

**Linking Preventable Hospitalisation Rates to Neighbourhood Characteristics
within Ottawa**

Geneviève Prud'homme

Master Thesis in Science in Health Systems

University of Ottawa

Supervisor: Kevin Brand – Telfer School of Management

Co-Supervisor: Elizabeth Kristjansson – School of Psychology

© Geneviève Prud'homme, Ottawa, Canada, 2012

Abstract

Enhancing primary care is key to the Canadian health care reform. Considered as an indicator of primary care access and quality, hospitalisations for ambulatory care sensitive (ACS) conditions are commonly reported by Canadian organisations as sentinel events signaling problems with the delivery of primary care. However, the literature calls for further research to identify what lies behind ACS hospitalisation rates in regions with a predominantly urban population benefiting from universal access to health care. A theoretical model was built and, using an ecological design, multiple regressions were implemented to identify which neighbourhood characteristics explained the socio-economic gradient in ACS hospitalisation rates observed in Ottawa. Among these neighbourhoods, healthy behaviour and - to a certain extent - health status were significantly associated with ACS hospitalisation rates. Evidence of an association with primary care accessibility was also signaled for the more rural neighbourhoods. Smoking prevention and cessation campaigns may be the most relevant health care strategies to push forward by policy makers hoping to prevent ACS hospitalisations in Ottawa. From a health care equity perspective, targeting these campaigns to neighbourhoods of low socio-economic status may contribute to closing the gap in ACS hospitalisations described in this current study. Reducing the socio-economic inequalities of neighbourhoods would also contribute to health equity.

Disclosure

“The research and analysis are based on data from Statistics Canada and the opinions expressed do not represent the views of Statistics Canada.”

“The research and analysis are also based on data from the Canadian Institute for Health Information - obtained from Ottawa Public Health - and the opinions expressed do not represent the views of these organisations.”

Acknowledgments

First and foremost, I offer my deepest gratitude to my two supervisors, Prof. Kevin Brand and Prof. Elizabeth Kristjansson from the University of Ottawa (UofO), who supported me throughout the process with their knowledge, expertise, patience and passion. Their encouragements and advices were more than crucial to the conception, analysis and writing of my thesis project, I am truly thankful for their contributions and I cherish the relationships we developed. I particularly enjoyed the on-going conversation with my supervisor Kevin, who continually pushed me further in my reflective and intellectual process. A special thank you goes to all, and in particular to Prof. Sawada from UofO, who were involved in the development of the Ottawa neighbourhoods and their health measures as part of the multidisciplinary research project of the Ottawa Neighbourhood Study, an initiative based at the Institute of Population Health of the UofO. I am more than pleased by having had the opportunity to be part of this project lead by my co-supervisor Elizabeth, thank you for having me as part of your team. The recommendations of Prof. Ted Schrecker from UofO in the early stages of my thesis were also most valuable to me and his recommendations as an external examiner were greatly appreciated. A special thank you goes to Eric Crighton from the Department of Geography of UofO for his thorough review of my thesis as an external examiner: his comments vastly improved the content of my work.

Numerous other people were indispensable for the completion of this thesis. Starting with Katherine Russell, Epidemiologist at Ottawa Public Health (OPH), who initiated the study of ACS hospitalisations in Ottawa and accepted that I pursue it for my thesis project. Cameron McDermaid, Epidemiologist at OPH, embraced my research project, organised my data sharing agreement with the institution and provided me with the customised hospitalisation and ER visit counts that I needed. His knowledge and expertise of the DAD and NACRS databases as well as his on-going advices were really valuable. I am also grateful to Leigh Ann Butler, Epidemiologist at OPH, and Amira Ali, Senior Epidemiologist at OPH, who thoroughly explored with me the

possibilities of using data from the Rapid Risk Factor Surveillance Survey to derive my neighbourhood compositional estimates.

The two analysts of the COOL RDC - Jean-Michel Billette and Darcy Hango - with whom I collaborated over the course of my thesis project were great support to me. A special and heartfelt 'merci' goes to Jean-Michel for his precious intellectual input on the methodological aspects of this study. I am beholden to Jean-Michel's expertise in statistics and infinite willingness to help.

I am also really grateful for the valuable assistance received from senior staff at Statistics Canada who accepted to meet with me and share their expertise on the Canadian Community Health Survey: Georgia Roberts, Chief of the Data Analysis Resource Centre; Marie-Claude Duval, Senior Methodologist and Steven Thomas, Senior Methodologist. Similarly, I want to thank Paul Villeneuve, Senior Research Scientist at Health Canada, for his insights on standardisation methods.

Goes without saying the valuable input from the following people who answered specific questions related to my thesis: Prof. Quon from Telfer Management School, Yanyan Gong from the Canadian Institute for Health Information, and Jackie Schultz and Dr. Hogg from the C.T. Lamont Primary Health Care Research Centre. My apologies if I am missing someone.

I could not omit my beloved parents who were always behind me throughout my studies at University; their love, presence, encouragements, and patience were most valuable to me. To them and to my brothers, I want to say thank you for believing in me. And, thank you to my dear friends for your support and for making sure I still had a life during my studies! A special mention goes to my classmate and friend, Marie-Andrée Cadieux, to whom I wish the best success.

Table of Contents

Chapter 1.	Introduction	1
1.1.	ACS Hospitalisations and Socio-economic Differences	2
1.1.1.	Primary Care Access Hypothesis	5
1.1.2.	Primary Care Utilisation Hypothesis	7
1.1.3.	Health Status Hypothesis	11
1.1.4.	Healthy Behaviour Hypothesis.....	12
1.1.5.	Other hypotheses	13
1.2.	Current Research Need	13
1.3.	Research Objectives and Questions	15
1.3.1.	Research Objective 1	15
1.3.2.	Research Objective 2	16
1.3.3.	Research Objective 3	16
Chapter 2.	Causal Model of Preventable Hospitalisation Rates.....	18
Chapter 3.	Methodology.....	26
3.1.	Overview	26
3.2.	Study Area	27
3.3.	Data Sources	31
3.3.1.	Discharge Abstract Database (DAD)	31
3.3.2.	National Ambulatory Care Reporting System (NACRS)	31
3.3.3.	Canadian 2006 Census (Statistics Canada)	31
3.3.4.	Canadian Community Health Survey (CCHS)	32
3.3.5.	Other.....	32
3.4.	Variables of Interest	33
3.4.1.	Dependant Variable.....	35
3.4.2.	Independent Variables.....	35
3.5.	Data Preparation	41
3.5.1.	Taking Age into Consideration.....	41
3.5.2.	Taking Time into Consideration	42

3.5.3.	Deriving Rates of ACS Hospitalisation and ER Visits	45
3.5.3.1.	Selection of ACS Conditions.....	45
3.5.3.2.	Database Extraction of ACS hospitalisations and ER Visits.....	48
3.5.3.3.	Methodology for Deriving Rates: Age- and Gender- Standardisation.....	51
3.5.3.4.	ACS Hospitalisation and ER Visit Rates – Neighbourhood Estimates in Ottawa	57
3.5.4.	Deriving Neighbourhood Estimates from the Canadian Community Health Survey	61
3.5.4.1.	Selection and Pooling of Neighbourhood Respondents.....	62
3.5.4.2.	Deriving Representative Neighbourhood Estimates Using Pooled CCHS Respondents	67
3.5.4.3.	Neighbourhood Level Imputations.....	71
3.5.4.4.	Further Comments on Measurement Quality of the Variables Derived from CCHS.....	73
3.6.	Analyses Models.....	75
3.6.1.	Analyses Model for Research Objective 1	75
3.6.2.	Analyses Models for Research Objectives 2 and 3	79
3.6.2.1.	Applicability and Usefulness of Linear Regressions.....	79
3.6.2.2.	Multiple Linear Regression Model – Estimation, Assumptions and their Diagnostics Methods	83
3.6.2.3.	Statistical Inference from Linear Regression Models	92
3.6.2.4.	Sensitivity Analyses – Population-weighted Regression Models.....	97
3.7.	Regression Model Preparation.....	100
3.7.1.	Exploration of the Variables of Interest.....	101
3.7.2.	Assessment of Model Validity using a Full Exploratory Regression Model.....	104
3.7.2.1.	Evaluation of the absence of multivariate outliers among the independent variables.....	105
3.7.2.2.	Evaluation of the absence of multicollinearity in exploratory model	105
3.7.2.3.	Evaluation of the normality, linearity, homoscedasticity and independence of the residuals.....	107
3.7.2.4.	Evaluation of the Potential Presence of Outliers in the Solution	107
3.7.2.5.	Evaluation of the ratio of cases to independent variables	109
3.7.3.	Selection of the Variables for the Final Regression Models.....	110

Chapter 4. Results	114
4.1. Results for Research Objective 1	114
4.2. Results for Research Objective 2	120
4.3. Results for Research Objective 3	125
Chapter 5. Discussion	131
5.1. Neighbourhood Preventable Hospitalisations and Social Deprivation in Ottawa	131
5.2. Strengths and Limitations of the Analyses Models: Interpreting the Results and Formulating Generalisations	134
5.3. About Hypotheses Explaining Differences in Preventable Hospitalisations in Ottawa	140
5.3.1. Interpretation of Lack of Significance of Hypothesis of Differences in Primary Care Access	141
5.3.2. Interpretation of the Unexpected Results of the Sequential Models Testing the Differences in Primary Care Utilisation Hypothesis	147
5.4. About Neighbourhood Compositional/Contextual Characteristics and Preventable Hospitalisations in Ottawa	153
5.4.1. Compositional Neighbourhood Characteristics Associated with ACS Hospitalisation Rates in Ottawa	153
5.4.1.1. Neighbourhood socio-economic status	153
5.4.1.2. Neighbourhood prevalence of unhealthy behaviour	154
5.4.1.3. Neighbourhood Prevalence of Poor or Fair health status	156
5.4.2. Contextual Neighbourhood Characteristics Associated with ACS Hospitalisation Rates in Ottawa	158
5.4.2.1. Neighbourhood Distance to Primary Care Facilities	158
5.4.3. Interpretation of Lack of Significance of Other Compositional and Contextual Neighbourhood Aspects	161
Chapter 6. Conclusion	163
6.1. Research Contributions	163
6.2. Future Research	166
References	169

Appendix 1		184
Appendix 2		185
Appendix 3		192
Appendix 4		194
Appendix 5		201
Appendix 6		202
Appendix 7		203
Appendix 8		205
Appendix 9		207
Appendix 10		209
Appendix 11		210
Appendix 12		212
Appendix 13		214
Appendix 14		216
Appendix 15		217

List of Figures

Figure 1: Neighbourhood Primary Care Access and Preventable Hospitalisations .	19
Figure 2: Neighbourhood Utilisation of Primary Care and Preventable Hospitalisations	20
Figure 3: Quality of Primary Care Used and Preventable Hospitalisations.....	21
Figure 4: Health Status & Risky Behaviour Norm and Preventable Hospitalisations	22
Figure 5: Place & Environmental Quality and Preventable Hospitalisations.....	24
Figure 6: Causal Model of Neighbourhood Compositional and Contextual Factors Influencing ACS Hospitalisations Rates	25
Figure 7 City of Ottawa: Features and Neighbourhoods	27
Figure 8: Crude ACS Hospitalisation Rates by Age Groups for Ottawa	52
Figure 9: Comparison of ACS Hospitalisation Derived Using 10- and 20- Years Age Intervals.....	56
Figure 10: Funnel Plot – ACS Hospitalisation Rates	58
Figure 11: Distribution of Percent Who Did Not Visit a GP by Neighbourhood Population Size	74
Figure 12: Ottawa Neighbourhoods by Quintiles of ACS Hospitalisation Rates	115
Figure 13: Ottawa Neighbourhoods by Deprivation Quintiles.....	116
Figure 14: Distribution of ACS Hospitalisation Rates by Neighbourhood Deprivation Quintiles in Ottawa	117
Figure 15: Final Model – Predicted vs Observed Values of ACS Hospitalisation Rates.....	128
Figure 16: Illustration of the Pattern of Relationship Predicting Negative Suppression	152

List of Tables

Table 1: Social and Economic Characteristics of Ottawa and Comparisons	30
Table 2: List of Variables of Interest	34
Table 3: Selected ACS Conditions and Related ICD-10-CA Codes	47
Table 4: Neighbourhood with Exceptional ACS Hospitalisation Rates	59
Table 5: Number of CCHS Respondents per Cycle in ONS Region	64
Table 6: Distribution of the Number of CCHS Respondents per Neighbourhood	64
Table 7: Variables Selected for Final Regression Model (Standard Regression) ..	112
Table 8: Variables Selected for Sequential Regression Models	112
Table 9: Reasons for the Exclusion of Variables of Interest in Final Models	113
Table 10: Descriptive Statistics of ACS Hospitalisation Rates by SES Quintiles...	116
Table 11: ANOVA Results – ACS Hospitalisation Rates by SES Quintiles	119
Table 12: Sequential Regression Models – Final Results (Weighted Equally)	123
Table 13: Sequential Regression Models – Sensitivity Analyses Results (Weighted by Neighbourhood Population Size)	124
Table 14: Full Model Final Results (Weighted Equally)	129
Table 15: Full Model – Sensitivity Analyses (Weighted by Neighbourhood Population Size)	130

List of Figures in Appendices

Figure A-1: ACS ER Visit and Hospitalisation Rates by SES Quintiles	184
Figure A-2: Causal Model Including All Variables of Interest.....	201
Figure A-3: Causal Model Representing Variables Selected for Final Model	202
Figure A-4: Relationship Diagram: Primary Care Access, Health Care Utilisation, ACS Hospitalisations and ACS ER Visits Rates.....	206
Figure A-5: Partition of the Variance in Dependent Variable in Multivariate Regression Model	211
Figure A- 6: Scatter Plot of SES Index vs ACS Hospitalisation Rates.....	214
Figure A- 7: Scatter Plot of ‘Percent with poor or fair health status’ vs ACS Hospitalisation Rates.....	214
Figure A- 8: Scatter Plot of ‘Percent without a family physician’ vs ACS Hospitalisation Rates.....	215
Figure A- 9: Scatter Plot of ‘Percent of women who never had a Pap test’ vs ACS Hospitalisation Rates.....	215
Figure A-10: Scatter plot of Studentized Residuals and Predicted Values – Exploratory Model	217
Figure A-11: Q-Q Plot of Residuals – Exploratory Model	217

List of Tables in Appendices

Table A-1: List of Ottawa Neighbourhoods and Selected Characteristics	185
Table A- 2: List of Variables of Interest and Related Information	194
Table A- 3: Descriptive Statistics – Variables of Interest.....	209
Table A- 4: Linear Model Assumptions and Diagnostic Methods	212
Table A- 5: Diagnostic Test for Multivariate Outliers, Leverage and Influence	213
Table A- 6: Univariate & Bivariate Regression Results – Variables of Interest.....	216

Chapter 1. Introduction

According to Health Canada (2004a, para. 1), "primary health care is key to maintaining and improving Canadians' health, and to the quality and sustainability of the health care system". In 2000, the federal, provincial and territorial Canadian governments agreed to renew their primary care system as it has the potential to address other health care system issues, including emergency room overcrowding (Romanow, 2002). In fact, a primary care reform was identified as the key to efficient, timely and quality care by the First Ministers of Canada (Health Canada, 2003). Enhancing access to appropriate primary care, via the supply of health care providers, was identified in the 10-year plan to strengthen the health care system in Canada (Health Canada, 2004b). Family physicians provide primary care by assuring direct first-contact services, coordinating continuity of care, easing reception of specialty care and responding to community needs (Health Canada, 2006; Hogg, Rowan, Russell, Geneu, & Muldoon, 2008; Starfield, 1998). Since primary care assures continuity of care, it can prevent the onset of severe conditions and reduce the need for tertiary care or hospitalisations (American College of Physicians, 2008; Marshall et al., 2006; Romanow, 2002).

In the literature, the rate of ambulatory care sensitive (ACS) hospitalisations – based on conditions for which timely and effective primary care can reduce the risks of hospitalisations – is used as an indicator of the accessibility and quality of the primary care system (Billings et al., 1993). In this Introduction Chapter, the concept of ACS hospitalisations and the tendency of neighbourhoods of low socio-economic status (SES) to have higher rates of ACS hospitalisation than better-off neighbourhoods within urban areas are presented (Section 1.1). Then, the four hypotheses formulated in the literature to explain the socio-economic gap in ACS hospitalisations in areas with a predominantly urban population are discussed (Sub-sections 1.1.1, 1.1.2, 1.1.3, 1.1.4 and 1.1.5). The need for additional research on the interpretability of the indicator of ACS hospitalisation rates and on the importance of these four hypotheses was expressed (Section 1.2). In the current study, these hypotheses are tested within an ecological framework. The specific research goals

are to: describe the rates of ACS hospitalisation across neighbourhoods of different socio-economic status, assess which hypotheses contribute significantly to differences in ACS hospitalisation rates between neighbourhoods and assess the contribution of neighbourhood characteristics to the rates of preventable hospitalisations for a Canadian environment with a predominantly urban population (Section 1.3).

1.1. ACS Hospitalisations and Socio-economic Differences

The concept of ACS conditions was developed by Billings et al. (1993) while studying hospitalisations in New York City. Using a modified Delphi approach¹, a medical advisory panel defined three basic categories to group causes of hospital admissions: 1) marker conditions, 2) ACS conditions and 3) referral-sensitive conditions. ACS conditions are those "for which timely and effective outpatient care can help to reduce the risks of hospitalisation by either preventing the onset of an illness or condition, controlling an acute episodic illness or condition, or managing a chronic disease or condition" (Billings et al., 1993, p.163). Basically, individuals suffering or at risk of suffering from an ACS condition should be managed in a timely manner on an outpatient basis by a primary care physician and, under most circumstances, should not be admitted to the hospital for this ACS condition. Therefore, hospitalisations for ACS conditions are also called 'potentially preventable hospitalisations' (Billings et al., 1996). For brevity, the shorthand 'preventable hospitalisation' refers to ACS hospitalisations in this study.

Billings et al. (1993) identified 26 different medical conditions as sensitive to ambulatory care. Subsequently, a number of groups were formed and studies were conducted to further investigate the identification of ACS conditions. In the US, the

¹ The Delphi approach is a methodology used to obtain "the most reliable consensus of opinion of a group of experts, by a series of intensive questionnaires interspersed with controlled feedback" (McKenna, 1994, p.1221). The conventional approach follows a few questionnaire-and-response feedback cycles. In the first cycle, the experts answer a questionnaire and are then provided with the anonymous responses from their peers. In the second cycle, the experts re-evaluate the same questionnaire while taking in consideration the responses from their peers. Normally, consensus is reached during the second cycle. In the event that it was not, a third cycle is implemented. There is no defined methodology for the modified Delphi approach, it may involve a face-to-face meeting between the experts or the experts may develop the questionnaire themselves. The results from studies using different modified Delphi methodologies are harder to compare. (McKenna, 1994)

Committee on Monitoring Access to Personal Health Care Services of the Institute of Medicine identified 9 causes of avoidable hospitalisation for chronic diseases and 7 causes of avoidable hospitalisation for acute conditions (Institute of Medicine, 1993). In the UK, panels of clinicians reviewed a set of 174 common hospital discharge diagnostics and identified 30 for which timely and efficient outpatient (or primary) care would prevent at least 70% of the hospitalisations. They also identified an additional 66 conditions for which 50-69% of the hospitalisations would be prevented (Sanderson & Dixon, 2000). In Canada, Brown et al. (2001) conducted three different panels (a Delphi panel, a modified Delphi panel and a survey based panel) where 13, 23 and 17 groupings of conditions were identified as ACS conditions by the different panels respectively. All three panels identified 8 groupings of conditions as being ACS conditions; these groupings strongly overlap with the ACS conditions identified in the US and the UK.

At the onset of studying ACS hospitalisations, large variations in rates of such hospitalisations were observed by socio-economic status (SES) status; where areas with low SES showed significantly higher rates than areas with high SES. For example, in Billings et al. study (1993), hospitalisation rates for ACS conditions such as asthma, diabetic ketoacidosis and hyperosmolar coma, bacterial pneumonia and congestive heart failure were 4.6 to 6.4 times higher in low income areas than in high income areas of New York City. When all ACS conditions were combined, the hospitalisation rate among low SES areas was 3.13 times higher than in high SES areas. These findings were consistent across urban settings in the US (Billings et al., 1996).

At that time, the hospitalisation rates for ACS conditions among low SES areas were also greater than in high SES areas in Canada, but the gap was generally not as large as in the US (Billings et al., 1996). More specifically, in Hamilton, Ottawa and Toronto respectively, the hospitalisation rates were 1.58, 1.79 and 1.39 times higher in low income than in high income areas (Billings et al., 1996). In a recent study from the Canadian Institute of Health Information (CIHI, 2008a), the pan-Canadian hospitalisations rate for ten groups of ACS conditions combined was 196 per

100,000 people in high SES areas, 285 in average SES areas and 458 in low SES areas, where a notable ratio between low and high SES areas (2.3) was still observed. Similar gaps between the rates of hospitalisations for ACS conditions between low and high SES areas were reported in a number of other Canadian reports and studies (Agha, Glazier, & Guttmann, 2007; Anderson, 1996; CIHI, 2009a; CIHI, 2009b; Disano, Goulet, Muhajarine, Neudorf, & Harvey, 2010; Fransoo, 2009).

The smaller SES gap in ACS hospitalisation rates in Canada compared to the US is probably mitigated by the publicly funded health care system, where differential access to primary care is reduced across the social groups and the socioeconomic gradient between groups may be less steep (Billings et al., 1996). Curtis (2004) demonstrated the influence of political structures on health care consumption, which may affect the ACS hospitalisation rates. However, there are still marked socioeconomic differences in ACS hospitalisation rates within Canada, suggesting that factors beyond the type of health care system are contributing to higher rates of ACS hospitalisation.

A number of aspects of place may have an impact on ACS hospitalisation rates. These are often categorised into two groups: compositional and contextual. Compositional aspects refer to the differences in the kinds of people who live in different places and contextual aspects refer to the differences in social and physical contexts among places (Macintyre & Ellaway, 2003). An example of a compositional factor which could influence the ACS hospitalisation rates between places would be the proportion of residents suffering from multiple chronic conditions. An example of a contextual factor would be the local supply of primary care physicians. However, it is important to mention that the distinction between compositional and contextual aspects of places may be artificial as "people create places, and places create people" (Macintyre & Ellaway, 2003, p.27). In fact, people in an area may be sicker due to poorer environmental conditions or poorer access to resources including health care.

Four different hypotheses are commonly formulated in the literature to explain differences in ACS hospitalisations between urban neighbourhoods. These hypotheses state that the differences are due to either: 1) differences in access to primary care, 2) differences in health care utilisation, 3) differences in health status and 4) differences in healthy behaviour (Billings et al., 1993; Bindman et al., 1995; Disano et al., 2010). These hypotheses are more fully described in the sub-sections below (Sections 1.1.1, 1.1.2, 1.1.3 and 1.1.4). Other hypotheses are also identified as potentially contributing to differences in ACS hospitalisations across neighbourhoods, yet their relative importance may be limited – as discussed in Section 1.1.5.

1.1.1. Primary Care Access Hypothesis

Hospitalisations for ACS conditions are recognised as an indirect measure of access to primary care; however its applicability in areas with universal health care was questioned (Disano et al., 2010). Access to health care is a "concept describing people's ability to use health services when and where they are needed", based on their availability, accessibility, accommodation, affordability and acceptability (Cromley & McLafferty, 2002, p. 234). The availability of the health care refers to the supply of the resources needed to provide health services (e.g. nurses, physicians, clinics). The accessibility of health care refers to the location where the services are provided as well as the building's facilities. Health care accommodation refers to how convenient the services are based on populations' needs (e.g. opening hours, language proficiency of health care providers). Health care affordability refers to the cost of the services borne by the population at the time of receiving care. The acceptability of health care refers to the way the population view the system (e.g. respectful of human dignity). (Cromley & McLafferty, 2002)

Primary care resources need to be delivered locally as many people need to use such services relatively frequently and their catchment area is relatively small compared to other health care services (Curtis, 2004). As conceptualised by Curtis (2004) in her theory on landscape of health care consumption, the accessibility and availability of health services is influenced by the social milieu, the

administrative/political structures and the physical infrastructures. Considering differences in administrative/political structures of places, primary care availability and accessibility is influenced nationally by ideologies on health care (e.g. universal access to care) and locally by differences in health care organisation and management (e.g. presence of community clinics versus private clinics) (Curtis, 2004). Also, the local distribution of physical health care infrastructures may be the result of historical investments or recent housing developments (Curtis, 2004). Research has shown that differences in the rates of ACS hospitalisation are associated fairly consistently with the availability and accessibility of primary care services (Ansari, 2007; Laditka, Laditka, & Probst, 2005). The marked differences in ACS hospitalisation rates between urban and rural areas in Canada also demonstrate the impact of reduced primary care supply and proximity on such hospitalisations (Fransoo, 2009; Shah, Gunraj, & Hux, 2003).

Having a family physician is considered as an important factor enhancing Canadians' access to primary care. Evidence are mixed as it relates to the association between having a regular source of primary care and ACS hospitalisation (Ansari, 2007). In a study at the individual level among Canadians diagnosed with an ACS condition, the individuals who were hospitalised for such conditions were more likely to have a regular medical doctor (Sanmartin & Khan, 2011). Yet, these individuals were also more likely to have lower socio-economic status and to report unmet health care need compared to those not hospitalised for their ACS condition. The authors suggested that individuals with lower economic status may experience limited access to the full range of primary care services required to reduce the risk of an ACS hospitalisation. These health services may extend beyond the publically provided services, (e.g. monitoring devices), which entail out-of-pocket expenditures. In the US, it was clearly demonstrated that the affordability of services influence the rates of ACS hospitalisation, where a number of studies demonstrated that higher rates of ACS hospitalisation prevail in areas with higher proportion of poor people who are more likely to be medically uninsured (Ansari, 2007; Billings et al., 1993). A similar relationship between out-of-pocket

expenditures and ACS hospitalisation may still prevail in a context with subsidised access to services provided by primary care physicians.

However, primary care access is more than the mere availability, accessibility and affordability of services. Access also implies that people have the ability to use the services. There are a number of barriers influencing the utilisation of health care services. These include language, culture, cost to get to services, administrative hurdles, convenience ... (Schreiber & Zielinski, 1997). Therefore, adequate supply and proximity of physicians does not guarantee that everyone has the capacity to get care since different people may be affected differently by barriers listed above. This highlights the potential importance of service accommodation and acceptability on differential ACS hospitalisation rates across areas – yet these were not specifically studied in the literature. In an effort to assess the capacity of people to use primary care services, as opposed to the availability and accessibility of primary care, some studies evaluated the impact of self-rated access to care on ACS hospitalisations. Such studies demonstrated that, after controlling for a number of other variables, self-rated access was negatively associated with ACS hospitalisations, therefore providing support to ACS hospitalisation rates as an indicator of primary care access (Ansari, Laditka, & Laditka, 2006; Bindman et al., 1995).

1.1.2. Primary Care Utilisation Hypothesis

It has been suggested that area-based differences in ACS hospitalisation rates can be explained by differences in health care utilisation. The concepts of health care utilisation relates to the consultations with health professionals, the medical procedures and quality of care received and well as people's propensity to use health care services. Even though the utilisation of services is directly linked with health care access, other factors besides access influence it.

Aday and Andersen developed a theoretical framework to examine the determinants of health care utilisation, both at the individual and community level (Aday & Andersen, 1974; Andersen, 1968; Andersen, 1995). This framework includes three

categories of determinants: a) predisposing, b) enabling and c) need. Predisposing determinants refer to the characteristics of individuals prior to the occurrence of an illness, characteristics which influence their propensity to use services. These characteristics include age, sex, education and occupation. Enabling determinants refer to the means by which people use health services; these include income, health insurances, place of residence and supply of services. Need refers to the health status of the individuals which is believed to be the immediate cause of service utilisation. The role of place in Aday and Andersen's framework (1974) was further assessed by Law et al. (2005) in a study conducted in four neighbourhoods of Hamilton. It was demonstrated that, after controlling for all three aforementioned categories of health care utilisation determinants, place still had an effect on reported health care utilisation and unmet health care need among neighbourhood residents with a regular source of care. Law et al. (2005) argued that the role of place in Aday and Anderson's health care utilisation model (1974) tends to be overshadowed by the discussion of SES, demographic and health system characteristics.

Evidence for the primary care utilisation hypothesis and its relationship with ACS hospitalisations is mixed. In an Australian study conducted at the neighbourhood level, the ACS hospitalisation rates were negatively associated with the rates of primary care visits (Ansari et al., 2006). This finding is aligned with the definition of ACS conditions since hospitalisations for those conditions should be prevented or reduced when timely and efficient primary care is received. Conversely, in a case-control study among urban American children, the number of primary care visits was not significantly associated with a reduction in the risk of ACS hospitalisation (Steiner et al., 2003). As acknowledged by the authors, this non-significant finding could be explained by the small sample size of their study, which reduced the power of the analysis at yielding a significant effect (Steiner et al., 2003). Just as controversially, in a study at the individual level among Canadians diagnosed with an ACS conditions, the individuals who were hospitalised for such conditions were more likely to be frequent users of care for both primary and speciality care (Sanmartin & Khan, 2011). However, when the risk of being hospitalised is adjusted

for demographic characteristics, health status and healthy behaviour, the association with frequent utilisation of care was diminished.

More interestingly, a study in Manitoba by Roos, Walld, Uhanova, and Bond (2005) demonstrated that urban people with low income visited primary care physicians more than their affluent counterparts and were also more frequently admitted to the hospital for ACS conditions. This later finding may appear contradictory for two reasons. First, low income residents are often less accommodated than high income residents by the operating hours of most primary care facilities – which are often opened only during the day (Stewart et al., 2005). This is because low income residents cannot always afford time off from work to visit a primary care physician. Second, the services received in a primary care setting should be sufficient to prevent a hospitalisation related to an ACS condition. Therefore, by visiting primary care physicians more often, low income residents should experience a lower risk of ACS hospitalisation than high income residents. Yet, lower income residents in that study were admitted to the hospital more frequently than higher income residents for an ACS condition (Roos et al., 2005). The authors explain this finding by suggesting that the effectiveness of primary care at preventing hospitalisation for an ACS condition may be greater in affluent neighbourhoods than deprived neighbourhoods (Roos et al., 2005).

The above findings suggest that other aspects of health care utilisation may explain the ACS hospitalisation gap between low and high income neighbourhoods. As proposed by Lemstra, Neudorf, and Opondo (2006), differential utilisation of primary care for preventative services, and chronic disease screening and management may contribute to the gap in ACS hospitalisations more than the mere utilisation of services. People with low SES may use health care services in a different way or for different reasons than people with high SES (Disano et al., 2010). Also, low SES people may navigate the health care system less efficiently or may receive primary care of poorer quality than high SES people. For example, low SES people are less likely to have visited a specialist - which requires a physician referral - than high

SES people (Dunlop, Coyte, & McIsaac, 2000). This may impact the quality of the chronic disease management received among low SES people.

The differences in ACS hospitalisations may be influenced by primary care quality. As stated in the definition of ACS conditions, it is not only the timely utilisation of primary care but also the effectiveness of the care provided that reduces the risk of hospitalisation for such conditions (Billings et al., 1993). In fact, hospitalisations for ACS conditions are recognized as an indirect measure of the quality of the primary care provided (CIHI, 2009a). In Canada, high income residents are significantly more likely to report that the quality of the health care they received is excellent and that they are very satisfied with the health care received than low income residents (Lasser, Himmelstein, & Woolhandler, 2006). Such differences may contribute to the SES gap in ACS hospitalisations.

Also, it has been suggested that communities with reduced access and quality of primary care may rely on emergency departments for many primary care services (Laditka et al., 2005). However, primary care services received in an emergency context lack continuity – since follow-up visits are uncommon – and may not prevent ACS hospitalisations as efficiently as services provided in a primary care setting. An increased reliance on emergency departments was associated with a non-significant increase in the rates of ACS hospitalisation among urban residents in the US (Laditka et al., 2005). However, the emergency visits considered in that study were not limited to ACS conditions.

In a context where primary care utilisation – for prevention or for curative services – is not negatively associated with ACS hospitalisations, higher ACS hospitalisation rates may be a "sign of primary care system problems" (Schreiber & Zielinski, 1997, p.283). Such problems include inappropriate primary care services and the impossibility of obtaining timely quality care. These are experienced among poorer residents, and could increase the rate of ACS hospitalisations among that group (Stewart et al., 2005). Another issue may be differences in primary care physicians referral patterns, where physicians may use different set of criteria when deciding to refer a patient to tertiary care based on his(her) SES (Agha et al., 2007).

Unfortunately, the impacts of these problems on the differences in ACS hospitalisations are hard to assess due to the lack of adequate predictor variables. No studies have incorporated variables representing such concepts in their analysis model.

1.1.3. Health Status Hypothesis

It is possible that the increased utilisation of health care services and higher ACS hospitalisation rates among low income residents be explained by their poorer health status (Agha et al., 2007). In fact, a person's need for health care – measured as the number of health care problems – is the most important determinant of a physician visit among Canadians (McIsaac, Goel, & Naylor, 1997). Furthermore, deprived Canadians suffer from lower life expectancy and more health problems (Dunlop et al., 2000; Roos & Mustard, 1997; Stewart et al., 2005; Wilkins, Berthelot, & Ng, 2002). The severity of diseases among low income residents may be greater than among high income residents; this would increase the risk of an ACS hospitalisation among this first group (Agha et al., 2007; Sanmartin & Khan, 2011). In fact, Sanmartin and Khan (2011) demonstrated that the risk of hospitalisation for an ACS condition among Canadians is enhanced by poor self-rated health status (by about 10 times) and by the presence of co-morbidity (by about 4.5 times).

Taking this into consideration, it is likely that low SES communities with poorer health status and higher utilisation of health services have a greater number of ACS hospitalisations because there are greater proportions of sick people, or severely sick people, within those communities (Disano et al., 2010). In New York City, the prevalence ratios² of asthma and diabetes between low and high income populations were 1.35 and 2.36, respectively (Billings et al., 1996). However, the ACS hospitalisation rate ratios³ for the same conditions were 13.88 and 11.90. Billings et al. (1993) concluded that, even if the prevalence of disease is important in

² The prevalence ratio is a ratio value comparing two probabilities, where the probability of interest (e.g. disease prevalence) in a group of people (e.g. low SES) is divided by the probability of interest in another group of people (e.g. high SES).

³ The ACS hospitalisation rate ratios represent the ratio of ACS hospitalisation rate among low SES people over the ACS hospitalisations rate among high SES people. This ratio serves the same purpose as the prevalence ratio.

explaining the gap in ACS hospitalisations, it only accounts for a small portion of the differences in ACS hospitalisation rates among low and high income places.

1.1.4. Healthy Behaviour Hypothesis

The increased prevalence of diseases and ACS hospitalisation rates among low income residents could be explained by higher prevalence of unhealthy behaviours among this group of people. People with lower SES are more likely to adopt unhealthy behaviour like smoking and excessive alcohol consumption than higher SES people (Billings et al., 1993; CIHI, 2008a). This increased adoption and/or maintenance of unhealthy behaviour can be related to a number of factors associated with poverty such as increased stress and lack of opportunity to behave healthily (Krueger & Chang, 2008; Lawlor, Frankel, Shaw, Ebrahim, & Smith, 2003).

A study in Australia studied this hypothesis using the percent of current smokers, percent using alcohol, and percent doing adequate physical activity. They found that only the percent of current smokers was associated with an increased ACS hospitalisation rate at the neighbourhood level (Ansari et al., 2006). In New York, 11.6% of people admitted to the hospital for an ACS condition had alcohol or drug dependence/abuse, with levels at 18.2% among low income residents and 4.3% among high income residents (Billings et al., 1993). In this study, the differences in unhealthy behaviours among the 34-44 year olds explained some of the differences between ACS hospitalisation rates in low and high SES areas, but not for the other age groups. In a study at the individual level in Canada, Sanmartin and Khan (2011) demonstrated that individuals who were hospitalised for an ACS condition were more likely to be daily or former smokers and to be physically inactive. Sanmartin and Khan (2011) also mentioned that the combination of poor nutritious habits, poor health and lower socio-economic status may place individuals, especially elderly women, at greater risk of hospitalisation for an ACS condition.

It seems likely that, the higher prevalence of unhealthy behaviours among low income people may account for a portion of the increased prevalence of diseases in this group and consequently for some of the ACS hospitalisations among low SES people.

1.1.5. Other hypotheses

Few authors mention that poorer neighbourhoods may be exposed to poorer environments, which might contribute to the SES gap in neighbourhood ACS hospitalisation rates. Based on Künzli et al. (2003) findings, Laditka et al. (2005, p.1156-1157) suggested that "areas with more days of unhealthy air should have higher rates of ACS hospitalisation, as poor air quality is associated with increased hospitalisation risk for several ACS conditions including asthma, congestive heart failure and chronic obstructive pulmonary diseases". However, this relationship was not significant in their study and no other study attempted to demonstrate the relationship between ACS hospitalisation and environmental exposure.

However, differences in environmental exposure are noted between neighbourhoods of low and high SES status in Canada. As it relates to air quality, neighbourhoods experiencing material and/or social deprivation also experience increased exposure to NO₂ in Montreal (Crouse, Ross, & Goldberg, 2009). In Hamilton-Burlington, both the mean air pollutant levels and the mortality rates were higher in lower income neighbourhoods; and air quality and SES level acted synergistically on the risk of death (Finkelstein et al., 2003). It is also hypothesised that environmental exposure, in addition to other characteristics of lower SES neighbourhoods, may increase population's vulnerability and may act as a psychosocial stressor, which can directly lead to illness (Gee & Payne-Sturges, 2004).

Other hypotheses based on the concept of place were not formulated in the literature to explain differences in ACS hospitalisation rates.

1.2. Current Research Need

There is some evidence of a relationship with ACS hospitalisation rates for each of the four hypotheses described above. Only Ansari et al. (2006, p.719), in a study in Australia, attempted to test those four hypotheses simultaneously and they concluded that: "independent of prevalence (of self-reported asthma, hypertension, diabetes and chronic obstructive pulmonary disease (COPD)), propensity to seek

care, disease burden and physician supply, better access was associated with lower ACS hospitalisation rates". This provides support for the validity of the ACS hospitalisation indicator in urban areas where universal access to care is available but further research is needed.

In Canada, Disano et al. (2010) identified the need to study the mechanisms and causes of the disparity in ACS hospitalisations as it suggests that low SES people may experience barriers to optimal primary care. The Canadian Population Health Initiative of CIHI identified the same research need and formulated the following research questions to guide future research:

"What lies behind hospitalisation rates for conditions for which hospitalisation are potentially avoidable? To what extent are hospitalisation rates for ambulatory care sensitive conditions, for example, an indirect measure of access to primary care? What other factors may be related to such hospitalisation rates?" (CIHI, 2008a, p.8)

Studying the differences in ACS hospitalisations between Canadian urban neighbourhoods and neighbourhood aspects related to this indicator are current research priorities. Enhancing our understanding of the meaning of the ACS hospitalisation rates indicator in Canada could contribute to the evaluation of the primary care reform implemented by all First Ministers in 2003, where enhancing primary care access was identified in the 10-year plan to strengthen the health care system (Health Canada, 2003, Health Canada, 2004b). Also, identifying strategies, via the health care system or not, to avoid such hospitalisation may yield savings in hospital care since the average acute care inpatient cost per visit for ACS conditions⁴ is estimated at \$5,700 in Canada (CIHI, 2012). Moreover, such research may also assist health services planning and other types of local planning in an effort to match expenditure to the notion of need. Depending on the relative role of compositional and contextual aspects of places, health policy may be directed towards individuals or health-damaging/promoting characteristics of places

⁴ The ACSC considered in this calculation were angina, asthma, COPD, diabetes, epilepsy, heart failure and pulmonary edema and hypertension (CIHI, 2012).

(Macintyre & Ellaway, 2003). This would address health inequities, as they relate to differences in opportunity such as access to health services, between different population groups (Whitehead, 1992). And, studying the factors contributing to preventable hospitalisations is important from a health equity perspective, where the goal is the "absence of unfair and avoidable or remediable differences in health among social groups" (Solar & Irwin, 2010, p.4).

