is the government’s responsibility to provide healthcare in an uneven fluctuating environment, which can be challenging. As the 2010 World Health Report acknowledged, “a high proportion of the world’s poor population has no access to health services because they can’t afford it” (Bauer, 2016, p. 1). Vital to this statement is that people cannot “afford it”. Though the ability to provide healthcare exists, income as a barrier remains a key obstacle especially for poorer countries and marginalized communities (Bauer, 2016). However, for wealthier countries such as Canada that have universal healthcare (UHC) provided, access to healthcare goes beyond merely affordability. Even where universal healthcare is provided, access is shaped by a range of factors and can be extremely complex due to the spatial and socio-economic uneven landscape.

According to Joseph et al. (1982) and Luo (2003), there are four different categories of spatial accessibility: potential spatial access, potential aspatial access, revealed spatial access, and revealed aspatial access (p. 867). While revealed access emphasizes the usage of a particular service, potential access emphasizes the potential use of healthcare services without guaranteed use of them (Luo, 2003). Furthermore, while spatial access uses distance variables such as impedance, aspatial access does not. Rather, aspatial access looks at “nongeographic barriers” such as age, sex, and income (Luo, 2003, p. 865). Penchansky (1981) specifically
outlines four main barriers of access: availability, accessibility (distance and time), accommodation, affordability and acceptability. In this case if Joseph et al. (1982) and Luo’s (2003) definition of accessibility is applied, availability and accessibility would be considered spatial barriers while the remaining three would be defined as aspatial. These barriers, whether they are spatial or aspatial, affect access and can alter an individual’s access changing it from potential to revealed access (Bauer, 2016). This study in particular will reveal multiple dimensions of accessibility at both an aspatial and spatial level. While the Ontario Marginalization Index (ON-Marg) will be used as an indicator of marginalization, the Proximity Measures Database (PMD) will act as an indicator of access across Toronto. By analyzing and comparing both data sources, we are able to assess how they relate over space and analyze access to healthcare in the city of Toronto.

A study conducted by Oosterveer and Young (2015) analyzes access to primary healthcare in the Northwest Territories (NWT), specifically for indigenous communities. Indigenous communities are often very isolated and therefore lack proper access to healthcare facilities (Oosterveer & Young, 2015). In their own study, Oosterveer and Young measure accessibility in accordance with the United States Institute of Medicine, that define access to healthcare as “the timely use of personal health services to achieve the best possible outcomes” (2015). Through 14 interviews with primary healthcare providers and indigenous service users throughout the NWT, there was a clear consensus between both groups that living remotely did affect their health (Oosterveer & Young, 2015). Both groups emphasized that emergency health care was “scary” and “unacceptable” due to the lack of training and number qualified workers (Oosterveer & Young, 2015). Interestingly, both groups recognized that due to the small population size and changing indigenous landscape equal access to NWT healthcare was “unrealistic” (Oosterveer & Young, 2015). This is certainly surprising, as there seemed to be an acceptance that the WHO’s proclamation of universal healthcare access would never accumulate in indigenous communities. However, both groups do call for change and for conditions to be improved. In their case, and in many other indigenous communities alike, the barriers they experience are both spatial and aspatial.
As previously discussed, when we look at accessibility, the term ‘social determinant of health’ often emerges. To recall, the World Health Organization (WHO) defines a ‘social determinant of health’ as “the conditions in which people are born, grow, live, work and age... and the fundamental drivers of these conditions” (Braveman, 2014, p. 19). Critical to healthcare access is the concept of ‘equity’. Are we all, no matter who we are and where we live, able to access healthcare when we need it? Unfortunately, this is not always the case due to the variability and distribution of these social determinants. Social determinants such as our income, education, and age present challenges that can affect our health-related behaviours (Braveman, 2014). Are we able to afford the medicine we need? Can we afford to travel to receive the best treatment? Marmot (2008) argues that these conditions and imbalance are caused by the unequal distribution of power and resources between countries, and the condition that we live. However, despite the prevalence of these uneven conditions, many scholars argue that they are also avoidable. Through proper planning and comprehensive action that addresses these determinants and removes structural inequality this will be achievable (Marmot, 2008). However, this will take many years of hard work and change. By acknowledging the impact these determinants have on our health and behaviour, we are able to conclude that healthcare access itself is not the only factor that influences our health.

There have been a number of studies analyzing healthcare accessibility. For this section of the study, however, two different studies will be analyzed that assess healthcare accessibility for the marginalized in the study area Toronto. The first study conducted by Wang (2011) uses the two-step floating catchment area (2SFCA) gravity model and ArcGIS Network Analyst tool to examine the spatial accessibility of recent and established immigrants in Toronto to primary care physicians. A total of eight ethnicities were selected for the study: Hong Kong, Iran, Mainland China, Pakistan, Russia, Sri Lanka, Italy and Portugal (Wang, 2011). Accessibility varied among the immigrant groups substantially. However, Italians were determined to have greater median access to same-language family physicians (SLFP), while Portuguese and Sri Lankan immigrants had poorer access (Wang, 2011). Despite the variability between these groups, Wang also acknowledges how low the accessibility scores assigned to all of the immigrant groups are. She states that these scores suggest “that newcomers are the most disadvantaged
in their settlement process [and] in terms of locating SLFPs who accept new patients, compared with both their long-standing counterparts and the general population” (Wang, 2011, p. 247). In other words, as a collective group the marginalized had poorer access to healthcare than other Torontonians.

Similar conclusions were also produced in Khandor’s (2011) study that measured access to primary healthcare for the homeless in Toronto. The study took place between 2006 to 2007 and incorporated both a cross sectional survey and multivariable logistic regression to examine access to primary healthcare for 366 homeless adults (Khandor, 2011). Questions related to their demographic, health status, and healthcare accessibility were recorded, and regression was used to examine the link between having a family doctor, proof of health insurance, health status, and substance use (Khandor, 2011). Khandor found that the odds of having a family doctor greatly increased if the individual had a health card, was part of the LGBT community, and had a chronic medical condition (Khandor, 2011). However, it was also found that less than half of participants (43%) had access to a family doctor. This is significant and could foreshadow the results of this study as again, like the Wang (2011) study, marginalized populations in Toronto seem to lack proper access to healthcare.

2.2 Health Indices

ON-Marg is a relatively new index with its oldest version developed for the census year 2001, and most recent version for census year 2016 (OCHPP, 2018). Therefore, it was imperative that other health indices professionals may have utilized prior to the development of ON-Marg were recognised. One of the most popular marginalization indices used is the Canadian Marginalization Index (CAN-Marg). The Canadian Marginalization Index, like ON-Marg, is an area-based deprivation index that was created to reflect different levels of marginalization in both rural and urban locations (Matheson, 2012). It currently contains four different versions – 1991, 1996, 2001, and 2006 – however a new 2016 version will be released soon (OCHPP, 2012). CAN-Marg also uses the same four dimensions of marginalization used in ON-Marg. However, it only provides data at the census tract and dissemination area level, unlike ON-Marg.
that offers data at a larger geographical scale - to produce values for these larger geographies original DA factor scores were derived (OCHPP, 2012; Matheson, 2012). Furthermore, CAN-Marg also provides data for Canada as a whole, while ON-Marg delivers data specifically for Ontario. Interestingly, in some of its previous versions ON-Marg data was derived specifically from CAN-Marg data (OCHPP, 2018; Matheson, 2012). However, the 2011 and 2016 versions of ON-Marg were derived from principal component factor analysis using 18 census variables from the 2011 and 2016 Canadian census (OCHPP, 2018; Matheson, 2012). Given that ON-Marg data is more recent and provided at a smaller scale in relation to the study area selected for this study (Toronto), ON-Marg was selected.

The Socio-economic Factor Index (SEFI) is another popular index that uses census data to indicate whether socio-economic conditions are favourable or unfavourable with a factor score (Chateau, 2012; Metge, 2009). Scores less than 0 are ideal, while a factor score greater than 0 mean the socio-economic conditions are poor (Chateau, 2012). SEFI contains two unique versions, SEFI and SEFI-2, each with a different number of variables that can be calculated at the DA level (Metge, 2009). SEFI contains 6 variables: age dependency ratio, rate of single parent households, rate of female single parent households, female labour force participation rate, unemployment rate composite, and a high school education rate composite (Chateau, 2012; Metge, 2009). On the other hand, SEFI-2, that was introduced in 2009, contains only four variables: average household income, percent of single parent households, unemployment rate, and high school education rate (Chateau, 2012; Metge, 2009). Due to data restrictions in previous census years SEFI was not able to include income as a variable like SEFI-2 did (Metge, 2009). Interestingly, some of these census variables from both versions are quite similar to those in the Ontario Marginalization Index.

Another popular index is the Health Utilities Index (HUI) that is used to describe health status and health-related quality of life (Furlong, 2001). HUI contains three different systems that analyze over 1 million health states for people ages 5 years and older: HUI1, HUI2 and HUI3 (Furlong, 2001). However, HUI2 and HUI3 are often used more frequently then HUI1, as they are quite similar in their foundation and generally function better than HUI1 (Furlong, 2001). Both of these standardized systems include a health classification and utility scoring system
that ranges from 0.00 to 1.00, where 0.00 means the individual is deceased and 1.00 means they are in perfect health (Furlong, 2001). Though this index does not focus specifically on demographic variables and marginalization, it is important to acknowledge its existence and role in measuring health and quality of life as it is the focus of this study as well.

2.3 Proximity Measures Database (PMD)

The Proximity Measures Database (PMD) was released by Statistics Canada in 2020 in collaboration with the Canada Mortgage and Housing Corporation (CMHC). Being the first ever national PMD, individuals are now able to easily access proximity data across Canada (Statistics Canada, 2020). It is currently available at all levels of geography including dissemination areas and provides proximity measures for the following 10 services/amenities: employment, grocery stores, pharmacies, healthcare, childcare, primary education, secondary education, public transit, neighbourhood parks, and libraries (Statistics Canada, 2020). For the purpose of this study, “proximity to healthcare” will be utilized. Moreover, the database also includes composite indicators that combine some of these measures into one (Statistics Canada, 2020). Proximity measures were calculated using a gravity model that measured the network distance (walking or driving) between a dissemination block (DB) centroid and other DB centroids where the facility is located (Statistics Canada, 2020). It also accounts for the size of the facility, which was measured using either total revenue or total employment, and presence of the facility in the DB (Statistics Canada, 2020). However, proximity of public transit was calculated using walking distance to the nearest transit stop rather than the DB centroid (Statistics Canada, 2020).

Proximity is provided as a “normalized index value”, where the computed values were changed to a scale that ranged from 0 to 1, with 0 representing the lowest proximity measure and 1 representing the highest (Statistics Canada, 2020). If a value was not assigned to a DB, that DB does not receive service from the amenity in question (Statistics Canada, 2020). On the other hand, data is also provided as an indicator assigned a value of either a 0 or 1, where 1 indicates the presence of at least 1 facility in the DB, and 0 indicates that no facilities are
present in the DB within the specified cut-off distance (Statistics Canada, 2020). **Table 2.1** below outlines the network distance cut-off used for each of the 10 facilities. Though this database is certainly beneficial and the first of its kind, Statistics Canada does also acknowledge a limitation related to inconsistency. Some of the sources used to access information about each facility, such as the Business Register (BR), at times provided inconsistent information (e.g., if a school exists in a certain DB) (Statistics Canada, 2020). Therefore, the accessibility score provided for a DB could be incorrect or slightly off. To avoid this issue, the proximity measure for that DB is suppressed (Statistics Canada, 2020).

Moreover, as previously stated, “proximity to healthcare” will be used in this study. According to Statistics Canada, “proximity to healthcare” was derived from the employment counts in the following NAICS codes in the Business Register: 6211, 6212, 6213, 621494, and 622 (Statistics Canada, 2020). This includes offices of physicians, offices of dentists, offices of other health practitioners, community health centers, and hospitals respectively. Though this variable is indeed through by including these major healthcare facilities, it is important to acknowledge a limitation in doing so when examining access. Though the database has included major hospitals such as Sick Kids Hospital, it does not account for the spatial influx of people that travel a far distance to receive specialized care. Therefore, it can be reasoned that a 3km network driving distance for a hospital in measuring access is inadequate. Consequently, it can also be argued that the PMD itself depicts an inaccurate representation of healthcare access.

**Table 2.1: PMD Network Distance Cut-off by Facility**

<table>
<thead>
<tr>
<th>Facility</th>
<th>Network Distance Cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grocery Store</td>
<td>1km walking</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>1km walking</td>
</tr>
<tr>
<td>Public Transit Stop</td>
<td>1km walking</td>
</tr>
<tr>
<td>Childcare Facility</td>
<td>1.5km walking</td>
</tr>
<tr>
<td>Primary School</td>
<td>1.5km walking</td>
</tr>
<tr>
<td>Library</td>
<td>1.5km walking</td>
</tr>
<tr>
<td>Health Facility</td>
<td>3km driving</td>
</tr>
<tr>
<td>Employment</td>
<td>10km driving</td>
</tr>
</tbody>
</table>

10
Due to the recent release of the Proximity Measures Database (PMD) on April 17th, 2020, there are currently no published studies that have used the database. Within the next few months there likely will be new articles published by scholars that have used the database in their own study. Regardless, this report will be one of the first few, if not the first, to analyze accessibility scores in the PMD, and compare them to levels of marginalization across Toronto using ON-Marg.

2.4 The Ontario Marginalization Index (ON-Marg)

The Ontario Marginalization Index is a tool created by research scientist Dr. Flora Matheson, St. Michael’s Hospital research coordinator Gary Moloney, and Analytic Services epidemiologist lead Trevor van Ingen. It incorporates a variety of demographic factors, combining them into four dimensions of marginalization (Ontario Community Health Profiles Partnership [OCHPP], 2018). These dimensions include residential instability, material deprivation, dependency, and ethnic concentration (OCHPP, 2018). Each category contains a number of health indicators, with a combined total of 18 indicators shared between the 4 dimensions (OCHPP, 2018; Matheson, 2012). Table 2.2 summarizes the indicators within each category. To create the 2016 and 2011 ON-Marg Index, a series of principal component factor analyses were completed with an initial total of 42 socio-economic indicators from the Canadian census. Variables with low factor loadings were removed to reduce the number of variables included in the index to 18, and the number of factors to four (OCHPP, 2017). The factors were then named based on the variables included in each grouping - which resulted in the four dimensions of marginalization.

ON-Marg is also available for the years 2001, 2006, 2011, and 2016, and can be mapped at a variety of different geographies. These geographies include dissemination areas, census tracts, and local health integration networks (LHINs) depending on the index year, using either principal component factor scores or quintiles (OCHPP, 2018). However, larger geographies such as LHIN’s and PHU’s are not assigned quintile values since they are too large to create significant quintile divisions (OCHPP, 2018).
<table>
<thead>
<tr>
<th>ON-Marg Dimension</th>
<th>Indicators</th>
</tr>
</thead>
</table>
| Residential Instability | • proportion of the population living alone  
• proportion of the population who are not youth (ages 5-15)  
• average number of persons per dwelling  
• proportion of dwellings that are apartment buildings  
• proportion of the population that are single/divorced/widowed  
• proportion of dwellings that are not owned  
• proportion of the population who moved during the past 5 years |
| Material Deprivation    | • proportion of the population aged 20+ without a high school diploma  
• proportion of families who are lone parent families  
• proportion of total income from government transfer payments for population aged 15+  
• proportion of the population aged 15+ who are unemployed  
• proportion of the population who are considered low income  
• proportion of households living in dwellings that are in need of major repair |
| Dependency              | • proportion of the population who are aged 65+  
• dependency ratio (total population 0-14 and 65+ / total population 15 to 64)  
• proportion of the population not participating in labour force (aged 15+) |
| Ethnic Concentration    | • proportion of the population who are recent immigrants (arrived in the past 5 years)  
• proportion of the population who self-identify as a visible minority. |
In order to compensate for the absence of the long form census in 2011, other data sources such as the Municipal Property Assessment Corporation (MPAC), and Registered Persons Database (RPDB) were used instead to create the 2011 ON-Marg Index (OCHPP, 2017). After the reinstitution of the long form for the 2016 census, the 2016 version of ON-Marg was able to exclusively use the short and long form census as data sources (OCHPP, 2018). Therefore, the 2016 version of the index will be used in this study for analysis.