1.3. Research Objectives and Questions

The purpose of the current research was identified in light of the research needs discussed above. This purpose is to study the neighbourhood factors that may explain the differences in ACS hospitalisation rates between neighbourhoods of a Canadian predominantly urban setting. The research is part of the Ottawa Neighbourhood Study (ONS), an initiative based at the Institute of Population Health of the University of Ottawa.

This investigation has three objectives presented and further developed below (Sub-sections 1.3.1, 1.3.2 and 1.3.3)

1.3.1. Research Objective 1

To illustrate the gap in ACS hospitalisation rates among low and high SES neighbourhoods in Ottawa and to identify its importance as well as its characteristics.

The research questions that will direct this objective is: Is there a gap in ACS hospitalisation rates between low and high SES neighbourhoods in Ottawa? If yes, what is the magnitude of the gap and is there a gradient of effect?

Using a different set of ACS conditions than the ones selected for the current investigation, the ONS demonstrated that there is a gap in ACS hospitalisation rates between low and high SES neighbourhoods in Ottawa (Appendix 1). It is therefore expected that there will be a significant SES gap between the rates of hospitalisation for the ACS conditions selected for this study, such that the lower SES neighbourhoods will experience significantly higher rates than the higher SES

neighbourhoods of Ottawa. It is further hypothesised that the ACS hospitalisation rates will be increasing across neighbourhood SES levels, therefore mirroring the socio-economic health gradient commonly reported in the literature.

1.3.2. Research Objective 2

To test the four different hypotheses formulated in the literature⁵ to explain differences in ACS hospitalisation rates between neighbourhoods of Ottawa.

The research question that will guide this objective is: which hypotheses commonly formulated in the literature explain the differences in preventable hospitalisations across the neighbourhoods of Ottawa?

For the current investigation, it is hypothesised that all four hypotheses formulated in the literature to explain differences in preventable hospitalisations will contribute significantly to neighbourhoods ACS hospitalisation rates in Ottawa. Also, it is believed that these hypotheses will help explain the gap in preventable hospitalisations across neighbourhoods of different socio-economic status.

1.3.3. Research Objective 3

To identify the importance of neighbourhood compositional and contextual factors as well as to assess the relative contributions of neighbourhood health status, healthy behaviour and access/utilisation of primary care services to the rates of ACS hospitalisation between neighbourhoods in Ottawa.

The research question that guides this last research objective is: in a universal health care system, what are the neighbourhoods compositional and contextual characteristics associated with ACS hospitalisation rates in the neighbourhoods of Ottawa? And, what are the relative contributions of each of the four hypotheses to these rates?

It is expected that the compositional and contextual aspects of neighbourhoods will be associated with neighbourhood preventable hospitalisation rates at different level

⁵ These hypotheses are neighbourhood differences in health status, in healthy behaviour, in primary care access and in primary health care services utilisation.

of significance. Differences in health status and healthy behaviour will have stronger associations with ACS hospitalisations rates than differences in primary care access and utilisation. In a system with universal access to health care, the utilisation of primary care services, especially preventative health services, will explain a greater proportion of differences in preventable hospitalisations than primary care access.

These three objectives will address the need for further research identified by Disano et al. (2010) and CIHI (2008a). The results will contribute to the interpretability of ACS hospitalisation rates as indicators of primary care access in an urban Canadian setting. Similarly, a better understanding of the characteristics of neighbourhoods influencing the rates of ACS hospitalisation between low and high SES areas will be sought.

Chapter 2. Causal Model of Preventable Hospitalisation Rates

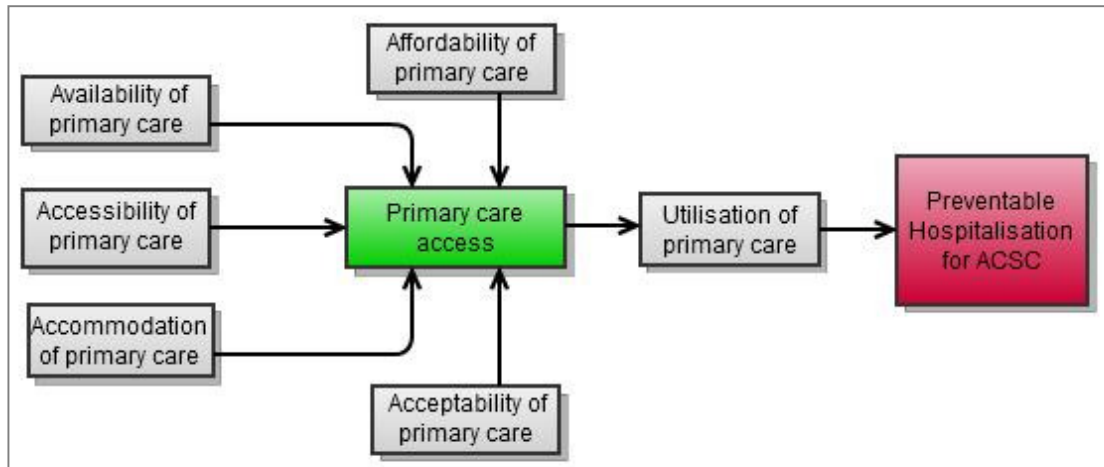
This chapter presents the causal model that was developed for the current investigation. It represents the different neighbourhood compositional and contextual factors identified in the literature and their theoretical relationships with neighbourhood preventable hospitalisations. This causal model also represents an effort to link the four different hypotheses formulated in the literature to explain the differences in neighbourhood ACS hospitalisations in an area where universal access to health care is available. Below, the causal model is introduced section by section, where each section represents one of the four hypotheses. All sections are then combined in the full causal model, which depicts all important relationships between the concepts and hypotheses of interest⁶. The full causal model also serves as an analytic framework for the current investigation.

Let us start with the first hypothesis stating that differences in primary care access explain the differences in ACS hospitalisations (Figure 1). In the literature, the rates of ACS hospitalisation are considered as a proxy for access to primary care services (Billings et al., 1993; Disano et al., 2010). This originates from the definition of ACS conditions which implies that timely and effective primary care can help to reduce the risk of hospitalisation for those conditions (Billings et al., 1993). However, people's access to primary care services is influenced by the availability, accessibility, accommodation, affordability and acceptability of primary care services (Cromley & McLafferty, 2002, p. 234). The above aspects influencing primary care access vary across places and between groups of people and can have an indirect impact on the local rates of hospitalisation for ACS conditions. For example, residents of low SES neighbourhoods may be located further from primary care facilities or may have greater difficulties travelling to those facilities than residents of high SES neighbourhoods. Similarly, the full range of primary care services may necessitate out-of-pocket expenditures which limits the affordability of services to people of lower SES status. Or, the acceptability of 'Western medicine' may be

⁶ The different concepts comprised in the causal model are greatly interlinked. To ease the interpretability of the model, only the most important relationships – as they relate to ACS hospitalisations - are depicted.

different for certain group of people living predominantly in low SES neighbourhoods, for example immigrants (Asanin & Wilson, 2008). These examples demonstrate that primary care access can vary across neighbourhoods. Also, access to primary care directly enables the utilisation of those services, which, in turn, is expected to reduce the risk of ACS hospitalisations.

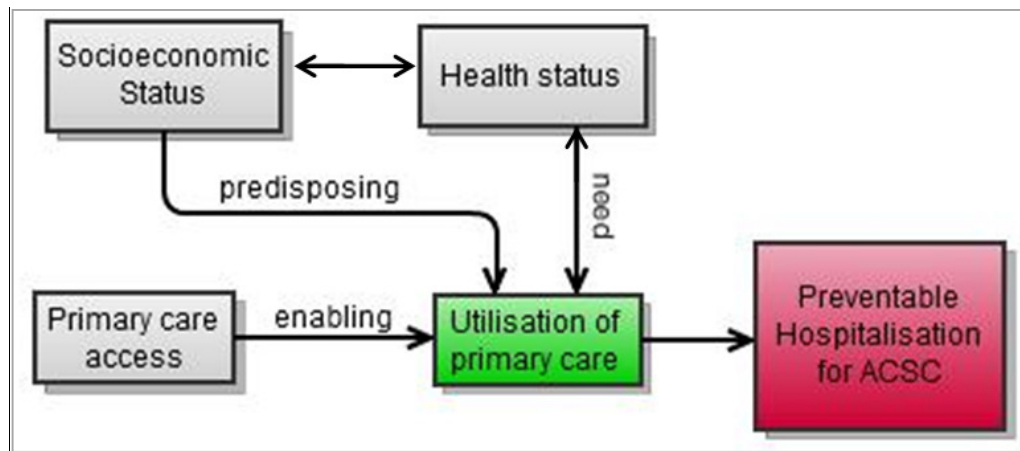
Figure 1: Neighbourhood Primary Care Access and Preventable Hospitalisations



Now, let us consider the second hypothesis – pertaining to the utilisation of primary care services – in order to enrich the causal model of neighbourhood preventable hospitalisation rates. Aday and Andersen (1974) developed a conceptual model listing three major determinants of health care utilisation: a) predisposing determinants like age, sex, education, occupation ... b) enabling determinants like health insurance, place of residence ... and c) need referring to the health status of individuals (Figure 2). Based on this model, for the same health condition and severity, people’s propensity to use primary care is influenced by their socio-economic status in concert with their access to primary care. For example, people with lower education may delay seeking care due to their unawareness of the importance of certain symptoms to one’s future health status (Disano et al., 2010). In such a context, residents of low SES neighbourhoods may not receive timely primary care as regularly as residents of high SES neighbourhoods due to their lower propensity for utilising those services. Therefore, a neighbourhood rate of preventable hospitalisation may be directly influenced by different patterns of primary care utilisation among their residents.

In addition to different rate of primary care utilisation, the quality of primary care received may be different across places and groups of people, which would also impact the rates of preventable hospitalisation in neighbourhoods (Figure 3). Since health status is improved, the risk of hospitalisation for an ACS condition is reduced when efficient primary care is provided – as stipulated in the definition of ACS conditions first established by Billings et al. (1993). Therefore, independently of access and utilisation of primary care, the quality of those services can directly influence the rates of ACS hospitalisation.

Figure 2: Neighbourhood Utilisation of Primary Care and Preventable Hospitalisations

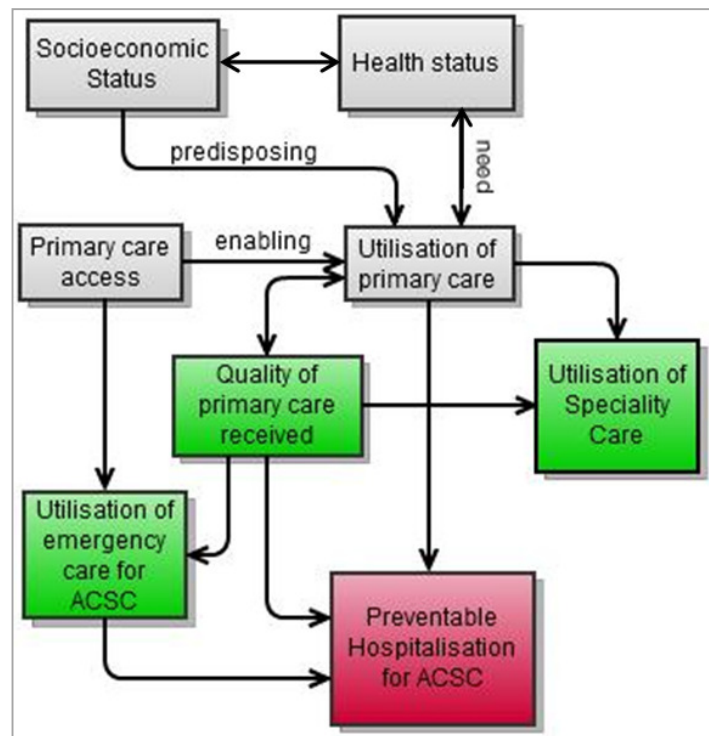


Lower quality of services can also influence other aspects of health care utilisation (Figure 3). For example, it can prompt additional visits to primary care physicians, which could have been avoided, or it can lead to a consultation in an emergency department. Neighbourhood reliance on emergency department for primary care services can also be the result of reduced access to primary care. Unfortunately, the quality of care tends to be lower in an emergency setting as opposed to a primary care setting due to the lack of continuity. Therefore, a visit to the emergency for an ACS condition may directly increase the risk of hospitalisation for that condition.

Because a referral from a physician is necessary to consult a medical specialist in Canada, the quality of primary care can influence the utilisation of specialty care. Low SES people tend to visit specialists less frequently than high SES people do (Dunlop et al., 2000). This may have an impact on the management of a chronic condition which, in turn, could lead to the onset of an ACS condition. However,

specialty care utilisation should not have a direct impact on the risk of preventable hospitalisations. This assumption is based on the definition of ACS conditions stipulating that the risk of hospitalisations for ACS conditions should be reduced by primary care services without the need for speciality care.

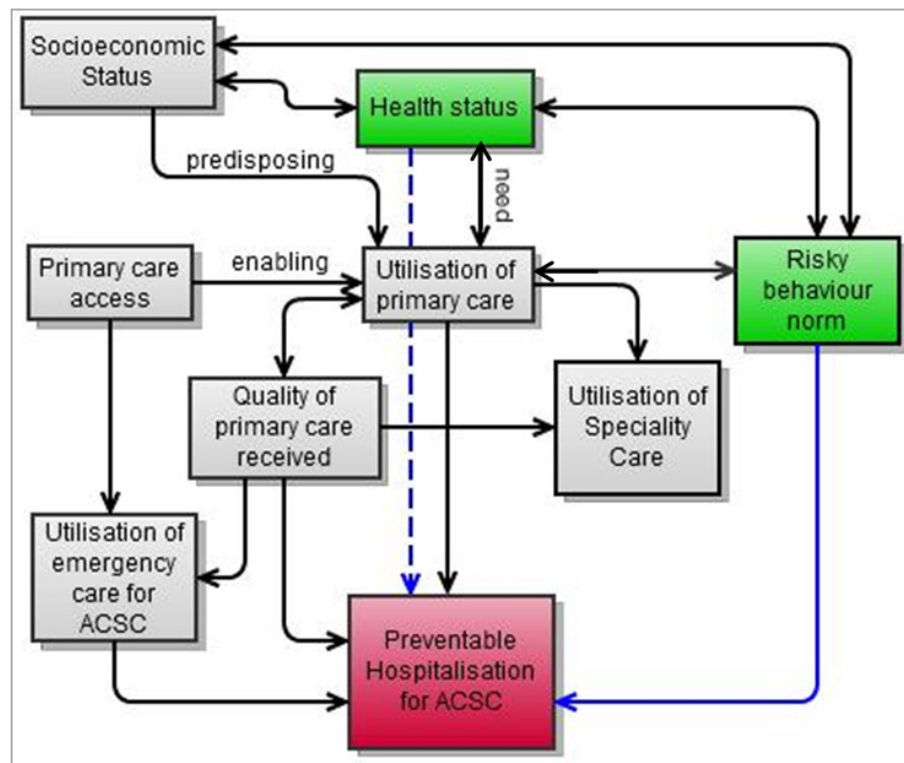
Figure 3: Quality of Primary Care Used and Preventable Hospitalisations



The third hypothesis which could explain the gap in preventable hospitalisations between neighbourhoods of low and high SES rests on the variations in health status across places (Figure 4). This assumes that a neighbourhood rate of preventable hospitalisation is directly influenced by the prevalence of diseases within that place, without being mitigated by the increased utilisation of primary care services by the residents suffering from those diseases or primary care access within that neighbourhood. In fact, primary care services are expected to reduce, but not eliminate, the risk of ACS hospitalisation. Therefore, neighbourhood with greater proportion of severely ill people may experience greater rates of ACS hospitalisation. The fourth hypothesis rests on the variations in healthy behaviours between neighbourhoods (Figure 4). It assumes that the rate of preventable hospitalisation is directly influenced by the prevalence of unhealthy behaviours

within a neighbourhood, mainly due to their associations as risk factors with most ACS conditions. It also acknowledges that unhealthy behaviour prevalence can impact the overall health status of that place and utilisation of primary care services. And, healthy behaviour norms vary across places and these variations may be associated by the socio-economic status of places (Ahern, Galea, Hubbard, & Syme, 2009).

Figure 4: Health Status & Risky Behaviour Norm and Preventable Hospitalisations

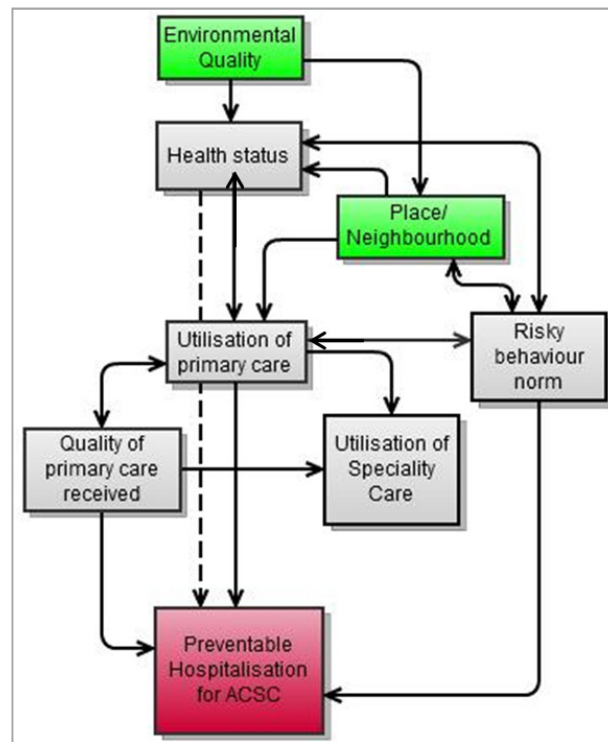


Less commonly formulated hypotheses rest on the role of place and environmental quality on the socio-economic gap in ACS hospitalisations (Figure 5). As mentioned by Ansari et al. (2006), environmental quality may influence the overall health status as well as the severity of certain conditions. For example, air quality affect the severity of asthma and respiratory infections (Lin, Stieb, & Chen, 2005). Both of those conditions are ACS, therefore neighbourhoods experiencing lower air quality may show higher rates of ACS hospitalisation.

As presented by Cresswell (2004), places such as neighbourhoods are the result of rich and complicated interplays between people and the environment, where meaning and experiences are created. Taking a social constructionist point of view, places are believed to be socially constructed in the context of unequal power relations and to generate different ways of being as well as ways of understanding, knowing and experiencing the world (Cresswell, 2004). Such outlook on place demonstrates its theoretical influence on ACS hospitalisation rates. A number of studies have highlighted the role of place itself on the health status and well-being of residents from different neighbourhoods within urban settings in Canada (Figure 5). It was demonstrated that measurable aspects of neighbourhoods, for example food and local resources, influence health outcomes over and above individual characteristics (Pouliou & Elliott, 2010). More specifically, Sanmartin and Khan (2011) suggested that the association between hospitalisation for an ACS condition, socio-economic status, poor health and poor nutrition among the elderly could be mitigated by enhancing community-based nutritional services. Beyond measurable aspects of neighbourhoods are people's perceptions of places. In Hamilton, neighbourhood stigma was shown to directly affect the health status and well-being of their residents (Elliott, Cole, Krueger, Voorberg, & Wakefield, 1999; Wakefield & McMullan, 2005).

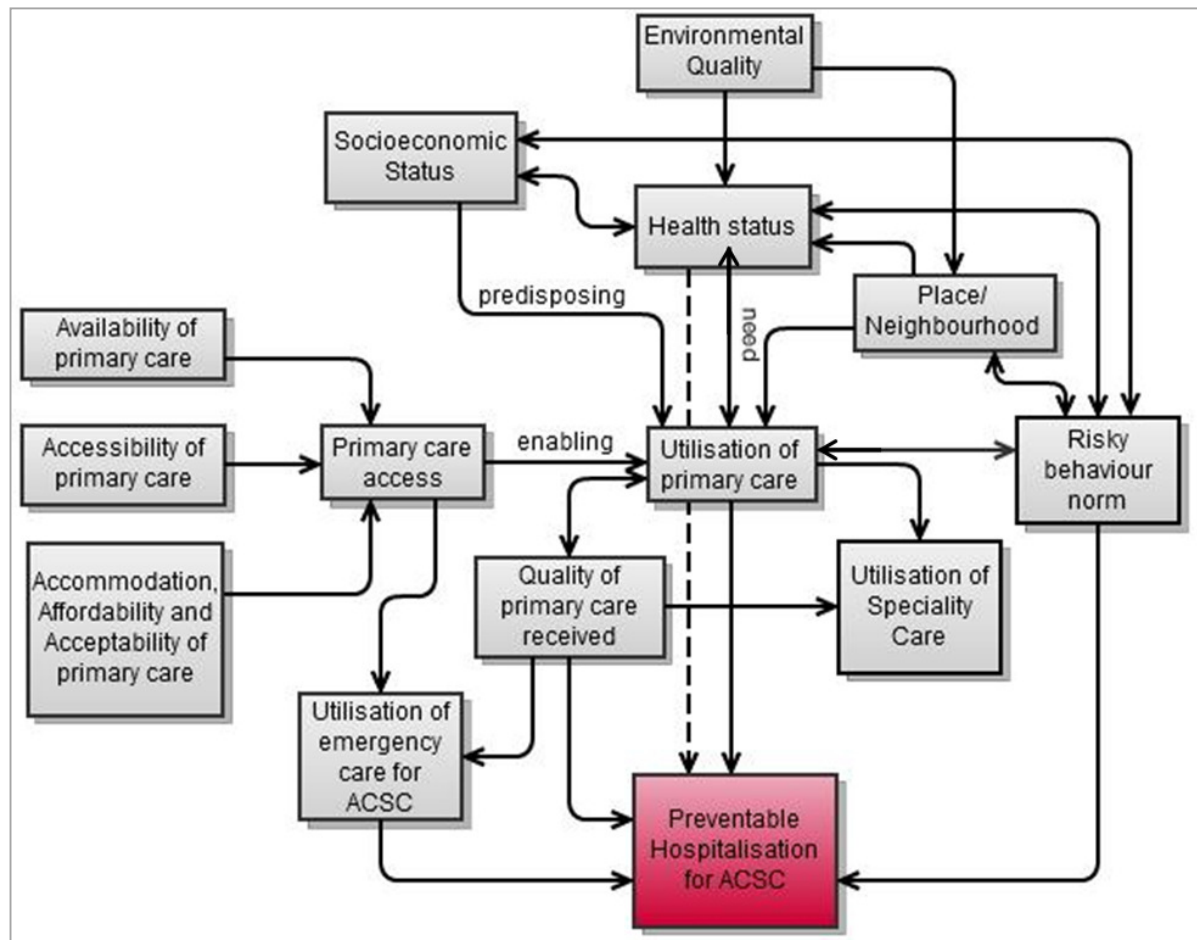
By influencing people's health, aspects of place such as the availability of recreation, food environments, environmental quality and social cohesion could indirectly impact the rate of ACS hospitalisation experienced in their neighbourhoods. Also, it was demonstrated that place influences other elements of this causal model. For example, social norms influence people's uptake and maintenance of risky health behaviours like smoking, which in turn may influence the ACS hospitalisation rates (Frohlich, Potvin, & Chabot, 2002). Moreover, place seems to have an impact on health care utilisation and people's propensity to use services over and above individual characteristics (Law et al., 2005).

Figure 5: Place & Environmental Quality and Preventable Hospitalisations



Overall, complex associations between neighbourhood primary care access and their residents health care utilisation patterns, health status, healthy behaviours and socio-economic status are thought to be at play. All those factors influence the rates of ACS hospitalisation of places directly and/or indirectly. The different hypotheses raised to explain the differences in preventable hospitalisation rates between neighbourhoods are also interlinked and their interactions influence their respective impact on the differences in preventable hospitalisations. Figure 6 summarises the relationships between those hypotheses as well as the interplay between compositional and contextual aspects of neighbourhoods that are thought to influence the rates of ACS hospitalisation between places.

Figure 6: Causal Model of Neighbourhood Compositional and Contextual Factors Influencing ACS Hospitalisations Rates



This causal model of neighbourhood hospitalisations for ACS conditions serves as a theoretical guide for the current research. The objectives and methodology of the current research hope to shed light on the theoretical relationships between the neighbourhood factors identified above and the rates of ACS hospitalisation among neighbourhoods of a Canadian setting with a predominantly urban population.

Chapter 3. Methodology

3.1. Overview

This study is interested in the particularity of neighbourhoods which can define people's experience of ACS hospitalisation rates. The research design of this study was non-experimental and ecological⁷; therefore the dependant variable (rates of ACS hospitalisation) and the predictor variables were defined at the neighbourhood level and derived using secondary source data. In it, analysis of variance and multiple linear regressions were used to assess the relationship between the neighbourhood predictors and ACS hospitalisation rates. This study recognised that the ecological design is subject to ecological fallacy, and that it cannot be used for making inferences about individuals based on grouped data⁸ (Diez Roux, 2003). However, the ecological design is congruent with the objectives and research need of the current investigation. In fact, the indicator of primary care access under investigation – the ACS hospitalisation rates – is conceptualised at the group level in the literature and is reported as such by Canadian institutions. Therefore, from a health planning perspective, studying the ecological associations with ACS hospitalisation rates is of interest, regardless of whether they are cofounded at the individual level (Diez Roux, 2003). Moreover, there are important spatial components to the distribution of primary care services that can be assessed via an ecological design. A multi-level study using both individual and neighbourhood level information would provide additional insights to the current research objectives, yet such design was not possible due data availability and linkage.

The Methodology chapter is divided in a number of sub-sections: where the study area (Section 3.2), the data sources (Section 3.3) and variables of interest (Section 3.4) are presented first. Following is the data preparation phase (Section 3.5), where the methods used to derive the rates and predictor variables are discussed along with their limitations and other considerations. The analysis models selected

⁷ Ecological studies use groups as the unit of analysis. All dependant and independent variables are measured at the group level and the variations in the outcome variable are also analysed at that level. Ecological studies are often used to investigate the relationship between characteristics of places and a health outcome conceptualized at the area level. (Diez Roux, 2003)

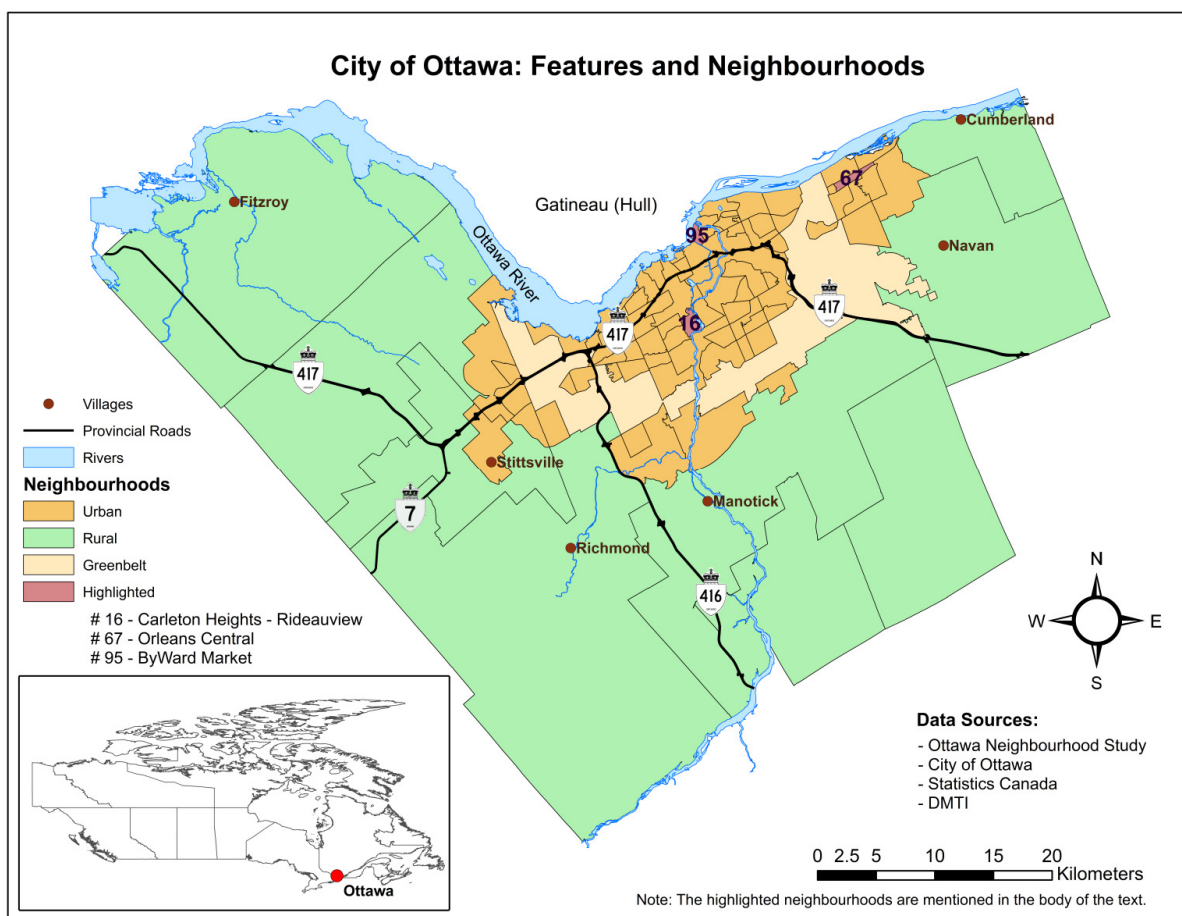
⁸ This ecological design was necessary due to a dearth of individual information in the data files.

including their assumptions and statistical tests are described in Section 3.6. And, lastly, the model preparation phase is presented (Section 3.7).

3.2. Study Area

The current study is conducted in the City of Ottawa, located in the province of Ontario in Canada (See Figure 7 for a map of its location and features). The City covers an area of more than 2,700 squared kilometers and stretches over a distance of 90 kilometers from East to West (Statistics Canada, 2007). About 80% of its land is rural and the overall population density is 292 people per squared kilometer. The majority of the 812,000 habitants of Ottawa live within the urban neighbourhoods identified in Figure 7 (Statistics Canada, 2007).

Figure 7 City of Ottawa: Features and Neighbourhoods



Compared to the Canadian population, Ottawa is characterised by a higher proportion of people with University certificate, diploma or degree (37% versus 23%) and a higher proportion of residents knowing both French and English (37% versus 17%) – based on the 2006 census (Statistics Canada, 2007). Although the median income is considerably higher in Ottawa compared to the rest of Ontario or Canada, the City has similar rates of unemployment as well as proportions of people and children under the low income (after-tax) cut-off. This may suggest that income disparity is greater in Ottawa than the rest of the country: among the neighbourhoods, the proportion of all persons living under the low income cut-off varies from 2% to 40%. Ottawa also has a greater proportion of row houses and apartments than the province of Ontario and Canada as a whole. See Table 1 for additional information on the context of Ottawa. Understanding the socio-economic context in which the current study is conducted is important and will be taken in consideration while interpreting the study results.

This study was conducted using the neighbourhoods delimited by the Ottawa Neighbourhood Study (ONS) and depicted on the map in Figure 7. Neighbourhoods units are defined as small areas having common predefined set of characteristics (Galster, 2001). At the local level, how place is conceptualised in health research is considered as the weakest theoretical aspects of health studies (Matthews, 2008, Parenteau & Sawada, 2011). Apart from the ONS, only a few Canadian studies are based on neighbourhoods as oppose to administrative areas (Parenteau & Sawada, 2011). The ONS neighbourhoods were operationalised specifically for health research, where the objective is to understand the physical and social pathways through which neighbourhoods in Ottawa affect health (ONS, 2010). In the first round of the ONS study, there were 89 neighbourhoods delimited based on economic, social and housing characteristics, physical barriers such as bridges and main streets, and the Ottawa Multiple Listing Service maps (Parenteau et al., 2008). Since then, some amendments to the neighbourhood boundaries have been made to be more congruent with local knowledge or city planning needs. In the second

revision of the neighbourhood delimitations, 97 neighbourhoods were defined⁹. To ensure data reliability, the majority of neighbourhoods are comprised of at least 4,000 persons and are formed by an aggregation of disseminations areas or postal codes (Parenteau et al., 2008).

In contrast to census tract delimitations or other administrative areas, the ONS neighbourhoods are thought to better represent natural areas in which health is constructed, promoted or damaged. In fact, these neighbourhoods were designed to increase internal homogeneity as well as external heterogeneity as it relates to socio-economic aspects of people and places. Therefore, they are delineated using processes known to be associated with the ACS hospitalisation rates and may better capture the neighbourhood aspects linked with this health outcome (Parenteau & Sawada, 2011). Also, using neighbourhoods comprised of relatively homogenous populations as opposed to census tracts ensures that the diversity between small areas is not attenuated when deriving group averages (Meade & Erickson, 2000). The scale and size of these neighbourhoods is believed to be adequate to study differences in health outcomes within a predominantly urban environment (Parenteau & Sawada, 2011).

For the variables derived as part of the current study, sufficient data was available for 90 of the 97 ONS neighbourhoods. Among the seven neighbourhoods that could not be used are two cemeteries (Notre-Dame and Beechwood), two industrial areas (Hunt Club South Industrial and Orleans Industrial), one university-based neighbourhood (Carleton University), one neighbourhood in development (LeBreton) and one rural neighbourhood (Stittsville – Bassewood): all sparsely inhabited at the time of study. Nine of the remaining 90 neighbourhoods are considered rural: they have low population density, are predominantly comprised of agricultural or forested areas even if a population centre or village falls within their boundaries and are

⁹ The rural neighbourhoods are currently being revised.

located outside the Greenbelt of Ottawa¹⁰. See Appendix 2 for the list of ONS neighbourhoods, where those identified with a star are considered rural.

Table 1: Social and Economic Characteristics of Ottawa and Comparisons

	Ottawa	Ontario	Canada	Neigh. Min	Neigh. Max
Social Demographics					
% lone-parent families	16%	16%	16%	6%	34%
% knowing both English and French	37%	11%	17%	19%	61%
% immigrants	22%	28%	20%	6%	55%
% recent immigrant (5yrs)	4%	5%	4%	0%	39%
% of visible minority	20%	23%	16%	1%	54%
% of population aged 70-79	5%	6%	6%	2%	16%
Education					
% no certificate, diploma or degree	15%	22%	24%	7%	32%
% high school certificate or equivalent	24%	27%	26%	13%	32%
% apprenticeship or trades certificate or diploma	6%	8%	11%	1%	11%
% college; CEGEP or other non-university certificate or diploma	18%	18%	17%	10%	24%
% university certificate, diploma or degree	37%	25%	23%	19%	68%
Economic Demographics					
Employment rate	65%	63%	62%	45%	78%
Unemployment rate	6%	6%	7%	2%	11%
Median earnings - Persons 15 years and over (\$)	34,343	29,335	26,850	22,737 *	81,390*
Median income in 2005 - Households (\$)	69,743	60,455	53,634	41,538*	160,565*
% low income after tax - All persons	12%	11%	11%	2%	40%
% low income after tax - Persons less than 18 years of age	15%	14%	13%	0%	57%
Housing					
% of owner	66%	71%	68%	20%	98%
% single-detached houses	43%	56%	55%	2%	96%
% semi-detached houses	6%	6%	5%	0%	34%
% row houses	19%	8%	6%	0%	63%
% apartments; duplex	2%	3%	5%	0%	17%
% apartments in buildings with fewer than five storeys	11%	11%	18%	0%	47%
% apartments in buildings with five or more storeys	19%	16%	9%	0%	81%

* The neighbourhood values represent the average income, not the median income.

Source: 2006 Census (Statistics Canada, 2007)

¹⁰ The Greenbelt is an arc of 20,350 hectare of open lands and forests surrounding the most densely populated area of the capital (National Capital Commission, 2009).

3.3. Data Sources

Five different data sources were used for the current investigation.

3.3.1. Discharge Abstract Database (DAD)

This database is hosted by the Canadian Institute for Health Information (CIHI) and contains records of hospital discharges – referring to patients who were hospitalised. It includes demographic, administrative and clinical information for all patients from hospitals in Canada except for the province of Québec (CIHI, 2005a). It was used to extract the counts of ACS hospitalisations by neighbourhood of residence and to derive the ACS hospitalisation rates themselves – see Section 3.5.3 for additional details.

3.3.2. National Ambulatory Care Reporting System (NACRS)

This database is also hosted by CIHI and contains records of all hospital-based and community-based ambulatory care patients who were treated either in day surgery, outpatient clinics and emergency departments (CIHI, n.d.a). It also contains demographic, administrative and clinical information and was used to extract the counts of ACS emergency visits by neighbourhood of residence and to derive the ACS emergency rates themselves – see Section 3.5.3 for additional details.

3.3.3. Canadian 2006 Census (Statistics Canada)

The ONS derived information on neighbourhood socio-economic characteristics from the 2006 census of Statistics Canada. From five items of the census questionnaire an index of material and social deprivation¹¹ was created by Kristjansson – see Section 3.4.2 for details. The neighbourhood population counts were also obtained from this census.

¹¹ Deprivation was defined by Townsend as "a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs" (Townsend, 1987. P.125). Material deprivation refers to the goods and conveniences of modern life and social deprivation refers to the relations within the family, workplace and community (Pampalon, Hamel, Gamache, & Raymond, 2009). Indices of social and material deprivation are used to represent numerically the socio-economic status of individuals or groups of people (e.g. neighbourhoods). Indices are a new and synthetic numerical variable that synthesises the information contained in a number of other measured variables and created using principal component analysis (Havard et al., 2008; Trochim & Donnelly, 2008).

3.3.4. Canadian Community Health Survey (CCHS)

A number of neighbourhood variables were derived from the Canadian Community Health Survey (CCHS) of Statistics Canada. The CCHS is a nationwide cross-sectional survey conducted every two years¹² since 2001 and it targets Canadians aged 12 years or older. It excludes the Canadians living on Indian Reserves, in some remote areas and in institutions as well as members of Canadian Forces Bases. The aim of the survey is to collect information on health status, health care utilisation and health determinants among the Canadian population. The questionnaires are comprised of core contents covering all Canada and different optional contents selected by the health regions at every cycle. For each cycle of the survey, a dataset containing more than 1,000 indicators is generated for more than 100,000 respondents across the country. The sample is designed to generate reliable estimates at the national, provincial and health region levels, where the Ottawa-Carleton health region (ID#3551) corresponds to the ONS region. In order to have a sufficient sample size for each neighbourhood of Ottawa, the respondents records from five cycles of the CCHS were combined (Cycle 1.1. in 2001, Cycle 2.1 in 2003, Cycle 3.1 in 2005, Cycle 4.1 in 2007 and Cycle 2009 in 2009) Access to the master files was obtained at the Carleton, Ottawa, Outaouais Local Research Data Centre (COOL RDC) – See Appendix 3 for details. Variables were selected from the CCHS based on their relevance with the concept of preventable hospitalisations and neighbourhood estimates were derived from the pooled sample – see Section 3.5.4 for additional details. (Statistics Canada, 2008)

3.3.5. Other

Contextual indicators of primary care access were derived from information on primary care resources and facilities provided by Dr. Hogg from the C.T. Lamont Primary Health Care Research Centre from the Élisabeth Bruyère Research Institute (E. Kristjansson, personal communication, February 2, 2010). It contained the address and number of physicians for every primary care facility of the Champlain Local Health Integrated Network in 2006, including private offices, community health

¹² Since 2007, yearly components of the CCHS are also released.

centers and walk-in clinics.¹³ M. Sawada, a Geography Professor at the University of Ottawa and a member of the ONS, used ArcGIS 9.0 to determine the location of the facilities and physicians within the Ottawa neighbourhoods and to derive the indicators.

3.4. Variables of Interest

Below is the description of the twenty-six variables of interest for the current study plus a short discussion of the concepts they represent in the context of preventable hospitalisations. Table 2 lists the variables of interest, the theoretical concept they represent as well as their data source; a detailed table containing all the information related to the variables derived is available in Appendix 4. For a graphical representation of the interactions between the variables of interest and the concepts they represented in the causal model of neighbourhood ACS hospitalisations, please refer to Appendix 5.