Although ON-Marg is a useful credited index with multiple unique versions, it does include a few limitations. While data is available at smaller levels of geography, Statistics Canada has suppressed some values to protect the privacy of residents living in sparser neighbourhoods (OCHPP, 2018). Therefore, some census tracts or dissemination areas in Ontario will contain missing data or 0 values were applicable (e.g., no immigrants living in that neighbourhood) (OCHPP, 2018). Moreover, factor analysis itself it also a method of data reduction. Throughout the process to create the ON-Marg Index valuable data was likely lost. This may therefore produce an inaccurate representation of marginalisation in Toronto. As previously stated, ON-Marg also contains a variety of different versions set during a certain census year (e.g., 2011 and 2016). Therefore, it is important to select the correct version of the index based on the data being used. Since this report will be comparing accessibility measures from the 2020 Proximity Measures Database to levels of marginalization in ON-Marg, the most recent version 2016 was selected. It is also important to acknowledge the limitations of the Canadian census itself. Though the Government of Ontario is able to retrieve information about the majority of the population, some populations such as indigenous or institutionalized people (e.g., care homes) are possibly excluded or under-counted (OCHPP, 2018). These limitations are important to consider as they could produce slightly inaccurate or skewed results.

To conduct this analysis suitably, it was essential to review past studies that have incorporated ON-Marg. Since it was first released, the Ontario Marginalization Index has been used to analyze Ontario in many different subject areas. For instance, a study conducted in Ontario used ON-Marg data to test whether there was an association between multi-morbidity and marginalization using socio-economic status (SES) (Moin, 2018). By using this data, they were able to perceive, for instance, a rise in multi-morbidity when using the ‘material
deprivation’ category of the Ontario Marginalization Index. Moreover, they were also able to
determine that there was a correlation when comparing levels of multi-morbidity to material
depression and housing instability (Moin, 2018). Overall, they concluded that the occurrence of
multimorbidity was greater in areas where material deprivation and residential instability was
higher. On the other hand, their results also exhibited very little to no correlation between
multimorbidity, and ethnic concentration and dependency. Despite these results, this study is
important as it supports that there is a correlation between health and marginalization.

Another study conducted by Silverman (2013) also supports this claim. However, they examine
the number of cyclist collisions in Toronto connecting them to socio-demographic
factors in the Ontario Marginalization Index rather than health and disease (Silverman, 2013).
This study is particularly interesting as it takes place in Toronto as well. On-Marg data was
applied to each intersection where the collisions occurred, and a logistic regression analysis was
performed to examine the association between the four ON-Marg dimensions and the cyclist
collisions (Silverman, 2013). Based on these dimensions, they concluded that local areal
dependency and material deprivation were not significant factors in the cyclist collision
(Silverman, 2013). However, they did determine that the number of collisions increased in
neighbourhoods where residential instability and ethnic concentration was high (Silverman,
2013). This is quite interesting as again, like the Moin (2018) study, two dimensions emerge
more prominently than the others. However, in this study ethnic concentration is determined
to be significant, while the previous study looking at multimorbidity concluded that this
dimension was not significant.

Going beyond health and safety, ON-Marg has also been used in crime analysis. A study
also completed in Toronto examined the link between the 4 dimensions of the Ontario
Marginalization Index and different types of crime. Crime is often linked to variables such as
low-income, therefore this study examines the distribution of property and violent crime in
Toronto along with the distribution of these 4 dimensions (Wang, 2018). To start, they used
LISA statistics to examine where in Toronto clusters of property and violent crime would appear
(Wang, 2018). Next, they used OLS regression to examine the link between crime and the ON-
Marg dimensions (Wang, 2018). It was determined that neighbourhoods with high levels of
material deprivation and residential instability were associated with larger concentrations of violent crime, and those with lower levels of ethnic concentration and higher levels of residential instability were associated with higher levels of property crime (Wang, 2018).

When analyzing the results of all three of the studies described above, it is evident that a few similarities exist. For instance, all three studies have found higher levels of residential instability to be a significant factor, compromising the health and safety of Ontario residents and also encouraging crime. This could possibly foreshadow the results of this study. Most importantly however, although the majority the results vary between each study, they all have concluded that correlation does exist between their field of study and marginalization.
Chapter 3: Methodology

3.1 Methodological Context

The goal of this study was to explore the association between area-level marginalization and healthcare accessibility, using the City of Toronto as a case study area. To explore these objectives a quantitative approach that focused on spatial analysis was used. This is a suitable approach as the study of geographic phenomena – such as the patterns of healthcare access, marginalization, and the relationship between them – employs methods to facilitate exploration and generate insight (Burt and Barber, 1995; Rogerson, 2010). This quantitative approach allows for hypothesis to be constructed prior to data collection and for these hypotheses to be tested and validated (Rogerson, 2010). This method also allows for the identification of cause-and-effect relationships and the ability to generalize research findings drawn from a large population (i.e. all DAs in Toronto). In addition, the research design may be able to eliminate the confounding influence of many variables or autocorrelation between variables (Burt and Barber, 1996).

The spatial analysis conducted in this study provides additional strengths. As Rogerson (2010) argues “understanding the effects of spatially dependent observations in statistical analysis provides an important motivation for learning more about spatial patterns,” which is a key concern in all three research objectives (p. 257). A primary component of spatial analysis is that of spatial autocorrelation. Spatial autocorrelation measures the degree of dependency among observations in a geographic space (Getis, 2008), and is based on one of the tenets of geography, Tobler’s First Law, where “everything is related, but near things are more related than distant things” (Tobler, 1970, p. 236). The measure of spatial autocorrelation can be identified at both a global scale and a local one. This method of analysis, therefore, provides an approach to identifying and understanding both broad and underlying structures to healthcare accessibility and marginalization in Toronto.

Additionally, regression analysis is used to further explore the relationship between healthcare access and marginalization. An assumption of most statistical analysis, however, is that the observations are independent of each other and that there is no spatial pattern within
the data. Observations that are not independent can affect statistical rigour and the ability to judge significance. For instance, spatial autocorrelation violates key assumptions of Ordinary Least Squares regression – specifically for variable independence and error correlation – that bias the estimates of the model (Abreu et al, 2005; Anselin, 2001). Taking the spatiality of marginalization and accessibility of healthcare into account allows for a clearer and more accurate ‘picture’ of the condition in Toronto to be captured. Furthermore, it provides a way to disentangle and understand complex relationships across space (Abreu et al, 2005; Pavlovskaya, 2006).

### 3.2 Study Area: The City of Toronto

Toronto is a city with a diverse and growing population. According to the City of Toronto, a total of 2,956,024 people were estimated to live in the city as of July 2018 (City of Toronto, 2020). With a total of 140 neighbourhoods and over half of its population identifying as a visible minority (51.5%), Toronto remains one of the most culturally diverse cities in Canada (City of Toronto, 2020; Statistics Canada, 2019). In the 2016 census, 47% of the total population also identified as immigrants (Statistics Canada, 2019). When compared to Ontario as a whole with a total of 29.1% of immigrants, it becomes clear that Toronto is a hot spot for new immigrants (Statistics Canada, 2019). According to Statistics Canada (2019), in 2016 the majority of people that lived in Toronto had European origins (47.9%), Asian origins (40.1%), and other North American origins (12.8%). However, the majority of Toronto’s immigrant population currently come from China and the Philippines with a total of 131,480 and 118,775 people respectively (Statistics Canada, 2019). This diversity has influenced Toronto’s landscape a great deal, leading it to be home to numerous ethnic enclaves such as Chinatown, Little Italy, Greektown, and Little Portugal - these neighbourhoods will likely produce higher index values in the ethnic concentration dimension of ON-Marg.

Toronto is also diverse in terms of its household income values. In 2015, the median total income of households jumped by 5.3% from $62,506 to $65,829 (Statistics Canada, 2019). Moreover, the majority of households (Toronto CMA) make under $100,000, with 23.8% of
households making between $60,000 and $99,999, 21.6% making $30,000 and $59,999, and 16.6% making under $30,000 (Statistics Canada, 2019). Interestingly, however, there is also 19.4% of the population that make over $150,000 (Statistics Canada, 2019). This large gap illustrates Toronto’s diversity very well, as it indicates that many people of different occupations and lifestyles live here. When compared with its neighboring cities such as Pickering (99,701) and Vaughn (105,351) where the median total household income is about $100,000, Toronto stands out significantly. In fact, it has the lowest 2015 median household income value of all its neighbours (Statistics Canada, 2019). This is quite odd, considering that many adults (25-64) have achieved a high school diploma (88.6%). Though, the completion of a bachelor’s degree or higher in 2016 was only 44.1% (Statistics Canada, 2019). The majority of Toronto’s labour force in 2016 were in the following three industry sectors: professional, scientific, and technical services; retail trade; healthcare and social assistance (Statistics Canada, 2019). When we consider dependency and material deprivation in Toronto, recognizing the difficulty of finding and keeping a job is very important. Based on these statistics, we can conclude that Toronto’s job market is very competitive, with many people working at jobs that pay below a livable wage. Therefore, we can expect to see lower levels of material deprivation and dependency in neighbourhoods where, for instance, a doctor or surgeon may live, and higher levels of dependency and material deprivation where a retail store employee may live.

Housing prices in Toronto have also increased in recent years, which has caused many residents to switch from home ownership to renting a home. According to Statistics Canada (2019), in 2016 52.8% of homes were owned and 47.2% were rented. Though home ownership remains more prevalent than renting, both housing prices have also rose steadily. In 2006 it cost individuals $1312/month to own a home and $931/month to rent, which increased to $1682/month and $1242/month in 2016 respectively (Statistics Canada, 2019). It is unsurprising, then, that only 24.2% of the population live in single-detached housing and 44.3% live in apartment buildings as they are much cheaper – though this can depend on the location (Statistics Canada, 2019). If we consider Toronto’s landscape, we can expect to see higher levels of residential instability where population density is high (such as downtown Toronto), and also where household income is lower.
The statistics discussed above are important as they help to depict how diverse and how uneven Toronto’s landscape truly is. Toronto is filled with many people of different backgrounds, income, occupations, and histories. These factors may also impact their ability to access healthcare. Are they able to leave their job in the event of a family medical emergency? Does their employment insurance cover the medication they may need? By comprehending these values and how they have shaped the Toronto landscape, we are able to develop a deeper understanding of the city and the people who live there.

3.3 Data and Data Sources

For this study there were two key data sources: the Ontario Marginalization Index and the Proximity Measures Database (described in detail in Sections 2.2 and 2.3). ON-Marg is available for the years 2001, 2006, 2011, and 2016, and includes a total of 18 indicators shared between four dimensions of marginalization: dependency, material deprivation, ethnic concentration, and residential instability (OCHPP, 2018). The 2016 index was derived from principal component factor analysis, with data retrieved from the short form and long form census for a total of 20,640 DAs and 2,376 CTs in Ontario (OCHPP, 2018). Data is provided in quintiles or as a factor score, where a lower factor score indicates lower levels of marginalization and a higher score indicates higher levels of marginalization (OCHPP, 2018). The PMD, on the other hand, is the first ever national PMD and provides proximity measures for 10 different services including healthcare. Proximity measures were calculated using a gravity model that measures the network distance between a dissemination block (DB) centroid and other DB centroids where the facility is located (Statistics Canada, 2020). This method also accounts for the size of the facility, which is measured using either total revenue or total employment, and presence of the facility in the DB (Statistics Canada, 2020). Accessibility in the PMD is provided as a normalized index value, where the computed values were changed to a scale that ranged from 0 to 1, with 0 representing the lowest proximity measure and 1 representing the highest (Statistics Canada, 2020). However, it is also provided as an indicator and assigned a value of 0 or 1, where 1 indicates the presence of at least 1 facility in the DB,
and 0 indicates that no facilities are present in the DB within the specified cut-off distance (Statistics Canada, 2020).

Both data sources are provided at different levels of geography, including DA’s and census tracts. Spatially, this study uses dissemination areas as the areal unit of analysis. DA’s were selected for this study as they provided the finest level of spatial detail for both ON-Marg and PMD. According to Statistics Canada, a dissemination area (DA) is “a small area composed of one or more neighbouring dissemination blocks, with a population of 400 to 700 persons... it is the smallest standard geographic area for which all census data are disseminated” (Statistics Canada, 2019a). They were first introduced by Statistics Canada in 2001 replacing the enumeration area (EA), and between 2001 and 2006, census tracts were surveyed and delineated if population size increased drastically (Statistics Canada, 2019a). This was also the case in 2016, however, the maximum number of dissemination blocks that could be assigned to each DA (99) was removed (Statistics Canada, 2019a). Currently, over 54,000 DA’s with their own four-digit code cover all of Canada (Statistics Canada, 2019a).

3.4 Analytical Approach

To address the objectives of this study, this research used multiple techniques including choropleth mapping, cluster analysis (both global and local), and spatial regression. Further guiding the research, based on the research context of healthcare access and the geographic context of the city of Toronto were two hypotheses:

H1: That positive spatial autocorrelation (clustering) is present in Toronto for both the PMD and ON-Marg data.

H2: That as marginalization increases, healthcare accessibility decreases.

More specifically, H1 addresses the first two research objectives, while H2 addresses research objective three. The approach used here – spatial/cluster analysis and regression – followed an analytical approach laid out by Anselin (1994, 2001, 2005), Cleave et al. (2020), and Rogerson (2010).
3.4.1 Choropleth Mapping

This research was conducted in three steps. First, both the Ontario Marginalization Index (ON-Marg) and the Proximity Measures Database (PMD) were mapped using a choropleth at the DA level. By doing so, patterns of marginalization and accessibility across Toronto could be visualized and interpreted. For each map, the data was organized using a quantile classification method, and divided into five distinct groups. A total of five initial maps were created: one depicting the PMD accessibility scores, and four others mapping the distribution of each of the four ON-Marg dimensions. This provided an initial exploration of the data, and a preliminary evaluation of any spatial patterns that exist.

3.4.2 Exploratory Spatial Data Analysis

The next step involved exploratory spatial data analysis (ESDA), focusing on local indicators of spatial autocorrelation (LISA). GeoDa was used to test for clustering and dispersion across all Toronto DA’s. For this analysis, a spatial weights matrix that operated under queen contiguity rules was employed. This created a simple binary connectivity matrix, where \( w_{ij} = 1 \) if regions \( i \) and \( j \) had any spatial contiguity, and \( w_{ij} = 0 \) otherwise. This matrix was then used to conduct a Univariate local and global Moran’s I test for both the four dimensions of ON-Marg as well as the PMD. However, isolates in the weights, specifically the Toronto island in the South, were removed.