The construct validity of the measures is an important aspect to consider when identifying variables of interest. Construct validity is the degree to which inferences can legitimately be made from the variables selected to the theoretical constructs they represent as well as to the causal model of interest (Trochim & Donnelly, 2008, p.56). Such validity rests on the way the theoretical constructs are operationalised, which is the translation of the constructs or concepts of interest into their representing variables. In any research, the measurement validity of the variables selected is limited to the information available and the variables selected may not be the most appropriate representation of the concepts identified in the causal model. The variables of interest listed below are the best available measures for the concepts represent in the causal model of preventable hospitalisations, yet their construct validity is varying. For this reason, construct validity was considered while selecting the variables for the analysis models (8 predictor variables and one

¹³ This list was created from information provided by Dr. Graham and supplemented by information originating from the Ontario Family Health Network, the Academy of Medicine Ottawa and the College of Physicians and Surgeons of Ontario as well as from internet searches for community health centers and telephone book searches. All the practices listed were contacted, or attempts were conducted to contact them, in order to validate the information (J. Schultz, personal communication, September 20, 2011).

dependent variable) as well as the interpretation of the results. The analytical model representing the variables selected for the final models in the causal model of preventable hospitalisation is provided in Appendix 6.

Table 2: List of Variables of Interest

Concept Represented	#	Variable Name	Data Source
Preventable hospitalisations	1	ACS Hospitalisation Rates	Discharge Abstract Database (DAD)
Socio-economic Status	2	SES Index	Census
	3	SES Quintiles	Census
Health Status	4	% with fair or poor self-rated health status	CCHS
	5	% with 1 or more ACS conditions	CCHS
	6	% with 2 or more ACS conditions	CCHS
Healthy Behaviour	7	% of people smoking	CCHS
	8	% of people binge drinking in past year	CCHS
	9	% of people binge drinking in past month	CCHS
Utilisation and Quality of Health Care Services	10	Mean number of contacts with a primary care physician in past yr	CCHS
	11	% who did not visit or talk with a primary care physician in past yr	CCHS
	12	Mean number of contacts with a nurse in past yr	CCHS
	13	% who did not visit or talk with a nurse in past yr	CCHS
	14	Mean number of contacts with a medical specialist in past yr	CCHS
	15	% who did not visit or talk with a medical specialist in past yr	CCHS
	16	% who never received a flu shot	CCHS
	17	% who never received a Pap test	CCHS
	18	% who never received a breast examination	CCHS
	19	% who never received a mammography	CCHS
	20	% who rated the quality of the care provided by physician as being fair or poor	CCHS
	21	Rates of ER visits for ACS conditions	NACRS
Primary Care Access	22	Average distance to closest four primary care facilities	Élizabeth Bruyère Research Institute
	23	Primary care physician-to-population ratio	Élizabeth Bruyère Research Institute
	24	% without a family physician	CCHS
	25	% who rated the availability of health care in the community as being fair or poor	CCHS
	26	% with unmet health care need in past yr	CCHS

3.4.1. Dependant Variable

The dependant variable was the neighbourhood ACS hospitalisation rates for people aged 20-79 years.¹⁴ These rates were age- and gender- standardised using the 2006 census population of Ottawa (direct method) and expressed per 100,000 person-year – see variable numbered 1 in Appendix 4.

3.4.2. Independent Variables

a. Neighbourhood Compositional Variables

a.1. Socio-economic Status of Neighbourhoods

In Appendix 4, the variables numbered 2 (SES index) and 3 (SES quintile) both represented the concept of socio-economic status of the neighbourhoods. Neighbourhood SES was based on Kristjansson's index of socio-economic disadvantage as described in Parenteau et al. (2008). This SES index consists of the following measures: percent of residents with less than a high school education, percent of single-parent families, percent of unemployed residents, percent of households below the low income cut-off and average household income. The neighbourhood values for these measures were generated from the 2006 census information. Principal component analysis was used to combine the five measures into one new and synthetic composite: the SES index (Vyas & Kumaranayake, 2006). From this numerical index, the neighbourhoods were divided into quintiles, where the first quintile (Q1) was comprised of the most affluent neighbourhoods and the fifth quintile (Q5) of the most deprived neighbourhoods.

a.2. Neighbourhood Health Status

Variables representing neighbourhood general health status (#4 in Appendix 4) and ACSC disease prevalence (#5 and #6 in Appendix 4) were derived. With these variables, it would be possible to test the hypothesis based on health status to explain neighbourhood differences in ACS hospitalisation rates.

¹⁴ All variables originating from the CCHS were also derived for the 20-79 years old only. See Section 3.5.1 for details on age considerations.

The ACSC disease prevalence variables were: a) percent of respondents with one or more ACS chronic conditions, b) percent of respondents with two or more ACS chronic conditions. These were derived from four items of the CCHS representing *'long-term conditions that have lasted 6 months or more and were diagnosed by a health professional'*.¹⁵ The ACS conditions selected were asthma, high blood pressure, diabetes and heart diseases. Chronic bronchitis and COPD were also considered, yet differences in the formulation of the CCHS questions across the different cycles prevented their utilisation.¹⁶

Neighbourhood self-rated health status was derived from the CCHS question *'in general, would you say your health is excellent, very good, good, fair or poor?'* The variable derived was the percent of neighbourhood respondents who rated their health as being 'poor' or 'fair'. In contrast with the disease prevalence variables, this variable represented a self-rated indicator of neighbourhood health status and did not rely on a diagnosis from a health professional.

a.3. Neighbourhood Healthy Behaviour

The neighbourhood health behaviours were represented by three variables originating from two items of the CCHS: smoking and binge drinking. The first variable represented the respondents who answered 'daily' or 'occasionally' to the question *'at the present time, do you smoke cigarettes daily, occasionally or not at all?'* (#7 in Appendix 4) The other two variables represented proportion of neighbourhood residents engaging in binge drinking based on the frequency at which this behaviour is adopted: the percent who reported binge drinking in the past month or in the past year (#8 and #9 in Appendix 4). The former variable represented the proportion of neighbourhood residents engaging in risky health behaviour frequently as it included the respondents who answered 'once a month' or more frequently to the question *'in the past 12 months, have you had 5 or more*

¹⁵ The wording of the CCHS items originates from the survey questionnaire available at the COOL RDC.

¹⁶ It was planned to create, using principal component analyses (PCA), a composite indicator from the neighbourhood prevalence of the four chronic diseases available. However, the solution was not amenable to PCA and the composite variable was discarded.

drinks on one occasion? The later variable represented the proportion of neighbourhood residents engaging in risky behaviour frequently and less frequently as it also included the respondents who answered 'less than once a month' to the CCHS question above. These variables would be used to test the hypothesis based on neighbourhood differences in unhealthy behaviour prevalence put forward to explain the differences in ACS hospitalisations.¹⁷

a.4. Neighbourhood Utilisation of Health Services and Quality

The variables below would be used to test the hypothesis speculating that variations in health care utilisation explain the differences in ACS hospitalisations between neighbourhoods.

For each health professionals (primary care physicians, nurses and specialists), two types of variables were derived to represent neighbourhood utilisation of these types of services: a) the mean annual number of contacts with the health professional and b) the percent of respondents who did not contact the health professional in the past year. For the primary care physicians, these variables were based on the CCHS question '*in the past 12 months, how many times have you seen or talked on the phone about your physical, mental or emotional health with a family doctor/general practitioner?*' (#10 and #11 in Appendix 4) Even if derived from the same CCHS item, the concepts represented by these two variables are slightly different. The mean variable is closely related to the concept of health care need, which is directly related with poor health status. Yet, the proportion variable could be considered as a proxy of neighbourhood propensity to use primary care services, since it represent only the respondents who were not sick in the past year, who were sick but chose to not consult a primary care physician and who did not have an annual physical examination even if recommended.

¹⁷ Individuals who are physically inactive were found to be at higher risk of ACS hospitalisation by Sanmartin and Khan (2011). A weak relationship between nutritious habit and the risk of ACS hospitalisation was also demonstrated in this study published in 2011. Yet, no variable representing the healthy behaviours of physical activity and nutritious habit were derived for the current investigation: at the time of the creation of the neighbourhood variables, no evidence of a link between these healthy behaviours and preventable hospitalisations was published in the literature.

The mean and proportion variables for the utilisation of nursing services were based on the CCHS question '*in the past 12 months, how many times have you seen or talked on the phone about your physical, mental or emotional health with a nurse for advice?*' (#12 and #13 in Appendix 4) In the Canadian context, a nursing act performed without the supervision of a medical doctor represents a primary care service. Examples of such services are providing a general health advice, addressing a minor ailment, providing home care and supporting prevention campaigns such as smoking cessation or vaccination campaigns. The frequency of nursing service utilisation could be a proxy of primary care utilisation.

The two variables representing neighbourhood rates of specialty care use were based on the CCHS question '*in the past 12 months, how many times have you seen or talked on the phone about your physical, mental or emotional health with any other medical doctors such as a surgeon, a allergist, a orthopaedist, a gynaecologist, a psychiatrist, ...?*' These can be indicators of health care need as well as of the quality of primary care received.

In addition, neighbourhood rates of use of preventive primary care services were represented by a number of variables (#16 to #19 in Appendix 4). Those variables were the percent of CCHS respondents who never received the following preventive health services: 1) flu shot, 2) Pap smear test among females aged 18 years or more, 3) breast examination among female respondents aged 18 years or more and 4) mammography among females aged 35 years old or more. The percentages reflected the proportion of CCHS respondents who responded 'no' to the question '*have you ever had (the above tests name) tested or checked?*'¹⁸ It is important to note that these variables mainly represented women utilisation of preventative services. To avoid under-representing male utilisation of preventative services, it was hoped to derive variables representing the following preventive services: 1) blood pressure reading, 2) PSA testing among males aged 40 years or more and 3)

¹⁸ For the sake of reducing dimensionality, it was planned to create an index of neighbourhood utilisation of preventive health services from these variables using principal component analysis (PCA). However, the solution was not amenable to PCA and this composite index was discarded. The original variables were retained as potential proxies of preventable health service utilisation for the final models.

complete physical examination. Yet, these items were part of the optional content of the CCHS and were selected for less than three of the five cycles for the health region of Ottawa-Carleton. The limited sample size of the pooled respondents who answered these items precluded the creation of these variables at the neighbourhood level.

The quality of physicians care received by the neighbourhood residents was derived from a self-rated measure of the CCHS (#20 in Appendix 4). It represented the percent of respondents who have seen a physician in the past year and answered 'fair' or 'poor' to the question: '*thinking of the most recent care you received from a physician, how would you rate the quality of the care you received?*'¹⁹

Lastly, neighbourhood rates of emergency (ER) visits for ACS conditions among the 20-79 years old were also derived (#21 in Appendix 4). They were age- and gender-standardised using the 2006 census population of Ottawa (direct method) and expressed per 100,000 person-year – see Section 3.5.3 for details. This variable was considered as a proxy of primary care access and quality.

b. Neighbourhood Contextual Variables

b.1. Neighbourhood Primary Care Access

The variables representing the different aspects of access to primary care, listed below, could be used to test the hypothesis speculating that differences in ACS hospitalisation rates are linked to differences in neighbourhood primary care access.

The variable representing primary care accessibility (proximity) was the average distance to the closest four primary care facilities (expressed in meters) for each neighbourhood (#22 in Appendix 4). This indicator was derived by Prof. Sawada using Pearce, Witten, and Bartie (2006) technique. Network analysis based on road network files was used to determine the nearest facilities, located inside or outside

¹⁹ Here, it is important to note that different people may judge the quality of the services received differently. For example, people highly educated may be more critical of the care they received as they may be more informed about standard procedures or alternative treatments.

the neighborhood boundaries, and distances were calculated from neighbourhood population-weighted centroid (Parenteau et al., 2008).

Two variables represented the availability of primary care. The first variable was the supply of physicians, captured by the indicator of primary care physician-to-population ratio (adjusted by 1,000 people) (#23 in Appendix 4). Prof. Sawada derived this indicator using the number of physicians per facility located within each neighbourhood and the 2006 census population. The second variable was the percentage of respondents who answered 'yes' to the CCHS question: '*do you have a regular medical doctor?*' (#24 in Appendix 4) Having a regular doctor can facilitate the utilisation of timely primary care when an acute and minor health problem is experienced. It also enhances the management of chronic health conditions and increases the likelihood of receiving preventative services like an annual physical examination as well as vaccination. These are all primary care acts that may prevent hospitalisation for ACS conditions. This later variable could also be perceived as a compositional aspect of neighbourhoods, embedding both notion of primary care accommodation and quality.

A variable was derived from the CCHS to capture the concept of self-perceived availability of health services; therefore some aspects of health care accommodation, affordability and acceptability could be contained in this variable. It was represented by the percent of neighbourhood respondents who answered 'fair' or 'poor' to the question: '*overall, how would you rate the availability of health care services in your community?*' (#25 in Appendix 4) It is important to note that this item was not specific to primary care only and that the community may not translate into neighbourhoods. This variable was an approximate measure of the concept it sought to represent.

A commonly used indicator of access to health care is the proportion of individuals who expressed having unmet health care needs. This variable was estimated as the percent of CCHS respondents who answer 'yes' to the following question: '*during the past 12 months, was there ever a time when you felt that you needed health care but you didn't receive it?*' (#26 in Appendix 4) This indicator was not specific to

primary care services, so it can reflect access issues experienced for other type of health care (e.g. surgery). This variable could also be perceived as a compositional aspect of neighbourhoods.

b.2. Other Neighbourhood Contextual Factors

Other contextual aspects of neighbourhoods and places, for example air pollution exposure or access to affordable nutritious food, may have an impact on the rates of ACS hospitalisation as indicated in the causal model (Section 1.1.5). These aspects were not tested in the current investigation, either due to the unavailability of the information or their distant relationship hypothesised with the concept of preventable hospitalisation.

3.5. Data Preparation

3.5.1. Taking Age into Consideration

To achieve the second and third objectives of this study, it was important to derive the ACS hospitalisations / ER visits rates and neighbourhood compositional variables based on the same age range. In this study, the age ranges were truncated on the low- and high- sides. The rationale for truncation involved some combination of concern over comparability of different data sources, concern over the homogeneity of the included age ranges, and concern over the type of ACS conditions selected by age ranges chosen by other investigators.

The different data sources used different age ranges: the ACS hospitalisations / ER visits were available for all ages and the CCHS data for respondents aged 12 years old or more. The demographic information – that was needed for standardisation - as well as the counts of ACS hospitalisations / ER visits were available by 10 years age intervals only. Due to these differences, it was decided to exclude the 0-19 years old from the population under investigation.

It was also decided to exclude the population older than 80 years due to differences in the type of ACS conditions relevant for this age group compared to younger adults (Steiner et al., 2003; Walker, Teare, Hogan, Lewis, & Maxwell, 2009). In fact, the

lists of ACS conditions identified by Billings et al. (1993) and Brown et al. (2001)²⁰ were developed for the general population and include conditions primarily affecting younger individuals. To make the ACS conditions more applicable to the elderly population, subsequent research based on medical expertise suggested excluding certain ACS conditions, for example congestive heart failure and pneumonia (Steiner et al., 2003). Steiner et al. (2003) also suggested adding certain conditions for the elderly, such as septicemia and stroke. It was also justified to exclude children and youth from the population studied for similar concerns over the relevance of the ACS conditions in this age group. In a study among 0-14 years old in Spain, no hospitalisation for angina, hypertension, diabetes and congestive heart diseases were identified - as expected due to these diseases' etiology (Casanova, Colomer, & Starfield, 1996). Moreover, the list of ACS conditions for the children and youth population was modified by a number of investigators (Steiner et al. 2003).

In summary, due to the age characteristics of the data sources available and the age dependencies of certain ACS conditions, the ACS hospitalisations and ER visits rates were derived for the 20-79 years old only. The neighbourhood compositional characteristics developed from the CCHS data were also based this age group only, for comparability reasons. The SES index was the only compositional variable derived from the whole neighbourhood population in the current investigation. The other variables of interest (e.g. distance to primary care and supply of primary care physician) were contextual aspects of neighbourhoods; therefore independent from the population subgroup assessed.

3.5.2. Taking Time into Consideration

Considering that causal relationships cannot be ascertained from ecological designs, temporal precedence²¹ was not a necessary criterion of the current investigation. However, to identify neighbourhood characteristics associated with

²⁰ These lists were used to identify the ACS conditions selected for the current investigation.

²¹ Temporal precedence is one of nine criterion identified by Hill (1965) for epidemiological research aiming at determining causal relationships between potential contributing factors and disease among individuals. This criterion states that the potential causal factor, representing a certain exposure, must precede the outcome chronologically to be considered as a causal agent (Bhopal, 2002). This criterion is necessary to established causality, but not sufficient.

the rates of preventable hospitalisation, it was important to rely on information covering a similar period of time. Due to inherent limitations associated with the data source and size of the neighbourhoods, relying on information originating from multiple years was a necessity. This information covers the period of 2000-2010 and was slightly different for the different data sources:

- The ACS hospitalisation and ER visit counts extracted from the DAD and NACRS were based on the fiscal years of 2002-2003 to 2009-2010.
- The predictor variables derived from the pooled CCHS cycles were based on interviews which occurred between 2000 and 2009.
- The SES index as well as demographic information were based on the 2006 census.
- The primary care availability and accessibility indicators were based on information from 2006.

Multiple years were selected for the extraction of the ACS hospitalisations since such events may be relatively rare in the neighbourhoods of Ottawa, especially for those having smaller population. Rates derived for rare events in small populations are unstable. Enlarging the population at risk of an ACS hospitalisation was achieved by selecting an eight-year exposure period, which contributed to the reliability of the rates derived for less populated neighbourhoods. Expanding the time frame has a potential drawback. Specifically, it may increase the chance of extracting recurrent hospitalisations for the same individual. This would be a form of double counting that may be best avoided only if occurring for the same ACS condition - considering that the concept of preventable hospitalisation is based on conditions for which the risk of a hospitalisation should be reduced, not on individual for which the risk should be reduced. The frequency of repeated hospitalisations for the exact same ACS condition was not expected to be high and was assumed to be roughly non-differential across neighbourhoods²²; therefore having limited impact on

²² The neighbourhood on the extreme poor end of the spectrum may have a differential proportion of chronically ill residents, and these people could be the ones accounting for most of the recurrent events. The impact of this possibility would affect the calculation of the rates only if occurring in neighbourhoods with small populations.

the calculation of the neighbourhood' rates. For comparability reasons between the rates, the same time period was used to extract the ER visits for ACS conditions, even if the potential drawback of repeated visits was also possible.

Multiple cycles of the CCHS were selected to provide sufficient sample size at the neighbourhood level. The independent cycles, taken separately, did not yield sufficient observations per neighbourhood to produce reliable estimates, especially for the neighbourhood comprised of small populations (≈ 2000 respondents / 90 neighbourhoods ≈ 22 respondents per neighbourhood – assuming an equal distribution of respondents across neighbourhoods). In order to provide a reasonable sample size for most neighbourhoods, the micro-level data of the first five cycles was pooled and treated as a sample drawn from one population. Pooling cycles was also necessary to ensure that a minimum of 15 respondents per neighbourhood and per variable be available to preserve the confidentiality of the respondents and to release the neighbourhood estimates from the COOL RDC.

By combining similar subsequent years for the CCHS and hospital-based data, the estimates derived originated from the same evolving population. However, the history threat²³ may be affecting the CCHS respondents' answers or the pattern of hospitalisations and ER visits for ACS conditions. Threat due to history on the validity of the estimates, if any, was assumed to be negligible. In fact, the estimates derived from the CCHS and from the hospital data were considered as period estimates rather than a point estimates for the current investigation and period estimates may mask any trend which occurred during the lapse of time they represent (Thomas & Wannell, 2009). Since trend analysis was not the objective of this study, the reliance on period estimates did not introduce bias and was acceptable. On the other hand, the indicators of primary care accessibility and

²³ The history threat is "a threat to internal validity that occurs when some historical event affects your study outcome" (Trochim & Donnelly, 2008, p.236). For example, the internal validity of the CCHS estimates could be threatened by history if an event occurred in 2004 and affected the health or health care access of CCHS respondents of Cycle 3.1 (2005) but not Cycles 1.1 and 2.1 (2001 and 2003). The extent at which such event would influence differently the neighbourhood estimates derived from the CCHS would depend on if different proportion of respondents from different cycles were forming the neighbourhood samples. The counts of ACS hospitalisations or ER visits from one year to the next could also be affected by changes in the health or health care access in Ottawa.

availability were point estimates derived from a single year. As 2006 is roughly the median year between the years included in this study (2000 and 2010), it was assumed to be representative of the study period. The context of primary care access during those years was assumed to be constant and not affected by history threat.²⁴

3.5.3. Deriving Rates of ACS Hospitalisation and ER Visits

As mentioned in the variable of interest (Section 3.4), the neighbourhood ACS hospitalisation rates for the 20-79 years old were age- and gender- standardised using the 2006 census population of Ottawa (direct method) and expressed per 100,000 person-year (based on person-year calculations). These rates represented the proxy for the rates of potentially preventative hospitalisation at the neighbourhood level and were the dependent variable. The neighbourhood rates of ER visits for ACS conditions, considered as a predictor variable of interest, were derived using a similar method.

The sub-sections below contain additional information on the ACS conditions and related identification codes selected (Section 3.5.3.1), the method used to extract the ACS hospitalisations and ER visits from the administrative databases (Section 3.5.3.2) and the method used to derive the age- and gender- standardised rates (Section 3.5.3.3). Differences between the different calculation methods will be identified and their limitations acknowledged. An overview of the rates derived is presented in Section 3.5.3.4. All data manipulations were performed in Microsoft Excel 2007.

3.5.3.1. SELECTION OF ACS CONDITIONS

In the literature, there is not a single universally accepted list of ACS conditions. Different lists have been offered by different researchers and among experts. For the current investigation, the selection of ACS conditions builds on previous work performed by Ottawa Public Health (OPH) and the Ottawa Neighbourhood Study

²⁴ In Ottawa, the openings of additional Appletree clinics as well as community care access centers in rural areas are plausible historical events which may alter the validity of the primary care access indicators.

(ONS) (Russell et al., 2010). In the ONS, OPH selected 23 ACS conditions based on articles and reports from Billings et al. (1993), Shah et al. (2003) and the Manitoba Centre for Health Policy (MCHP) (2009). Four conditions of these ACS conditions were removed for the current investigation:

- 1) Peptic ulcer diseases were excluded since they were used only in the analytic investigation of Shah et al. (2003) and not identified as ACS conditions in Shah et al.'s reference - which was Brown et al. (2001).
- 2) Hemorrhoidectomy was excluded as this condition is identified in the intervention field of the DAD and NACRS, which were not extracted in the current query performed by OPH. Also, this ACS condition was not identified by Billings et al. (1993).
- 3) Hysterectomy for cervical cancer was excluded for the same reasons as hemorrhoidectomy.
- 4) Dental conditions were excluded since most preventative and curative dental health services are not covered by the universal health insurance and the factors influencing a hospitalisation for such these conditions may be different than for the other ACS conditions identified.

OPH also identified the ICD-10-CA codes²⁵ identifying the different ACS conditions in the databases. The list of ICD-10-CA codes identifying the 19 ACS conditions selected was compared to the codes included in the indicators of ACS hospitalisation rates developed by CIHI (2007a; 2009c; 2010a) and MCHP (2009).²⁶ Some differences were noted²⁷, but considering that the Ottawa Public Health

²⁵ ICD-10-CA codes are based on the 10th revision of the international standard diagnostic classification of diseases and related health problems endorsed by the World Health Organization and used by all its Member States. Since fiscal year 2002-2003, these codes are used to identify the cause(s) of hospitalisations and ER visits in the DAD and NACRS databases. In previous years, the ICD-9-CA codes were used in the databases. To avoid any misclassification of ACS conditions due to the different coding versions, no hospitalisation or ER visit from the fiscal years prior to the implementation of the ICD-10-CA in the databases were extracted for the current investigation. (CIHI, n.d.b)

²⁶ It was not possible to compare the list of ICD-10-CA codes selected by the OPH with the ones identified by Billings et al. (1993): the ICD-9-CA codes are used for that study.

²⁷ For some conditions, the list of ICD-10-CA codes identified by OPH omits certain sub-codes identified by CIHI and MCHP. Examples are for angina (3 sub-codes not included) and urinary tract infections. For four conditions, additional sub-codes were included in the OPH list: ear, nose and

database query was already created for a previous study, it was decided to use the same list of ICD-10-CA codes identified for the 19 ACS conditions selected for the current investigation. The list of conditions and their related ICD-10-CA codes selected for this research are in Table 3.

Table 3: Selected ACS Conditions and Related ICD-10-CA Codes

Ambulatory Care Sensitive Conditions:	ICD-10-CA
Angina	I20
Asthma	J45-J46
Cellulitis	L03
Chronic obstructive pulmonary disease	J40-J44
Congestive heart failure	I50.0
Iatrogenic hypoglycaemia	E15, E16.0
Immunizable conditions (measles, mumps, rubella, diphtheria, tetanus, diphtheria, whooping cough, acute poliomyelitis)	A33-A35, A36, A37, A80, B05, B06, B26
Infections of the ear, nose, and throat	J00-J06, H60-H62, H65-H67
Malignant hypertension	I10.1
Pelvic inflammatory disease	N70-N77
Pneumonia (including bacterial)	J12-J18
Status epilepticus	G41
Urinary tract infection	N39.0
Tuberculosis (pulmonary and other)	A15-A19
Gastroenteritis	A09, K52
Diabetes (including hyperosmolar coma and diabetic ketoacidosis)	E10-E14
Hypertension	I10-I15
Iron deficiency anaemia	D50
Epilepsy	G40

In the technical notes of CIHI (2007a; 2009c; 2010a) and MCHP (2009), hospitalisations for ACS conditions were not considered as sensitive to ambulatory care if they resulted in the death of the patient; therefore these events are excluded from their rates. Similarly, certain conditions, if associated with certain surgeries or interventions, are not considered as preventable by primary care services. CIHI and MCHP performed the later exclusion for the conditions of angina, congestive heart failure, ear, nose and throat infections, pelvic inflammatory diseases and

throat infections (9 additional sub-codes included), pelvic inflammatory diseases (7 additional sub-codes included), gastroenteritis (1 additional code included) and pneumonia (1 additional code included). These additional codes represented specific forms of the conditions themselves; therefore they were representing ACS conditions.

hypertension. Pilot studies conducted by CIHI demonstrated that about 10% of the hospitalisations for angina and 2% of the hospitalisations for congestive heart failure were excluded based on previous interventions (Y. Gong²⁸, personal communication, September 27, 2010).²⁹ However, exclusions based on previous interventions could not be performed for the current investigation: the databases queries did not include the intervention field. No exclusions were performed if the patient's hospitalisation resulted in death.

3.5.3.2. DATABASE EXTRACTION OF ACS HOSPITALISATIONS AND ER VISITS

As requested of OPH³⁰, the ACS hospitalisation discharges were extracted from the Discharge Abstract Database (DAD) for the fiscal years starting in 2002 to 2009 (April 1st to March 30th). The DAD was queried using the field '*most responsible diagnosis*' applying the list of ICD-10-CA codes identified for the ACS conditions selected. The '*most responsible diagnosis*' is defined as "the one diagnosis or condition that can be described as being the most responsible for the patient's stay in hospital" (CIHI, 2008b, p.8).³¹ The variables extracted in the query included the patient's age, sex and residential postal code. Unfortunately, there is no field in this database indicating if the patient had a family physician.

All hospitalisation discharges with a residential postal code located within the ONS area were selected from the above DAD extraction of ACS hospitalisations. This selection is reliant upon valid postal codes attached to each record. Based the DAD's data quality documentation (CIHI, 2004a; 2004b; 2005a; 2006a; 2007b; 2008c; 2009d), the percentage of discharge abstracts³² across Canada with missing or invalid postal codes was less than 0.001% for the fiscal years 2004-2005 to 2008-

²⁸ In September 2010, Yanyan Gong was the Program Lead of the Health Indicators at CIHI.

²⁹ To the author's knowledge, no pilot study was performed to assess the proportion of ACS hospitalisations resulting in death.

³⁰ The Ontario Ministry of Health and Long Term Care provides OPH with access to the NACRS and DAD through intelliHEALTH.

³¹ In the event that multiple diagnoses are listed, the most responsible diagnosis is selected as the condition associated with the longest length of stay or that is the most resource intense (CIHI, 2008b).

³² This percentage includes discharges for acute hospitalisations and day surgeries.

2009³³. A similar proportion of ACS hospitalisations missing due to postal code non-response or invalidity for the region of Ottawa could be expected; therefore it was assumed that this was a negligible source of error.³⁴ Hospitalisations among Ottawa residents in hospitals located in Québec – institutions which are not participating to the DAD³⁵ – were believed to be uncommon and to have negligible impact on the different neighbourhood rates derived from the information available. Additional information the data quality of the DAD can be found in Appendix 7.

The ER visits for ACS conditions were extracted by OPH from the National Ambulatory Care Reporting System (NACRS) for the fiscal years starting in 2002 to 2009.³⁶ Only the unscheduled emergency room visits were selected by restricting the query to these types of visits using the 'ambulatory case type' field.³⁷ The NACRS was queried by the '*main problem*' field using the list of ICD-10-CA codes identified for the ACS conditions selected.³⁸ The '*main problem*' represents the "problem that is deemed to be the clinically significant reason for the client's visit, which requires evaluation and/or treatment and/or management" (Gibson, Richards, & Chapman, 2008, p.102). The variables extracted in the query were the patient's age, gender, residential postal code and disposition status.

Exclusions from the query were performed according to the disposition status, which refers to the outcome of the visit (e.g. discharged at home, not seen or left, admitted). The ER visits which resulted in an admission at the hospital or a transfer

³³ Information not available for fiscal years 2002-2003 and 2003-2004.

³⁴ According to CIHI data quality summaries for the fiscal years of 2006-2007, 2007-2008, 2008-2009 (CIHI, 2007b; 2008c; 2009d), there were discrepancies in patient residence as identified by the postal code and by the health care number (HCN). For the patients with a HCN from another province than the institution reporting the hospitalisation, approximately 30% reported a residential postal code within the province. CIHI reported the following reason as an explanation: "a patient who has relocated may have sought care using the HCN issued by the original province / territory and the postal code in the current province or territory" (CIHI, 2009d, p.27). For this study, it was assumed that the postal code provided represented the residential address of the patient, even if the HCN associated with the hospitalisation may be from another province than Ontario.

³⁵ Hospitalisation discharges from institutions located in the province of Québec are recorded in the database 'MED-ÉCHO' and managed by the Ministère de la santé et des services sociaux.

³⁶ In 2000, all the ER departments of Ontario joined the NACRS (CIHI, 2010b).

³⁷ The NACRS contains records on emergency visits (scheduled and unscheduled), day surgeries and visits to outpatient clinics.

³⁸ The '*main problem*' can be a diagnosis, condition, problem or circumstance.

to another facility were excluded from the query. The exclusions due to an admission were performed to ensure that the ER visits and hospitalisations for the same patient and event be extracted only in the hospitalisation counts. Also, the exclusions due to a transfer were performed to ensure that only one ER visit for the same event - as oppose to two visits in two different ER departments - be extracted in the ER visit counts. However, the completeness and accuracy of the deposition field is difficult to assess. This field is mandatory in the data entry system of emergency departments and some facilities are using default values (CIHI, 2010b; Gibson et al., 2008). For these facilities, the completeness of the mandatory field is 100%, yet it is not possible to differentiate non-response from defaults values (K. Russell, personal communication, February, 2010)³⁹. Due to the reliance on default values, there may be a proportion of the ER visits which resulted in an admission or an inter-facility transfer that could not be excluded from the query. For similar concerns regarding the family physician field, also mandatory in NACRS database, the ER visit rates for ACS conditions were not derived based on the family physician criterion (CIHI, 2010b; Gibson et al., 2008).

All ER visits with a residential postal code located within the ONS area were selected from the above NACRS extraction. According to the NACRS documentations and executive summaries (CIHI, 2004c; 2005b; 2005c; 2006b; 2007c; 2008d; 2009e; 2010b), the percentage of records⁴⁰ across Canada with missing or invalid postal codes was between 0.62% and 1.9% for the fiscal years of 2004-2005 to 2009-2010⁴¹. A similar proportion of ER visits missing due to postal code non-response or invalidity for the region of Ottawa could be expected; therefore it was assumed to be a negligible source of error. All emergency departments located in the Ottawa region are included in NACRS, yet the facilities in Québec are not included in addition to some other facilities in Canada. This means that ER visits among residents of Ottawa in facilities not participating to the NACRS

³⁹ Based on personal communication between Katherine Russell and Arran Shemmans from CHIM, Ministry of Health & Long Term Care, Health System Information Management & Investment Division, February 2010

⁴⁰ NACRS contains records on ER visits, day-surgeries and visits to outpatient clinics. The percentage of missing or invalid postal codes refers to all type of records.

⁴¹ Information not available for the fiscal years of 2002-2003 and 2003-2004.

were not included in the extraction. These omissions were assumed to be non-differential across the neighbourhoods.

The counts of hospitalisations (and ER visits) for all 19 ACS conditions combined were provided by OPH for the different ONS neighbourhoods by age (20-29, 30-39, 40-49, 50-59, 60-69 and 70-79), gender (male and female) and fiscal year (starting in 2002 to 2009). The age and gender counts from the different fiscal years were subsequently pooled. The counts provided reflect the number of ACS hospitalisations (or ER visits), not the number of persons who were hospitalised (or visited the ER) for ACS conditions independently from the number of times such event occurred (see Section 3.5.2 for interpretation details).

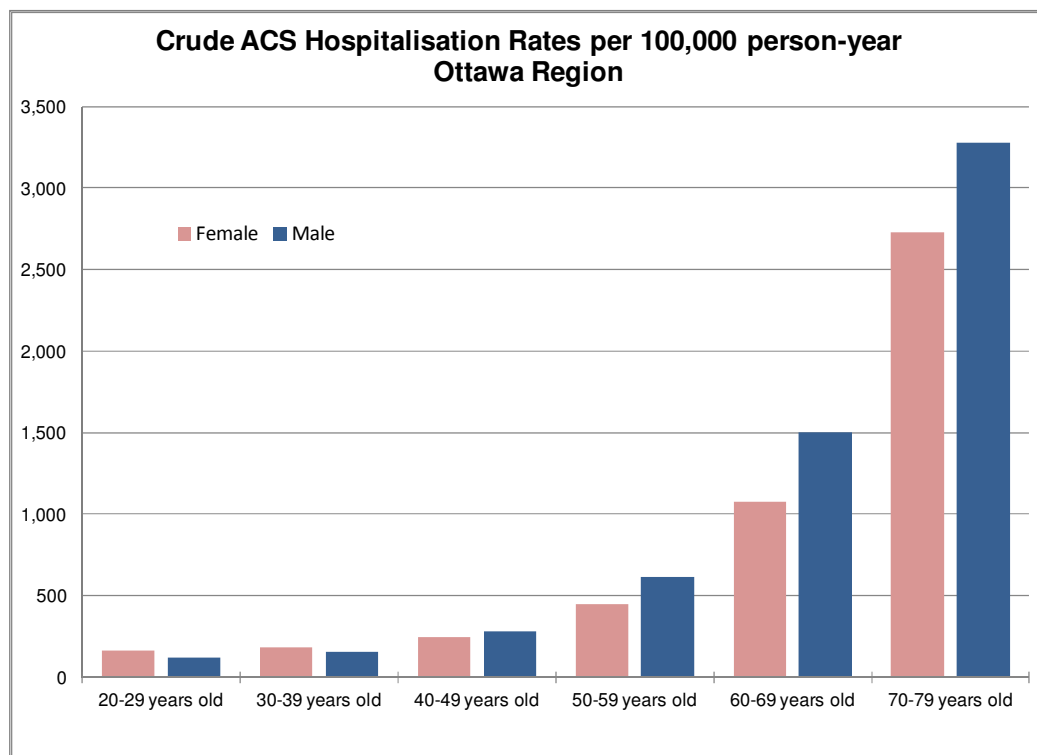
3.5.3.3. METHODOLOGY FOR DERIVING RATES: AGE- AND GENDER-STANDARDISATION

The counts of ACS hospitalisations obtained for the different neighbourhoods of Ottawa required adjustment for population size and population composition before comparisons could be made. Simply dividing the neighbourhood counts by the population at risk counts⁴² would achieve the proper adjustment for population size, but it would not adjust for potential differences in the age and gender compositions of the neighbourhood populations. Ahmad et al. (2001, p.3) demonstrated that comparisons of rates derived for age-dependent conditions between neighbourhoods "may be very misleading if the underlying age composition differs in the populations being compared". Indeed, most ACS conditions are age dependent, where the risk of suffering from an ACS condition either: 1) increases with age for a number of conditions such as angina, hypertension, congestive heart failure and pneumonia or 2) decreases with age for some other conditions such as measles. Moreover, the risk of being hospitalised for ACS conditions typically increases with age due to co-morbidity and fragility or due to the individual's cumulative exposure to an ACS chronic disease (e.g. diabetes-related ACS conditions such as diabetic ketoacidosis) (Sanmartin & Khan, 2011).

⁴² This calculation would represent crude prevalence rates and would be expressed per 100,000 person-year.

For the current investigation, the necessity of age-standardising the neighbourhood ACS hospitalisation rates was shown by the crude hospitalisations rates derived for all Ottawa residents by 10-years age intervals (Figure 8). This figure demonstrates that the rate among the 70-79 year olds was more than 10 times greater than among the 20-29 year olds. Therefore, a higher crude ACS hospitalisation rate for a neighbourhood could simply be the consequence of a greater proportion of older residents in this neighbourhood. Among the neighbourhoods of Ottawa, the proportions of 70-79 years old vary from 1.8% to 15.9%, which demonstrates the difference in neighbourhood age breakdown and further support the need to age-standardise the rates. Additionally, the validity of adjusting for the potential effect of gender composition on the crude ACS hospitalisation rates was demonstrated in Figure 8: the crude rates among males and females were different within and across age groups. Sanmartin and Khan (2011) also demonstrated that the risk of an ACS hospitalisation is different among males and females in Canada, especially among the older age groups. In all, standardisation was used to remove the effect of age and gender composition on the rates.

Figure 8: Crude ACS Hospitalisation Rates by Age Groups for Ottawa



Standardisation is an adjustment procedure based on weighted averages of age- and gender- specific rates in which the weights are chosen to provide a standard basis for comparison (Schoenbach & Rosamond, 1999). There are two standardisation methods commonly use in social and epidemiological sciences: the direct method and the indirect method. The later method is often implemented in studies where the neighbourhood age- and gender- specific rates are not available or the count of events for an important proportion of the age and gender sub-groups is zero; two issues not at play in the current investigation (Andersen & Keiding, 2006)⁴³. Considering that the standardisation weights used in the indirect method would be different across the neighbourhoods, comparisons of absolute indirect rates from one neighbourhood to the next could be problematic. In fact, when the age and gender distributions differ importantly across neighbourhoods, as demonstrated above, comparing indirectly-standardised rates could not be better than comparing crude rate s(Gale Encyclopedia of Public Health, 2012; Schoenbach & Rosamond, 1999).⁴⁴ For this reason, the direct method was selected for the current study.

In the direct method, the standardised rate of a population is obtained by applying the age- and gender- specific rates for each group to the age- and gender- composition of a standard population. Therefore, the rate derived with this method represents the rate that would have been observed if the neighbourhood population shared the same age and gender structure as the standard population. The direct method applies the same weights to the age- and gender- specific rates of the neighbourhood populations, therefore allowing comparisons across neighbourhoods and between the neighbourhoods and the overall rate observed in the standard population. (Schoenbach & Rosamond, 1999, p.132)

Using the direct standardisation method, the formula implemented to standardise the ACS hospitalisations and emergency visits rates of the current study was:

⁴³ The percent of age and gender groups, for all the neighbourhoods, with no ACS hospitalisation was 0.01% (for the rates derived using 10-years age interval and the direct method).