In the ESDA, a Univariate Global Moran’s I was calculated to determine if there were any broad patterns of spatial autocorrelation within the five variables. A Global Moran’s I measures spatial autocorrelation based on both feature locations and feature values simultaneously. Given a set of features and an associated attribute, it evaluates whether the pattern expressed is clustered, dispersed, or random. The null hypothesis for Global Moran’s I is that patterns detected are due to random chance (Anselin, 2001; Rogerson, 2010). This hypothesis can be appraised in two ways. First, the Moran’s I scatterplot provides insight into the relationship of variables across space (i.e. the City of Toronto). In this case, the upper right and lower left quadrant of the graph represent clustering of like values and positive spatial autocorrelation,
while the lower right and upper left quadrants indicate the presence of spatial outliers and negative spatial autocorrelation. Second, spatial autocorrelation was tested through a z-value computed by comparing the observed patterns of variables against random distributions of variables (99999 permutations of spatial distribution, in GeoDa). By using this method, it was determined whether the Moran’s I value calculated was significant by comparing it to random chance. Using a significance level of 0.05, or a 5% risk of concluding that a difference exists when there is no actual difference, the critical z-value is +/- 1.96. If the z-value calculated is outside +/-1.96 then there is significant spatial autocorrelation and we reject the Moran’s I null hypothesis. If it is within +/-1.96 there is no spatial autocorrelation and the null hypothesis is accepted.

The next step in the exploratory analysis looked at evidence of local spatial autocorrelation or clustering. For each variable, a LISA univariate local Moran’s I test was conducted to test for local spatial autocorrelation between the DA’s, identify clusters of similar neighbouring DAs (i.e. high-high, low-low), as well as the significance of each cluster (in GeoDa, these are presented visually in map form). Any clusters identified are assigned a probability (p-values) – with probabilities less than 0.05 indicating a significant cluster that is not expected to occur by random chance. Similar to the Global Moran’s I, LISA analysis uses the null hypothesis that the patterns are random, so areas with p-values exceeding 0.05 will reject that hypothesis and indicate local autocorrelation (i.e. clustering). The presence of high-high and low-low clustering on each LISA map will help determine whether positive spatial-autocorrelation exists. Furthermore, those assigned a low-high or high-low classification will be identified as spatial outliers. This means that they are different from their neighbouring DA’s and therefore exhibit negative local spatial autocorrelation. It is important to acknowledge however when interpreting these results that each Univariate LISA map and Global Moran’s I scatter plot can be misleading. This is because they focus solely on one variable at a time. Though they are able to suggest potential covariance between the DA’s, they are entirely exploratory in nature. To properly examine these dimensions and their associations regression is used following their examination.
3.4.3 Spatial Regression Analysis

The final stage of analysis compared the influence of the ON-Marg variables to healthcare access in the PMD. While the ESDA and LISA analysis identify spatial patterns, they do not provide any direct insight into interaction between the variables. To explore this, and achieve a key research goal of the study, spatial regression analysis was utilized. Non-spatial regression analysis typically utilizes an ordinary least squares (OLS) estimator to model the relationship between dependent and independent variables. The regression model can be defined as:

\[ y = x\beta + \varepsilon \]  

(3.1)

where \( y \) is the response variable, \( x \) is the explanatory variable, and \( \beta \) is the estimate of the coefficient. The error term \( (\varepsilon) \) is assumed to be normally distributed \( (\varepsilon \sim N(0, \theta^2)) \). However, a key assumption of OLS is that there is no autocorrelation between observations of the variable. This estimator implicitly assumes that each region is independent of all other regions (Baumont et al., 2001). However, if there is spatial dependence, the endogeneity of the regressors causes the OLS estimator to be biased and inconsistent (Abreu et al., 2005). Therefore, it produces inefficient estimators and the potential for invalid statistical inference.

If spatial autocorrelation exists, three different spatial regression models be used to illustrate intensity. First is spatial lag, where dependent variables are influenced by observed independent variable values in neighbouring areas. Second is spatial error, where there is correlation in the error between neighbouring regions. For the first two instances when OLS assumptions are broken, a maximum likelihood (ML) estimator is needed to produce a valid model between dependent and independent variables. Anselin’s (2004) taxonomy of spatial econometric models argues that the model can be global or local, as well as modelled in the error term or un-modelled. Since it has been assumed that \( \varepsilon \sim N(0, \theta^2) \), error has been modelled. The process can also be global if the model contains endogeneity and can be local if the model is exogenous. To show these patterns within the data ML error and spatial lag were modelled. The model for spatial lag is:

\[ y = \rho W y + \beta x + \mu \]  

\[ \mu \sim N(0, \theta^2) \]  

(3.2)
where $\rho$ is the lag coefficient and $W_y$ is the weighted matrix. The model for spatial error is:

$$y = \lambda W_\varepsilon + \beta x + \mu, \quad \mu \sim N(0, \theta^2)$$

(3.3)

where $\lambda$ is the error coefficient and $W_\varepsilon$ is the weighted matrix.

The third way in which the assumptions of an OLS are broken is when there is a combination of both spatial lag and spatial error influencing the dependent variable. In this case, a generalized method of moments (GMM) estimator is needed to produce valid variable coefficient estimates for the independent variables. As opposed to ML estimators which require assumptions about the distribution of variables and the error term, GMM allows models to be specified while avoiding often unwanted or unnecessary assumptions, such as specifying a particular distribution for the errors. This lack of structure means GMM is widely applicable, particularly when there is potential spatial lag and error.

The approach used here followed Cleave et al (2020). First, a standard OLS regression was implemented with the PMD heath accessibility index as the dependent variable, and the four ON-Marg indices as the independent variables. The previously constructed spatial weights matrix (see section 3.4.2) was also integrated into the ‘classic’ OLS model to allow diagnostic tests for spatial dependency to be conducted. These tests will be used to determine whether the regression is to be run a second time under different stipulations. Specifically, the values for the Lagrange Multiplier lag and error were assessed and classified as significant or not significant. If neither the LM lag nor LM error are significant, the results produced from the classic OLS regression model will be analyzed, as there was no spatial dependency within the model (so no assumptions of OLS were broken; Anslin, 1994; 2004). If only one is significant a second regression, ML-regression, will be run selecting spatial-error or spatial-lag as appropriate (Anslin, 1994). However, if both LM lag and LM error are significant then the values for the Robust LM lag and error tests will also be evaluated. If only one Robust score is significant then then an ML-regression model will be used. Lastly, if both Robust LM lag and error were significant, a GMM estimator for a lag-error regression model using GeoDa Space would be necessary (Cleave et al., 2020).
Chapter 4: Results

The results will be presented in the following order. First, the choropleth maps for each of the four dimensions of ON-Marg and the PMD will be introduced, and their visual spatial pattern will be discussed. Second, the z-value for each Univariate Global Moran’s I will be evaluated to determine whether spatial autocorrelation exists in the model, and the results of the Univariate Local Moran’s I will then be analyzed through clustering in the LISA maps. Thirdly, the OLS Regression Lagrange Multiplier and Robust LM probability values are discussed in relation to spatial dependence. These results help to determine the final step of the analysis – which happens to be a lag-error regression. The statistics produced in the lag-error regression, specifically the probability of each variable, their coefficient value, and pseudo r-squared, are used to analyze the relationship between the ON-Marg dimensions and accessibility.

4.1 Choropleth Mapping

After mapping each of the ON-Marg dimensions it became evident that each one contained some clustering. In some cases, however, clustering would exist in similar locations. For each map, the presence of marginalization is explained by positive values or negative values close to 0, while the absence of marginalization is explained by negative values that are distant from 0. In Figure 4.1 it appears that higher levels of dependency, specifically DA’s with an index value between -0.105559 and 11.463103, are quite dispersed but exist more prevalently around the periphery.
Though there are a few smaller DA’s in the highest quantile located in downtown Toronto, most DA’s there contain low to moderate levels of dependency with an index value between -0.105559 and -1.789239. To recall, this dimension focuses on the presence of retired or unemployed seniors (65+). By living in the periphery, they are able to remove themselves from hectic city life in exchange for a proper peaceful life of retirement. Housing is also much more affordable here for retirees compared to downtown Toronto which is often expensive.