⁴⁴ Indirectly-standardised rates can always be compared safely between the study populations and the standard population, but it may be misleading across study populations.

$$\frac{\sum N_i r_i}{N}$$

Where:

- **N_i** was the population in age and gender group *i* of the standard population
- **N** was the total 20-79 years old population of the standard population
- **r_i** was the event rate in the age and gender group *i* of the neighbourhood

The population of Ottawa was selected as the standard population for this investigation. In fact, the sum of the study populations (in this case: neighbourhoods) is commonly selected as the standard population in the literature for similar studies (Andersen & Keiding, 2006). The neighbourhood population data (**N_i** and **N**) originated from the 2006 census and was available for age groups of ten years intervals⁴⁵. Since the neighbourhood counts of ACS events were extracted over a period of 8 years (the fiscal years of 2002 to 2009), the population at risk of an ACS related event was the cumulative number of people in the neighbourhood for that 8-year period. This was the sum of the population for the fiscal years starting in 2002, 2003, 2004, 2005, 2006, 2007, 2008 and 2009. However, only the neighbourhood population for the year 2006 was available. Therefore, the event rate in the age and gender sub-groups of the neighbourhood (**r_i**) were calculated on the 2006 neighbourhood age and gender population multiplied by 8 – therefore reflecting person-year calculations⁴⁶.

⁴⁵ At the time of the study, the population of the age groups was not available by gender at the neighbourhood level. The gender breakdown by 10-year age intervals of the whole region of Ottawa were applied to the 10-year age interval neighbourhood populations in order to derive the age- and gender- specific population counts needed for standardisation. This assumed that the neighbourhood gender composition by age group was the same across the different neighbourhoods of Ottawa, which was expected not to differ substantially and not to affect the rates derived. The Ottawa population information by gender was obtained from the 2006 census on the website of the City of Ottawa (n.d.).

⁴⁶ Person-year calculations ensure that the rates derived reflect the proportional number of the event of interest in the standard population based on a period of one year, and not the time period of the study (Bhopal, 2002).

The rates of ACS hospitalisation (ER visits) that were derived represented the number of hospitalisations (ER visits) per 100,000 person-year, not the number of people who were hospitalised (or visited the ER) for an ACS condition per 100,000 person-year. Such an indicator was valid considering that the concept of preventable hospitalisation is based on conditions for which the risk of a hospitalisation should be reduced, not on individual for which the risk should be reduced.

Although direct standardisation enabled comparisons across neighbourhoods, the stability of the rates may have been affected by the low occurrence of hospitalisations or ED visits in the different age and gender sub-groups. In fact, "if the age- and gender- specific rates in the neighbourhood population are zero for a number of age and gender groups, then the directly standardized rate is poorly estimated and can have a large standard error" (Andersen & Keiding, 2006, p.431). As rules of thumb, at least 10 events in each important⁴⁷ age and gender group and a population denominator of 100 should be sufficient to ensure the stability of the rates (Schoenbach & Rosamond, 1999). In order to explore the stability of the rates standardised using the direct method, the rates were derived using 10-year age intervals⁴⁸ as well as 20-year age intervals⁴⁹. The rules of thumb were fulfilled for the ACS ER visit rates; both for the estimates derived using the 10- or 20- year age intervals. For the ACS hospitalisation rates, no age and gender sub-group had cumulative populations below 100 people. However, the 'at least 10' rule pertaining to the numerator was not fulfilled regardless of the age intervals used. Based on further examination of the counts of hospitalisations by sub-groups, the rates derived with the 20-year age intervals may have been slightly more stable than the rates derived with the 10-year age interval, especially for the neighbourhoods with small populations.

On the other hand, the 20-year age interval ACS hospitalisation rates may not account adequately for the effect of age on ACS hospitalisation. When the event

⁴⁷ An important sub-group refers to groups with a substantial weight, therefore a large proportion of the neighbourhood residents.

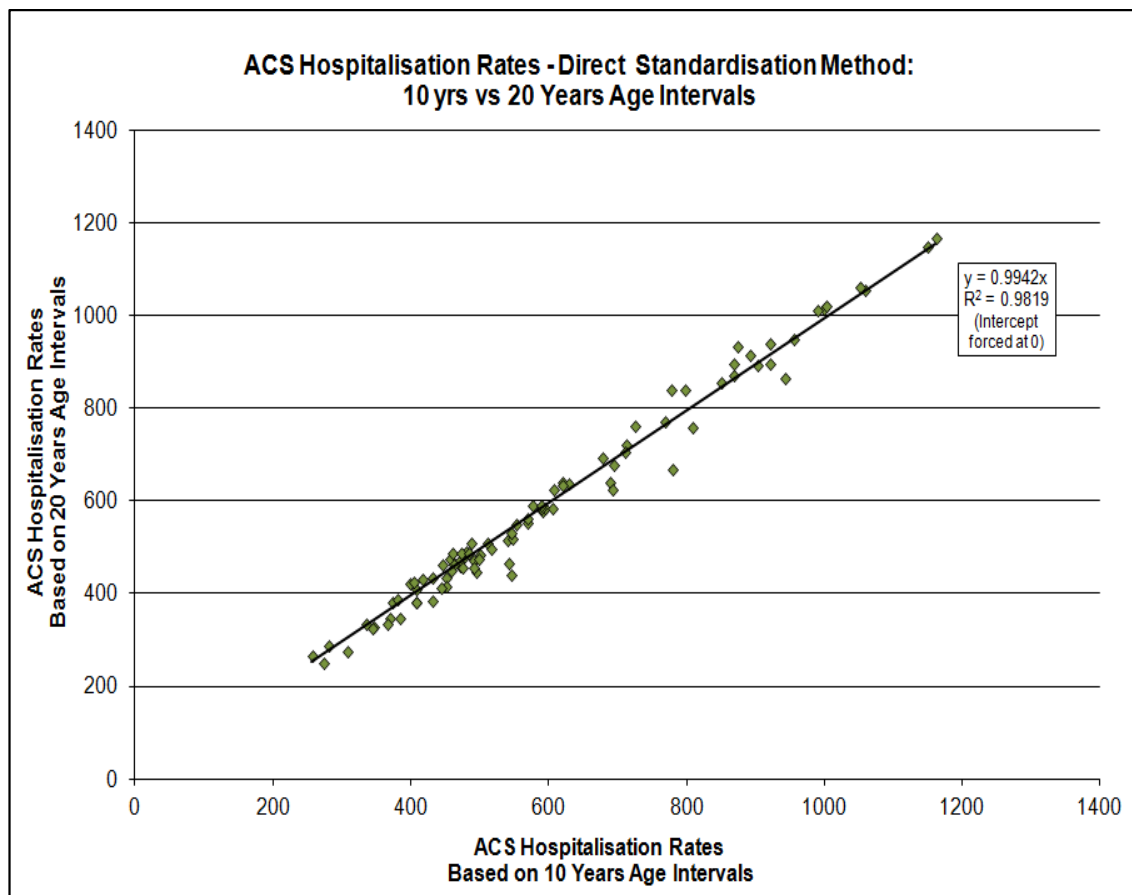
⁴⁸ The age intervals were 20-29, 30-39, 40-49, 50-59, 60-69 and 70-79 years old.

⁴⁹ The age intervals were 20-39, 40-59 and 60-79 years old.

rate in the age and gender sub-group of the neighbourhood (r_i) is derived, it is assumed that the pattern of ACS hospitalisations among this sub-group is constant for the age interval used. This assumption was harder to defend for the rates derived using the 20-year age intervals, as demonstrated by the crude ACS hospitalisation rates shown in Figure 8. Differences in the ACS hospitalisation rates for the 60-69 and 70-79 years old were substantial and differences between the 40-49 and 50-59 years old were also apparent, suggesting that combining these age groups may not be adequate.

In order to compare the two types of rates, a scatter plot of the standardised neighbourhood rates derived from the 10-year versus 20-year intervals was generated in Figure 9.

Figure 9: Comparison of ACS Hospitalisation Derived Using 10- and 20- Years Age Intervals



This figure shows that the neighbourhood standardised rates derived using these two methods were very similar for the vast majority of neighbourhoods. Both estimates may suffer from instability due to small ACS hospitalisation counts in some age and gender sub-groups. However, considering that the 10-year age intervals were more commonly used in the literature and that these rates better addressed the effect of age on the risk of suffering from an ACS hospitalisation, the rates based on 10-year age intervals were selected for the current investigation. There was no suspicion that the ER visit rates derived from 10-year age interval suffer from instability.

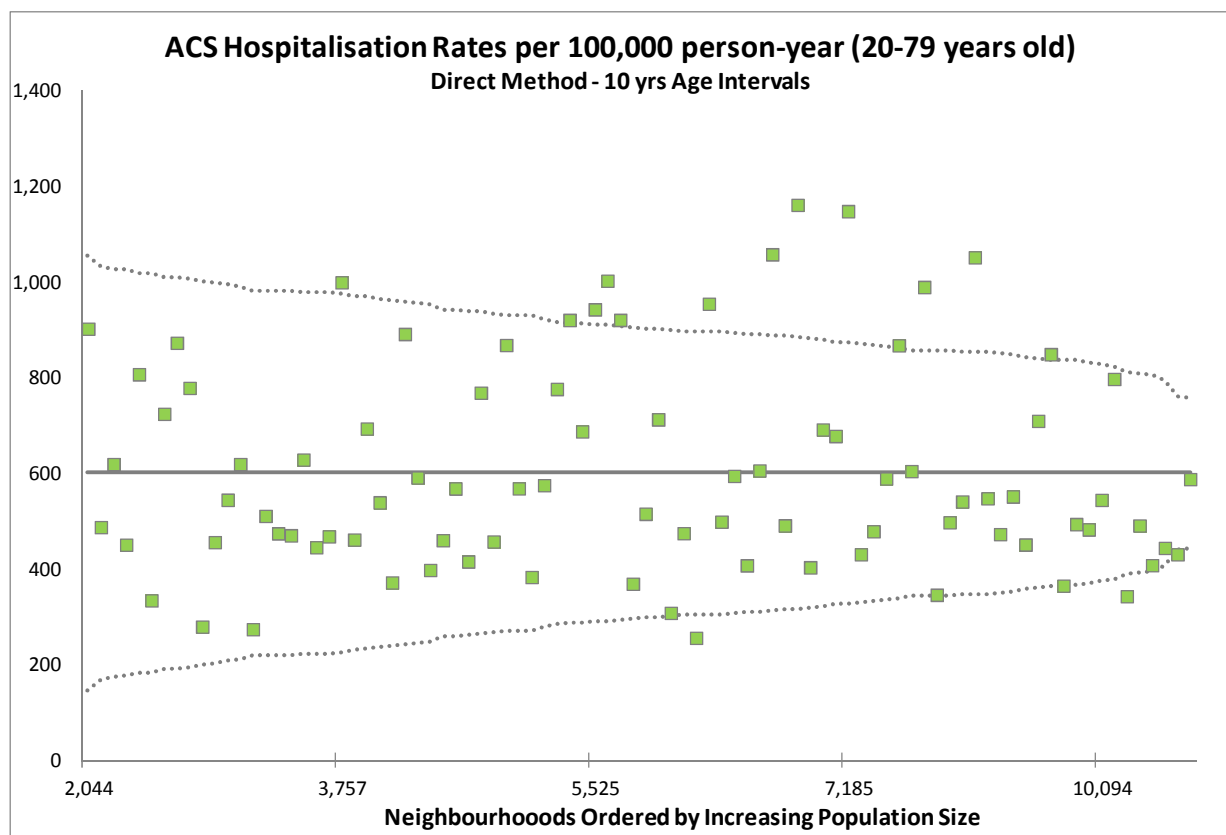
3.5.3.4. ACS HOSPITALISATION AND ER VISIT RATES – NEIGHBOURHOOD ESTIMATES IN OTTAWA

From this point forward, all rates presented in this document were age- and gender-standardised based on the method described in Section 3.5.3.3 (direct method – 10 year age interval). Among the 20-79 years old, the ACS hospitalisation rate in Ottawa was 600.70 per 100,000 person-year (CI = 594.18 – 607.22). Across all 90 neighbourhoods, the ACS hospitalisation rates ranged from 256.1 to 1,161.6 per 100,000 person-year and the coefficient of variation was 36.2%. The ER visit rate for the 20-79 years old for the whole region was 3,519.49 per 100,000 person-year (CI = 3,505.59 – 3,533.39). Across all 90 neighbourhoods, the ACS ER visit rates ranged from 1,447.7 to 6,174.5 per 100,000 person-year and the coefficient of variation was 30.7%. In Appendix 2 are the ACS hospitalisation and ACS ER visit rates for all the neighbourhoods and the description of the calculation methods used to derive the confidence intervals.

The neighbourhood ACS hospitalisation rates were depicted in a funnel plot (Figure 10). The funnel plot is a scatter plot of the neighbourhood rates (vertical axis) ordered by neighbourhood population size (horizontal axis). On this graph, the overall rate for the region of Ottawa is represented by a line and its 95% confidence interval, based on neighbourhood population size, is captured by the zone delimited

by the dotted lines⁵⁰. Taking into consideration chance variation expected given neighbourhood population size, a neighbourhood with a rate falling outside the zone is characterised as having a rate that is exceptionally high or low compared to the average rate for Ottawa. If more than 5% of the neighbourhood rates are falling outside the expected zone, it is indicative that an underlying pattern, other than luck, is affecting the neighbourhood ACS hospitalisation rates. In the funnel graph (Figure 10), 18% of the neighbourhoods are falling outside the expected zone, suggesting that the observed variation between neighbourhoods is manifesting real differences above and beyond differences that would be compatible with chance. In Table 4 are the neighbourhoods, based on the funnel plot, whose rates were exceptionally high or low.

Figure 10: Funnel Plot – ACS Hospitalisation Rates



⁵⁰ The 95% confidence intervals were calculated based on the methodology presented by the consulting firm Kurtosis (2010).

Table 4: Neighbourhood with Exceptional ACS Hospitalisation Rates

Neighbourhood ID	Neighbourhood Name
Neighbourhoods with exceptionally high ACS hospitalisation rates	
#95	ByWard Market
#6	Bayshore
#77	Glen Cairn - Kanata South Business Park
#54	Vanier South
#94	Britannia Village
#71	Cummings
#42	Lowertown
#63	Vanier North
#18	Carlington
#33	Hintonburg – Mechanicsville
#57	Whitehaven - Queensway Terrace North
#44	Overbrook – McArthur
#55	West Centertown
Neighbourhoods with exceptionally low ACS hospitalisation rates	
#79	Kanata Lakes - Marchwood Lakeside - Morgan's Grant - Kanata North Business Park
#83	Orleans Avalon - Notting Gate - Fallingbrook - Gardenway South
#69	East Industrial

The ACS hospitalisation and ER visit rates were highly correlated, with a Pearson correlation coefficient of 0.845 ($p < 0.0001$). In population health research, such high correlation is rare, especially for ecological data. An overlap between the two measures is suspected, even if precautions to avoid it were implemented. In the data sources, patients who were admitted to the hospital from a visit to the emergency department are recorded both as an ER visit in the NACRS and as a discharge in the DAD. In the extraction of the ACS ER visits counts, the visits resulting in a hospitalisation were excluded using the disposition status field. However, this field may be incompletely or inappropriately coded in the NACRS database. Indeed, 13 out of 15 facilities studied in Ontario indicated that default values were used for various data elements, including disposition status (Gibson et al., 2008). Consequently, it is possible that the ER visit counts included cases which were hospitalised via the emergency room, therefore explaining the unusually high correlation between the ACS hospitalisation and ER visit rates.

Studying the relationship between neighbourhood utilisation of emergency services and the rates of preventable hospitalisation was not feasible due to the suspicion of an overlap between the two measures. Unfortunately, the linear regression model⁵¹ is inadequate to test the effect of a factor closely related to the concept under investigation. Also, concerns about the construct validity of the ACS ER visit rates would further complicate the interpretation of the results of a model in which this variable was included as a predictor. For additional details on the interpretability challenges and construct validity concerns associated with this variable, please refer to Appendix 8. Based on the above reasons, the ER visit rates for ACS conditions were excluded from the list of potential predictor variables for the analyses of the current investigation.

On the other hand, the rates of ER visits for ACS conditions could be considered as a potentially valuable indicator of inadequacies in primary care access, independently of hospitalisation rates, and could be analysed separately (as a dependent variable). In Canada, the emergency departments was shown to be used as an alternative source for primary medical care by Roberge et al. (2007) as cited in McCusker et al. (2010, p.972). It was also shown that Canadians with an ACS condition are more likely to use the emergency department if they have low SES or live in rural areas (CIHI, 2012). Among Canadian children, low supply of primary care physicians was associated with a reduction in primary care utilisation, an increase in ER visits and, to a lesser extent, an increase in ACS hospitalisations (Guttmann, Shipman, Lam, Goodman, & Stukel, 2010). This may indicate that ER visit rates could be a more direct proxy of primary care access than ACS hospitalisation rates; as suggested by the series of arrows connecting primary care access, ACS ER visits rates and ACS hospitalisation rates in the causal model of Figure 6. However, the ER visits among people with ACS conditions were not shown to be more sensitive to characteristics of primary care compared to ER visits in those with other or no chronic condition (McCusker et al., 2010).

⁵¹ Linear regression model is the method selected to test the relationships between neighbourhood characteristics and the rates of preventable hospitalisation in the current investigation.

In all, studying neighbourhood rates of ER visits for primary care services, independently of ACS conditions, would be valuable for the purpose of health system planning. People's reasons to use emergency services for non-urgent care may be numerous and identifying them would provide new insights on patterns of health care utilisation in Canada. Yet, this was not identified as an objective for the current investigation. Also, the causal model explaining the concepts believed to be related to ACS ER visit rates, or non-urgent ER visit rates, would be different than the one conceptualised for the study of ACS hospitalisation rates. And, the concern over the possibility that certain ER visits included in the current rates lead to hospitalisations would still be at play for such investigation. For these reasons, the ACS ER visit rates were discarded from the current research.

3.5.4. Deriving Neighbourhood Estimates from the Canadian Community Health Survey

Variables representing neighbourhood health status and healthy behaviour as well as health care utilisation and access were selected from information gathered by Statistics Canada in the Canadian Community Health Survey (CCHS). The CCHS is not ideal for deriving neighbourhood estimates: it is designed to provide population estimates at the health region level only. This generated two limitations for the creation of neighbourhood level population estimates: the sampling weights were inadequate to derive estimates representative of the neighbourhood population and one cycle did not provide sufficient sample size to derive reliable neighbourhood estimates. Yet, no other adequate data source was available for the neighbourhoods of Ottawa.⁵² Two measures were undertaken to address the limitations associated with the inappropriateness of the data source: pooling the first five CCHS cycles

⁵² The Rapid Risk Factor Surveillance Survey (RRFSS), a health-based survey administered by York University, was considered as an alternative data source. Unfortunately, this survey is also designed to provide population estimates at the regional level. Pooling the different yearly cycles available yielded similar sample size than the one available from the pooled CCHS cycles. Considering that the RRFSS contains fewer data elements than the CCHS, the later source of data was selected for the current investigation.

available to increase the sample size at the neighbourhood level and age-standardising the neighbourhood estimates derived to avoid sampling bias.⁵³

Below is a discussion of the methodology that was undertaken to derive the neighbourhood estimates from the CCHS data. The process followed to select the CCHS respondents residing in Ottawa as well as the method used to pool the respondents and questionnaire items from the five CCHS cycles is presented in Section 3.5.4.1. Then, a short overview of the inappropriateness of using the sampling weights to derive representative neighbourhood estimates is provided, which leads to the explanation of the favoured method based on age-standardisation to avoid sampling bias (Section 3.5.4.2). The method followed to perform imputations at the neighbourhood level is in Section 3.5.4.3. And, lastly, further comments on the measurement quality of the neighbourhood estimates are provided in Section 3.5.4.4. All data manipulations were performed using SAS 9.2.

3.5.4.1. SELECTION AND POOLING OF NEIGHBOURHOOD RESPONDENTS

The CCHS respondents living in the ONS region can be identified from the CCHS health region named Ottawa-Carleton (ID#3551)⁵⁴. However, the 6-digit postal code was necessary to determine the residential neighbourhood of the CCHS respondents. In the publically available microdata files of the CCHS, refined level of geographical information on the respondents' residence is not available. At the individual level, such information is confidential and only available in the master files of the CCHS. Secure access to confidential information collected by Statistics Canada is made available at the Carleton, Ottawa, Outaouais Local Research Data Centre (COOL RDC) located in the Morisset Library of the University of Ottawa. Free access to the data files was granted to the investigator, following a successful application to the Social Sciences and Humanities Research Council (SSHRC).⁵⁵

⁵³ Sampling bias occurs when marked differences between the sample composition and the population composition are observed, which may affect the representativeness of estimates derived from such sample (de Vaus, 2002, p.152-157).

⁵⁴ The correspondence of geographical boundaries of these two regions will be demonstrated in sub-section d.

⁵⁵ Additional detail on the application process and approval of the current project is available in Appendix 3.

Access was granted for the first five CCHS cycles available at the COOL RDC at the time of the current investigation. For each master data files of the CCHS (one per cycle), the respondents living in the health region of Ottawa-Carleton (ID#3551) were identified and extracted. Using the list of postal codes per neighbourhoods provided by Prof. Sawada⁵⁶, the residential neighbourhood of each Ottawa-based respondent was identified using the variable containing their 6-digit residential postal code⁵⁷. Only 23 respondents living in the health region of Ottawa-Carleton, across all cycles, could not be assigned to a neighbourhood – see Table 5 for sample sizes by CCHS cycles. The pooled sample for the whole ONS region was 8,865 respondents, from which 1,127 respondents aged between 0-19 years and 408 respondents aged 80+ years were excluded – see Section 3.5.1 for justification. Overall, the pooled CCHS sample available for the current investigation was 7,313 respondents distributed unequally across the 97 neighbourhoods. From this pooled sample, the number of respondents per neighbourhood ranged from 8 to 367 –see Table 6 for details. For confidentiality reasons, only estimates derived from neighbourhoods with more than 15 respondents could be released from the COOL RDC. Ninety of the 97 neighbourhoods satisfied this condition.

⁵⁶ None of the postal codes area is overlapping with the neighbourhood boundaries; all postal codes are located in only one neighbourhood.

⁵⁷ A Statistics Canada reviewer, for the COOL RDC application process, mentioned that a proportion of residential postal codes for the respondents selected from the telephone frame were imputed and may not be verified during the interview. The validity of these imputed postal codes relies on the accuracy of the internal administrative conversation files used to assign a postal code to the telephone numbers listed in this frame (Canadian Phone Directory). To the researcher's knowledge, there is no publication or information available on the accuracy of the administrative conversion files or on the proportion of unverified imputed postal codes among the CCHS respondents. Among the pooled sample, 40.1% of the respondents were selected from the telephone frame. Among those, only a fraction may have an erroneous and unverified imputed postal code which may lead to an inaccurate assignment of neighbourhood for those respondents. It is important to mention that the vast majority of neighbourhoods in Ottawa are comprised of more than 50 postal codes. Therefore, if the inaccurate respondents' postal codes are geographically close to the imputed ones, there are good chances these respondents would still be assigned to the neighbourhood in which they live. The risk of assigning respondents to a different neighbourhood due to the unverified and erroneous imputation of postal codes for the respondents selected from telephone frame was considered as limited and acceptable.

Table 5: Number of CCHS Respondents per Cycle in ONS Region

CCHS Cycle	Respondents of Ottawa-Carleton Region* (no exclusion)	Respondents not Matched to a Neighbourhood
Cycle 1.1	1936	5
Cycle 2.1	2047	7
Cycle 3.1	1975	5
Cycle 4.1	1967	4
Cycle 2009 ⁵⁸	940	2
Total	8865	23

Table 6: Distribution of the Number of CCHS Respondents per Neighbourhood

N	90
Minimum	16
5th Percentile	23
10th Percentile	27
Lower Quartile	39
Median	70
Upper Quartile	106
90th Percentile	144
95th Percentile	176
Maximum	367

The method used to combine the cycles was the ‘pooled approach’ proposed by Thomas and Wannell (2009) as it is a way for increasing sample size and improving the stability of the estimates. Although pooling yielded satisfactory sample size at the neighbourhood level, the validity of this strategy had to be assessed. Thomas and Wannell (2009) recommended examining carefully any change in survey content, coverage, geography and data collection methods prior to pooling CCHS cycles. This four-step assessment is performed below.

⁵⁸ In 2007, Statistics Canada changed the survey data collection methodology from one two-year period to two one-year periods (Statistics Canada, 2008). At the time of the current investigation, the 2010 respondents of Cycle 5.1 were not available at the COOL RDC. This explains the smaller sample size for the Cycle 2009.

a. Assessment of Survey Content – Cycle 1.1 to Cycle 2009

In terms of content, the CCHS questionnaire has undergone continual modifications during the first five cycles. Such modifications were the addition or removal of questions and changes in the wording of the questions. The researcher consulted the CCHS documentation and identified no modification in the formulation of the questions for the variables selected for this study. However, the addition or removal of items as well as changes to the wording of questions precluded the selection of some variables of interest (e.g. COPD and emphysema).

b. Assessment of Survey Coverage – Cycle 1.1 to Cycle 2009

The population targeted by certain items or modules of the CCHS – namely their coverage – can vary across CCHS cycles and health regions. Certain items of interest were part of the optional content and selected only for certain cycles by the region of Ottawa-Carleton. For example, the module of Pap test was not selected for Cycle 2009. Items of interest were discarded from the pooled sample if their coverage was insufficient to generate neighbourhood estimates for a majority of neighbourhoods. This occurred principally for items representing preventative health services utilisation, such as *‘having had a blood pressure check’*, *‘having had a PSA test’* among males aged 40 or more and *‘having had a physical exam with a general practitioner’* – see Section 3.4.2 for details. Also, some questionnaire items were not asked to all respondents depending on their age or gender. For example, *‘having had a mammography test’* is only asked to female respondents aged 35 years or more. From Cycle 1.1 to Cycle 2009, the exclusions in coverage for the items selected did not change.

Differences in coverage generated different sample size per item of interest, therefore the neighbourhood estimates derived from the pooled CCHS data were based on different number of respondents across the variables. The pooled sample size at the regional level for each CCHS items is specified in Appendix 4. In this Appendix are also identified the cycles for which the different CCHS items were not available as well as any exclusion based on age or gender.

c. Assessment of Survey Geography – Cycle 1.1 to Cycle 2009

Even if the name of the health region changed in Cycle 2.2⁵⁹, the boundaries of the health region ID#3551 used to select the ONS respondents were not modified for the cycles of the pooled sample (Statistics Canada, 2002a, 2003, 2005a). Also, the boundaries of this health region are mostly the same as the ones of the ONS region: there are roughly 14-20 meters offset between the two geographical areas and this is not expected to have any impact on the selection of the respondents (M. Sawada, personal communication, February 7, 2011).

d. Assessment of Survey Data Collection – Cycle 1.1 to Cycle 2009

In order to improve the CCHS data collection effectiveness and flexibility, Statistics Canada started to collect data over a period of one year as of Cycle 4.1 compared to a period of two years as in previous cycles (Statistics Canada, 2008). The respondents from Cycles 4.1 and Cycle 2009 were included in the pooled sample even if changes in the data collection methodology were implemented. The impact of such data collection modification on the estimates derived was assumed to be minimal, especially in comparison to the impact of deriving estimates based on smaller sample size if the later cycles were excluded.

Considering the items of interest for the current investigation, the above assessment of the CCHS cycles demonstrated the acceptability of pooling the respondents of the first five cycles. For each of the data files containing the respondents of the ONS region only (one per cycle), the variables of interest for the current investigation were selected. Common variable names were given to the same items across the cycles and the data files were merged. The final dataset included only the CCHS items of interest and the CCHS respondents aged 20-79 years old living in the ONS region for which a residential neighbourhood was assigned. Having this database in hand, the process to derive neighbourhood estimates could be undertaken. For a discussion of the assumptions associated with deriving period estimates from a pooled cycle covering a 9-year period, refer to Section 3.5.2.

⁵⁹ The name changed from 'Ottawa-Carleton Public Health Unit' to 'City of Ottawa Health Unit'.

3.5.4.2. DERIVING REPRESENTATIVE NEIGHBOURHOOD ESTIMATES USING POOLED CCHS RESPONDENTS

The neighbourhood samples generated above were comprised of a group of respondents (sample units) for which measurements were available. These respondents formed sub-groups of the populations of interest, which were the entire 20 to 79 year old residents of the different neighbourhoods. Therefore, due to the sampling design, the neighbourhood samples may not be representative of these neighbourhoods composition. Deriving variable estimates from measurements obtained on sampled respondents yield statistics directly generated from the data, and the degree at which these statistics can be considered as population estimates is relative to the representativeness of the samples. Population estimates are the values of the statistics that would be obtained if the sample was comprised of the entire population of interest. (Trochim & Donnelly, 2008, p.34-38)

Statistics Canada recommends that users apply the sampling weights in order to generate estimates representing the population of interest and taking into consideration the sampling design of the survey (Statistics Canada, 2002b). However, for the current investigation, the utilisation of the CCHS sampling weights to derive neighbourhood estimates was inadequate, as demonstrated below in point a). A method based on direct age-standardisation was selected as a suitable alternative and is presented in point b).

a) Inappropriateness of Using Regional Sampling Weight at Neighbourhood Level

The main purpose of the CCHS weights is to take into consideration the survey sampling design in the generation of population estimates. This purpose becomes clearer when recognising that sampling weights are also defined as the reciprocal of the respondent's selection probability based on the sampling methodology of the survey (Lee & Forthofer, 2006). The CCHS is a complex survey for which the sample design is not self-weighting; therefore the sampling weights are not identical for all individuals in the sample. Using the weights while deriving population estimates ensures that the effect of unequal probabilities of selection of the

respondents based on the CCHS sampling methodology is reduced (Sturgis, 2004). This way, estimates representative of the population of interest are generated by applying the sampling weights. Unfortunately, the CCHS population of interest is the health region, not the neighbourhoods. The available weights therefore, though well suited for the region, are not suitable for obtaining representative estimates at the neighbourhood level.

In Appendix 9 is a short description of the CCHS sampling strategy and methodology employed to derive the weights, both implemented at the regional level. The appropriateness of the sampling weights at the neighbourhood level is questionable considering that:

- Non-response patterns may be different from one neighbourhood to the next and neighbourhood non-response pattern may not reflect the regional pattern used to derive the CCHS weights (S. Thomas, personal communication, June 13, 2011).
- The proportion of respondents selected from the area frame and from the telephone frame in the neighbourhood samples would vary and may not reflect the regional proportion. Therefore, the adjustment factor selected for the integration of the area and telephone weights would be different for the neighbourhood samples than for the regional sample.
- At the neighbourhood level, the identification of the outlier weights having an impact on the neighbourhood estimates could be different than at the regional level. Therefore, it is possible that some weights included in the neighbourhood samples, which may or may not be already adjusted downward to reduce their impact on the regional estimates, need further adjustment to reduce their impact on the neighbourhood level estimates. (S. Thomas, personal communication, June 13, 2011).
- Post-stratification is based on the age and gender composition of the population of Ottawa-Carleton, therefore the relative sizes of the respondents' weights are representative of the proportion of male and female as well as the

different age groups of the Ottawa-Carleton population, not the neighbourhood populations.

In all, the regional sampling weights may not capture the unequal sampling probability of the respondents at the neighbourhood level. Moreover, these weights may not ensure that the sample age and gender characteristics represent the neighbourhood population ones. The appropriateness of the sampling weights is also further questionable considering that five cycles were pooled for the current investigation. Statistics Canada's Senior Methodologists, Steven Thomas and Marie-Claude Duval, strongly discouraged the researcher to adjust the weights provided in the master files in order to suit the particular need of the current investigation since important information on the data collection process is not available in those files (personal communication, June 13, 2010). Based on the above discussion, neighbourhood estimates that would be derived from weighted data were believed to be non-representative of the neighbourhood populations.

b) Standardisation as a Method to Derive Neighbourhood Estimates
Representative of Neighbourhood Age Composition

The decision to disregard the weights in the creation of the neighbourhood estimates assumes that the neighbourhood residents had an equal probability of being sampled. Estimates produced from such samples are not necessarily representative of the populations from which they are drawn due to the potential presence of sampling bias (Tronchim & Donnelly, 2008). Sampling bias occurs when marked differences between the sample composition and the population composition are observed, either due to the flaw in the sampling process or the luck of the draw. The potential impact of these differences on the estimates arises from the assumption that respondents with certain demographics in the sample are more representative of the people with the same demographics in the population of interest compared to people with different demographic characteristics. For example, it assumed that the responses provided by younger adults in the sample will be closer to the reality of younger people in the neighborhood compared to the reality of the elderly of the same neighbourhood. (de Vaus, 2002, p.152-157)

Standardisation is an attractive alternative method to derive neighbourhood estimates from the CCHS samples which takes into consideration age-related differences between the sample composition and the neighbourhood composition. The objective of standardisation is to adjust for potential effect of such age differences in the calculation of the neighbourhood estimates.

For the current investigation, the standardisation method selected adjusted for difference between the age composition of the neighbourhood samples and the neighbourhood populations. Analyses of the differences between the proportions of 20-39, 40-59 and 60-79 years old among the neighbourhood samples and the neighbourhood populations acknowledged the presence of sampling bias, which may have an impact on the estimates derived for certain neighbourhoods.

The formula used to age-standardise the neighbourhood estimates for the variables of interest derived from the pooled CCHS neighbourhood samples was:

$$X = \sum [(n^*_i / n) (N_i / N)]$$

Where:

- **n^{*}_i** was the number of CCHS respondents in age group *i* of the neighbourhood sample with the characteristic of interest (e.g. 'do not have a family physician')
- **n_i** was the number of CCHS respondents in aged group *i* of neighbourhood sample (e.g. including the respondents with and without a family physician)
- **N_i** was the neighbourhood population in age group *i*
- **N** was the neighbourhood population aged 20-79 years old

The neighbourhood population characteristics used in the standardisation were based on the 2006 census.⁶⁰ 20-year age intervals were selected for the age groups

⁶⁰ For the variables covering only women, the female population characteristics of the neighbourhoods were used to standardise the neighbourhood estimates. For the variables not covering all respondents based on their age, the neighbourhood population used for the standardisation only included the relevant age groups.

(20-39, 40-59 and 60-79 years old). Such age interval, as oppose to 10-year, was selected to limit the number of age-specific estimates (n^*_i / n) derived from a small number of respondents, which may be more unstable. For the variables covering all cycles and respondents of the pooled sample (see Appendix 4), 6 neighbourhoods had estimates generated from an age-specific sample size of less than 5 respondents for at least one of the age groups averaged. For these same variables, 28 neighbourhoods had estimates generated from an age-specific sample size of less than 10 respondents for at least one of the age groups averaged. For the variables not covering every cycle or respondent pooled, the number of neighbourhoods with estimates generated from small age-specific samples size was greater.⁶¹ The situation would have been worse if 10-year age intervals were used for this standardisation.

In addition to age composition, other differences between the sample and neighbourhood composition – such as gender, immigration status or income level – could be incorporated in the standardisation process, if believed to have an impact on the representativeness of the estimates. For the current investigation, preserving the socio-economic characteristics of the neighbourhood populations could be argued. To do so, detailed information on the proportions of neighbourhood residents and sampled respondents by SES level – such as income category or education level – would be necessary. Such information was available, yet suffering from missing data in the sample. Due to the limited sample size of certain neighbourhoods and the lack of complete information on samples composition, it was not possible to address other compositional differences between the sample and neighbourhood population in the standardisation process.

3.5.4.3. NEIGHBOURHOOD LEVEL IMPUTATIONS

As mentioned in Section 3.5.2, Statistics Canada restricted the release of neighbourhood estimates generated from a total of 15 respondents or less. This

⁶¹ The percentage of women who never had a Pap test is the only variable selected for the analyses models which did not cover all cycles or respondents. For this variable, 30 neighbourhoods had estimates generated from an age-specific sample size of less than 5 respondents for at least one of the age groups averaged. And, 58 neighbourhoods had estimates generated from an age-specific sample size of less than 10 respondents for at least one of the age groups averaged.

condition was not satisfied among all neighbourhoods for the variables with smaller sample size due to their reduced coverage (see Appendix 4).⁶² Yet, to avoid dropping neighbourhoods with missing estimates for some variables of interest, especially in the data preparation and model preparation phases of the study, imputations at the neighbourhood level were performed.

The mean substitution was selected as the imputation strategy. This commonly used method consists of assigning the mean value of all cases whose value on the imputed variable is available to any case which has a missing value for that item (Levy & Lemeshow, 1999, p.409). Imputations of the mean ensure that the missing values are replaced by an 'expected' value with relatively high degree of stability (Levy & Lemeshow, 1999, p.409). This assumes that the distribution of the estimates around the mean is the same throughout all neighbourhoods. The disadvantages of this method are: 1) inserting the mean value to a number of cases reduces the variance for the imputed variable, and 2) imputing mean values to two or more correlated variables understate the strength of association between them when cross-analysed (Dorofeev & Grant, 2006, p.132).

In an effort to minimise such drawbacks, subgroup means can be imputed rather than grand means (Levy & Lemeshow, 1999, p.409). This assumes that the distribution of the estimates of the imputed variable is related to some common characteristics of neighbourhoods. Socio-economic status is a strong predictor of health-related characteristics of neighbourhoods (Ansari, 2007; Dunn, Schaub, & Ross, 2007). Therefore, the imputations performed for the current investigation consisted of assigning the mean of all neighbourhoods of the same SES quintile to any neighbourhoods of this SES quintile with a missing estimate. For example, for the variable of percent of women who never had a Pap test, the mean percent for all neighbourhoods of the least deprived quintile was assigned to the neighbourhoods of this quintile with a missing estimate.

⁶² No imputation at the individual level was performed to account for coverage differences across the variables.

Imputations were performed on the following variables:

- Percent who never received a Pap test⁶³ (13 neighbourhoods imputed)
- Percent who never received a breast examination (21 neighbourhoods imputed)
- Percent who never received a mammography (23 neighbourhoods imputed)
- Percent who rated the quality of the care provided by physician as being fair or poor (31 neighbourhoods imputed)
- Percent who rated the availability of health care in the community as being fair or poor (9 neighbourhoods imputed)
- Percent with unmet health care need in past year (7 neighbourhoods imputed)

Following the neighbourhood imputations, a dataset containing 90 neighbourhoods and estimates for the 24 variables of interest which could be derived from the pooled CCHS sample was released from the COOL RDC. This dataset also included flags identifying the estimates imputed or derived from small age-specific sample size. Descriptive statistics about these variables are available in Appendix 10.

3.5.4.4. FURTHER COMMENTS ON MEASUREMENT QUALITY OF THE VARIABLES DERIVED FROM CCHS

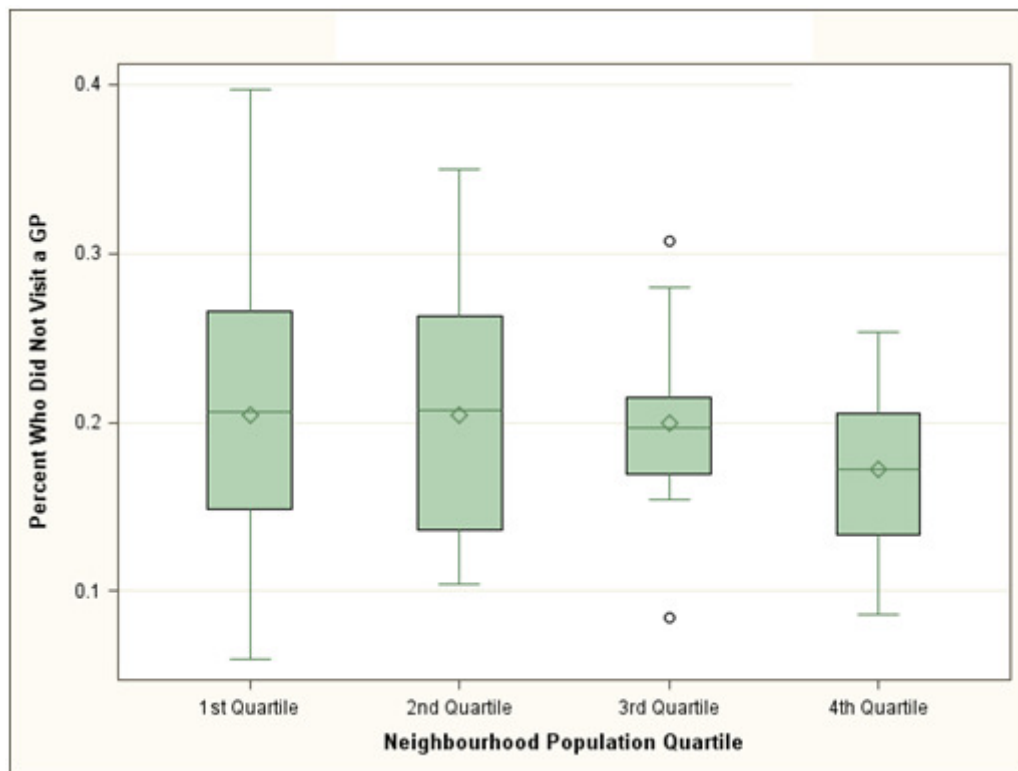
In addition to the limitations discussed and addressed previously, the quality and representativeness of the neighbourhood estimates derived from the pooled CCHS samples may be affected by measurement error⁶⁴. In fact, the variability in the neighbourhood estimates derived from small sample size was greater than the

⁶³ This variable is the only imputed variable which was selected for the analyses models. Sensitivity analyses were undertaken to evaluate the effect of neighbourhood imputations on the results of the analyses models. The sequential model and the final model including this variable were run without the neighbourhoods with imputed estimates (n=76). The results of these models did not change from the ones of the final models presented in this document. Imputations did not affect the results of the current investigation.

⁶⁴ Measurement error is the variation between the measurements of the same quantity on the same case, here neighbourhoods (Bland & Altman, 1996). Measurement error may arise from random error, where any factor affecting the measurement is different across cases in the sample, or systematic error, where a factor affect the measurement similarly for all the cases in the sample (Lattin, Carroll, & Green, 2003, p.82). Measurement error associated with small population at risk or small sample size is subject to random error, which is considered as noise.

variability of the estimates derived from large sample size. For the ONS neighbourhoods, the pooled sample size of neighbourhoods comprised of small population tended to be smaller than for the neighbourhoods with large population. For some of the variables derived, this generated greater variability in the estimates derived for neighbourhoods with smaller populations than for the estimates derived for neighbourhoods with larger populations. An example is shown in Figure 11, where the variability of percent of people who did not visit a general practitioner in the past year is greater among the neighbourhoods with the smaller populations than among the neighbourhoods with the larger populations. In the analyses models based on linear regression, this may affect the model assumption of constant error variance (Section 3.6.2.2). In order to address the potential impact of measurement error linked with neighbourhood population size, regressions analyses weighted by neighbourhood population size were performed as sensitivity analyses (Section 3.6).

Figure 11: Distribution of Percent Who Did Not Visit a GP by Neighbourhood Population Size



3.6. Analyses Models

Three analysis models were built to answer the three research questions of the current investigation: an ANOVA model for the first objective and two different types of linear regression models for the second and third objectives. These models are presented in Sections 3.6.1 and 3.6.2 below, along with their strengths and limitations. The models assumptions and diagnostics methods used to assess them are also discussed in addition to the statistical tests available for these models. Sensitivity analyses models are also presented .

3.6.1. Analyses Model for Research Objective 1

The first objective of the current investigation was to illustrate the gap in ACS hospitalisation rates among low and high SES neighbourhoods in Ottawa and to identify its importance as well as its characteristics.

To illustrate the gap in ACS hospitalisation rates among low and high SES neighbourhoods in Ottawa: descriptive statistics and mapping was used. The distribution of ACS hospitalisation rates by socio-economic quintile was presented in tables and box plots generated using SAS EG 4.3. Following Parenteau et al. (2008) methodology, the quintiles of ACS hospitalisation rates and the quintiles of socio-economic status among the neighbourhoods of Ottawa were represented by different colours on choropleth maps using ArcGIS 10.0.

To assess the importance of the gap in ACS hospitalisation rates among low and high SES neighbourhoods in Ottawa, analysis of variance (ANOVA) was used as in Macintyre, Macdonal, and Ellaway (2008). ANOVA comprises a number of "statistical techniques designed to test if different groups differ in terms of a continuous dependent measure" (Proctor & Badzinski, 2002, p.289). This method is applicable to one continuous dependent variable (e.g. ACS hospitalisation rates) and one categorical independent variable (e.g. SES quintiles) (Lattin et al., 2003, p.386-387). For the current investigation, a single-factor ANOVA model was built to test whether the means of ACS hospitalisation rates differ across the five SES

quintiles. Using this technique, it was also possible to characterise the gap in ACS hospitalisations rates among the neighbourhoods' deprivation levels.

The ANOVA performs well under a balanced design, meaning that each group (SES quintile) is represented by the same number of observations (neighbourhoods) (Lattin et al., 2003, p.390). To yield reliable results, ANOVA also requires a minimum of 15 observations per group, among which the dependent variable is distributed normally (Proctor & Badzinski, 2002, p.293). The model assumes that the observations are independent and that the variances in the dependent variable are equal among the groups (Kerr, Hall, & Kozub, 2002, p.90). The homogeneity of the variance was tested using the Levene statistic, where a highly insignificant value ($p > 0.15$) refers to homogeneous variances among the groups (Kerr et al., 2002, p.90). In the event that homoscedasticity was not satisfied in the ANOVA model, transformations on the rates of ACS hospitalisation can be performed (square root and natural log) to improve the reliability of the results (Forthofer, Lee, & Hernandez, 2007, p.324).

Once the assumptions of the ANOVA model were met (or no important violation was reported)⁶⁵, the following three statistical tests were performed: ANOVA F test, Tukey's HSD comparisons and test for linear trend.

a) ANOVA F Ratio Test – Testing the Significance of Differences in Mean ACS Hospitalisation Rates across Neighbourhood Quintile

All ANOVA F tests were computed using the 'One-Way ANOVA' module in PASW 18.0. This test assessed the significance of the differences across the means of ACS hospitalisation rates by neighbourhood quintiles, where the null hypothesis assumed that all quintile's means are equal. The F test was computed from the ratio of the variance in ACS hospitalisation rates within the quintiles on the variance in ACS hospitalisation rates across the quintiles. The variance within the quintiles was estimated from the variation of the neighbourhood rates around their quintile's

⁶⁵ Most authors consulted reported that the ANOVA tests were not affected by small departures from the model assumptions, only extreme violations should be avoided (Forthofer et al., 2007, p.323-324; Kerr et al., 2002, p.91; Proctor & Badzinski, 2002, p.293).

means. The variance across the quintiles was estimated from the variation of the quintiles means around the overall mean. Greater variance across quintiles was associated with greater meaningfulness of differences. The F test value was compared to a tabled F based on $m-1$ and $n-m$ degrees of freedom, where m is the number of groups and n is the number of cases. A significant F test would signify that ACS hospitalisation differ significantly by neighbourhood deprivation level, yet no information about how the quintiles' means differ would be provided by this test. (Lattin et al., 2003, p.420-421)

b) Tukey's HSD Test – Post Hoc Comparisons of Quintile's Means of ACS Hospitalisation Rates

In the event that the ANOVA F test is significant⁶⁶, post hoc comparisons can be performed to characterise the differences observed. The Tukey's HSD test is a multiple comparison method used to compare the means ACS hospitalisation rates by SES quintile amongst each other. The objective of this test is similar to a t test used for comparing the means of two independent samples, yet Tukey's HSD test control for the number of comparisons⁶⁷ to be made across the quintiles which reduces the possibility of making Type I error (rejecting the null hypothesis of a statistical test when the null hypothesis is true) (Proctor & Badzinski, 2002, p.304).

The Tukey's HSD test involves "comparing the differences between any two means to a critical value following a sampling distribution called the studentized range statistic q " (Proctor & Badzinski, 2002, p.305). This critical value is calculated as $q * \sqrt{\text{variance within groups} / \text{number of cases per group}}$, where q is found on a statistic table of this distribution and has $m-1$ and k ⁶⁸ degrees of freedom. Any difference in the quintile means greater than the critical value is significant. This test also assumes that the variance across the neighbourhood quintiles is constant.

⁶⁶ As for the current investigation (see Section 4.1)

⁶⁷ A total of 10 pairwise comparisons would be performed among the quintile's means for the current investigation. As the number of t tests performed increases, the likelihood of making Type I error increases. Therefore, concluding to a significant difference between two quintile means using t tests could be the result of sampling error rather than true differences among the populations. (Proctor & Badzinski, 2002, p.304)

⁶⁸ K represents the number of groups (e.g. SES quintiles).

Using Tukey's HSD test, it was possible to identify which quintiles means in ACS hospitalisation rate were significantly different and describe the magnitude of the gap in ACS hospitalisation rates among neighbourhoods of different SES status. All Tukey's HSD comparisons were performed by prompting the 'Post Hoc' option in the 'One-Way ANOVA' module of PASW 18.0. (Proctor & Badzinski, 2002, p.305)

c) Test for Linear Trend – Testing the Presence of a Socio-Economic Gradient in ACS Hospitalisation Rates

To further characterise the gap in ACS hospitalisation rates by neighbourhood deprivation level, a statistical tool was available to gain insight on whether the relationship was following the socio-economic gradient of health. To do this, the differences between the ACS hospitalisation rates by neighbourhood deprivation quintile can be tested for linearity – assuming that the ANOVA F test is significant – using the test for linear trend. This test was prompted by selecting the 'Constrats' option and specifying 'linear' in the polynomial box of the 'One-Way ANOVA' module in PASW 18.0. Two components formed this test: 1) the linear component, a measure of the probability that the means are following a linear trend and 2) the deviation component, a measure of any deviation from the linear trend, such as a quadratic or a cubic trend (Kerr et al., 2002, p.157). The significance of these two components were tested using an F distribution where a significant linear term ($p < 0.05$) and an insignificant deviation term ($p > 0.05$) would mean that the relationship between the ACS hospitalisation rates and the neighbourhood deprivation quintiles was linear (Garson, 2009; Kerr et al., 2002, p.158). This test assumed that the distance between the groups was ordered and equal. In the context of deprivation level, measured by SES quintiles, the steps between one quintile to the next were not equal: they were dependent on the distribution on the SES index values of the neighbourhoods. Considering that the test was used to verify if an increasing relationship between SES quintiles and ACS hospitalisation rates was present, therefore supporting the socio-economic gradient theory, it was acceptable that the distance between the quintile were not equal.

In the event that no linear relationship was demonstrated between ACS hospitalisations and neighbourhood deprivation, a monotonic increasing relationship can be tested. This test was performed by assigning an overall rank value to the neighbourhoods based on their ACS hospitalisation rates. Then, the linearity test described above was performed on the ranks of ACS hospitalisation rate by deprivation quintile. A significant linear relationship, in this case, would demonstrate that the relationship between ACS hospitalisation rates and neighbourhood deprivation quintile was monotonically increasing, therefore supporting the theory of the socio-economic gradient of health.

3.6.2. Analyses Models for Research Objectives 2 and 3

Linear regression analyses were used to achieve the second and third objectives of the current investigation. In the sub-sections below, the applicability and usefulness of this method to achieve these two objectives will be discussed (Section 3.6.2.1). Then, the linear regression model, estimation method and assumptions are presented in addition to the diagnostics methods used to assess them (Section 3.6.2.2). The statistical tests available for linear regression analyses and the process of making statistical inferences are described (Section 3.6.2.3). And lastly, the sensitivity analyses performed to assess the robustness of the models are presented (Section 3.6.2.4).

3.6.2.1. APPLICABILITY AND USEFULNESS OF LINEAR REGRESSIONS

Regression analyses are commonly used statistical methods for the analyses of dependence; they are used to assess the relationship between a set of independent variables and a single dependant variable. Regression analyses can be applied to observational data in which the independent variables are correlated amongst each other and with the dependent variable. Therefore, such technique is appropriate for an ecological research design as the one in the current investigation. (Tabachnick & Fidell, 2007, p. 111-112)

Multiple linear regression was selected for the current investigation. This method assumes that the rates of preventable hospitalisation are linearly associated with a

combination of neighbourhood characteristics – which was verified in Section 3.7.2.3. This method is normally applied to a continuous dependent variable and to continuous and/or dichotomous independent variables. Yet, the ACS hospitalisation rates were censored: they cannot take negative values or values greater than 100,000 and applying a linear regression model on this dependent variable could predict values outside its plausible range. However, considering that the mean of the rates was largely above 0 (599 ACS hospitalisations per 100,000 person-year) and that no rate was either 0 or 100,000 among the neighbourhoods of Ottawa, linear regression models could be used without concern of the model predicting negative values of ACS hospitalisation rates (J.M. Billette⁶⁹, personal communication, August 24, 2011). In the literature, linear regression was the method of choice to model the relationships between ACS hospitalisations rates and different characteristics of places (Ansari et al., 2006).

The result of a regression model is an equation that represents the best predictions of the dependent variable from the independent variables. There are two goals to regression: 1) to estimate the best set of regression equation parameters - using the information available on the independent variables - to generate predicted values as close as possible to the measured values of the dependent variable and 2) to obtain the proportion of the variance in the dependant variable that can be predicted by the different independent variables included in the model⁷⁰ (Tabachnick & Fidell, 2007, p.112; Warner, 2008, p.428).

To achieve the 2nd and 3rd objectives of the current investigation, regression analysis was a suitable method as it could be used to describe the relationships between neighbourhood ACS hospitalisation rates and neighbourhood characteristics as well as to make inferences on the importance of these relationships. In fact, the proportions of variance in the dependent variable explained by the independent variables in the model suggest their relative explanatory power and could be

⁶⁹ At the time of the study, Jean-Michel Billette (PhD) was a statistician at Statistics Canada, working as an analyst at the COOL RDC.

⁷⁰ More about the regression model, its estimation method and their assumptions is presented in Section 3.6.2.2.

assessed with statistical tests⁷¹. Therefore, the importance of each neighbourhood compositional and contextual characteristics to the rates of preventable hospitalisation could be assessed using regression. Moreover, based on the direction of the regression coefficients estimated from the model, it could be possible to describe the nature of the relationships signaled. Also, testing the importance of aspects of neighbourhoods, such as primary care access represented by a number of neighbourhood characteristics (variables), could also be performed using regression analyses. (Tabachnick & Fidell, 2007, p. 113)

There are three types of multiple linear regression methods: standard multiple regression, sequential multiple regression and statistical multiple regression⁷². The differences between these types of regression rest on the interpretation of the variance in the dependent variable explained commonly by the independent variables included in the model. A proportion of the variance in the dependent variable can be explained commonly by a number of variables included in the model considering that these variables are correlated, therefore explaining a common facet of the dependent variable – See Appendix 11 for greater details about the concept of explained variance in linear regression analyses. (Tabachnick & Fidell, 2007, p. 131-132)

In standard multiple regression, all variables are entered at once in the model and their relationships with the dependent variable are assessed simultaneously: the shared variability explained by all the variables in the model is not attributed to any predictors and its significance is not tested independently. In sequential regression models, the variables are entered in the model following a pre-established order determined by theory and their relationships with the dependent variable are assessed sequentially: the shared variance is attributed to the variable(s) entered first in the regression. Therefore, hypotheses about the proportions of variance attributable to some independent variables could be tested after the variance due to other variables has already been accounted for in the model (Tabachnick & Fidell,

⁷¹ The statistical test will be described in Section 3.6.2.3.

⁷² This later method was not used in the current investigation, as the selection of variables entered in the final models is based solely on statistical criteria (Tabachnick & Fidell, 2007, p. 133).

2007, p. 138). Further details on the how these two methods were applied to achieve the research objectives are provided below.

Four sequential linear regression analyses were conducted to test the four different hypotheses formulated in the literature (Objective 2). The variable representing socio-economic status, the SES index, was the variable entered first in the model. Then, the variables representing the hypothesis tested were entered all together (as a block) in the model.⁷³ Changes in R^2 and adjusted R^2 between the two models were assessed and an incremental F ratio test was conducted to test the significance of the additional variables included in the model⁷⁴. These models were performed in SAS using PROC REG.

In all, four different sequential models were conducted to test the four hypotheses. These four models closely parallel the key hypotheses of the causal model described in Chapter 2.

- A sequential model where the variable(s) representing the concept of differences in health status were entered (as a block) after the variable of SES index
- A sequential model where the variable(s) representing the concept of differences in healthy behaviour were entered (as a block) after the variable of SES index
- A sequential model where the variable(s) representing the concept of differences in primary care access were entered (as a block) after the variable of SES index
- A sequential model where the variable(s) representing the concept of differences in preventative health services utilisation were entered (as a block) after the variable of SES index

⁷³ See Section 3.7.3 for details about the variables selected.

⁷⁴ In the event that the added variables are significantly improving the proportion of variance in ACS hospitalisation rates explained, it could be concluded that the hypothesis formulated has an impact on the rates of ACS hospitalisation over and above the impact of differences in socio-economic status. Details on these statistical tests are in Section 3.6.2.3.

Standard linear regression was conducted to identify the importance of neighbourhood compositional and contextual factors as well as to assess the relative contribution of neighbourhood health status, healthy behaviour and access/utilisation of primary care services to the rates of hospitalisation for ACS conditions among neighbourhoods in Ottawa (Objective 3). A full model including all the variables representing the neighbourhood characteristics selected for the current investigation was run in SAS using the PROC REG procedure. These variables were selected in order to represent the causal model of preventable hospitalisation, see Section 3.7.3 for greater details. An F test was conducted to assess the validity of the model and *t* tests were conducted to assess the individual importance of the predictor variables included in the model. Incremental F ratio tests were conducted to assess the relative contributions of the variables representing neighbourhood health status, healthy behaviour as well as access and utilisation of primary care services. Details on these statistical tests are provided in Section 3.6.2.3.

3.6.2.2. MULTIPLE LINEAR REGRESSION MODEL – ESTIMATION, ASSUMPTIONS AND THEIR DIAGNOSTICS METHODS

The regression model is the theoretical statistical concept from which the regression equation presented below is estimated. Prior to conducting significance tests on the solution of the regression models, it is important to assess if the data fit the assumptions of the linear regression model. These assumptions are stemming from the theoretical aspects of the linear regression model itself and the method used to estimate the solution, which are presented in this section along with the diagnostic methods used to assess them in this research.

The multiple linear regression model with *k* predictor variables is expressed as:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

Where *y* is the dependant variable and the *X*s represent are the independent variables. The parameters, or the regression coefficients, of the equation are β_j (*j* = 0, 1, ..., *k*) where the intercept of the model is β_0 . These terms are unknown and must be estimated from the data. For the current investigation, ordinary least

squares (OLS) estimation was used to derive estimates of the regression coefficients. The parameters estimated from the model represent, taken individually, the expected change in the response variable y per unit change in the predictor variable X when all of the remaining independent variables are held constant. The symbol ε represents the random error component of the model, which can be thought of as a statistical error, meaning that it is a random variable accounting for the failure of the model to fit the data perfectly. (Montgomery, Peck, & Vining, 2006)

The estimation of the regression model for the current investigation was based on OLS, which was the most often selected estimation method in the literature on ACS hospitalisation rates (Ansari et al., 2006). This method is also one of the most commonly used in linear regression analyses, therefore the readers tends to be familiar with this estimation method. The rational of this method is to choose the estimates of the regression coefficients in order to minimise differences between the observed values of the dependent variable and its predicted values by the model equation. A measure of these differences in the model is the sum of squared errors (SSE), which is the sum of the squared differences between the observed values of the dependent variable and its predicted values of the model (DeMaris, 2004, p.47). This estimation method generates the most accurate fitted values and has the following desirable properties:

- The estimates of the regression coefficients are unbiased, meaning that the estimated coefficients are equal to the population (true) coefficients and
- The estimates of the regression coefficient are efficient, meaning that their variance is the lowest (Lattin et al., 2003, p.45)

The multiple linear regression model derived using OLS estimation has a number of assumptions which could be summarised as follows (Lattin et al., 2003, p.44):

- The predictor variables are fixed and the matrix of their values is full rank
- The error term ε is independently and identically distributed with mean 0 and constant variance σ^2

When the predictor variables are fixed, meaning that they must not be stochastic, they are measured without error. This assumption is important since, when the predictor variables are fixed, they cannot be correlated with the error term ε . If this assumption could be violated, the researcher must be explicit about the possible relationship between the measurement error in the predictor variables and the structure of the error term. For the current investigation, measurement error in the predictor variables may be at play and it may be affecting the assumption of constant variance of the error term (homoscedasticity). Sensitivity analyses were conducted in order to avoid breaching this important assumption (see Section 3.6.2.4 for details). (Lattin et al., 2003, p.44)

Another assumption of the linear regression model is that the values of the predictor variables are full rank; meaning that they are not too redundant or too highly related to each other (multicollinearity). Perfect redundancy occurs when a variable is the linear combination of one or more variables included in the model, if so the variables are said to be collinear. In such model, there is no unique solution to the regression equation since the different effects of the collinear variables included in the model cannot be identified. Perfect redundancy is rare in regression models, but near-collinearity among the predictors is possible. In such model, the full rank matrix assumption is not violated: the model can be solved using OLS and the estimates can be produced. Yet, the model is suffering from multicollinearity and the parameter estimates as well as their standard errors may be inflated (DeMaris, 2004, p.226). The presence of multicollinearity could therefore reduce the significance of the model's parameters. Due to this important limitation, multicollinearity was avoided in the final models of the current investigation. (DeMaris, 2004, p.87)

For the current investigation, multicollinearity was assessed using three indicators: the condition index (CI), the variance proportion and the variance inflation factor (VIF). The condition index, calculated with the COLLIN model statement in SAS' PROC REG procedure, are a measure of the seriousness of the linear dependencies in the matrix of the predictor variables included in the model. CI values above 30 indicate that multicollinearity involved among the predictor

variables may have a moderate effect of parameter variance estimated in the model (Lattin et al., 2008, p.57). The proportion of the variance in the parameter estimates attributable to multicollinearity was assessed by the variance proportions estimated with the COLLIN model statement. When the variance proportions of two or more variables are high for a component with a CI value above 30, these variables are approximately dependant linearly (DeMaris, 2004, p.228). Removing those variables reduces the presence of multicollinearity in the model. The diagnostic test of VIF, prompt using the VIF statement in PROC REG, indicates "how many times the variance of a coefficient is magnified as a result of the collinearity compared to the ideal case of perfectly orthogonal predictors" (Myers as cited in DeMaris, 2004, p.228). Any value of VIF for a parameter coefficient above 10 represents a collinearity problem (DeMaris, 2004, p.228). Due to the potential effect of multicollinearity on the significance of the statistical tests, the CI, variance proportions and VIF were assessed in the exploratory model and avoided in the selection of the final models.

The assumptions related to the error term of the model are based on the notion that this term is "capturing the effects of all unobserved and unmeasureable factors that might also influence the dependant variable" (Lattin et al., 2003, p.45). These assumptions are verified using the residuals of the model. The residuals are calculated for each case as follows:

$$e_i = y_i - \bar{y}_i, \quad i=1,2,\dots,n.$$

Where y_i is the observed value of dependent variable for a case and $\bar{y}_i$ is its predicted value calculated from the model equation. The residuals are a measure of the variability in the dependent variable not explained by the regression model (Montgomery et al., 2006, p.123). Therefore, analyses of the residuals could demonstrate any departure from the assumptions pertaining to the error of the model. In SAS, the model statement R prompts the analyses of the residuals for the model.

Residuals should follow a bell-shaped distribution (Lattin et al., 2003, p.45). If the normality assumption is grossly violated, the parameter estimates, their confidence interval and significance tests may be affected. Yet, small departures from normality do not affect the validity of the model greatly (Montgomery et al., 2006, p.129). For every model in the current investigation, four methods were used to assess the normality of the residuals distribution: the histogram of residuals, the scatter plot of studentized residuals against the predicted values and the Q-Q plot of residuals. The bell-shape of the residual distribution can be examined from the histogram. In the graph of studentized residuals and predicted values, if about 95% of the errors are located between ± 2 standard deviations, the assumption regarding the normal distribution of the error is supported (Lattin et al, 2003; Montgomery et al., 2006). The Q-Q plot is designed in such a way that normally distributed residuals are plotted along a straight line. Important departure from the straight line is an indicator that the distribution of residuals is not normal (Montgomery et al., 2006, p.129).

The assumption of independence of the errors means that the error for one case is not influenced or correlated to the error of other cases. In models where the independence of the errors is violated, the residuals are auto-correlated. In such models, the parameters estimated using OLS are unbiased, yet their variance is biased downward which might enhance the significance of the parameters (Lattin et al., 2003, p. 61). The Durbin-Watson statistic (DW) is a way to detect auto-correlation in the residuals, where values close to 2.0 represent no evidence of auto-correlation (Lattin et al., 2003, p. 61). The DW statistic was verified for every model of the current investigation using the model statement DW in SAS' PROC REG procedure.⁷⁵

⁷⁵ In regression models involving a spatial component, e.g. neighbourhoods as the unit of analysis, the residuals can be spatially auto-correlated according to their geographical ordering, which would violate the assumption of the independence of the errors (Cliff and Ord, 1981). Spatial auto-correlation can be assessed by calculating a Global Moran's I test in ArcGIS on the residuals of the multivariate model. This statistic measures the 'self-similarity of a spatial variable's value as a function of adjacency' (Parenteau & Sawada, 2011, p.4). In the event that the residuals are spatially auto-correlated, a spatial lag regression model can be conducted in GeoDa using a weighted matrix summarizing all the pairwise spatial relationships between the neighbourhoods (Anselin and Rey, 1991, Cliff and Ord, 1981, Ward and Gleditsch, 2007). In a study on respiratory health outcomes using the ONS neighbourhoods, Parenteau and Sawada (2011) demonstrated that the regression

Another assumption is homoscedasticity, which means that the errors of the regression model are assumed to have the same variance σ^2 (homoscedastic), regardless of the value of dependent variable predicted by the OLS. In the case of heteroscedasticity (where the variances are not the same), the parameters estimated using OLS are unbiased but inefficient, meaning that they are not estimated with the smallest variance and the significance of the variables in the models may be reduced (DeMaris, 2004, p.202). Heteroscedasticity often occurs when the variance of the error terms is related to the variance either the dependant or independent variables and its presence was tested using the following methods:

- The shape of the scatter plot of residuals and predicted values (R statement PROC REG in SAS),
- The shape of the scatter plot of residuals and independent variables - partial regression plots (PARTIAL statement PROC REG in SAS), and
- The White test (SPEC statement in PROC REG in SAS).

The vertical spread of the scatter plot of residuals and predicted values, as well as of the partial regression plots, should be constant to reflect constant variance of the residuals. In the case of heteroscedasticity, these plots demonstrate fan-shaped or U-shaped patterns (Lattin et al., 2003, p.58). When the null hypothesis of the White test is satisfied, it means that the errors are homoscedastic, that the errors are independent of the predictor variables and that the model is correctly specified (DeMaris, 2004, p. 203). Performing weighted regression⁷⁶ as sensibility analyses allow to assess the potential impact of measurement error on the constant variance of the errors – see Section 3.6.2.4.

Another important implicit assumption of the regression model is that the predictor variables are linearly related to the dependent variable. This means that the values of the dependant variable are linearly related to the additive function of the predictor

model based on these natural neighbourhoods was properly fit and did not suffer from spatial auto-correlation, compared to the models based on census areas. For the current investigation, spatial auto-correlation was not considered as a potential threat to the model's assumption.

⁷⁶ In the event that a particular variable was identified as the cause of heteroscedasticity, based on the partial regression plots, transformations on that variable could be performed as a remedy.

variables (DeMaris, 2004, P.163). Departure from linearity of the dependent variable with the independent variables can be revealed by the partial regression plots. If the scatter of the points around the residuals' 0 line demonstrates a non-linear shape, it may signal non-linearity (Montgomery et al., 2006, p.135). Moreover, non-linearity can be identified from the scatter plots of the predictor variables with the dependent variable. Any non-linear shape on these plots, for example an exponential curve, could be a sign of non-linearity between the dependant and independent variables (Tabachnick & Fidell, 2007, p. 77). For the current investigation, any departure from linearity was assessed in the Section 3.7.2.3.

Other than the above theoretical and statistical assumptions of the model, the results of linear regressions using OLS may be strongly influenced by one or few cases in the data. It is important, in addition to the assessment of the assumptions, to identify multivariate outliers and evaluate the cases with leverage and/or influence on the regression solutions.

A multivariate outlier is a case, here a neighbourhood, with extreme values on a combination of two or more independent variables which may distort the parameters estimated in the regression model (Tabachnick & Fidell, 2007, p. 66). There are many diagnostic indices that help identify putative outliers, where Mahalanobis distance is one of them. It represents the "distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of the means of all the variables" (Tabachnick & Fidell, 2007, p. 68). These distances follow the chi-square distribution and have a degree of freedom equals to the number of variables included in the model. Any cases with a Mahalanobis distance with a probability value inferior to 0.001 could be considered as a multivariate outlier (Tabachnick & Fidell, 2007, p. 68). In SAS, the Mahalanobis distance of the neighbourhoods included in the regression models were calculated manually from the leverage values⁷⁷.

⁷⁷ Mahalanobis distance = $(N-1) \times (\text{leverage} - 1/N)$, where N is the number of cases in the regression models.

Multivariate outliers in regression analyses may have leverage or influence on the regression results. A case with leverage is an outlier for which the observed value of the dependent variable has an impact on the fitted values derived from the model and may potentially have impacts on the parameter estimates, standard errors, predictor values and model summary statistics derived by the model (Montgomery et al., 2006, p.190). Thus, for each case in the data, a leverage value was calculated from the matrix of values omitting this case using the INFLUENCE statement in PROC REG. Any case with a leverage value greater than $2p/n^{78}$ is a leverage point in the regression model (DeMaris, 2004, p.220; Montgomery et al., 2006, p.191).

An outlier with influence is a case with an unusual combination of independent values which has an impact on the parameters estimated in the regression model. This case or a combination of cases pull(s) the regression in its (their) direction and the parameters estimated may depend more on this (these) case(s) than the majority of the cases in the dataset. There are three measures of influence in regression: the Cook's Distance, the DFBETAs and the DFFITS, which were all generated using the INFLUENCE statement in PROC REG in SAS. Cook's Distance is an overall measure of the difference in the distance of the parameter estimates generated with the case included in the regression model and the parameter estimates generated when this case is excluded from the regression. Cook's Distance assesses the overall influence of the case on the collection of parameters included in the model. Any case with a relatively high Cook's Distance compared to the other cases in the dataset should be further investigated. (Montgomery et al., 2006, p.194)

DFBETAS are a statistic indicating the degree of change in standard deviation unit of a regression coefficient when the case is deleted from the observations (Montgomery et al., 2006, p.195). A DFBETA value is computed for each case and each variable included in the model using the leverage values. This way, it is possible to determine which case is potentially influencing which regression coefficient. A $|DFBETA|$ greater than the cut-off of 2 identifies a case with potential

⁷⁸ For all the cut-off values of diagnostic methods testing for the presence of leverage and influence, p represents the number of parameters in the model and n the number of cases in the model.

influence on a regression coefficient (Lattin et al., 2008, p.65). As cited in Montgomery et al. (2006, p.195), Belsley, Kuh, and Welsch suggested a cut-off of $2/\sqrt{n}$.

The impact of the cases on the fitted values can also be assessed using the DFFITS statistics. A case's DFFITS value represents the change in standard deviation unit of the fitted values when this case is excluded from the regression, and it is calculated from its leverage value. One DFFIT value is calculated for every case in the regression. Any case with a $|DFFIT|$ value greater than $2/(\sqrt{p/n})$ warrants further attention. (Montgomery et al., 2006, p.196)

In the event that extreme multivariate outliers or cases with extreme influence/leverage are identified in a regression model, transformation on the variables generating these cases – if they could be identified – could be performed. This may reduce their impact on the regression results. However, transformation may not always change the status of a multivariate outlying case or its impact on the regression results. In such models, the researcher could decide to discard the case(s) identified as multivariate outlier(s) with influence and leverage: they may influence the results in any direction. (Tabachnick & Fidell, 2007, p. 71)

In the current investigation, the criteria to remove an influential outlying neighbourhood were 1) being identified as a severe multivariate outlier by the Mahalanobis distance, 2) being identified as having influence in the exploratory model and 3) having outlying characteristics not eliminated by variable transformations. The drawback of deleting a case from the population sampled, assuming it is not suffering from a data entry error or an anomalously high/low value, is a reduction of the generalisation of the model results to the population of interest (Montgomery et al., 2006, p.199). Therefore, when dealing with multivariate outliers, the researcher must weigh the impact of the case on the regression results with a reduction in the possibility to generalise the regression results to the population of interest. For this reason, no neighbourhood were excluded based on its potential impact on the final regression models, decisions to remove a case from the analysis

dataset were taken in the model exploration phase only (Sections 3.7.2.1 and 3.7.2.4).

A summary of the linear model assumptions and the diagnostic method employed in the current investigation is available in Appendix 12. Any gross violations of the model assumptions were taken into considerations while selecting the final models in the model regression preparation phase (Section 3.7). In general, the results of the regression model estimated using OLS are not influenced by small deviations from the assumptions (Weisberg, 2005). Such deviations, in addition to the cases with influence and/or leverage on the final models results, will be discussed in the results (Chapter 4) and any threatening deviations will be discussed in the Discussion (Chapter 5).

3.6.2.3. STATISTICAL INFERENCE FROM LINEAR REGRESSION MODELS

Statistical inference is the process of making conclusions from data that is subject to random variation by testing the different parameters of the model applied to the sample (Trosset, 2006, p.199). It is performed using different statistical tests on the importance of the model and on the parameters included in that model. The conclusions are based on probabilistic evidence, which is captured in the concept of significance. Significance is interpreted as the probability that chance only produced the 'extraordinary' results observed in the model⁷⁹ (Trosset, 2006, p.206). Therefore, statistical inference requires that the researcher have hypotheses to test (Tabachnick & Fidell, 2007, p. 142). In the current investigation, hypotheses about the significance of the model as a whole as well as the importance of single variables and groups of variables in predicting neighbourhood preventable

⁷⁹ For the regression models included in this investigation, to conclude that a phenomenon is not coincidental, a significance probability of 0.05 was selected, which represents a 5% probability of making an inaccurate conclusion based on the sample characteristics. Such level of significance is the most commonly used in epidemiological and social research involving linear regression. Considering the ecological research design and the potential presence of measurement error in the variables (two aspects of the current models affecting their capacity at identifying moderate relationships), the regression results significant at a probability of 0.1 were also discussed. (Forthofer et al., 2007, p.218)

hospitalisations were conducted using the set of statistical tools available in linear regression⁸⁰. These tests answer the questions:

- Is the model significant? Are all the predictor variables combined explaining a significant proportion of the dependent variable, in this case the ACS hospitalisation rates?
- Is the predictor variable – or group of variables – included in the model important, therefore predicting significantly the differences in ACS hospitalisation rates when the other variables in the model are held constant?
- Is the addition of variable(s) important for the prediction of the dependent variable in a model already accounting for the total variance explained by other variable included in it, here the socio-economic status of neighbourhoods? (Tabachnick & Fidell, 2007, p. 113-114)

The model goodness-of-fit test, an F test, assesses the significance of the total variance explained in the dependent variable by all the variables included in the model (total variance = $a+b+c$ in Appendix 11). To examine the significance of the unique variance explained by the predictor variables individually (unique variance = a or b), t tests are used. For models with three or more predictor variables, incremental F ratio tests are available to test the significance of the joint contribution of two variables or more. And, incremental F ratio tests can also be conducted between consecutive models in sequential regressions to assess the significance of the contribution of the additional independent variables to the prediction of ACS hospitalisation rates (Kleinbaum, Kupper, Nizam, & Muller, 2008, p378). Below is a description of these tests and how they were implemented to achieve the second and third objective of this study.

⁸⁰ In linear regression, to make statistical inferences, the probability distribution of the errors is assumed to be normal (Montgomery et al., 2006, p.80).

a) Goodness-of-Fit Test – Testing the Significant of All Variables Included in Model

The model goodness-of-fit test is based on the overall proportion of the variance in the dependent variable accounted by the variables in the model, which is R^2 (area $a+b+c$ in Appendix 11). When using OLS, R^2 is calculated as follows⁸¹ (Lattin et al., 2003, p.53):

$$R^2 = 1 - \frac{\text{Sum of Squares of Errors}}{\text{Total Sum of Squares}}$$

R^2 is tested for its significance using the following statistical test (the null hypothesis is $R^2 = 0$ and the alternative hypothesis is $R^2 \neq 0$):

$$F = \frac{R^2 / k}{(1-R^2) / (N-k-1)}$$

Where k is the number of predictor variables and N is the number of cases included in the model. This F statistic is tested against a tabled F value, based on k and $(N-k-1)$ degrees of freedom. When the absolute value of the F statistic is below the tabled F value, the null hypothesis is accepted. When $R^2 = 0$, it means that none of the predictor variables are significant in predicting the dependant variable. (Tabachnick & Fidell, 2007, p. 143)

The adjusted R^2 is an indicator of the proportion of the variance in the dependent variable explained by the combination of all the variables included in the model, which is adjusted based on the number of variables in the model. Such indicator is obtained by multiplying R^2 with the factor $(n-1) / (n-k-1)$ (DeMaris, 2004, p.90). This indicator is not tested for its significance, but it can be used to qualify the model adequacy and to compare models amongst each other. Using the adjusted R^2 ,

⁸¹ The sum of squares of the errors is the sum of the squared differences between the observed values of the dependent variable and the predicted values by the model. The total sum of squares is the sum of the squared differences between the observed values of the dependent variable and their mean. (Lattin et al., 2003, p.53).

comparisons between the four different sequential models built to test the hypotheses formulated in the literature could be done.

b) Incremental F Ratio Test – Testing Significance of a Group of Variables

For the sequential models, built to test the hypothesis formulated in the literature to explain the differences in ACS hospitalisation, changes in R^2 were tested between the two successive models in the sequence⁸². This tests if the addition of the variable(s) in the second model significantly increased R^2 from the simple model. It is assessed with the incremental F ratio test, with m and (N-k-1) degrees of freedom⁸³, calculated as follows:

$$F_{inc} = \frac{(R^2_{\text{model with added variable}} - R^2_{\text{model without added variable}}) / m}{(1 - R^2) / (N - k - 1)}$$

Where m is the number of added variable(s) in the subsequent model, k is the number of predictor variables and N is the number of cases included in the model. The null hypothesis stipulates no increase in R^2 between the subsequent models (R^2 change = 0). The null hypothesis is rejected if the absolute value of the F_{inc} is superior to the tabled F value. Such conclusion, for the sequential models built in the current investigation, was interpreted as follows: the differences in the hypothesis tested, for example differences in primary care access, contributed significantly to the differences in ACS hospitalisations once the total variance explained by neighbourhood socio-economic status was accounted for in the model. This test was also conducted to assess the relative importance of the concept of health status, healthy behaviour, primary care access and utilisation in the full model representing the causal model of ACS hospitalisation rates. (Tabachnick & Fidell, 2007, p. 144-145)

⁸² The simple model contained the variable of SES index only and the subsequent model included both the SES index variable and the variable(s) representing the hypothesis tested.

⁸³ Considering the method used to calculate the degrees of freedom, the incremental F ratio test is sensitive to the number of variables added in the subsequent model; where the test is more stringent as the number of added variables increases. (Tabachnick & Fidell, 2007, p. 144-145)

c) *t* Test – Testing the Significance of One Variable

In the full model containing the variables representing the causal model of preventable hospitalisations, the importance of the variables was tested separately. By doing so, it is possible to assess their relative importance for the prediction of preventable hospitalisations. This achieved the third objective of the current investigation which was to test the importance of neighbourhood compositional and contextual factors to the rates of hospitalisation for ACS conditions in Ottawa neighbourhoods.