On the other hand, clustering in the material deprivation dimension is much more noticeable. In Figure 4.2 material deprivation, seen through higher index values between 0.235981 and 7.853994, exists primarily in DA’s located in the East and West end of Toronto. Interestingly, deprivation takes on a ‘U’ shape, traveling from the Northwest, down South through downtown Toronto, and then back up to Northeast Toronto.
This dimension of ON-Marg acknowledges the presence of adults who are unemployed and lone parents, with little education, and a low income living in inadequate housing conditions. This pattern is predictable given that the majority of DA’s with lower material deprivation values (between -0.277167 and -2.323528) are wealthier neighbourhoods such as Rosedale, and those with high material deprivation values are poorer neighbourhoods like Rexdale. These poorer neighbourhoods are also populated by mostly immigrants and visible minorities.

Clustering is also quite apparent in Figure 4.3 for the ethnic concentration dimension, that recognises those who identify as visible minorities and recent immigrants. This spatial pattern is also similar to the material deprivation dimension previously discussed, with clustering of marginalization towards the Northeast and Northwest. On the other hand, lower levels of ethnic concentration, seen through index values between 0.268526 and -1.160816, are present primarily in and around downtown Toronto as well as along the Toronto waterfront.
These neighbourhoods in the North are known to frequent the presence of visible minorities. Many new immigrants looking to start a new life in Canada have very little in their possession and cannot find a job to support them. Therefore, to comfort themselves (and likely due to housing affordability as well), they often surround themselves with those in the same condition. This has resulted in the formation of various ethnic enclaves across Canada, resulting in a higher index value between 0.904590 and 4.749944 in these locations. Downtown Toronto also contains a few ethnic enclaves such as Chinatown and Little Italy, which explains why a few of the DA’s located there contain greater levels of ethnic concentration as well.

Interestingly, the clustering of residential instability is quite different from the other dimensions. In Figure 4.4 it is clear that higher levels of instability, between an index value of 0.540782 and 4.599305, exist downtown. However, there are also a few other DA’s in the North and West end of Toronto that are also in the highest quantile.
Figure 4.4: Spatial Distribution of the ON-Marg Residential Instability Dimension.

On the other hand, lower index levels between -0.152451 and -1.535368 exist in central Toronto and along the periphery. This dimension consists of a number of variables focused around home ownership, age, the number of people in each household, and marital status. Of importance, it includes those living in an apartment building, those who are not youth (age 5-15), and dwellings that are not owned. Due to high population density and the value of the land downtown, many apartment buildings have been built. Furthermore, downtown is also home to many renowned universities such as the University of Toronto and Ryerson University who provide housing for their students. Both these factors could partially explain why the spatial distribution of this dimension is inverted.

The choropleth map produced in Figure 4.5 using the data from the Proximity Measures Database (PMD) on the other hand was anticipated. In this case, higher index values indicate greater access to healthcare, while lower index values indicate less access.
In a health system based on the principles of equity and universal healthcare, healthcare facilities are generally placed where they can best serve the public. Therefore, it is expected that more facilities be placed where population density is high. Downtown Toronto is populated by many people as it is a location that many Torontonians travel to every day for work, shopping, leisure, etc. For this reason, many health facilities and hospitals have been located there making accessibility to healthcare here tremendously high. Specifically, higher access in the PMD is represented by an index value between 0.031301 and 0.715300. However, as many scholars have recognized, access to healthcare in the suburbs and rural areas is not as strong. This is evident in Toronto as well. The majority of Toronto’s periphery are assigned lower index scores between 0.00000 and 0.006900 indicating that their access to healthcare is relatively lower. This is often the case in most suburban neighbourhoods. However, there are clusters of DAs in the periphery that do contain higher levels of accessibility to healthcare. This is largely due to the presence of a hospital in these locations. For instance, North York General Hospital is
located in the Northern cluster of high access, while Birchmount Hospital and Centenary Hospital are located in the Northeastern clusters, and Etobicoke General in the Northwestern cluster. Without these hospitals as key points of access for healthcare Toronto’s periphery would likely only consist of the two poorest quantiles of access.

4.2 Univariate Global and Local Moran’s I

To start, the Univariate Global Moran’s I was calculated. The results for each variable are seen in Figure 4.6 below, followed by Table 4.1 comparing the z-value (99999 permutations) for each variable.

![Figure 4.6: Univariate Global Moran’s I all variable scatter plots (top row: dependency, material deprivation, ethnic concentration; bottom row: residential instability and proximity)]
Since each z-value is above 1.96 it can be concluded that there is significant spatial autocorrelation in each variable – especially proximity to healthcare at a high of 99.6767. This supports H1 as it demonstrates spatial autocorrelation across all five variables.

**Table 4.1: Z-value and Moran’s I for each ON-Marg dimension and the PMD**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Moran’s I</th>
<th>Z-value (99999 permutations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependency</td>
<td>0.228</td>
<td>24.3798</td>
</tr>
<tr>
<td>Material Deprivation</td>
<td>0.482</td>
<td>51.3021</td>
</tr>
<tr>
<td>Ethnic Concentration</td>
<td>0.596</td>
<td>63.5902</td>
</tr>
<tr>
<td>Residential Instability</td>
<td>0.588</td>
<td>62.4480</td>
</tr>
<tr>
<td>Proximity to healthcare (PMD)</td>
<td>0.937</td>
<td>99.6767</td>
</tr>
</tbody>
</table>

To test for the significance of clustering in each dimension the Univariate Local Moran’s I Local Indicators for Spatial Association (LISA) maps were also generated. The LISA for dependency in **Figure 4.7** below shows significant clustering of high levels of dependency towards the North, and clustering of low levels of dependency in downtown Toronto as suggested in **Figure 4.1** (significance maps for these clusters can be found in **Appendix A**). However, there are also some smaller DA’s scattered around Toronto that are assigned low-high and high-low cluster values, meaning that they are spatial outliers and different from their neighbours.
Figure 4.7: Dependency dimension Univariate LISA cluster map.

On the other hand, clustering is much more pronounced in the material deprivation and ethnic concentration dimensions. In Figure 4.8, the material deprivation LISA map shows significant clustering of low-low DA’s in central and Southwest Toronto, as well as clustering of high-high DA’s in Northwest and Northeast Toronto. Interestingly, the ethnic concentration LISA in Figure 4.9 exhibits a clear North-South divide, with higher levels of ethnic concentration in the North and lower levels in the South. If these three dimensions are compared side by side a clear spatial pattern emerges. Though there are a few outliers in each of the three maps, it can be concluded that DA’s in North Toronto contain a greater number of clusters of significant high-level marginalization. As previously discussed in section 4.1, North Toronto is home to many visible minority populations and poorer immigrant families making it highly marginalized. In comparison, South Toronto remains relatively free of marginalization, except for a few smaller DA’s, as the majority of DA’s are of the low-low cluster classification. However, the residential instability dimension presents an entirely different spatial pattern (see Figure 4.10). Rather than significant clusters of high-level marginalization existing in the north of the city, they are present in the south. However, this pattern is largely due to the variables the dimension
includes – such as the number of apartment buildings, individuals who are not youth (ages 5-15), and dwellings that are not owned. As previously discussed, the increased presence of apartment buildings is largely due to high population density and the value of the land South along the waterfront. Furthermore, the existence many famous universities such as the University of Toronto and Ryerson University who provide housing for their students accounts for the number of dwellings that are not owned. This makes residential instability quite spatially different from the other three dimensions.

Figure 4.8: Material deprivation dimension Univariate LISA cluster map.
Figure 4.9: Ethnic Concentration dimension Univariate LISA cluster map.