In standard regression model, *t* tests assess the significance of the variables taken individually by evaluating its estimated coefficient (β) as well as its squared semi-partial correlation with the dependent variable (sr^2 , represented by the area *a* for X_1 in Appendix 11). The squared semi-partial correlation is the amount by which R^2 is reduced if the independent variable tested is deleted from the regression equation, therefore it represents the unique contribution of that variable in explaining the variation in preventable hospitalisation among the set of variables included in the full model (Tabachnick & Fidell, 2007, p. 140). Therefore, this statistical test assesses if, for an individual variable, its $\beta=0$ or if its $sr^2=0$. The test statistic can be computed from the coefficient or from the squared semi-partial correlation. The calculation method based on the coefficient is:

$$t = (\beta - 0) / se(\beta)$$

Where the degree of freedom is $N-k-1$ and the standard error of the estimated regression coefficient is calculated from the diagonal elements of OLS estimation matrix (Montgomery et al., 2006, p.84).

This test is considered to be a ‘marginal’ test: it is testing the contribution of the independent variable given the other variables in the model. Therefore, this test is sensitive to the other variables included in the model. As discussed by Tabachnick and Fidell (2007, p. 144), "a very important independent variable that shares variance with another independent variable in the analysis may be non-significant although the two independent variables in combination are responsible of a large

part of the size of R^2 . This should be taken into consideration prior to disregarding the importance of a variable in a regression model based on its t test only.

As defined by Trosset (2006), statistical inference is the process of making conclusions from data that is subject to random variation based on statistical tests and the theory of probability distribution. Such conclusions can be reached when the model assumptions are met and they are based on the level of significance of the associations revealed by the model. Any threat to statistical inference can undermine the generalisability of the results. This was acknowledged by the sensitivity analyses conducted as part of the current investigation – further details in following sub-section 3.6.2.4. Other threats to the generalisability of the statistical results may also be at play, such as threats associated with the concepts of construct validity and external validity, and they will be addressed in the discussion (Chapter 5).

3.6.2.4. SENSITIVITY ANALYSES – POPULATION-WEIGHTED REGRESSION MODELS

Sensitivity analyses are often performed in the development of statistical models in order to assess their validity. To the extent that the researcher is unsure of certain assumptions, steps, or choices in an analysis, sensitivity analysis can be used to explore the degree to which conclusions would be altered had alternative assumptions, steps, or choices been made. If the results of sensitivity analysis show that the conclusions are relatively unchanged regardless of the altered assumptions, steps, or choices, then the conclusions can be described as robust to the suspected unknowns. The information gathered by performing sensitivity analyses is valuable to the process of making statistical inference (Pannell, 1997). Conducting sensitivity analyses should be an integral part of any statistical solution where the researcher is uncertain about the parameters of the models.

For the current regression models, the researcher was uncertain about the potential effect of measurement error in the variables on the results of the regression models. Consequently, two assumptions of the regression model could be breached: the predictor variables are fixed (measured without error) and the variance of the

model's errors is constant. By developing a model reducing this uncertainty, it was possible to assess the robustness of the statistical results. If the solution of the model performed as sensitivity analyses is robust, that is insensitive to changes in the model characteristics, the researcher's confidence in the initial results is enhanced (Pannell, 1997).

Measurement error in the predictor variables may be at play, especially among the neighbourhoods with small population size. The neighbourhood estimates derived from small samples or small populations at risk may have greater variance, therefore hold greater measurement error than the estimates derived for the neighbourhoods with larger populations.⁸⁴ When the predictor variables included in a linear regression model are not fixed, they may be correlated with the residuals of the model (Lattin et al., 2003, p.43). For the current models, the possible measurement variability associated with the neighbourhood of small populations may affect the assumption of constant variance of the residuals: there may be greater variability among the residuals of the neighbourhoods with small populations than among the neighbourhoods with large populations. Therefore, the conclusions drawn from the statistical tests performed on the final models may be subject to error. The uncertainty in the measurement of the estimates for the neighbourhoods of smaller populations and the uncertainty about its potential impact on the variance of the models' residuals highlight the necessity of performing sensitivity analyses.

Weighted linear models were built in order to reduce the potential impact of greater variability associated with the estimates of neighbourhoods with small populations on the results of the statistical tests. In linear regression, weights can be attributed to the different cases included in the model. In SAS, this is performed by specifying a weight variable using the WEIGHT statement in PROC REG. In these models, prior to estimating the model parameters, each of the estimates of the neighbourhoods are multiplied by their corresponding weight and then the model is solved using OLS on the transformed data (Montgomery et al., 2006, p. 180-181). By including the

⁸⁴ See section 3.5.4.4 for greater details on the possibility that estimates derived for neighbourhoods with smaller populations may hold more variability than the estimates derived from neighbourhoods with larger population size (Figure 11).

weights in the regression models of the current investigation, the effects expected were to place low weight on neighbourhoods with estimates believed to have large variance (i.e. small neighbourhoods) and to reduce their potential impact on the variance of the residuals (Lattin et al., 2003, p.60).⁸⁵

The weights used in the sensitivity analyses were derived from the neighbourhood population size. These were calculated as followed, so that the sum of all weights equalled 1:

$$\text{Wgt} = \frac{\text{Neighbourhood population 20-79 years old}}{\text{Total population 20-79 years old in Ottawa}}$$

Using these weights, the neighbourhoods with smaller populations had smaller weights in the regression models built as sensitivity analyses.

In linear regression using weights, the SAS output of the *t* tests and the incremental F ratio tests are valid.⁸⁶ Assuming the final models would be suffering from measurement error, the significance of these tests may be greater in the population-weighted models considering that the variance of their residuals may be reduced. In fact, OLS estimation is more efficient when the variance of the error term is reduced: smaller standard errors are calculated which enhances the significance of the statistical tests (DeMaris, 2004, p.202).⁸⁷

⁸⁵ This method is similar to regression models estimated using weighted least square to address the presence of heteroscedasticity in the OLS regression models (DeMaris, 2004, p.205). However, for the current investigation, the weights were not derived from the variance of the residuals themselves but from the neighbourhood population size; since the later are believe to be a potential source of variance in the residuals, if any.

⁸⁶ On the other hand, any statistics computed from the residuals are slightly erroneous for the population-weighted regression models. In fact, the residuals calculated by SAS in weighted regression models are based on the transformed data as opposed to the original data (DeMaris, 2004, p.205). Therefore, the values of R^2 and sr^2 must be computed manually using the original data (these calculations were performed for the final full model, but not for the sequential models). Also, this has an impact on the analyses of the residuals to assess the model assumptions. Due to this limitation of the statistical software, the population-weighted analyses were selected for the sensitivity analyses as opposed to the final models.

⁸⁷ The model assumptions of the population-weighted regression models were verified using the same method as for the final models. The assessment of leverage and influence was also performed. However, these verifications were performed on the residuals calculated from the transformed data.

The results of the statistical tests performed in the population-weighted regression models were compared to the results of the models among which all neighbourhoods were weighted equally. If the solutions of the two methods were equivalent - meaning that the significance of the tests performed from the models weighted equally or weighted by population size were roughly equivalent - the statistical inferences would be considered as robust to the potential violation of the assumption pertaining to the predictor variables being fixed (Pannell, 1997). This would reduce the uncertainty in the statistical inferences formulated for the 2nd and 3rd objectives of this study.

3.7. Regression Model Preparation

The variables of interest represented different aspects of neighbourhoods which were thought to be theoretically linked to the rates of preventable hospitalisation. Lattin et al. (2003, p.50) described the process of selecting variables for inclusion into a regression model as involving as much art as science. A mixture of exploratory methods - for examples descriptive statistics, patterns of correlation, scatter plots and univariate/bivariate regressions – in addition to theoretical judgement were used to select the variables for the final models of the current investigation. And, before building the final models, the practical limitations of multivariate linear regression – such as presence of multicollinearity or outliers – were taken into consideration by evaluating a full exploratory regression model.

The regression model preparation phase involved following three steps: a) the exploration of the variables of interest (Section 3.7.1), b) the assessment of a full exploratory regression model (Section 3.7.2), and c) the selection of the variables for the final models (Section 3.7.3). Details about their implementations for the current investigation are below.

3.7.1. Exploration of the Variables of Interest

Descriptive statistics, scatter plots and univariate/bivariate regressions were performed as part of this exploration⁸⁸. Appendix 10 shows the descriptive statistics of the variables of interest, grouped by the concept they sought to represent, for the 90 neighbourhoods. Appendix 10 also contains their pairwise correlations with the rates of ACS hospitalisation and the SES index. The dependent variable (ACS hospitalisation rates) was highly correlated with the variables of SES index, 'percent with poor or fair self-rated health status' and 'percent of smokers' (correlation above 0.5). The correlations between socio-economic status and these variables, as well as the variables of 'percent who never had a Pap test' and 'percent without a family physician', were also above 0.5. During the process of variable selection, it was tempting to exclude variables not correlated significantly with the dependent variable. However, this criterion alone did not fully justify the exclusion of variables: theoretical considerations and multivariate association also had to be acknowledged (Le, 2003, p.86).

The scatter plots between the variables of interest and the rates of ACS hospitalisation were also generated and consulted to enrich the investigator's understanding of the relationships among the variables. For brevity, only four of these scatter plots were included in Appendix 13 of this report. It is noteworthy that the relationships observed in these scatter plots were consistent with linearity (to the extent that a relationship was signaled) and thus lent further support to the use of linear regression.

⁸⁸ Exploratory spatial data analyses could have been performed to gain insight on the spatial dependencies between the variables of interest. Bivariate Moran's I could have been calculated between pairs of variables using the software GeoDa. This statistic provides information on the strength of spatial association similarly to correlation coefficient (-1 = perfect spatial dispersion to +1 perfect spatial correlation). (Bailey and Gatrell, 1995, Cliff and Ord, 1981). In a study among the ONS neighbourhoods, Parenteau and Sawada (2011) assessed the bivariate Moran's I correlation between respiratory health outcomes and census variables. They found spatial correlations between these variables. Global Moran's I analysis could have also been performed on the dependent variable, ACS hospitalisation rates, to assess its spatial autocorrelation (Bailey and Gatrell, 1995, Cliff and Ord, 1981). In the study of Parenteau and Sawada (2011), no spatial autocorrelation was detected for the respiratory health outcome – suggesting that "natural neighbourhood are internally homogeneous in terms of SES and adequately depict the spatial scale of the respiratory health outcome rate, thus reducing spatial dependence between neighborhoods" (p.11). The same could be expected for the ACS hospitalisation rates.

Appendix 14 shows the results of the univariate regression models between the variables of interest and the dependant variable among the 90 neighbourhoods. These regression results demonstrated significant association, at the alpha level of 0.05, between ACS hospitalisation rates and the following variables (individually):

- SES index,
- 'Percent who rated their health status as poor or fair',
- 'Percent of smokers',
- 'Percent of women who never had a Pap test',
- 'Percent without a family physician',
- 'Percent with unmet health care need',
- 'Primary care physicians per 1,000 people',
- 'Percent with 1 or more ACS chronic conditions',
- 'Mean number visits to a general practitioner in previous year', and
- 'Mean number of visits to a nurse in previous year'.

The table in Appendix 14 also includes the bivariate regression results including the SES index variable in the model. In these models, the statistical significance flipped to non-significant for the variables mentioned above, with the exception of 'percent of smokers' and 'percent with poor/fair self-rated health status'. On the other hand, the significance of two variables flipped from non-significant to significant association with the rates of ACS hospitalisation: 1) distance to primary care and 2) 'percent of women who never had a breast examination'. The significant results of the univariate and bivariate regression models were in the expected direction based on the causal model, with the exception of variable representing physician supply, the mean variables of health care utilisation and the 'percent of women who never had a breast examination'. These simple regressions demonstrated the patterns of association between the variables of interest and the dependent variable, contributing to the investigator's understanding of the data and to the process of variable selection for the final models.

Also, among the variables of interest, pairs were derived from the same CCHS items. These pairs of variables were small variants of the same construct. Selecting one variable from each pair would be recommended for the exploratory model, considering that the pairs contain redundant information. The exploratory analysis helped complement theoretical knowledge in the task of informing the choice of which variable to retain. Below is outlined the investigator's reasoning for selecting the most informative variable among these pairs:

- The variables of 'percent of people binge drinking in past year' and 'percent of people binge drinking in past month' both represented the same concept and were non-significant in the univariate/bivariate regressions. The variable 'percent of people binge drinking in past year' was discarded considering its unusual negative and significant correlation with socio-economic status ($p=0.02$). The retained variable ('percent of people binge drinking in past month') was also negatively correlated with SES, yet this association is not significant ($p=0.26$).
- Regarding the variables representing the percent of neighbourhood residents diagnosed with ACS-related chronic diseases, the variable 'percent with 2 or more ACS conditions' was selected over the variable 'percent with 1 or more ACS conditions'. The former is more commonly used in the literature to represent disease burden in a population. This variable also had a greater coefficient of variation, which tend to be an advantage in multivariate linear regression (Weisberg, 2005).
- For the variables representing the utilisation of health services (family physician, nurses or specialists), two types of variables were derived: the proportions represented the percent of people who did not visit or talk with the health professional of interest in the past year and the means represented the mean number of contacts with the health professional of interest in the past year. Although the mean variables had higher coefficients of variation, greater correlations with the dependent variable and were significant in the univariate

regressions, the proportion variables were selected for the three types of health professionals. This decision was directed by the theoretical interpretations⁸⁹ of these variables as well as concerns⁹⁰ on the appropriateness of the means as valid summary measures in the context of group estimation and health care utilisation (Goddard & Smith, 2001).

3.7.2. Assessment of Model Validity Using a Full Exploratory Regression Model

Prior to the selection of the final regression models, the applicability of the analysis method selected as well as its practical limitations should be carefully examined by the investigator using exploratory models. For the current investigation, the 18 variables which were included in the exploratory model are identified in the table of Appendix 4. Tabachnick and Fidell (2007) suggested verifying the most important practical aspects of the exploratory regression model following the verification process outlined below:

- "Evaluation of the absence of multivariate outliers among the independent variables,
- Evaluation of the absence of multicollinearity,
- Evaluation of the normality, linearity, homoscedasticity and independence of the residuals,
- Evaluation of the potential presence of outliers in the solution, and

⁸⁹ As demonstrated in Andersen's theoretical frameworks (1968; 1995), the utilisation of health services is determined by predisposing, enabling and need factors. The mean number of contacts with health professionals is a measure of health care utilisation which is directly associated with health care need (Sanmartin & Khan, 2011). On the other side, the variables of 'percent who did not visit or talk with a health professional in past year' are disassociated from the notion of health care need: the proportions excluded those who sought health care, either for preventative services such as annual exam or for a health care need. Therefore, this later variable benefited from greater construct validity in representing the concept of propensity to seek health services.

⁹⁰ While generating the mean estimates, it was observed that the mean variables are influenced by extreme individual responses in certain neighbourhoods. In fact, there were 7 neighbourhoods with a respondent who mentioned contacting a primary care physician more than 52 times per year (more than once a week). And, there were 16 neighbourhoods with a respondent who mentioned contacting a nurse more than 52 times per year (up to 365 times/yr for some neighbourhoods). These extreme values alone enlarged considerably these neighbourhood estimates, especially for the neighbourhoods with small CCHS sample size. This may have introduced measurement error in the values derived. Also, it could be simply due to the luck of the drawn that a respondent with an extreme number of annual visits be within the sample of a neighbourhood as opposed to another.

- Evaluation of the ratio of cases to independent variables" (Tabachnick & Fidell, 2007, p. 117-122)

No statistical inference was performed on this exploratory model; therefore there was no need to present the model results in this document (available upon request).

3.7.2.1. EVALUATION OF THE ABSENCE OF MULTIVARIATE OUTLIERS AMONG THE INDEPENDENT VARIABLES

Multivariate outliers were determined based on Mahalanobis distances. Three neighbourhoods were considered as multivariate outliers at the alpha level of 0.001: Orleans Central (#67), Navan - Vars (#89) and Crystal Bay - Lakeview Park (#93). The chi-square statistics of the Mahalanobis distances for these neighbourhoods were 60.9, 44.8 and 43.3 respectively, where the reference chi-square value was 42.3. Based on these distances, Orleans Central was considered as an extreme multivariate outlier and was nominated for exclusion as it met the first criteria established in Section 3.6.2.2. Its outlying characteristics can be explained by its estimate on the variable representing physician supply: 14.45 primary care physicians per 1,000, which is 4 times greater than the second largest value for this variable in the dataset. Further investigations demonstrated that this value is not a data entry error.⁹¹ The other two neighbourhoods were not nominated for exclusion as the chi-square values of their Mahalanobis distances were not as high as for Orleans Central, and were not considered as extreme multivariate outliers.

3.7.2.2. EVALUATION OF THE ABSENCE OF MULTICOLLINEARITY IN EXPLORATORY MODEL

Based on the VIF values, the exploratory model did not show signs of multicollinearity. In fact, the VIF values of the variables included in this model were

⁹¹ It was suspected that the extreme value of 14.45 primary care physicians per 1,000 for the neighbourhood of Orleans Central may be a data entry error. Considering the small population of this neighbourhood (3,461 people), 50 primary care physicians would yield the observed supply. The accuracy of this number of primary care physicians was verified by phoning some medical clinics located in this neighbourhood. At the Orleans Urgent Care Clinic, there were about 20 registered primary care physicians. At Orleans Medical Centre, there were 8 physicians and a number of other private offices were also located in the same building. At St. Joseph Family Medicine Clinic, there were 10 primary care physicians. And, in Centrum Medical Centre, there were about 7 primary care physicians. In all, the value of 14.45 physicians per 1,000 is plausible, yet highly unusual for the neighbourhoods of Ottawa.

all below 10 (the SES index variable was associated with the highest value of 3.84). However, two condition index values were well above the reference value of 30 suggested by Lattin et al. (2003, p.57). These CI values were 54.12 and 94.48, suggesting the presence of moderate to important collinearity among the independent variables. The near collinear variables can be identified based on their high variance proportion values on the components with CI values above 30. For the component with the CI of 54.12, the variables 'percent who did not visit a medical specialist in the last year', 'percent who never had a flu shot' and 'percent who did not visit a nurse in the last year' had elevated variance proportion values (0.83, 0.26 and 0.16 respectively). For the component with the CI of 94.48, the variables 'percent who did not visit a nurse in the last year' and the intercept had elevated variance proportion values (0.82 and 0.88 respectively).

The exploratory model was re-run without the 'percent who did not visit a medical specialist in the last year' and the 'percent who did not visit a nurse in the last year' and a CI value of 39.74 was generated, which was still above the reference value of 30. By exploring the variance proportion values, potential collinearity between the variables 'percent who never had the flu shot', 'percent who rated the availability of health care in the community as being fair or poor' and the intercept was observed (0.43, 0.15 and 0.98 respectively). The variable of 'percent who never had a flu shot' was removed from this subsequent model and the presence of multicollinearity was still demonstrated by a CI value of 31.09. By exploring the variance proportion values, potential collinearity between the variables 'percent who rated the availability of health care in the community as being fair or poor', 'percent who never had a mammography' and the intercept was observed (0.31, 0.18 and 0.99 respectively). By further excluding the variable 'percent who rated the availability of health care in the community as being fair or poor', the highest CI value generated was 25.95 and the highest proportion of variance was 0.18; suggesting mild multicollinearity among the remaining variables and limited impact on the regression results. In the process of selecting the final variables (Section 3.7.3), the observed multicollinearity between the four variables identified above was considered.

3.7.2.3. EVALUATION OF THE NORMALITY, LINEARITY, HOMOSCEDASTICITY AND INDEPENDENCE OF THE RESIDUALS

The normality of the residuals of the exploratory model was demonstrated by the bell-shaped exhibited by the histogram plotted for the residuals. Also, 95% of residuals were within ± 2 standard deviations in the scatter plot of studentized residuals and predicted values (Appendix 15). In fact, there were only four neighbourhoods with residuals between ± 2 to 4 standard deviations, which represented 4.4% of the cases. And, there was no material departure from the straight line in the Q-Q plot of residuals other than one case at both end of the distribution which may be indicative of the presence of outliers (Appendix 15) (Montgomery et al., 2006, p. 129).

The linear relationships between the predictor variables and the rates of ACS hospitalisation depicted in the scatter plots (Appendix 12) were also demonstrated in the partial regression plots of the exploratory model: there was no departure pattern for any of the variables included in the model. This further supported the decision to model the relationship between preventable hospitalisations and neighbourhood characteristics using linear multiple regression and there was no need to include non-linear terms in the final models.

The independence of the errors in the exploratory model was demonstrated by the results of the Durbin-Watson test, which was 1.763 with a non-significant first order autocorrelation. Similarly, the homoscedasticity of the residuals was demonstrated by the scatter plot of residuals and predicted values in Appendix 15, where a constant vertical spread was observed. This constant vertical spread was also depicted in the partial regression plots. Moreover, the White test was highly non-significant ($p=0.52$), indicating that the model was well-specified.

3.7.2.4. EVALUATION OF THE POTENTIAL PRESENCE OF OUTLIERS IN THE SOLUTION

As discussed in Section 3.6.2.2, the presence of outliers was assessed based on the Cook's Distances, the leverage values and the DFFITs values for each neighbourhood included in the exploratory model. Although each diagnostic method

identified a number of neighbourhoods as potential outliers, the following five were identified by all methods:

- ByWard Market (#95)
- Crystal Bay - Lakeview Park (#93)
- Fitzroy Harbour - West Carleton (#2)
- Orleans Chapel Hill South (#84)
- Orleans Central (#67) (already nominated for deletion – see Section 3.7.2.1)

Also, a number of neighbourhoods were identified – based on their DFBETAs values – as potentially having influence on one or a number of regression coefficients of the exploratory model. Among the neighbourhoods identified as potential outliers, ByWard Market (#95) may have influence on nine regression coefficients and Crystal Bay – Lakeview Park (#93) on eight regression coefficients. The neighbourhood of South Keys - Heron Gate - Greenboro West (#52), even if not documented as a potential outlier, was identified as potentially having potential influence on 13 regression coefficients.

As of now, Orleans Central met two of the three criteria for exclusion (Section 3.6.2.2). In addition to being considered as an extreme multivariate outlier, this neighbourhood had important influence on the variable of physician supply (DFBETA value for this coefficient was 1.3383, where the cuff-off reference value was 0.2108). Transformation on this variable did not reduce the impact of Orleans Central on the regression results, which satisfies the third criteria for exclusion. For these reasons, it was decided to exclude Orleans Central from all final analyses of the current investigation. No other neighbourhood was excluded from the analysis models considering that they did not meet all three criteria.⁹² Not excluding additional neighbourhood preserved the representativeness of the sample as well as the number of cases in the dataset. In all, the final models were based on a sample of 89 neighbourhoods.

⁹² The neighbourhood of ByWard Market, identified as having influence on the full final model, was not excluded considering that it met only one criteria for exclusion (potential influence on the results). No variable specific transformations would reduce its potential influence on the regression results.

3.7.2.5. EVALUATION OF THE RATIO OF CASES TO INDEPENDENT VARIABLES

The number of cases to variables in a multivariate regression model influences the statistical power of the model. Statistical power is the probability that the statistical test performed rejects the null hypothesis when this hypothesis is actually false (Forthofer et al., 2007, p.215). The greater the ratio of independent variables to cases, the greater is the statistical power of the tests performed. Therefore, regression model based on a ratio of cases to variables of 5:1 – like the current exploratory model – has limited statistical power compared to a model based on a ratio of 20:1. Statistical power is also affected by the effect size between the dependent variable and the independent variable as well as the presence of measurement error in the variables (Tabachnick & Fidell, 2007, p.117). The smaller the relationship is believed to be between the variables tested, the higher preponderance of cases to variables is required to attain the same level of power. Similarly, the greater the measurement error expected in the values, the higher the preponderance of cases relative to variables is required.

For the current investigation, selecting an acceptable ratio of variables to cases involved a trade-off between the desired statistical power for the models and the inclusion of theoretically relevant variables. The statistical power was limited by the 89 neighbourhoods available and the potential presence of measurement error among the predictor variables. Therefore, some non-significant variables in the exploratory model had to be excluded from the final models to increase the statistical power of the retained models. However, variables representing theoretically important aspects could not be excluded, at the risk of the models suffering from the omitted-variable bias⁹³ (DeMaris, 2004, p.98). Also, to achieve the second research objective, all four hypotheses formulated to explain the differences in preventable hospitalisations had to be represented by variables in the final models. But, the variables representing certain aspects of these hypotheses, for example the quality of health services for the hypothesis of differences in health care

⁹³ Such bias occurs when an important causal factor is not included in the model and the model compensates for the missing factor by over- or under- estimating one of the other predictor variables. (DeMaris, 2004, p.98)

utilisation, could be excluded from the models without jeopardising their ability to fulfill the research objective.

In all, the selection of variables involved a trade-off between the objectives of the study, the desired statistical power of the model and the inclusion of variables representing the aspects of neighbourhoods believed to be important causal agent of the rates of preventable hospitalisation. From the predictor variables included in the exploratory model, it was decided to keep a maximum of 9 variables in the final full model; yielding a ratio of variables to cases of 10:1⁹⁴. It was believed that this ratio would afford the model a reasonable power to capture relationships with the dependant variable among the neighbourhoods of Ottawa. Also, 9 variables were believed to be sufficient to represent the four hypotheses as well as all the important causal factors in the model, enabling the achievement of the second and third objectives of the current investigation.

3.7.3. Selection of the Variables for the Final Regression Models

A total of 8 variables were selected for the full regression models. Among these variables, one represented the hypothesis of differences in health status, one represented the hypothesis of differences in healthy behaviour, three represented the hypothesis of differences in primary care access and two represented the hypothesis of differences in primary care utilisation (Table 7). The variables included in the sequential regression models are presented in Table 8. All the models included a variable representing neighbourhood socio-economic status, which was identified as an important theoretical concept. Two variables represented contextual aspects of neighbourhoods in the final full model: distance to primary care facilities and primary care physician supply.

⁹⁴ Numerous rules of thumb have been proposed to determine the number of independent variables to the number of cases in linear regression. Tabachnick and Fidell (1989) suggested a bare requirement of 5 cases for each independent variable included in the model, where a ratio of 20:1 would be ideal. Schmidt (1971) recommended ratios ranging from 15:1 or 25:1. Other, as Green (1991), suggested equations as $n > \text{or} = 50 + 8m$ (where m is the number of independent variables and n the number of cases) to test R^2 and $N > \text{or} = 104 + m$ to test the importance of individual predictor.

Table 9 details the reasons why the other 10 predictor variables included in the exploratory model were discarded from the final analyses. Although most variables excluded demonstrated lower explanatory power in the exploration analysis phase, only two variables were excluded solely on this basis: ‘percent with 2 or more ACS-related chronic conditions’ and ‘percent engaging in binge drinking in past month’.

The variables with imputed estimates were excluded from the final models due to their reduced variance⁹⁵ and low reliability⁹⁶. The ‘percent of women who never received a Pap test’ was the only imputed variable not excluded from the final models. This decision was based on a theoretical consideration: the concept of neighbourhood preventative health services utilisation should be represented in the final model. Considering its greater explanatory power and least number of imputations, this variable was selected over the other two variables also representing this concept (‘percent who never had a breast examination’ and ‘percent who never had a mammography’). Unfortunately, just as the other alternative variables, the ‘percent who never had a Pap test’ was based on a preventive service performed on women only, which may compromise its construct validity at the neighbourhood level.

The variables ‘self-perceived availability of care in community’ and ‘unmet health care need’ were excluded due to concerns about their construct validity: the CCHS questions on which they were based referred to the whole health care system as oppose to the primary care system only. In the current study, it is access to primary care services which is under investigation, not access to the whole health care system.

And, finally, four variables were excluded based on their multicollinearity in the exploratory model.

⁹⁵ Imputations by sub-group means reduce the variance of the variable (Dorofeev & Grant, 2006, p.132).

⁹⁶ The neighbourhood estimates of these variables were derived from smaller samples size – see Appendix 4.

Table 7: Variables Selected for Final Regression Model (Standard Regression)

Concept	#	Variable Name	Variable Label
<i>Preventable Hospitalisations</i>	1	ACS Hospitalisation Rates per 100,000 person-yr (<i>dependent variable</i>)	Hosp_Rate
<i>SES</i>	2	SES Index	SESindexcont
<i>Health Status</i>	4	Percent with fair or poor self-rated health status	health_status4_5_ag_st
<i>Healthy Behaviour</i>	7	Percent of people smoking	smoker_ag_st
<i>Primary Care Utilisation</i>	11	Percent who did not visit or talk with a primary care physician in past year	didnt_visit_gp_ag_st
	17	Percent who never received a Pap test	pap_input
<i>Primary Care Access</i>	22	Average distance to closest 4 primary care facilities (meter)	dist_4_physician
	23	Primary care physician-to-population ratio (per 1,000)	physician_1000
	24	Percent without a family physician	fam_doc_ag_st

Table 8: Variables Selected for Sequential Regression Models

Hypothesis Testing Models Y = ACS Hospitalisation Rates (per 100,000 person-yr) (hosp_rate)				
	<i>Hypothesis of Difference in Health Status</i>	<i>Hypothesis of Difference in Healthy Behaviour</i>	<i>Hypothesis in Difference in Primary Care Access</i>	<i>Hypothesis of Difference in Primary Care Utilisation</i>
Simple Model	SES Index (SESindexcont)	SES Index (SESindexcont)	SES Index (SESindexcont)	SES Index (SESindexcont)
Successive Model*	Percent with fair or poor self-rated health status (health_status4_5_ag_st)	Percent of people smoking (smoker_ag_st)	Average distance to closest four primary care facilities in meter (dist_4_physician) Primary care physician-to-population ratio – per 1,000 (physician_1000) Percent without a family physician (fam_doc_ag_st)	Percent who did not visit or talk with a primary care physician in past year (didnt_visit_gp_ag_st) Percent of women who never had a Pap test (pap_input)

* All variables in the successive model were entered as a block (group). The variable labels are in parenthesis.

Table 9: Reasons for the Exclusion of Variables of Interest in Final Models

Variables	Reasons for Exclusions			
	Construct Validity Concerns	Lack of Explanatory Power	Involved in Multicollinearity	Neigh. Level Imputation
% with 2 or more ACS conditions		X		
% of people binge drinking in past month		X		
% who rated the availability of health care in the community as being fair or poor	X	X	X	X
% who did not visit or talk with a nurse		X	X	
% who did not visit or talk with a medical specialist		X	X	
% who never received a flu shot		X	X	
% who never received a breast examination	X			X
% who never received a mammography	X	X		X
% with unmet health care need in past yr	X			X
% who rated the quality of the care provided by physician as being fair or poor		X		X

Chapter 4. Results

Below are the results of the different analyses carried for the current investigation, presented by research objectives. All the results from the final models and sensitivity analyses are presented in tables and figures, but only statistically significant findings are highlighted in the text. Every model was assessed for its validity as per methods described in Section 3.6.2.2. Any possible threat to the model validity is reported along with a discussion of its potential impact on the results. Also, any unusual statistical findings are highlighted in this section. All interpretations of the results are presented in the discussion (Chapter 5).

4.1. Results for Research Objective 1

To illustrate the gap in ACS hospitalisation rates among low and high SES neighbourhoods in Ottawa and to identify its importance as well as its characteristics:

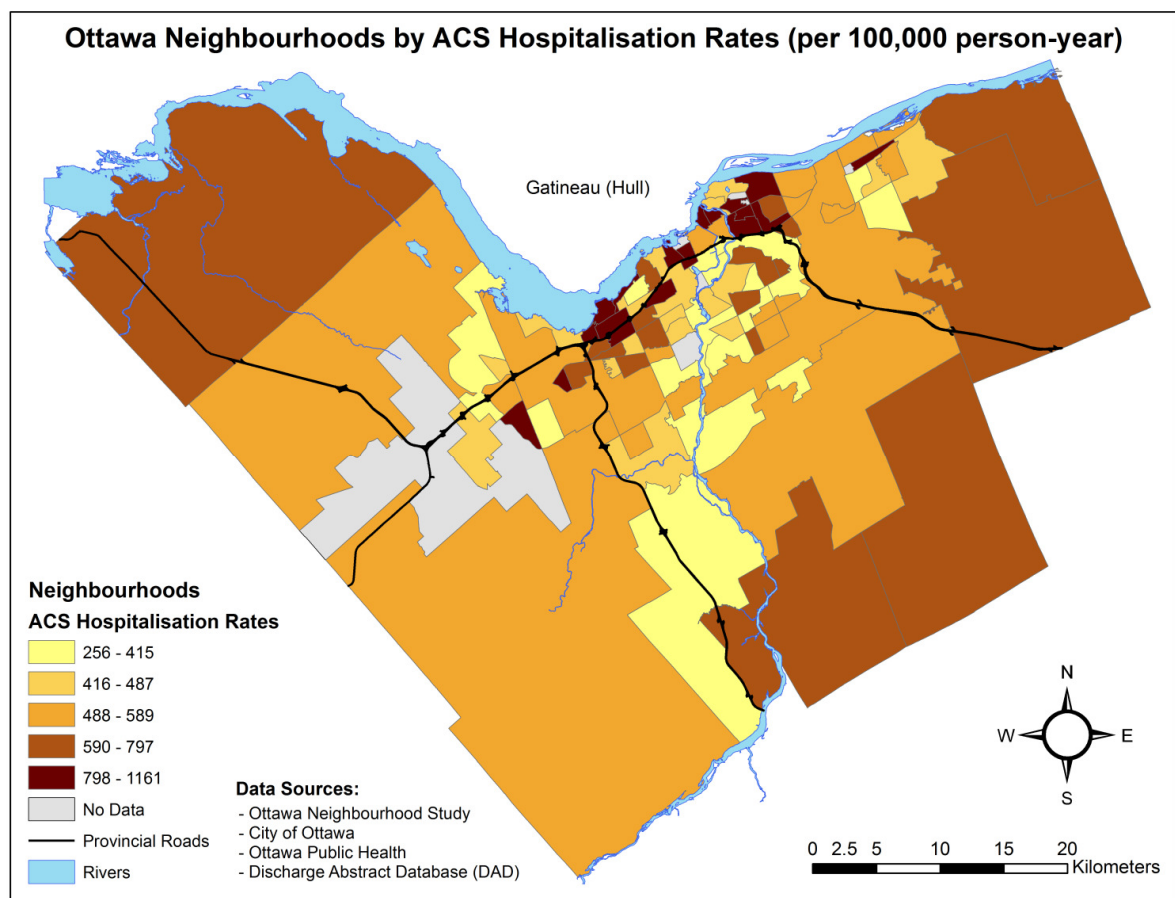
The maps illustrating the geographic distribution of Ottawa neighbourhoods by ACS hospitalisation rates and by deprivation quintiles are in Figure 12 and in Figure 13 respectively. By observing these maps, one can note that the inner city neighbourhoods (located between the Ottawa River and the 417 highway) tended to be in the higher quintiles of ACS hospitalisation rates as well as in the most deprived quintiles. Although the neighbourhoods located across the highway 417 were generally of lower SES status, they experienced moderate rates of ACS hospitalisations. As for the rural neighbourhoods located outside the Greenbelt, most were among the least deprived neighbourhoods yet they seem to experience rates of ACS hospitalisation amongst the higher quintiles.⁹⁷

Table 10 displays the descriptive statistics of ACS hospitalisation rates by neighbourhood deprivation quintile (roughly 18 neighbourhoods per quintile). Among the most affluent neighbourhoods, the mean of ACS hospitalisation rates was 458 per 100,000 person-year while the mean was 875 per 100,000 among the most deprived neighbourhoods. The box plot in Figure 14 shows that the rates of ACS

⁹⁷ For a representation of the location of the features of Ottawa mentioned here, refer to the map in Figure 7.

hospitalisation appear to increase as neighbourhood deprivation increased and an exponential-like trend may be observed as highlighted by the quintiles' medians. Also, the interquartile range (IQR)⁹⁸ increased as the level of socio-economic deprivation increased, which demonstrates that greater variability in the rates of ACS hospitalisation is observed among the most deprived neighbourhoods. Four neighbourhoods, as part of their respective deprivation level, were identified as having outlying rates of ACS hospitalisation in the box plot of Figure 14.⁹⁹

Figure 12: Ottawa Neighbourhoods by Quintiles of ACS Hospitalisation Rates



⁹⁸ The IQR is the distance between the 25th and the 75th percentile cases; it is a measure of statistical dispersion. It is represented by the rectangular boxes in the box plot.

⁹⁹ Their neighbourhood rate was larger or lower than 1.5 times their deprivation quintile's IQR, to be further discussed in Section 5.1.

Figure 13: Ottawa Neighbourhoods by Deprivation Quintiles

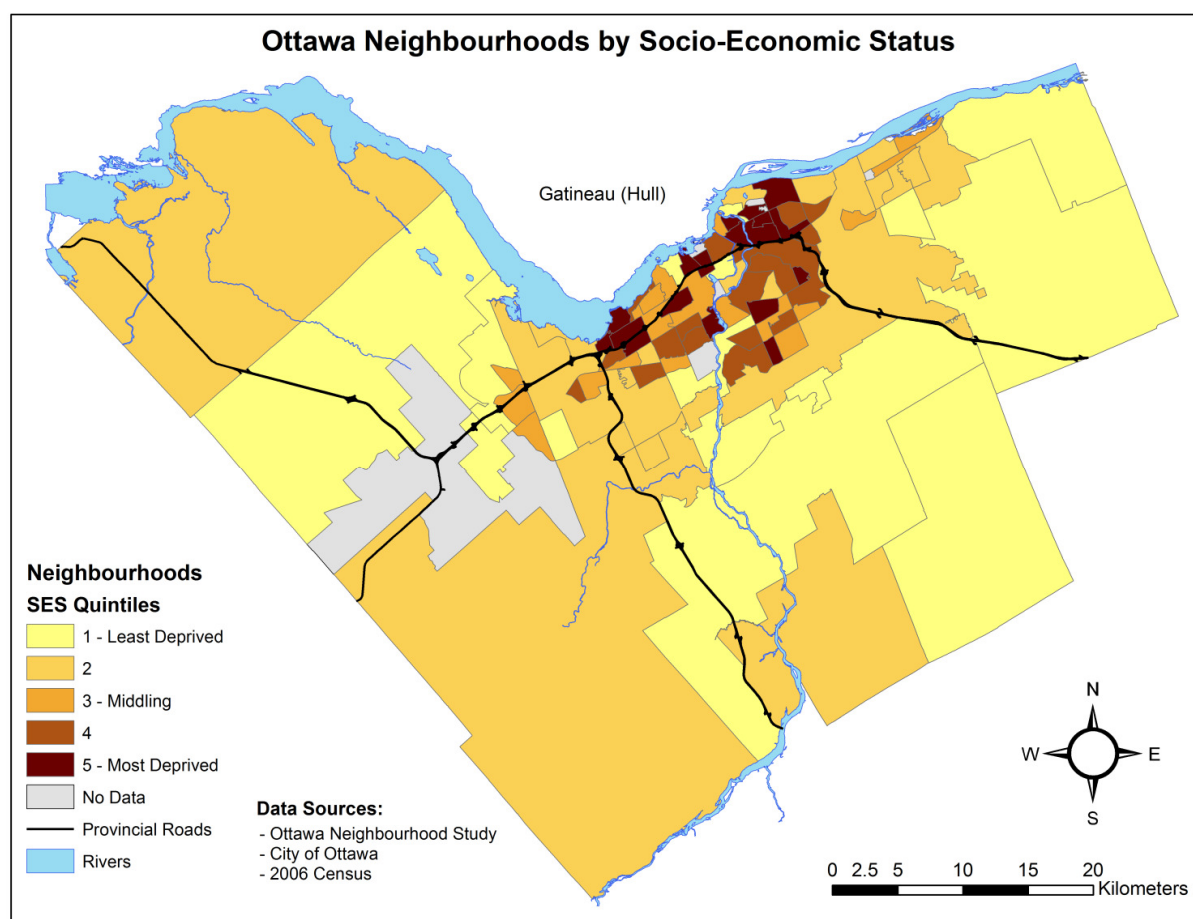
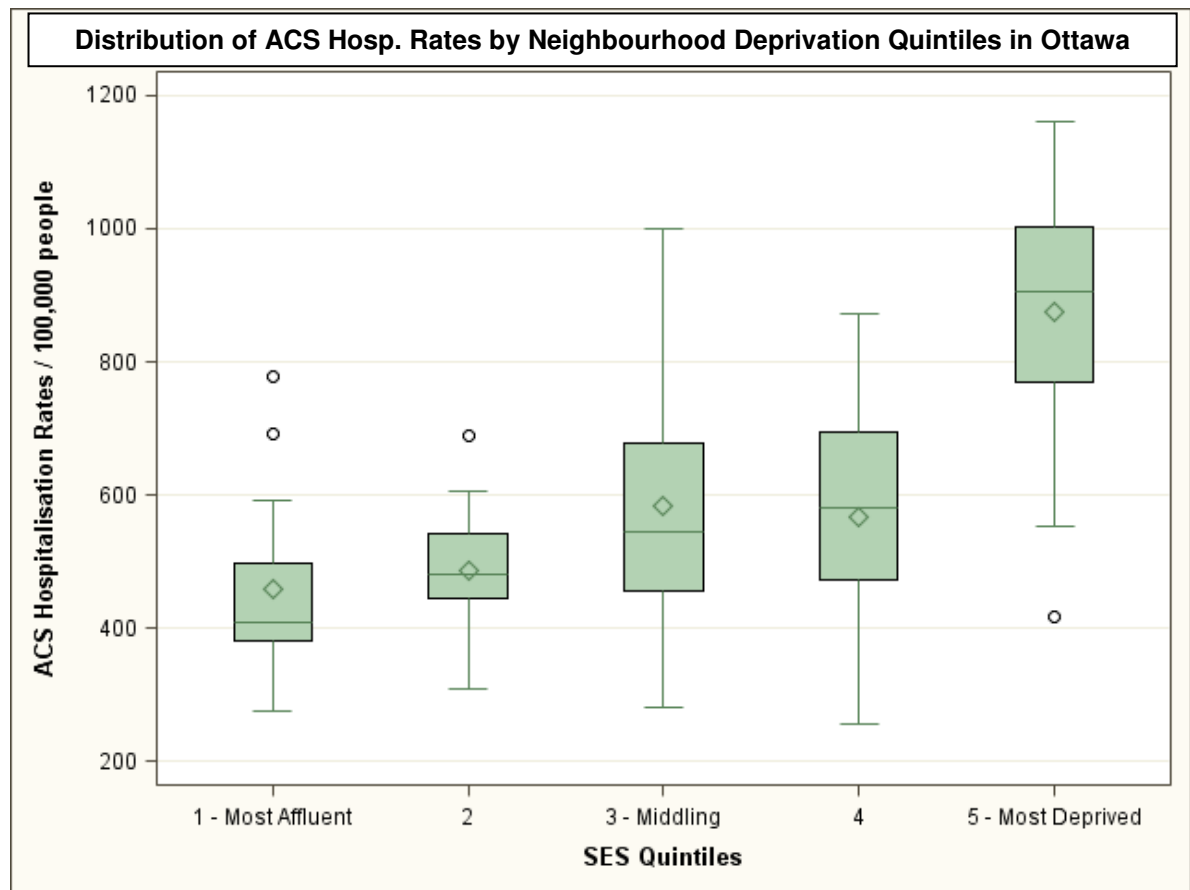


Table 10: Descriptive Statistics of ACS Hospitalisation Rates by SES Quintiles

Neighbourhood Deprivation Quintiles (Rates per 100,000 person-yr)	Q1 - Most Affluent	Q2	Q3 - Middling	Q4	Q5 - Most Deprived	All Neigh.
N	17	18	18	18	18	89
Min	274.04	308.28	279.75	256.10	415.82	256.10
P5	274.04	308.28	279.75	256.10	415.82	334.48
Q1	380.58	444.04	456.11	473.10	768.64	450.97
Median	408.11	481.12	544.74	581.70	906.25	541.36
Q3	497.54	541.36	678.19	693.74	1002.95	712.99
P95	778.81	687.93	999.43	872.98	1161.56	1002.95
Max	778.81	687.93	999.43	872.98	1161.56	1161.56
Range	905.46	504.77	379.65	719.68	616.89	745.73
Mean	458.28	487.15	582.72	567.22	875.12	595.62
CV	28.17	18.39	31.70	29.57	22.89	36.24
Standard Deviation	215.84	129.09	89.61	184.74	167.70	200.29

Figure 14: Distribution of ACS Hospitalisation Rates by Neighbourhood Deprivation Quintiles in Ottawa



The results of the analyses of variance are in Table 11. The minimum of 15 cases per group was met and the balance of the design was acceptable, where four quintiles were comprised of 18 neighbourhoods and one quintile was comprised of 17 neighbourhoods. For the model based the untransformed values of ACS hospitalisation rates, the Levene statistic was 2.184 ($p=0.078$) indicating that the variance of the variable was heteroscedastic across the neighbourhood quintiles. To meet the model assumptions, the square root and natural log transformations were performed on the ACS hospitalisation rates, yielding valid ANOVA models with Levene statistics of 1.383 ($p=0.247$) and 1.12 ($p=0.335$) respectively. These later models had a statistically significant F test ($p=0.000$ in both models), thus establishing that the means of ACS hospitalisation rates across the neighbourhood deprivation quintiles did significantly differ.

Based on the Tukey's HSD comparison tests, the mean of ACS hospitalisation rates of the most deprived neighbourhood quintile was significantly higher than the means of ACS hospitalisation rates for the deprivation quintiles 1, 2, 3 and 4 ($p=0.000$ for all comparisons), testifying to the presence of a gap in ACS hospitalisation rates between the neighbourhoods of low and high SES level. On the other side, the differences in the means of ACS hospitalisation rates between the quintiles 1, 2, 3 or 4 were not significant. The comparisons based on the transformed and untransformed values of ACS hospitalisation rates yielded the same results.

According to the linear trend tests, the relationship between the means of ACS hospitalisation rates and SES quintiles illustrated in the box plot was increasing, yet not linear: the linear terms were significant, but the deviation terms were also significant in the different models of untransformed and transformed values. This means that the gaps in ACS hospitalisation rates by SES quintile may increase more rapidly between the 4th and 5th quintiles than between the other quintiles.¹⁰⁰ However, a monotonic increasing relationship – even if not linear – was demonstrated by the test performed on the ranks of ACS hospitalisation rates (the linear term was significant and the deviation term was not significant).

¹⁰⁰ Assuming that the distance between the neighbourhood SES quintiles was equal. Yet, this is not guaranteed considering that the quintiles were created based on the distribution of the SES index.

Table 11: ANOVA Results – ACS Hospitalisation Rates by SES Quintiles

Variable	Neighbourhood SES Quintiles	N	Mean	Variance	Min	Max	Levene Statistic (p value) ¥	ANOVA F Test (p value)	Linearity Test €
ACS Hospitalisation Rates (untransformed)	1 - Most Affluent	17	458.3	16664	274.0	778.8	2.18 (0.078)	19.17 (0.000)	Linear Term = 57.55 (p= 0.000) Deviation Term=6.14 (p=0.001)
	2	18	487.1	8029	308.3	687.9			
	3 – Middling	18	582.7	34129	279.8	999.4			
	4	18	567.2	28124	256.1	873.0			
	5 - Most Deprived	18	875.1*	40114	415.8	1161.6			
Square Root of ACS Hospitalisation Rates	1 - Most Affluent	17	21.2	8.28	16.6	27.9	1.38 (0.247)	17.33 (0.000)	Linear Term = 53.50 (p= 0.000) Deviation Term=5.10 (p=0.003)
	2	18	22.0	4.15	17.6	26.2			
	3 – Middling	18	23.9	13.86	16.7	31.6			
	4	18	23.6	13.04	16.0	29.5			
	5 - Most Deprived	18	29.4*	13.03	20.4	34.1			
Natural Log of ACS Hospitalisation Rates	1 - Most Affluent	17	6.1	0.07	5.6	6.7	1.12 (0.335)	15.9 (0.000)	Linear Term = 47.75 (p=0.000) Deviation Term=4.06 (p=0.010)
	2	18	6.2	0.04	5.7	6.5			
	3 – Middling	18	6.3	0.10	5.6	6.9			
	4	18	6.3	0.10	5.5	6.8			
	5 - Most Deprived	18	6.7*	0.07	6.0	7.1			
Rank of ACS Hospitalisation Rates	1 - Most Affluent	17	26.9	450.06	2	71	1.462 (0.221)	13.511 (0.000)	Linear Term = 46.36 (p= 0.000) Deviation Term=2.56 (p= 0.060)
	2	18	33.0	285.65	4	63			
	3 – Middling	18	45.6	545.67	3	85			
	4	18	45.1	565.58	1	77			
	5 - Most Deprived	18	74.1*	313.11	18	90			

*: Pairwise comparisons with ACS hospitalisation means for quintile 1, 2, 3 and 4 significant at $\alpha=0.05$ based on Tukey's HSD tests

¥: A highly not significant Levene statistic support the model assumption of homoscedasticity (constant variance across groups)

€: A significant linear term ($p=0.05$) and an insignificant deviation term ($p=0.05$) means that the relationship between the ACS hospitalisation rates and the neighbourhood deprivation quintiles was linear (Garson, 2009).

4.2. Results for Research Objective 2

To test the four different hypotheses formulated in the literature to explain differences in ACS hospitalisation rates between neighbourhoods of Ottawa:

Table 12 displays the results of the sequential regression models to test the four hypotheses pertaining to differences in health status, healthy behaviour as well as primary care access and utilisation. Table 13 shows the results of the sensitivity analyses, where weights based on neighbourhood population size were used to address the potential impact of measurement error on the models' results (weighted equally). In the tables, significant results at $\alpha = 0.05$ are presented in bold and the results significant at $\alpha = 0.1$ are underlined.

From both the final model as well as the sensitivity analysis model, the variable representing the hypotheses of differences in neighbourhood health status and differences in healthy behaviour have significant incremental F ratio tests. The p values of the incremental F ratio test for the variable of 'percent with poor/fair health status' were 0.04 (F statistic = 4.57) in the final model and 0.01 in the population weighted model (F statistic = 6.65). The p values of the incremental F ratio test for the variable of 'percent smoking' were less than 0.0001 (F statistic = 19.69) in the final model and equal to 0.0001 in the population weighted model (F statistic = 15.89). Thus, differences in health status and healthy behaviour at the neighbourhood level were significantly related to the rates of preventable hospitalisation, over and above the impact of neighbourhood socio-economic status. Based on the R^2 of these two final models¹⁰¹, the differences in healthy behaviour explained a greater proportion of the variance in ACS hospitalisation rates than the differences in health status ($R^2 = 0.55$ and 0.48 respectively).

On the other hand, the addition of the three variables representing the hypothesis of differences in primary care access was not significant ($p=0.16$ and $p=0.25$ for the equally and population weighted models respectively).

¹⁰¹ The R^2 of these models can be compared since they are comprised of two variables each; the R^2 values did not have to be adjusted for differences in the number of variables between the models.

For the sequential models testing the importance of differences in utilisation of primary care on the rates of preventable hospitalisation, the results of the final model weighted equally and the model weighted by population size were different; highlighting a problem with the robustness of the models. In the final model, the incremental F ratio test assessing the importance of the variables representing primary care utilisation was not significant, where $p = 0.15$. Yet, in the model weighted by population size, this statistical test was significant at the 0.05 alpha level. The significance in this later model may be associated with a reduction in the potential effect of measurement error on the model's residuals. In this model, any impact of measurement error caused by the variability of the estimates derived for small neighbourhoods on the heteroscedasticity of the residuals would be reduced; therefore yielding more significant results. However, based on the three diagnostic methods used, no sign of heteroscedasticity was observed in this final sequential model (weighted equally).

Alternatively, the inconsistent results between the models representing the hypothesis of primary care utilisation may be a sign of spurious relationships between the variables representing this concept and the rates of ACS hospitalisation. It appeared that the significance of the incremental F ratio test in the sensitivity analysis model was driven by the variable 'percent of women who never had a Pap test'. In the sequential model, this variable was significantly and negatively associated with the rates of ACS hospitalisation, which was opposite to what was expected from theory as well as from the descriptive exploration of variables (see Appendix 10 and Appendix 14). This relationship may be spurious and will be further discussed in Section 5.3.2.

All sequential models were assessed for multicollinearity, normality, linearity, homoscedasticity and independence of the residuals as well as the potential presence of outliers in the solutions. None of the models showed signs of multicollinearity as the condition index values were all below 10 and the VIF values were all below 2. Also, no departure from linearity was observed and the DW tests demonstrating independence of the errors were also met for all sequential models.

The normality of the models was also demonstrated by the scatter plots of studentized residuals and predicted values, where between 4 to 6 residuals were outside ± 2 standard deviations (representing 4.5 – 6.7% of the residuals), which was acceptable. The normality of the residuals for the different sequential models was also supported by the alignment of the residuals in the Q-Q plots and the bell-shape of the residuals' histograms.

Based on the scatter plots of studentized residuals and predicted values, none of the models suffered from heteroscedasticity (the spread of the residuals was similar to the one generated by the exploratory model in Appendix 15). This was further supported by the partial regression plots and the non-significant results of the White tests¹⁰².

Diagnostics for outliers found one potential outlying neighbourhood (ByWard Market (ID #95)) in the four final models weighted equally. However, subsequent analyses (re-running the models with this neighbourhood dropped) revealed only minor sensitivity to its exclusion: significance (or none-significance) at the alpha level 0.05 for the incremental F ratio tests of all four models remained the same.¹⁰³

¹⁰² With the exception of the final model testing the hypothesis of differences in primary care access where the White test was significant ($p=0.04$). However, the White test was non-significant ($p=0.18$) for its respective population weighted model, which was designed to reduce the potential effect of measurement error on the spread of the residuals. As the incremental F ratio tests were non-significant for the sequential models weighted equally and weighted according to neighbourhood population size, the robustness of the final model testing the hypothesis of differences in primary care was demonstrated.

¹⁰³ The largest difference observed was for the model testing the differences in health status. The significance of the incremental F ratio test for this model increased from $p=0.04$ to $p=0.002$ when ByWard Market was excluded.

Table 12: Sequential Regression Models – Final Results (Weighted Equally)

Variable [‡]	Coefficient	Standardised Coefficient	Standard Error	t Test Statistic	P value	sr ² *	Incremental F ratio test (p value)
Hypothesis of Difference in Socio-Economic Status							
(Intercept)	593.9	0.00	17.1	34.8	<0.0001	0.45	
SES	144.2	0.67	17.1	8.4	<0.0001		
<i>R</i> ² = 0.45							
Hypothesis of Difference in Health Status							
(Intercept)	514.5	0.00	40.8	12.6	<0.0001	0.20	F = 4.57
SES	118.8	0.552	20.6	5.8	<0.0001		
% Poor/Fair Health Status	781.3	0.204	365.4	2.1	0.04	0.03	(p = 0.04)
<i>R</i> ² = 0.48		Adj. <i>R</i> ² = 0.46		(Shared variability = 0.25; Total unique variability = 0.23)*			
Hypothesis of Difference in Healthy Behaviour							
(Intercept)	398.5	0.00	46.7	8.5	<0.0001	0.15	F = 19.69
SES	100.4	0.466	18.4	5.5	<0.0001		
% Smoking	984.2	0.380	221.8	4.4	<0.0001	0.10	(p < 0.0001)
<i>R</i> ² = 0.55		Adj. <i>R</i> ² = 0.54		(Shared variability = 0.26; Total unique variability = 0.29)*			
Hypothesis in Difference in Primary Care Access							
(Intercept)	541.7	0.00	47.2	11.5	<0.0001	0.34	
SES	153.7	0.714	20.8	7.4	<0.0001		
PC Physician Supply	13.1	0.069	16.2	0.8	0.42	0.00	F = 1.79
Distance to PC	0.0162	0.199	0.0	2.3	0.02	0.03	
% Not Having Physician	64.2	0.020	303.4	0.2	0.83	0.00	(p = 0.16)
<i>R</i> ² = 0.48		Adj. <i>R</i> ² = 0.46		(Shared variability = 0.11; Total unique variability = 0.37)*			
Hypothesis of Difference in Primary Care Utilisation							
(Intercept)	669.3	0.00	63.8	10.5	<0.0001	0.41	
SES	166.2	0.772	20.5	8.1	<0.0001		
% Didn't Visit PC Physician	-74.3	-0.023	261.4	-0.3	0.78	0.00	F = 1.95
% Never Had Pap Test	-893.6	-0.183	453.6	-2.0	0.05	0.02	(p = 0.15)
<i>R</i> ² = 0.47		Adj. <i>R</i> ² = 0.46		(Shared variability = 0.04; Total unique variability = 0.43)*			

Table 13: Sequential Regression Models – Sensitivity Analyses Results (Weighted by Neighbourhood Population Size)

<i>Variable[¥]</i>	<i>Coefficient</i>	<i>Standardised Coefficient</i>	<i>Standard Error</i>	<i>t Test Statistic</i>	<i>P value</i>	<i>Incremental F ratio test (p value)</i>
<i>Hypothesis of Difference in Socio-Economic Status</i>						
(Intercept)	590.1	0.00	15.6	37.9	<0.0001	
SES	147.0	0.72	15.1	9.75	<0.0001	
<i>Hypothesis of Difference in Health Status</i>						
(Intercept)	492.9	0.00	40.6	12.1	<0.0001	
SES	115.9	0.570	19.0	6.1	<0.0001	F = 6.65
% Poor/Fair Health Status	957.4	0.240	371.4	2.6	0.01	(p = 0.01)
<i>Hypothesis of Difference in Healthy Behaviour</i>						
(Intercept)	414.2	0.00	46.4	8.9	<0.0001	
SES	108.0	0.531	17.0	6.3	<0.0001	F = 15.89
% Smoking	873.6	0.334	219.2	4.0	<0.0001	(p = 0.0001)
<i>Hypothesis in Difference in Primary Care Access</i>						
(Intercept)	556.4	0.00	44.4	12.5	<0.0001	
SES	157.4	0.774	19.1	8.2	<0.0001	
PC Physician Supply	5.3	0.028	15.9	0.3	0.74	F = 1.41
Distance to PC	0.0132	0.168	0.01	2.1	0.04	(p = 0.25)
% Not Having Physician	18.8	0.006	304.2	0.1	0.95	
<i>Hypothesis of Difference Primary Care Utilisation</i>						
(Intercept)	694.9	0.00	60.4	11.5	<0.0001	
SES	174.5	0.857	18.4	9.5	<0.0001	
% Didn't Visit PC Physician	-205.2	-0.057	273.7	-0.8	0.46	F = 3.14
% Never Had Pap Test	-980.9	-0.209	410.8	-2.4	0.02	(p = 0.05)

* Appendix 11 defines the concepts of sr^2 , shared and total variance. These values were not reported for the weighted models considering that they would have to be calculated manually; SAS provides erroneous estimates based on the weighted data as opposed to the original data.

¥ The variables representing the hypotheses tested were entered as a block in the successive model, where the simple model was comprised of the SES variable only.

4.3. Results for Research Objective 3

To identify the importance of neighbourhood compositional and contextual factors as well as to assess the relative contributions of neighbourhood health status, healthy behaviour and access/utilisation of primary care services to ACS hospitalisation rates between neighbourhoods in Ottawa:

Table 14 displays the results of the regression model representing a simplified version of the causal model of preventable hospitalisations. Table 15 shows the results from the sensitivity analysis for this regression model, where the model was weighted by neighbourhood population size.

In the full model where the neighbourhoods were weighted equally, 59% of the variance in ACS hospitalisation rates was explained by the model components. When adjusted for the number of variables included in the model, this proportion amounted to 55%, which appears to constitute an acceptable fit: Lattin et al. (2003, p.53) mentioned that, when dealing with social science data, typical R^2 values range between 0.1 and 0.5. The predictive capacity of the full model is represented in Figure 15, where the model's predicted values for ACS hospitalisation rates are plotted in comparison with the observed ACS hospitalisation rates. This graph demonstrated an acceptable fit of the regression model.

From these full models, the importance of the variables representing the different neighbourhood compositional and contextual factors believed to be associated with the rates of ACS hospitalisation were tested using t tests. These tests assessed the significance of the relationship between the variable tested and the rates of ACS hospitalisation over and above all the variance in the rates explained by the other variables included in the model. The results of the t tests from the final model weighted equally and from the sensitivity analysis model weighted by population size were somewhat different.

The SES index and 'percent smoking' were significant at the alpha level 0.05 in both models. However, at the alpha level 0.1, the variable of 'percent with poor or fair self-rated health status' was associated with the rates of ACS hospitalisation in the

model weighted by population size ($p=0.06$), but not in the model weighted equally ($p = 0.12$). And, the variable of distance to primary care physicians was significant at the alpha level 0.1 in the final model weighted equally ($p=0.07$), yet this association was no longer significant in the model weighted by population size ($p=0.17$)¹⁰⁴. In all, the sensitivity analysis demonstrated the robustness of the results at the alpha level 0.05, but not at the alpha level 0.1. This could be explained by the increased likelihood of rejecting the null hypothesis at this probability level, and the results of the statistical tests performed at this probability level are more subject to the sample characteristics (Forthofer et al., 2007, p.218).

All the variables significantly associated with the rates of ACS hospitalisation, at alpha 0.05 and 0.1, had a regression coefficient in the expected direction, adding further plausibility to the models. For example, as the proportion of neighbourhood residents smoking increased, the neighbourhood rates of ACS hospitalisation also increased. Similarly, the rates of ACS hospitalisation also increased when the deprivation level of the neighbourhood increased. When considering all urban and rural neighbourhoods of Ottawa, as the distance to primary care physicians increased, the rates of preventable hospitalisation also increased. And, for the weighted regression, as the proportion of neighbourhood residents who expressed having a poor or fair health status increased, the rates of ACS hospitalisation for these neighbourhoods also increased. The directions of these relationships were as expected by the relationships described in the causal model of preventable hospitalisations (Figure 6).

In the full models, the importance of the variable(s) representing the hypotheses believed to explain differences in preventable hospitalisations was tested statistically using incremental F ratio tests. This was done to test the relative contribution of the different hypotheses in a model acknowledging a greater complexity of the context in which preventable hospitalisations occur. Only one incremental F ratio tests was significant at the alpha level 0.05: namely the one for the variable of 'percent smoking' which represented the hypothesis of healthy behaviour ($p=0.01$ and

¹⁰⁴ The full model, weighted equally, was run among the urban neighbourhoods only. In this model, the variable of distance to primary care physicians was not significant ($p = 0.17$).

$p=0.005$ in the equally and population weighted models respectively). In the model weighted by population size, the incremental F ratio test for the hypothesis of difference in health status, represented by the variable of 'percent with poor or fair health status', was significant at the 0.1 alpha level ($p=0.06$).

As with the sequential models, the full models were assessed for multicollinearity, normality, linearity, homoscedasticity and independence of the residuals as well as the potential presence of outliers in the solutions. All model assumptions were met for both full models. The diagnostic tests of multicollinearity were all within acceptable range, where the condition index values were all below 20 and the VIF values below 4 (the acceptance reference values are 30 and 10 respectively). The normality of the residuals was also demonstrated in the scatter plot of studentized residuals by predicted values, where four residuals were outside the ± 2 standard deviations for both models (representing 4.5% of the neighbourhoods). Also, there was no material departure from the diagonal line in the Q-Q plots of the residuals. The independence of the residuals was demonstrated by DW values which were not different from 2 for both models (where $DW = 1.73$ in the model weighted equally and $DW = 1.82$ in the model weighted by population size). Linearity of the model was also demonstrated by the partial regression plots depicting the residuals by the predictor variables: the scatter of the points followed a linear shape. And, homoscedasticity of the residuals was demonstrated by the constant spread of the residuals on the scatter plot of studentized residuals by predicted values and by the White tests.

With regard to the potential presence of outliers in the solution, the neighbourhood of ByWard Market (ID #95) was identified in both full models. In the final model weighted equally, this neighbourhood had a studentized residual of 3.1^{105} , a Cook's Distance of 0.46 (four times greater than the second largest value) and a DFFIT value of 2.15 (above the unadjusted cut-off of 2 as well as the adjusted cut-off of 0.599). Also, based on the DFBETA values, this neighbourhood may have had influence on five regression coefficients.

¹⁰⁵ Cases with studentized residuals greater than ± 3 standard deviation could be outliers (Montgomery et al., 2006, p.123).

Subsequent analyses (re-running the model weighted equally with this neighbourhood dropped) revealed some sensitivity to its exclusion: two additional variables became significant at the alpha level of 0.05. These variables were the 'percent with poor or fair health status' as well as the distance to primary care facilities ($p = 0.03$ and $p = 0.04$ respectively). The incremental F ratio test for the hypothesis of differences in health status was also significant (F test = 5.1, $p = 0.03$). According to some authors, the influence of this neighbourhood on the results of the final model could have justified its exclusion (Tabachnick & Fidell, 2007, p. 119-122), yet it did not meet all three criteria for exclusion set in this study (see Sections 3.6.2.2 and 3.7.2.4). The assessments of the potential presence of outliers in the final models were performed to examine the impact of their inclusion on the results, but not to justify exclusion. The specificity of the neighbourhood of ByWard Market and its influence on the results will be presented in the discussion (Sections 5.2, 5.4.1.3 and 5.4.2.1).

Figure 15: Final Model – Predicted vs Observed Values of ACS Hospitalisation Rates

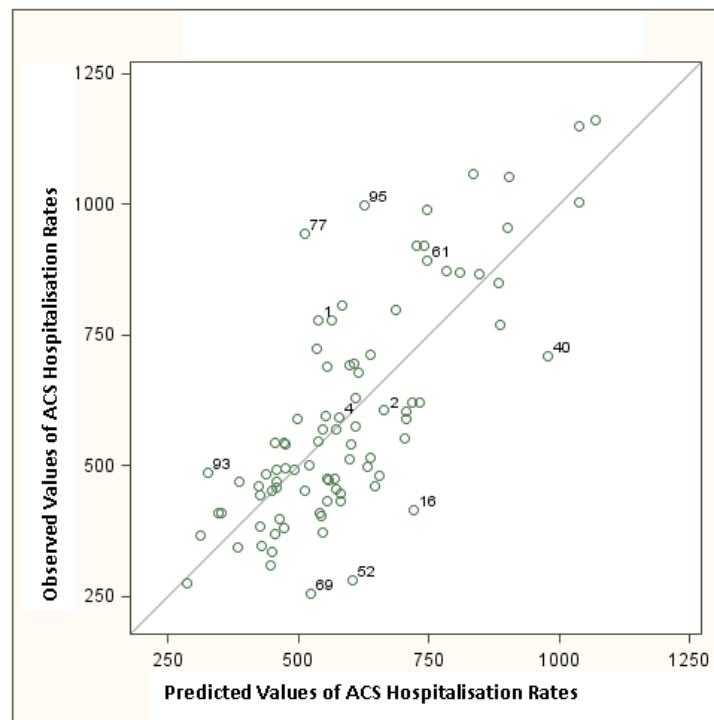


Table 14: Full Model Final Results (Weighted Equally)

<i>Variable</i>	<i>Coefficient</i>	<i>Standardised Coefficient</i>	<i>Standard Error</i>	<i>t Test Statistic</i>	<i>P value</i> [‡]	<i>sr</i> ^{2*}	<i>Partial F test (p value)</i>
<i>(Intercept)</i>	341.24		95.4	3.6	0.001		
SES	100.67	0.468	27.7	3.6	0.001	0.067	
% Poor/Fair Health Status	554.51	0.145	348.7	1.6	0.12	0.013	F = 2.53 (p = 0.12)
% Smoking	826.14	0.319	231.1	3.6	0.001	0.065	F = 12.78 (p = 0.001)
PC Physician Supply	10.03	0.053	15.1	0.7	0.51	0.002	
Distance to PC	0.0121	0.148	0.01	1.9	<u>0.07</u>	0.018	F = 1.43 (p = 0.24)
% Not Having PC Physician	239.65	0.073	281.0	0.9	0.40	0.004	
% Didn't Visit PC Physician	-4.24	-0.001	243.56	0.0	0.99	0.000	F = 0.58 (p = 0.56)
% Never Had Pap Test	-457.15	-0.094	427.0	-1.1	0.29	0.006	
<i>R</i> ² = 0.59 Adj. <i>R</i> ² = 0.55 (Shared variability = 0.42; Total unique variability = 0.18) [*]							

[‡] Significant results at alpha = 0.05 are presented in bold and the results significant at alpha = 0.1 are underlined.

^{*} Appendix 11 defines the concepts of *sr*², shared and total variance.

Table 15: Full Model – Sensitivity Analyses (Weighted by Neighbourhood Population Size)

<i>Variable</i>	<i>Coefficient</i>	<i>Standardised Coefficient</i>	<i>Standard Error</i>	<i>t Test Statistic</i>	<i>P value</i> [‡]	<i>sr</i> ² *	<i>Partial F test (p value)</i>
<i>(Intercept)</i>	405.24		92.1	4.4	<0.0001		
SES	110.81	0.545	26.8	4.1	<0.0001	0.044	
% Poor/Fair Health Status	705.39	0.177	361.7	2.0	<u>0.06</u>	0.011	F = 3.80 (p = 0.06)
% Smoking	662.66	0.253	229.2	2.9	0.005	0.056	F = 8.36 (p = 0.005)
PC Physician Supply	3.89	0.020	14.9	0.3	0.79	0.006	
Distance to PC	0.0083	0.106	0.01	1.4	0.17	0.011	F = 0.79 (p = 0.50)
% Not Having PC Physician	191.07	0.059	283.7	0.7	0.50	0.000	
% Didn't Visit PC Physician	-137.48	-0.038	263.5	-0.5	0.60	0.001	F = 1.15
% Never Had Pap Test	-580.28	-0.123	402.6	-1.4	0.15	0.007	(p = 0.32)
<i>R</i> ² = 0.58 <i>Adj. R</i> ² = 0.54 (<i>Shared variability = 0.45; Total unique variability = 0.14</i>) *							

[‡] Significant results at alpha = 0.05 are presented in bold and the results significant at alpha = 0.1 are underlined.

* Appendix 11 defines the concepts of *sr*², shared and total variance. These were calculated manually for the weighted model, considering that SAS provides erroneous estimates based on the weighted data as opposed to the original data.

Chapter 5. Discussion

The objectives of the current investigation were to contribute to the interpretation of the indicator of preventable hospitalisation in Ottawa – an area with a predominantly urban population and where universal access to health care is available. As expected, the results demonstrated a gap in ACS hospitalisation rates between neighbourhoods of low and high socio-economic level. This gap is further described below (Section 5.1). Then, the strengths and limitations of the regression analyses models are discussed in parallel with the issue of generalisability (Section 5.2). In light of this discussion, the interpretations of the results of the sequential models - testing the four hypotheses put forward to explain neighbourhood differences in ACS hospitalisation rates - are presented (Section 5.3). Subsequently, the results of the full model testing the associations between neighbourhood compositional and contextual aspects with preventable hospitalisations are interpreted (Section 5.4).

5.1. Neighbourhood Preventable Hospitalisations and Social Deprivation in Ottawa

The descriptive statistics, maps and ANOVA tests provide good avenues to answer the first research question of this investigation on the existence and size of a socio-economic gap in ACS hospitalisations in Ottawa.

As hypothesised, the current investigation demonstrated the presence of a significant gap between neighbourhoods of low and high SES status in ACS hospitalisation rates among the 20-79 years old people. In fact, the mean rate of ACS hospitalisation among the neighbourhoods of the most deprived quintile (5th quintile) was 1.9 and 1.8 times greater than the mean rates among those in the 1st and 2nd quintiles respectively (significant at $\alpha=0.05$). This gap is also observable with the neighbourhoods of the 3rd and 4th deprivation quintiles: the mean ACS hospitalisation rate of the most deprived neighbourhood quintile (5th quintile) was 1.5 times greater than the mean rates of these two neighbourhood quintiles (significant at $\alpha=0.05$). These findings on the magnitude of the SES gap in ACS hospitalisations rates in Ottawa were similar to the ones published in CIHI's report 'Reducing Gap in

Health: A Focus on Socio-Economic Status in Urban Canada' (2008a). For the region of Ottawa-Gatineau, the dissemination areas of low SES status experienced rates of ACS hospitalisation that were 1.9 times greater than the dissemination areas of high SES status. This demonstrates that, although some limitations associated with deriving the rates were at play in the current study (see Section 3.5.3), the magnitude of the socio-economic differences in ACS hospitalisation rates were comparable with other studies in the region.¹⁰⁶

One interpretation for the above gap in ACS hospitalisation rates in Ottawa is that SES level only affects the most deprived group of neighbourhoods, but that SES exerts no effect beyond this. Yet, the linear trend tests suggest that the impact of SES was not confined to the highest SES deprivation quintile and that SES exerted an indirect effect across the range of SES quintiles, demonstrating the presence of a social gradient in preventable hospitalisations in Ottawa. The social gradient of health refers to "the fact that health status is directly related to social status" and that "inequalities in population health status are related to inequalities in social status" (Kosteniuk & Dickinson, 2003, p.263). From this perspective, an increasing gradient of ACS hospitalisation rates from neighbourhoods of higher socio-economic status to neighbourhoods of lower socio-economic status should be observed. Indeed, such a monotonic increasing relationship was confirmed by the linearity test in the ANOVA model based on the ranks of ACS hospitalisation rates¹⁰⁷. This means that neighbourhood SES exerted an indirect effect on the rates of ACS hospitalisation across the range of SES quintiles; its effect was not confined to the highest deprivation level. The results of the linear trend test based on the rates themselves suggested that the burden of preventable hospitalisation may be increasing in an exponential-like manner¹⁰⁸, where the multi-faceted aspects of socio-economic status comprised in the SES index could have a synergetic effect on

¹⁰⁶ The interpretations of the comparison between the gap in ACS hospitalisation rates for the current investigation and for CIHI study is limited by differences in ACS conditions included in the rates, differences in the deprivation categories and differences in the region considered.

¹⁰⁷ The linear term was significant and the deviation term was not significant.

¹⁰⁸ The linear term was significant and the deviation term was also significant. This is a valid interpretation only if the level of socio-economic status across the quintiles is assumed to be equally distributed.

health. These aspects could be social (e.g. lone-parent families, education) or economical (e.g. income or poverty level).

The results of the current investigation provide additional support to the body of evidence demonstrating that SES exerts an effect on ACS hospitalisation rates in countries with different ideologies on health care universality (Ansari, 2007, Curtis, 2004). The effect of SES being observed in Ottawa – a region with higher economic status compared to the rest of Ontario and Canada (see Section 3.2 for more details) – suggests that a SES gap and a social gradient in ACS hospitalisation rates may be present in urban regions across the country.

The layout of neighbourhood by ACS hospitalisation quintiles depicted on the map in Figure 12 is also aligned with literature-based findings, where rural neighbourhoods and inner-city neighbourhoods tend to experience higher ACS hospitalisation rates (Ansari et al., 2006, CIHI 2008a, Steiner et al., 2003). In Ottawa, three neighbourhoods had ACS hospitalisation rates larger than 1.5 times the IQR of their respective deprivation quintile: the neighbourhoods of Navar-Vars (#89) and Cumberland (#1) among the first quintile (most privileged) and the neighbourhood of Kars-Osgoode (#3) as part of the second quintile. Interestingly, these three neighbourhoods are rural ones and may have reduced health care access compared to urban neighbourhoods in the same deprivation quintile. Only one neighbourhood (Carleton Heights – Rideauview (#16)) had a lower ACS hospitalisation rate than its associated socio-economic quintile, and this area was among the most deprived neighbourhood quintile. This neighbourhood may be an example of the healthy immigrant effect¹⁰⁹, as 43% of this neighbourhood population were immigrants who came to Canada between 1996 and 2006 (ONS, n.d.a).

The larger IQR of the most deprived neighbourhood quintile may indicate that aspects of neighbourhoods, other than socio-economic status, may be at play when interpreting preventable hospitalisation rates among most deprived neighbourhoods.

¹⁰⁹ The healthy immigrant effect is associated with immigrants who are in "relatively better health on arrival in Canada compared to native-born Canadians, and immigrant health often converges with years in Canada to native-born levels" (McDonald & Kennedy, 2004, p.1613).

Below is a discussion on the usefulness of the analyses models at identifying neighbourhood aspects associated with the rates of ACS hospitalisation.

5.2. Strengths and Limitations of the Analyses Models: Interpreting the Results and Formulating Generalisations

The current investigation was one of the first known attempts to study the associations between characteristics of places and ACS hospitalisation rates in neighbourhoods as opposed to census tracts or dissemination areas. One of the strengths of these analyses is that they are based on 'natural neighbourhoods', which are comprised of relatively homogenous populations and are thought to better represent communities in which health is constructed, promoted or damaged (Meade & Erickson, 2000, Parenteau & Sawada, 2011). It was demonstrated that the ONS neighbourhoods were adequately conceived to study health outcomes and that they properly capture the SES of places, where the neighbourhoods are homogenous internally and heterogeneous externally (Parenteau & Sawada, 2011). Neighbourhood based studies are, to the extent that the neighbourhoods are well defined, also less subject to misspecification and to the modifiable areal unit problem (MAUP), which arises when different results are generated depending upon how the areal units investigated are divided (Gatrell & Elliott, 2008; Law et al., 2005).¹¹⁰ Parenteau and Sawada (2011) demonstrated that health analysis using the ONS neighbourhoods were less subject to MAUP than the same analysis conducted using census areas in the region of Ottawa.

While interpreting the statistical tests of the analysis models, it is expected that inferences could be formulated at three levels: a) to the concepts and causal model represented by the statistical models, b) to the population represented by the sample included in the model, and c) to other similar populations, such as other places and other times. Any threat to statistical inference, such as violation of their assumptions or lack of robustness, has an impact on the possibility of making further generalisations about the model. Also, interpreting the results of statistical tests involves a good understanding of the strengths and limitations of the models on

¹¹⁰ This was not tested in the current investigation.

which they are based. Threats to the interpretations of the models are either practical (such as models' samples size, variable estimates reliability or model mechanics) or theoretical (such as construct validity of the measures or the external validity of the sample). The following summary of limitations, organised by inference level, is provided with the premise that an understanding of the limitations is an important pre-requisite to the appropriate interpretation of the results.

- a. Limitations associated with making inferences on the concepts and causal model represented by the statistical models

In the current investigation, there are three main limitations associated with this inference level. The first and most important limitation is that, due to the ecological design and the observational nature of the data¹¹¹, causal relationships between neighbourhood factors and the rates of preventable hospitalisation could not be ascertained: the models of this study could only reveal associations. Such design put researchers at the risk of making inferences suffering from the ecological fallacy, which occurs when "inferences regarding individual-level association are based on group-level data" (Diez Roux, 2009, p.50). No inferences suffering from this fallacy were formulated in this discussion.

The second limitation is that the regression models were limited in their ability to ascertain the relative importance of different aspects of places believed to contribute to the rates of preventable hospitalisation. Below are the main practical and theoretical aspects of the current models that were associated with this limitation:

- The reduced statistical power of the models
- Concerns over the reliability of estimates derived for some variables (as a notable example, the 'percent of women who never had a Pap test')
- The inclusion of strong predictors in the models (e.g. SES)
- The construct validity of some measures

¹¹¹ Observational data is used in non-experimental design, where there is no control group. The information used in non-experimental studies was not collected specifically for their investigation (Trochim & Donnelly, 2008).