Figure 4.10: Residential Instability dimension Univariate LISA cluster map.
The LISA for the PMD in Figure 4.11 on the other hand was interesting as it contained no spatial outliers. Clusters of low-low DA’s are prevalent in Toronto’s periphery, while clusters of high-high DA’s confine themselves to downtown Toronto. Again, this is unsurprising given the population density here and the number of hospitals located downtown. As previously discussed in section 4.1, the majority of healthcare facilities are primarily located where population density is the greatest. That is, where they can serve the public most effectively. Therefore, suburban areas, though they do contain healthcare facilities, contain significantly less in comparison since their population density is much lower. This impacts one’s ability to quickly access healthcare when needed, consequently explaining the clustering of low-low access around the periphery. With this map, we are truly able to see the significant lack of healthcare access in suburban Toronto. Most importantly, the proximity LISA map and the four ON-Marg LISA maps also further support H1 as there is clustering of the data (i.e. high positive spatial autocorrelation).

Figure 4.11: Proximity Measures Database (PMD) Univariate LISA cluster map.
4.3 OLS Regression

To test for spatial dependence and bias that may exist, an OLS regression model was conducted. The probability values of the Lagrange Multiplier and Robust LM values seen in Table 4.2 below were of particular importance.

Table 4.2: OLS Regression Results

<table>
<thead>
<tr>
<th>Test</th>
<th>MI/DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagrange Multiplier (lag)</td>
<td>1</td>
<td>7343.4022</td>
<td>0.00000</td>
</tr>
<tr>
<td>Robust LM (lag)</td>
<td>1</td>
<td>694.2180</td>
<td>0.00000</td>
</tr>
<tr>
<td>Lagrange Multiplier (error)</td>
<td>1</td>
<td>6924.6702</td>
<td>0.00000</td>
</tr>
<tr>
<td>Robust LM (error)</td>
<td>1</td>
<td>275.4861</td>
<td>0.00000</td>
</tr>
<tr>
<td>Lagrange Multiplier (SARMA)</td>
<td>2</td>
<td>7618.8883</td>
<td>0.00000</td>
</tr>
</tbody>
</table>

Since all probabilities for both the LM and Robust tests are less than 5% (0.05) they are all significant. This means that there is spatial dependence, and that spatial lag and spatial error influence the regression. Therefore, it was necessary to run a lag-error regression model using GeoDa Space.

4.4 Lag-error Regression

For the lag-error regression the probability of each variable, as well as their coefficient value and the pseudo- r-squared will be used to explain the relationship between the variables (see Table 4.3). Again, if a variable has a probability value below 5% (0.05) it is significant, which will indicate which relationships are the most important. Coefficient values will then be assessed based on whether they are positive or negative, and the pseudo r-squared will help to identify how much variance in the proximity score is explained by the ON-Marg variables.

Based on the pseudo r-squared value, it was concluded that 95% of the variability in the proximity variable was explained by the four ON-Marg variables – which is quite a high percentage. The results for W_prox_idx_h and Lambda are also important, as W_prox_idx_h represents the spillover of the dependent variable (the spatial lag of the PMD) and Lambda represents the spillover of the error term. The probability values for both W_prox_idx_h and
Lambda indicate that there is significant spatial spillover for both lag and error because their probability of 0.0000 is less than 0.05 (and their z-statistics are outside +/-1.96). Since error is negative, this can suggest that there is something not captured by the Ontario Marginalization Index that negatively impacts accessibility.

### Table 4.3: Lag-error Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependency</td>
<td>0.0003611</td>
<td>0.0002856</td>
<td>1.2642102</td>
<td>0.2061546</td>
</tr>
<tr>
<td>Material Deprivation</td>
<td>-0.0004092</td>
<td>0.0002723</td>
<td>-1.5026536</td>
<td>0.1329284</td>
</tr>
<tr>
<td>Ethnic Concentration</td>
<td>0.0006265</td>
<td>0.0002443</td>
<td>2.5649742</td>
<td>0.0103183</td>
</tr>
<tr>
<td>Residential Instability</td>
<td>0.0016183</td>
<td>0.0004085</td>
<td>3.9619886</td>
<td>0.0000743</td>
</tr>
<tr>
<td>W_prox_idx_h</td>
<td>0.9828726</td>
<td>0.0086544</td>
<td>113.5689204</td>
<td>0.0000000</td>
</tr>
<tr>
<td>Lambda</td>
<td>-0.4494985</td>
<td>0.0594790</td>
<td>-7.5572627</td>
<td>0.0000000</td>
</tr>
</tbody>
</table>

Pseudo r-squared: 0.9491

The probability values for both dependency and material deprivation indicate that they are not significant. Based on their coefficient values (positive or negative), it was concluded that as material deprivation increases healthcare access decreases, and as dependency increases access also increases. However, since both variables are not significant, this relationship may be false and caused by random chance. On the other hand, ethnic concentration and residential instability are significantly associated with accessibility since their probability values are less than 0.05. Moreover, since both their coefficients are positive, this indicates that accessibility to healthcare increases as ethnic concentration and residential instability increase. This relationship is quite strange and counter intuitive. However, if their choropleth maps are re-visited we can see that there is an ethnic enclave located where accessibility to healthcare is also the highest – Chinatown. Only one street over west of Chinatown lies Toronto Western Hospital, and two streets over east of Chinatown lie three major hospitals: Mt Sinai Hospital, Toronto General Hospital, and The Hospital for Sick Children. Similarly, we can see that there is a lot of instability downtown, perhaps due to the rising prices...
of the condo/housing market and the increased presence of Universities who provide housing to their students. Overall, these results suggest that H2 is not correct for Toronto - instead greater marginalization is associated with greater healthcare access. Though this relationship is significant, it is important to acknowledge that it is very slight. For instance, an increase of 1 on the ethnic concentration portion of the ON-Marg index only increases the PMD index by 0.0006265. This is quite small, considering this index ranges only from 0 to 1.
Chapter 5: Discussion & Conclusion

This section of the research paper is organized in the following order. First, the methods and results of the study are reviewed, followed by a discussion of how these results compare to those found in recent literature. The implications and strengths of the study are then examined. Finally, the limitations of the study are considered, and the paper is concluded with final thoughts and recommendations.

5.1 Summary of the Research

The objectives for this study were the following: 1) analyze the distribution of each of the ON-Marg dimensions across Toronto’s dissemination areas (DA); 2) examine spatial patterns of access and marginalization in Toronto and identify any clusters (high or low) within the city; 3) Explore the relationship between accessibility to healthcare calculated in the PMD and the four dimensions of marginalization in ON-Marg. By mapping and statistically analysing each ON-Marg dimension as well as the PMD, objective 1 and 2 were completed. Exploratory choropleth mapping allowed for an initial look into the spatial distribution of each variable, while the LISA mapping determined where significant clustering and spatial autocorrelation was occurring. The results produced from the Univariate Global Moran’s I also determined whether spatial autocorrelation was occurring for each variable. On the other hand, objective 3 was completed using regression analysis. By initially completing an OLS regression and then a Lag-error regression, the relationship between marginalization, defined through the four dimensions of ON-Marg, and accessibility to healthcare was examined.

Prior to running the analysis two hypothesis were stated: H1) That positive spatial autocorrelation (clustering) is present in Toronto for both the PMD and ON-Marg data; H2) That as marginalization increases, healthcare accessibility decreases. In this case, H1 addressed objectives 1 and 2, while H2 addressed research objective 3. The results of the LISA and Univariate Global Moran’s I confirmed that H1 was correct. The presence significant low-low and high-high clustering, as well as the fact that all variables had a z-value above 1.96 demonstrated that there was significant spatial autocorrelation in each variable. However, the results of the lag-error regression determined that H2 was possibly incorrect. The results
indicated that as material deprivation increased healthcare access decreased, and that as dependency increased access also increased. However, for both of these variables their results were not significant - which could indicate that this relationship is false and caused by random chance. On the other hand, ethnic concentration and residential instability were significantly associated with healthcare accessibility – though the significance of the relationship was very slight. Since both their coefficients were positive, it indicated that ethnic concentration and residential instability increased as accessibility to healthcare also increased. Based on these values, it was concluded that greater marginalization is associated with greater healthcare access in Toronto.

5.2 Discussion

The majority of studies that incorporate ON-Marg often conclude similarly: that their field of study and marginalization exhibit some correlation. However, like this study, not all ON-Marg variables prove to be significant and vary. For instance, if the results of this study are compared to the multimorbidity study conducted by Moin (2018), a few similarities and differences present themselves. Rather than finding both ethnic concentration and residential instability significant, Moin finds that the occurrence of multimorbidity was greater is areas where material deprivation and residential instability is higher. In other words, only one of the significant variables (residential instability) from Moin’s study was also significant in this study. On the other hand, Wang (2018) finds that all ON-Marg variables are significant in their study but for certain types of crime. Specifically, neighbourhoods with high levels of material deprivation and residential instability are associated with larger concentrations of violent crime, and those with lower levels of ethnic concentration and higher levels of residential instability are associated with higher levels of property crime (Wang, 2018). This result is different from this study as only two variables were found to be significant here - not all of them. Interestingly, however, Silverman’s (2013) study correlates with this study quite nicely as they find both ethnic concentration and residential instability to also be significant. They conclude that local areal dependency and material deprivation were not significant factors in the cyclist collision,
and that the number of collisions increased in neighbourhoods where residential instability and ethnic concentration was high (Silverman, 2013). As previously discussed in the literature review, all three studies have found higher levels of residential instability to be a significant factor. Therefore, it was not surprising to find that this study produced this result as well.

On the other hand, this study contradicts the results of recent healthcare accessibility studies. For instance, a study conducted by Wang (2011) recognizes the inadequate level of healthcare access for various ethnic immigrant populations. Despite the variability between each group, Wang acknowledges how low the accessibility scores assigned to all of the immigrant groups are compared to the general population. Khandor (2011) argues similarly by recognizing the disconnect between healthcare access and the homeless population in Toronto. Out of the 366 homeless adults participating in their study, only 43% claimed to have access to a family doctor. There seems to be a consensus in recent literature that marginalized communities in Toronto do not have adequate access to proper healthcare. However, the results of this study produce different findings. Rather than having less access to healthcare as these three studies suggest, areas exhibiting higher levels of marginalization seem to be associated with greater access to healthcare – which was not expected making this study quite unique.

However, it is important to recognize when making this deduction that though accessibility in this study is defined spatially through the PMD as the dependent variable, there are other barriers to healthcare accessibility. As discussed in the literature review, accessibility is also affected by barriers such as income, education, and other unique circumstances not captured by the PMD. Though ON-Marg does recognize some of these barriers, it measures marginalization more broadly in the province rather than access specifically. While a marginalized individual may live close to healthcare, there may be other barriers limiting their access (e.g., work or family support). Therefore, it is important to recognize that there is variation within these broad groups to avoid ecological fallacy. Assuming that everyone living in a certain DA experiences the same level of marginalization or has the same circumstance is likely incorrect and problematic.
Lastly, it is important to recognize that the PMD and ON-Marg are predictors of access, rather than measures of access themselves. Though an individual may live close to a healthcare facility, this does not necessarily mean that they will use the service available. In other words, access does not necessarily equal utilization. Our behaviours, circumstance, and social status influence our ability to access these services. Therefore, the practicality of both data sources, ON-Marg and the PMD, as well as the validity of the results of this study come into question.

5.3 Implications

The findings of this study imply that each individual and the city they live in is different. Toronto’s diversity and growth have made it an irregularity, especially when diversity is thought about in the context of healthcare access and marginalization. Without this diversity, the results of this study would likely have been similar to that found in recent literature - that marginalized populations lack proper access to healthcare. Therefore, when interventions look to increase healthcare access for marginalized communities, it is critical that they acknowledge who lives there and the history they may have. What makes the city different? What makes it the same as other cities? By recognizing each city’s own unique history, the interventions that aim to improve healthcare access will reach all people including those who are marginalized.

5.4 Strengths and Contributions of the Research

The first strength of this study is an obvious one. Due to the recent release of the Proximity Measures Database (PMD) in 2020, this study will be one of the first to use these proximity measures, as well as analyze them in combination with the Ontario Marginalization Index. The results of this study could therefore inform the use of the PMD in future studies. The second strength of this study is that it incorporates a robust form of spatial analysis. By incorporating exploratory choropleth mapping, cluster analysis using both global and local Univariate Moran’s I, OLS regression, and a Lag-error regression model each variable and the relationships between them was fully explored. A third strength of this study is the insight it provides into healthcare accessibility in Toronto, and its contribution to Canadian research.
Although many studies related to healthcare accessibility exist in Canada, very few have used ON-Marg. The last strength of this study is its ability to address questions that emerge from the WHO about healthcare access – such as “is everyone getting some?”. This study proves that there are people who continue to lack the proper access to healthcare, despite their socio-economic status and locating where universal healthcare is provided.

5.5 Limitations and Future Research

Though the study was thorough it did have a few limitations. The first limitation stems from the data itself. While ON-Marg is certainly beneficial, it does contain suppressed values at smaller levels of geography, and a variety of different versions set at a specific time which could skew the results of this analysis should one use data during a different year. This was the case for this study, as the PMD was created during the year 2020 while ON-Marg contained census data from 2016. The census is also limiting, as some populations such as indigenous or institutionalized people (e.g., care homes) are excluded or under-counted. Moreover, the use of multiple principal component factor analyses (a data reduction technique) to create the four dimensions, has also likely resulted in the loss of valuable data. Therefore, this method may have produced an inaccurate representation of marginalization in Toronto. The PMD, on the other hand, is also limited due to source inconsistency which could cause the accessibility score to be incorrect. To avoid this issue however, the proximity measure for some DA’s were suppressed. Furthermore, the inclusion of major healthcare facilities in the PMD may have resulted in an inaccurate representation of access as it does not account for the spatial influx of people, that travel a far distance, to receive specialized care. Lastly, as discussed in section 5.2, both the PMD and ON-Marg are indicators of access, meaning that their scores do not necessarily guarantee the use of a particular healthcare facility. Therefore, the practicality of both data sources and the validity of the results of this study come into question.

To improve this research, further investigation into why greater marginalization can lead to greater healthcare access is needed. This could be achieved by going to neighbourhoods where marginalization and access is high and distributing a survey to local residents. Questions
about their demographic, and whether they believe they have proper access to healthcare could be enquired. Moreover, the study area could also be expanded. Rather than focusing on Toronto, the research area could be focused to include only the greater Toronto area (GTA,) or broadened to include all of Ontario (or other major cities) to see if similar patterns hold. For instance, if the methods used in this study were applied to another city of a more suburban or rural nature where population density is lower, the accessibility measures would likely be extremely different and could produce different findings. By focusing on Toronto specifically, this limited the generalizability of the final results of the study. Lastly, it would be interesting to see how the results of this study would vary if alterations to the data were made, or different measures of accessibility and marginalization were used. For example, removing hospitals from the analysis would also be interesting as accessibility to healthcare could be assessed without their influence.

5.6 Conclusion

These results indicate that City of Toronto has done well in distributing healthcare facilities where they are needed the most – in marginalized communities. This is quite rare, as recent literature has described the inadequate condition of healthcare in most of these communities. Nevertheless, it can be argued that improvement is still needed. Many neighbourhoods in Toronto still lack proper access and would certainly benefit from the increased presence of healthcare facilities. However, this is easier said then done, as the majority of neighbourhoods in Toronto that lack access to healthcare are those in the periphery where population density is lower. This is unsurprising, as literature has also recognized the disproportionate level of healthcare access between urban, suburban, and rural areas. However, as Marmot (2008) has recognized, achieving equal access to healthcare is not impossible. Through proper planning and comprehensive action that addresses the social determinants of health and aims to remove structural inequality, this will be achievable (Marmot, 2008).
Appendix A – LISA Significance Maps

ON-Marg Dependency Dimension LISA Significance Map:

ON-Marg Material Deprivation Dimension LISA Significance Map:
ON-Marg Ethnic Concentration Dimension LISA Significance Map

ON-Marg Residential Instability Dimension LISA Significance Map
Proximity Measures Database (PMD) LISA Significance Map
References