As mentioned in the model exploration phase (Section 3.7.2.5), the small case-to-variable ratio of the final model and the potential presence of measurement error affecting the reliability of the estimates reduced the models' power to capture moderate or small effect size (Tabachnick & Fidell, 2007, p.117). Therefore, the statistical models may be subject to Type II errors, or so-called false negatives occurring when a hypothesis test fails to detect a real relationship. This may affect the models' capacity at identifying important relationships (Forthofer et al., 2007, p.215). These concerns over insufficient statistical power would be further aggravated if geographical dependencies in the structure of neighbourhoods were adjusted for in the regression models. Considering that adjacent neighbourhoods may not be truly independent observations, the effective sample size of the dataset could be affected. Due to the context, it is possible that groups of adjacent neighbourhoods have similar characteristics in comparison to neighbourhoods that would be located further. This may affect the correlation structure among the neighbourhoods or the regression residuals, yet this was not considered in the current investigation. It was assumed that the external heterogeneity of the ONS neighbourhoods would palliate to this possible issue.¹¹²

Similarly, the presence of strong predictors among the independent variables (e.g. SES) may, due to the mechanics of the regression model, influence the capacity of the models at identifying the importance of other neighbourhood aspects on the rates of preventable hospitalisation. In fact, the solution of a regression model is sensitive to the combination of the variables, where the importance and significance of a predictor variable is dependent on the other independent variables included in the model. As discussed by Tabachnick and Fidell (2007, p. 116), "if the independent variable of interest is the only one that assesses some important facet of the dependent variable, the independent variable will appear important; if the independent variable of interest is only one of several that assess the same important facet of the dependent variable, it usually will appear less important". Using the statistical tests available, the importance of the variables included in the

¹¹² The regression models of another ecological health study using the ONS neighbourhoods did not suffer from spatial auto-correlation (Parenteau & Sawada, 2011).

model can only be assessed over and above the importance of the other predictor variables. Therefore, the variables' importance should be interpreted in light of the other variables included in the model. This concern is particularly heightened when the explanatory variables are correlated with one and another.

This is particularly relevant for the current investigation, as a variable representing neighbourhood SES level was included in all models. In order to avoid the omitted-variable bias, controlling for SES in the regression models is a necessity in population health studies based on an ecological design (DeMaris, 2004, p.98). Indeed, the literature and the descriptive analyses strongly indicate that socio-economic status would be associated with most of the concepts/variables included in the regression models of this study (Ansari et al., 2006; Ansari, 2007 and see Section 3.7.1 for descriptive statistics). However, neighbourhood socio-economic status – even if associated with a gap in preventable hospitalisation rates - is considered to have an indirect effect on such hospitalisations (see causal model in Figure 6). Alternatively, the other concepts included in the models are considered to have a direct impact on ACS hospitalisations¹¹³. So, it is only through its associations with these other aspects of places that SES is believed to exert an effect on the rates of preventable hospitalisation. In this situation, the variable representing neighbourhood SES in the models may be absorbing a proportion of the variance in ACS hospitalisation rates which could also be explained by differences in health status, health behaviour or primary care access and utilisation.¹¹⁴ If this is the case, the significance of the variables representing these concepts could appear (based on model estimation) to be less important than they really are. For these reasons, Brown et al. (2001) cautioned researchers about interpreting the results of models adjusted for socio-economic status in studies on preventable hospitalisations.

¹¹³ With the exception of primary care access; this concept is believed to have an impact on the rates via its effect on the utilisation of primary care services (based on causal model (Figure 6) and conceptual model of Andersen (1968; 1995)).

¹¹⁴ In the sequential models, the variance in ACS hospitalisation rates explained commonly between the predictor variables representing the concept of interest was attributed to the variable of SES, possibly reducing the significance of the other variables. Similarly, in the full model, the variance explained commonly by all variables was not attributed to any variable, which may reduce the significance of their statistical tests.

Teasing out the importance of the different variables in the model may be further affected by the construct validity of the measures themselves. If the variables are inadequate representations of the theoretical concepts in the causal model, their capacity to capture the importance of their relationships with the rates of preventable hospitalisation in the regression model would be limited. In such event, the non-significance of a variable may be more a reflection of the imperfectness of the proxy, whereas another proxy – if available – could better represent the concept and capture the importance of its relationship with the rates of preventable hospitalisation.

The third limitation pertaining to this inference level is the restrained capacity of the models at discerning the importance of the complete set of theoretical relationships suggested in the causal model of preventable hospitalisations (Figure 6). The regression models represented simplified versions of this causal model: some aspects of the theoretical concepts of interest were not represented by variables in the regression models (see Section 3.7.3 and Appendix 5 and Appendix 6). This limitation was particularly relevant for the sequential models, which aimed to test the relevance of certain hypotheses, for example primary care access, at explaining the differences in neighbourhood preventable hospitalisation rates. Yet, if some aspects of this concept – for example the acceptability or accommodation of services – were not represented in the model, it may reduce the overall significance of the variable (or group of variables) representing it in the model (i.e., primary care access). Therefore, the absence of variables representing certain aspects of the concepts tested in the model has to be taken into consideration while interpreting the significance or non-significance of the regression results.

- b. Limitations related to making inferences to the population represented by the sample

For the inferences formulated to be applicable to the neighbourhoods of Ottawa, the sample on which the regression models are based must be representative of the population of neighbourhoods in that region (Lattin et al., 2003, p.34). For the current investigation, the representativeness of the sample was reduced by the

seven neighbourhoods that could not be included in the dataset for data availability reasons (see Section 3.2). However, it may not compromise the applicability of the inferences to the context of Ottawa considering that most of these neighbourhoods were uninhabited. According to some authors, the exclusion of the neighbourhood of Orleans Central from the models, which met all criteria for exclusion in the model exploration phase (See Section 3.6.2.2 and 3.7.2.4), may limit the generalisability of the results to the neighbourhoods of Ottawa (Montgomery et al., 2006, p.199). Yet, the decision to remove this exceptional multivariate outlier eliminated its influence on the results and on the inferences formulated at the causal model level (Tabachnick & Fidell, 2007, p.119-122). In all, the regression models were based on 89 neighbourhoods out of the 91 inhabited neighbourhoods of Ottawa¹¹⁵. Therefore, inferences from the sample to the neighbourhoods of Ottawa can be considered as valid and adequate for the current investigation.

c. Limitations related to making inferences to other similar populations

External validity refers to the degree to which the inferences formulated in the study on a certain population of interest would hold in other places or times (Lattin et al., 2003, p.34). Here, other places would refer to neighbourhoods with similar characteristics, therefore predominantly urban neighbourhoods located in Canada or in other countries with universal access to health services. Although the objectives of the study were not to project the study results to other places, but rather to describe neighbourhood experiences in Ottawa, such generalisations could be valid only if supported by the results of similar investigations published in the literature.

Having identified ByWard Market as exerting an impact on the significance of the results of the full final model (Section 4.3) may impact the external validity of this model, yet this will be taken into consideration while formulating generalisations. This neighbourhood was retained in the dataset as it did not meet the exclusion criteria (see Sections 3.6.2.2 and 3.7.2.4). ByWard Market is characterised by an

¹¹⁵ This count excludes the neighbourhoods which are mainly not residential: Hunt Club South Industrial (population = 691), Orleans Industrial (population = 114), Carleton University (population=440), LeBreton (population=15), Notre-Dame and Beechwood Cemeteries (population at risk of hospitalisation = 0).

unusual combination of socio-economic status, health status, healthy behaviour and ACS hospitalisation rate. The presence of places with unusual characteristics as ByWard Market is relatively common in studies on socio-economic inequality and population health (Dunn et al., 2007). As stressed by these authors, evaluating the specific aspects of this neighbourhood may enrich the understanding of the phenomenon studied in the population of interest. The recent revitalisation of ByWard Market into a commercial and residential hub attracted a lot of young healthy professionals and students who are now cohabitating with the mainly low income population already established in this area (ONS, n.d.b). The changing characteristics and heterogeneity of this neighbourhood support its influence on the study of ACS hospitalisation rates in Ottawa. Its inclusion in the final models could be defended as having enhanced the representativeness of the study results to the neighbourhoods of Ottawa, yet it may compromise the possibility of applying the inferences formulated to other populations.

In summary, formulating inferences about the associations between neighbourhood characteristics and preventable hospitalisations from the current analyses implied interpreting the results in light of the limitations discussed above. These may have restricted the models' capacity at discerning some relationships. On the other hand, the inferences are representative of the inhabited neighbourhoods of Ottawa and they are based on natural neighbourhoods, which enhanced their theoretical value. Yet, the possibility of projecting these inferences to other places depends on supporting evidence in the literature.

5.3. About Hypotheses Explaining Differences in Preventable Hospitalisations in Ottawa

In the literature and the suggested causal model (Figure 6), health status, healthy behaviour as well as access and utilisation of primary care are considered as having a direct influence on differences in preventable hospitalisations.¹¹⁶ Testing these hypotheses in the context of Ottawa neighbourhoods was the second objective of this study. Four sequential regression models were built to identify which of these

¹¹⁶ With the exception of primary care as discussed in footnote 113.

pre-established hypotheses have independent explanatory power on ACS hospitalisation rates over and above variation explained indirectly by neighbourhood socio-economic status.

It was expected that all four hypothesis would help explain the variance in ACS hospitalisation rates. Yet, from the results of the sequential models, only two hypotheses did: differences in health status and differences in healthy behaviour. These results were aligned with the expected relationships (i.e., regarding the sign on the relevant coefficients) laid out in the causal model and in the literature. Further discussion about these results and their interpretations in light of the findings of the final full model will be provided in Section 5.4.1. However, the results of the sequential models representing the hypothesis of differences in primary care access and utilisation were not aligned with the expected relationships hypothesised in the causal model. Such results are not unusual: unexpected and inconsistent findings have been published in the literature. The lack of significance of these two hypotheses is discussed below (Sections 5.3.1 and 5.3.2).

5.3.1. Interpretation of Lack of Significance of Hypothesis of Differences in Primary Care Access

Although ACS hospitalisation is often considered as a proxy of primary care access, primary care access was not related to the rates of ACS hospitalisation among the neighbourhoods of Ottawa, once differences in socio-economic status were taken into account. One plausible explanation could be that universal access to health care dampens or obviates the anticipated relationship between neighbourhood access to primary care and the rates of preventable hospitalisation. Evidence from the literature is inconsistent in this regard and will be discussed following a brief overview of the limitations specific to this sequential model.

Primary care access is a multi-faceted "concept describing people's ability to use health services when and where they are needed", based on the services' availability, accessibility, accommodation, affordability and acceptability (Cromley & McLafferty, 2002, p.234). In the sequential model, only two aspects of primary care access were represented: availability and accessibility. The former was represented

by the variables of 'primary care physician per 1,000 people' and 'proportion of residents without a family physician'.¹¹⁷ The latter was represented by the 'average distance to the closest four primary care facilities'.

The appropriateness, or construct validity, of these variables could be influencing the lack of statistical significance of the incremental F ratio test. Gauvin et al. (2007, p. S25) argued the importance of "delineating the most appropriate territorial units for the specific exposure of interest". The neighbourhoods were appropriate for the health outcome studied (ACS hospitalisation rates), yet they may not be as appropriate for measures of primary care accessibility and availability (Parenteau & Sawada, 2011). At the neighbourhood level, physician supply may not be the most representative of primary care availability. People may be willing to use primary care resources located outside of their immediate residential environment and it may be more important for people to have access to primary care services in their work environment. A physician supply variable taking into consideration the activity space¹¹⁸ of the neighbourhood residents as well as primary care service capacity constraints could be more appropriate (Cromley and McLafferty, 2002, p. 49). Also, the variable of physician supply could be subject to the 'container' fallacy or size of the neighbourhoods, considering that it was calculated based on the number of physicians practicing within the neighbourhood boundaries.¹¹⁹ In fact, the number of primary care physicians per 1,000 was zero for 18% of the neighbourhoods. As pointed by Schreiber and Zielinski (1997, p.281), the absence of primary care physicians in a small area "does not necessarily mean that people in that area are underserved, they may be obtaining health care in an adjacent or nearby ZIP code [area]". But, the importance of providing primary care services at the local level was demonstrated by Curtis (2004). As for the accessibility variable, a measure

¹¹⁷ To a certain extent, the variable of 'proportion of residents without a family physician' could encompass notion of service accommodation.

¹¹⁸ Activity space consists on the set of locations that people visit regularly during everyday life such as workplaces, school, and shopping center ... (Cromley and McLafferty, 2002, p. 49).

¹¹⁹ The variable of 'distance to closest four primary care facilities' was not subject to the 'container' fallacy considering that the closest four facilities from the neighbourhood population centroid were selected independently of the neighbourhood boundaries.

incorporating the notion of access opportunity could be more appropriate (Curtis, 2004, p.237). For example, distances calculated based on public transportation network as opposed to road network, or based on the relative travel times by public and private transportation methods as well as car ownership data, could be more appropriate proxies of primary care accessibility. Furthermore, the objectively measured variables representing availability and accessibility may not be the most appropriate for this study. Relying on variables assessing the perceived availability and accessibility experienced by the local residents, which may have an impact on their primary health care seeking behaviour, could be more appropriate – if they were available.

Although the case-to-variable ratio (22:1) of the models was good, two additional reasons may explain the lack of a significant relationship between primary care access and preventable hospitalisation rates in Ottawa. First, three of the five aspects of primary care access could not be represented in the models due to the unavailability of appropriate measures. Second, from the review of the literature performed below, the concept of socio-economic deprivation appears to be closely related to the concept of primary care access. Therefore, it is plausible that the inclusion of neighbourhood socio-economic status in the reduced model absorbed a proportion of the variance in preventable hospitalisations explained by differences in primary care access. This would reduce the relative importance of primary care access in the subsequent model. Because of the above issues, concluding that there is no association between primary care access and preventable hospitalisations is not possible. Interestingly enough, the literature on this topic is inconsistent.

This inconsistency in the literature could partly be explained by the different operationalisations of primary care access used in the different studies. Self-assessed measures of access to primary care services (not available for this study) are considered to be one of the best proxies of primary care access since all five aspects of the concept are embedded in it. It was used in two studies on ACS hospitalisations and both concluded that self-assessed primary care access, after

controlling for a number of other variables, is negatively associated with preventable hospitalisation rates (Ansari et al., 2006; Bindman et al., 1995).

However, in most studies on ACS hospitalisations, primary care access was represented by objectively-measured proxies of its availability only (e.g. having a primary care physician or physician supply) – yielding inconsistent findings. For example, in twelve studies reviewed by Ansari (2007), six showed evidence of a negative relationship between physician supply and the rates of preventable hospitalisation, two yielded positive significant relationship (contrary to expectations), and four yielded non-significant findings. As for the studies representing primary care availability by variables such as having access to a regular source of care, inconsistent associations with preventable hospitalisations were also observed (Ansari, 2007). This conflicting evidence may indicate that these measures of primary care availability could lack construct validity for the study of preventable hospitalisations, except for areas experiencing important shortage of primary care physicians.

The accessibility of primary care – defined by the location where the services are provided – and its association with preventable hospitalisations has not been previously studied among predominantly urban areas in the literature (to the author's knowledge). In the sequential model, the significant association between the variable representing the accessibility of primary care services was not sufficient to yield a significant incremental F ratio test, but its contextual role will be further discussed in the interpretations of the full models' results (see Section 5.4.2.1).

The other three aspects of primary care access, with the exception of affordability in the United States, have not been well studied in the literature. In the sequential model, proxies of primary care affordability, accommodation and acceptability were also not studied. Therefore, it is difficult to assess how the significance of the incremental F ratio test could have been influenced by the inclusion of such measures in the model. The full impact of primary care access on the rates of preventable hospitalisation could not be captured by the model. Yet, barriers pertaining to these aspects of primary care services are experienced in the

Canadian population, and this may theoretically influence the rates of ACS hospitalisation.

In fact, in a recent investigation of the Health Council of Canada (2010), affordability of services was identified as a barrier to accessing care. In Canada, examples of primary care costs not covered by universal health care are those associated with certain medical equipment, treatments and prescription drugs or travelling to medical appointments. Moreover, this report states that "9% of the lowest-income Canadians reported not consulting a doctor because of cost, compared with 1% of the highest-income Canadians, 10% of the lowest-income Canadians reported skipping a medical test or treatment due to cost, compared with 3% of the highest-income Canadians, and 21% of the lowest-income Canadians reported not filling a prescription because of cost, compared with 2% of the highest-income Canadians" (Health Council of Canada, 2010, p.15). These findings support Sanmartin and Khan's (2011) suggestion that Canadians of lower economic status may experience limited access to the full range of health care services required to reduce the risk of an ACS hospitalisation. In the literature, it seems that out-of-pocket expenditures associated with health care have a disproportional impact on the most deprived residents. For the current investigation, the SES index may have captured the impact of primary care affordability on the rates of ACS hospitalisation in the sequential model.

The same recent Health Council report (2010) demonstrated that Canadians experience barriers pertaining to primary care access due to the lack of services' accommodation: 55% mentioned not being able to get an appointment on the same or next day when needed and 65% reported difficulties assessing medical care in the evenings or weekends. Since Canadians expressed difficulties in accessing primary care services in ways ensuring their timeliness, an important aspect of the ACS hospitalisation concept, it may affect the rates of ACS hospitalisation and differences at the neighbourhood level could be observed (Billings et al., 1993).

The acceptability of primary care services is also a barrier to access experienced among certain groups of Canadians, such as immigrants or Aboriginal people

(Asanin & Wilson, 2008; Shah et al., 2003). An interesting study demonstrated that the rates of ACS hospitalisation in First Nation reserves located in Manitoba decreased the longer community health services have been under community control (Lavoie et al., 2010). Although these findings are not transferable to the context in Ottawa, the perceived acceptability of primary care services may vary based on neighbourhood composition, which may affect the rates of ACS hospitalisation. In the Ottawa map of ACS hospitalisation quintiles (Figure 12), it was observed that the neighbourhoods with greater proportion of francophone population (located in the Eastern portion of neighbourhoods within the Greenbelt area) experience high ACS hospitalisation rates. In these areas, there is a high concentration of deprived francophones with poor health status and a growing community of ethno-cultural francophones for which the health needs are not well-understood or addressed by policy makers (LeBlanc, 2003). Although Ottawa has community health centers providing services in French, their supply may be insufficient and the acceptability of anglophone services may be an issue affecting the ACS hospitalisation rates in the francophone community (LeBlanc, 2003). Further research in this regard is needed, especially considering the historical and recent changes in the provision of French health services in Ottawa (Curtis, 2004, LeBlanc, 2003).

Overall, the statistical non-significance of the variables representing primary care access in the sequential model may be indicative of the absence of a relationship between primary care access and neighbourhood preventable hospitalisation in Ottawa, when socio-economic status is taken into consideration. Yet, the inferential process is compromised by some theoretical limitations as well as practical limitations. Unfortunately, evidence from the literature is conflicting and cannot provide support for the statistical results generated for the neighbourhoods of Ottawa. Thus, it is difficult to draw firm conclusions on the validity of preventable hospitalisation as an indicator of primary care access in a predominantly urban setting with universal access to health care services. Further research on the role of services accommodation, affordability and acceptability is recommended.

5.3.2. Interpretation of the Unexpected Results of the Sequential Models Testing the Differences in Primary Care Utilisation Hypothesis

The results of the sequential models testing the role of primary care utilisation on the rates of ACS hospitalisation among the neighbourhoods of Ottawa generated unexpected and inconsistent results, attesting to the lack of robustness of this sequential model. In the results of the final sequential model, the lack of significance of the incremental F ratio test was indicative of the absence of a relationship between this hypothesis and the rates of preventable hospitalisation. Yet, the results of the sensitivity analysis model weighted by population size did not support those of the final sequential model: a significant and negative association was identified. Moreover, this relationship is contrary to what was expected by the suggested causal model (Figure 6). A spurious relationship caused by a variable included in these subsequent models ('percent of women who never had a Pap test') was considered as the source of the inconsistency between models' results for the incremental F ratio tests.¹²⁰ Below is a discussion of the limitations associated with these sequential models and a review of the literature showcasing the difficulties associated with assessing the impact of primary care utilisation on ACS hospitalisations. The phenomenon of negative suppression will also be presented to demonstrate the presence of a spurious relationship generating the inconsistencies in the results of these two sequential models.

Primary care utilisation rates are complex measures to interpret as they are determined by predisposing, enabling and need factors (Aday & Andersen, 1974; Andersen, 1968; Andersen, 1995). Two aspects of the concept of primary care utilisation, other than the mere utilisation of services, are believed to have an impact on the rates of ACS hospitalisation: propensity to use services for curative reasons or for preventative reasons and the efficiency of the services received (Billings et al., 1993; Disano et al., 2010; Roos et al., 2005). The two variables included in the sequential models represented the former aspect: the variable of 'percent who did

¹²⁰ The assumptions of both sequential models were tested and supported; therefore the sensitivity analysis model weighted by population size did not correct a misspecification of the final sequential model – which could have explained the difference in the results.

not visit a primary care physician in the past year' as a proxy neighbourhood propensity to seek health services and the variable of 'percent of women who never had a Pap test' as a proxy of neighbourhood propensity to seek preventative services.

There are a number of reasons which may explain the inconsistency in the findings of the sequential models, or the lack of a significant relationship in the direction expected by the causal model. First, the construct validity of the variable of 'percent of women who never had a Pap test' could be questioned, as it based on women only. This may not be representative of neighbourhood propensity to seek preventative services as women's tendencies may not mirror men's tendencies. Also, the use of the birth control pill was identified as a health care need contributing to the reception of Pap test – which may enhance women's tendency to seek that preventive service as oppose to other screening tests (Worthington, McLeish, & Fuller-Thomson, 2012). The undermined construct validity of this variable may have limited the models' capacity at capturing the true relationship between neighbourhood primary care utilisation and ACS hospitalisation rates. Second, since the neighbourhood estimates for this variable were derived from smaller samples comprised of female respondents only, they may have reduced the models' reliability – therefore explaining the inconsistency of the results. Third, from the review of the literature below, the concept of primary care utilisation appears to be related to SES. It would be plausible that the inclusion of neighbourhood SES in the reduced model absorbed a proportion of the variance in preventable hospitalisations that could be explained by differences in primary care utilisation – which may explain the lack of a significant relationship in the expected direction. Or, the association between primary care utilisation and SES may have lead to the phenomenon of negative suppression – therefore explaining the unexpected significant results.

Unfortunately, no clear relationship between the concept of primary care utilisation and preventable hospitalisation has been demonstrated in the literature, and unexpected results or non-significant results are often depicted. In two Canadian ecological studies, as the frequency of primary care utilisation increased in a

population, the rates of ACS hospitalisation also increased, but the opposite relationship was found in an Australian study (Ansari et al., 2006; Roos & Mustard, 1997; Roos et al., 2005). At the individual level, frequent reception of primary care services was associated with an increased risk of an ACS hospitalisation among Canadian adults (Sanmartin & Khan, 2011). It is possible that frequent users of health care services have greater health care need, such as more complex health problems or greater fragility. Consequently, ACS hospitalisations among sicker individuals may not be avoidable through primary care. This could explain the unexpected association between primary care utilisation and preventable hospitalisations (Sanmartin & Khan, 2011). Differences in the efficiency of the primary care services received may also explain these unexpected relationships (Roos et al., 2005). Although commonly referenced as potential aspects of primary care utilisation explaining differences in preventable hospitalisations, the quality and continuity of primary care received were not studied in the context of ACS hospitalisations in the literature or in this study (to the author's knowledge). Similarly, the propensity to use primary care in a population – which may account for the timely reception of care and a reduced risk of ACS hospitalisation – could be related to lower ACS hospitalisation rates in a population. However, in urban areas, the propensity to seek health services was not found to be significantly associated with the rates of preventable hospitalisation in the two studies which evaluated it in the literature (Ansari et al., 2006; Bindman et al., 1995).

Also, differences in primary care utilisation and socio-economic status may be at play, which further complicates the interpretation of the above literature. In Canada, equitable contact with primary care services based on income has been shown, however lower income Canadians are more likely to be frequent users of primary care even when severity of health status is taken into consideration (Dunlopp et al., 2000; Glazier, Klien-Geltink, Kopp, & Sibley, 2009). Therefore, it is possible that primary care services do not generate the same health benefits among lower socio-economic status people compared to higher socio-economic status people and that primary care services may be more effective in preventing hospitalisations among most affluent people compared to most deprived people (Roos et al., 2005). Also,

socio-economic differences in the utilisation of preventive health methods are observed in Canada, even when adjusted by measure of needs or ease of access (Demeter, Reed, Lix, MacWilliam, & Leslie, 2005; Gupta, Roos, Walld, Traverse, & Dahl, 2003; Roos & Mustard, 1997). In the sequential models, SES index may have captured the differences in ACS hospitalisations rates due to differences in preventive health services utilisation or in efficiency of care received based on socio-economic status.

The lack of robustness of the sequential models could be explained by a spurious relationship - caused by the variable 'percent of women who never had Pap test' - which generated the phenomenon of negative suppression, therefore explaining unexpected negative and significant association between this variable and the ACS hospitalisation rates in these two models. Negative suppression occurs when an independent variable increases the predictive validity of another variable included in a regression model but its multivariate association with the dependent variable is the reverse of what is predicted by its univariate association (Conger, 1974).¹²¹ Negative suppressor variables are characterised by: 1) a positive correlation with the dependent variable, 2) a positive correlation with the other predictor variable involved in suppression and 3) a negative regression coefficient when entered in the multivariate model (Pandey & Elliott, 2010). In the regression, the primary 'role' of the suppressor variable would be to reduce the error variance of the other independent variable, rather than explaining a portion of the variance in the dependent variable. This leads to a reversion of the regression coefficient of the suppressor variable and an enhanced regression coefficient for the other independent variable, considered as the suppressed variable.

All three criteria identifying the presence of a negative suppressor were met for the two sequential models, where the variable of 'percent of women who never had a Pap test' was identified as the suppressor and the SES index was identified as the

¹²¹ Negative suppression represents the same phenomenon than the Simpson's Paradox in the analyses of categorical variables.

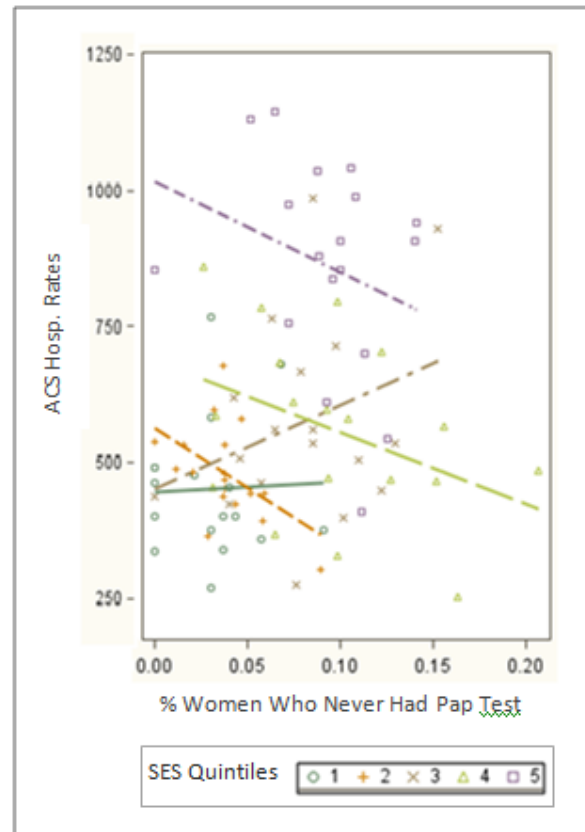
suppressed variable.¹²² Negative suppression is a common statistical situation in multiple linear regression involving continuous predictors, yet its presence is not as commonly acknowledged in the literature due to its controversy (Tu, Gunnell, & Gilthorpe, 2008). Controversies occur when causal inferences are formulated on the conditional relationship, therefore implying that the adjustment variable is a genuine confounder. In the current model, no causal relationship can be ascertained; therefore the identification of a variable acting as a negative suppressor in the statistical model is not a sign of causal confounding.

Tu et al. (2008) demonstrated that when the conditional relationship between two variables is the reverse than their marginal relationships in a three dimensional scatter plot, it is predictive of negative suppression in a regression model. Such pattern is demonstrated in the scatter plot of the 'percent of women who never had a Pap test' and the hospitalisation rates by SES quintiles (Figure 16): although the overall relationship between the percents and the rates is positive, there are apparent downward trends within most of the SES quintiles. Therefore, the significant and negative association between the 'percent of women who never had a Pap test' and ACS hospitalisation rates in the sequential models may be the result of a spurious relationship, which generated the statistical phenomenon of negative suppression. The relationship may be spurious considering the reduced construct validity of this variable, the measurement error of its estimates and/or the small number of neighbourhoods included in the model – all conditions making the results

¹²² In the sequential models: 1) the suppressor variable was correlated positively with the rates of ACS hospitalisation (0.25) and its univariate regression coefficient was positively significant (Appendix 10 and Appendix 14); 2) the suppressor variable had a greater bivariate association with the socio-economic index variable – their correlation was 0.52 – than with the rates of ACS hospitalisation; and 3) by including SES index in the sequential models, the regression coefficients of the suppressor variable became significantly negative and the regression coefficients of the variable SES index increased. In the final full model, all three conditions of negative suppression were present yet the 'percent of women who did not have a Pap test' variable was not significant: therefore this variable was not acting as a true negative suppressor. The variable 'percent who did not visit a primary care physician' was also involved in negative suppression, but not acting as a true suppressor in all models including it.

of the regressions more susceptible to unusual pattern in the data.¹²³ However, this is not warrant of a theoretical cofounding relationship.

Figure 16: Illustration of the Pattern of Relationship Predicting Negative Suppression



The above demonstrates the complex relationships between the utilisation of primary care services, health care need, socio-economic status and preventable hospitalisations. These relationships and the concept of primary care utilisation may be further complicated when evaluated in an ecological design or at the neighbourhood level (Law et al., 2005). Conclusions about the relationship between preventable hospitalisation rates and neighbourhood utilisation of primary care in Ottawa based on the current analysis and existing literature are not possible. Further research on the role of primary care quality and efficiency is recommended.

¹²³ It was suspected that neighbourhood imputations based on SES quintile means for the variable of 'percent of women who never had a Pap test' be at the origin of the presence of negative suppression. However, true negative suppression was still at play in models excluding the imputed neighbourhoods (n=76).

5.4. About Neighbourhood Compositional/Contextual Characteristics and Preventable Hospitalisations in Ottawa

A regression model representing a simplified version of causal model of preventable hospitalisation was built to achieve the third objective of this study: to identify neighbourhood compositional and contextual factors which may be associated with the rates of ACS hospitalisation in Ottawa. Two neighbourhood compositional aspects were found to be significant predictors of ACS hospitalisation rates: socio-economic status and prevalence of smoking. To a lesser extent, the prevalence of poor/fair health status was also associated with ACS hospitalisation rates. As for the contextual aspects of neighbourhoods, the variable of 'distance to primary care facilities' was slightly associated with the rates of ACS hospitalisation. The full model was also built to identify the relative contributions of each of the hypotheses formulated in the literature in a model acknowledging greater complexities of the concept of preventable hospitalisations. The hypothesis of differences in healthy behaviour was identified as the most important contributing factor to the rates of ACS hospitalisation ($p = 0.001$ in final model; $p = 0.005$ in sensitivity analysis model). The hypothesis of differences in health status was identified as significant in the sequential models, but its importance in the full models was limited ($p=0.12$ in final model; $p=0.06$ in sensitivity analysis model). The hypotheses of primary care access and utilisation were not found to be significantly contributing to the rates of preventable hospitalisations. Below is a discussion of these findings.

5.4.1. Compositional Neighbourhood Characteristics Associated with ACS Hospitalisation Rates in Ottawa

5.4.1.1. NEIGHBOURHOOD SOCIO-ECONOMIC STATUS

As predicted by the ANOVA results, the material and social deprivation of neighbourhoods in Ottawa (SES index) was highly associated with the rates of ACS hospitalisation in the final and sensitivity analysis models ($p = 0.0001$ and $p < 0.0001$ respectively). In a critical interpretative synthesis of the literature, Ansari (2007) identified socio-economic factors as the most important ones to explain variations in ACS admissions in small area analyses. In almost all studies reviewed by Ansari (2007), socio-economic measures were statistically significant and strong

predictors of ACS hospitalisations, independently of the study design and the other variables included in the analyses. Addressing the gap in neighbourhood socio-economic status would most probably softened the differences in ACS hospitalisation rates observed in Ottawa (Section 5.1). If the affordability of services was to be identified as an important contributor to the SES gap in ACS hospitalisations in Ottawa, programs funded within the health system could be implemented in low SES neighbourhoods, such as subsidies for prescription drugs, food supplements, equipment for the management of ACS chronic conditions, non-work-day hours of accessibility, transportation and child care options to attend medical appointments, etc (CIHI, 2012, Steward et al., 2005). However, the current study design could not demonstrate such relationship.

Yet, the strong association between neighbourhood SES and ACS hospitalisation rates calls for the implementation of systemic interventions, involving numerous governmental players, to soften the socio-economic disparities experienced in Canada. Solutions to address the socio-economic gradient of health do not rest solely on the health care system: SES is a multi-faceted issue and may influence ACS hospitalisations via the synergetic effect of education, income, social networks, reduced opportunities, environmental milieu and many more (Evans & Stoddart, 1990). This calls for the integration of governmental departments working in social, economic and environmental spheres - as they all relate to the socio-economic determinants of health - for the formulation of policies addressing SES disparities across areas. The health impact of policies implemented to reduce the SES differences across places could be assessed using population health intervention research (Hawe & Potvin, 2009). It is important to note that addressing the socio-economic gradient of neighbourhoods may not eliminate neighbourhood differences in ACS hospitalisation rates: other factors such as healthy behaviour and health status may be at play.

5.4.1.2. NEIGHBOURHOOD PREVALENCE OF UNHEALTHY BEHAVIOUR

The prevalence of unhealthy behaviour represented by the prevalence of cigarette smokers was significantly associated with the rates of ACS hospitalisation, over and

above the impact of other neighbourhood characteristics including socio-economic status. This neighbourhood compositional characteristic had the strongest association, after socio-economic status, with the rates of preventable hospitalisation among the neighbourhoods of Ottawa. In the literature, the association between unhealthy behaviour prevalence and ACS hospitalisation rates received only limited attention but all studies evaluating it supported the current results (Ansari, 2007). At the group and individual level, unhealthy behaviour were positively associated with ACS hospitalisations in Australia, US and Canada (Ansari et al., 2006; Billings et al., 1993; Sanmartin & Kan, 2011). These findings were expected considering that lifestyle factors such as smoking, alcohol consumption and physical activity are important predictor of morbidity (WHO, 2002). Moreover, smoking status exacerbates the risk of hospitalisation for a number of ACS conditions included in the indicator derived in the current investigation: for examples, COPD, asthma, infection of throat, pneumonia and tuberculosis (Almirall, Bolivar, Balanzo, & Gonzalez, 2001; Coker et al., 2006; Coultas, 1998; van Gageldonk-Lafeber et al., 2007).

In light of the findings of the current investigation, efforts to reduce the burden of ACS hospitalisations among all neighbourhoods of Ottawa, either via primary care interventions or health promotion interventions, should include smoking prevention and cessation components. Smoking cessation interventions can target individuals, communities or entire populations and can include media campaigns, competitions or quit contests, group- or individual- behaviour counselling by peers or health care providers, drug or nicotine replacement therapies, nursing or telephone counselling and many more, all having different effectiveness in different contexts (Valery, Anke, Inge & Johannes, 2008). Unfortunately, there is evidence that "health promotion interventions frequently increase rather than decrease socio-economic inequalities in health" due to differential uptake across social groups (Jepson, Harris, Platt, & Tannahill, 2010, p.11). As people of lower socio-economic status are more likely to adopt unhealthy behaviours, the prevalence of unhealthy behaviour in a neighbourhood may contribute to the gap in ACS hospitalisation between neighbourhoods of low and high socio-economic status in Ottawa (CIHI, 2008a;

Gulliford, Sedgwick, & Pearce, 2003). From an equity perspective, increased effort and support could be provided to enable residents of the most deprived neighbourhoods quit smoking or to prevent smoking initiation. In Ontario, low income population may not have access to the full range of smoking cessation aid, such as nicotine replacement therapy; as such resources are not covered by the provincial drug and benefits program (Ontario Smoker's Helpline, personal communication, July 23, 2012). In a review of 48 systematic reviews on individual, community or population-wide programs targeting smoking prevention or cessation, Jepson et al. (2010) identified only one study that evaluated the effectiveness of these interventions at targeting health inequalities. Further research would be needed to identify the health promotion interventions successful at addressing the barriers of low socio-economic populations in adopting and maintaining healthy behaviours. This may have an impact on the social gradient of preventable hospitalisations.

5.4.1.3. NEIGHBOURHOOD PREVALENCE OF POOR OR FAIR HEALTH STATUS

The prevalence of neighbourhood residents with self-reported poor or fair health status had a weak association with ACS hospitalisation rates, after controlling for the other concepts included in the full model. In fact, this variable was not significantly associated with ACS hospitalisation rates in the final model and was significant only at alpha 0.1 in the sensitivity analysis (model weighted by population size).

Therefore, it may be wrong to conclude that health status is an important neighbourhood compositional predictor of preventable hospitalisation rates in Ottawa¹²⁴, even if its importance was demonstrated in the simplified sequential models.

However, the lack of significance of the health status variable in the full model may be specific to the neighbourhoods of Ottawa and may be related to the influence of the neighbourhood of ByWard Market on the results. Excluding this neighbourhood from the dataset for the full model weighted equally yielded a significant *t* test for this

¹²⁴ Due to the risk of making a Type I error, which involves rejecting the null hypothesis of a statistical test when the null hypothesis is true.

variable ($p=0.03$).¹²⁵ Therefore, it may be possible that health status contribute to differences in ACS hospitalisations in predominantly urban neighbourhoods. However, neither the prevalence of ACS conditions nor the burden of disease was significantly related to the rates of preventable hospitalisation in Australia, a similar country than Canada (Ansari et al., 2006). The role of health status on preventable hospitalisation rates was not evaluated in other studies at the population level, except in the US where it was significant in certain age groups only (Laditka et al., 2005).

The weak association between neighbourhood health status and preventable hospitalisation rates in Ottawa may be mediated through a relationship at the individual level: people with greater disease severity/complexity are at greater risk of an ACS hospitalisation, as demonstrated by Sanmartin and Khan (2011). In fact, not all ACS hospitalisations are believed to be avoidable through primary care, especially for severe or complex cases (Billings et al., 1993; Sanderson & Dixon, 2000). At the neighbourhood level, the association between health status and preventable hospitalisation rates could be the group-level reflection of the impact of disease severity on the risk of suffering from an ACS hospitalisation at the individual level. To avoid such impact, Brown et al. (2001) and Liu and Wallace (2011) suggested using standardisation methods adjusting for the composition of co-morbidity prevalence in the population studied as a way to derive ACS hospitalisation rates acknowledging the difference in health status severity across places.

The results of the current investigation may suggest that sicker or more fragile population in Ottawa do not receive the most appropriate type or level of health care to ensure that equity in preventable hospitalisations is reached. Ensuring vertical equity in health care may be a solution, where people with greater health needs would have enhanced opportunities to access and use the most appropriate type and level of primary care they need. For example, Sanmartin and Khan (2011)

¹²⁵ As two other neighbourhoods in the dataset, ByWard Market had a value of 0% for the proportion of people with poor or fair self-reported health status. Yet, its ACS hospitalisation rate was in the 90th percentile. This explained the influence of this neighbourhood on the coefficient of the health status variable – see Section 3.6.2.2.

identified community nutrition programs for elderly women as a potential strategy to reduce their risk of an ACS hospitalisation. Considering its association with improved health outcomes, another potential strategy would be to incorporate self-management support as part of routine primary health care delivered to people having an ACS condition (Health Council of Canada, 2012). Additional research to identify other type of primary care services having the potential to reduce the risk of ACS hospitalisation among sicker populations is recommended.

5.4.2. Contextual Neighbourhood Characteristics Associated with ACS Hospitalisation Rates in Ottawa

5.4.2.1. NEIGHBOURHOOD DISTANCE TO PRIMARY CARE FACILITIES

Distance to primary care physician was mildly significant in the final full model ($p=0.07$). Yet, as with self-rated health, the lack of significance of this contextual aspect may be related to the influential neighbourhood of ByWard Market: excluding this neighbourhood from the dataset yielded a significant t test ($p=0.04$, final model weighted equally). However, in the sensitivity analysis model weighted by population size, distance to primary care facility was no longer significant. The difference in significance of this variable between the final and sensitivity models can be explained by the different impact of rural neighbourhoods on the regression results of these two models. All the rural neighbourhoods had relatively large distances to primary care facilities and experienced high rates of ACS hospitalisation (see map in Figure 12). Considering that the sensitivity analysis model was based on a weighting method proportional to neighbourhood population size, the rural neighbourhoods were assigned small regression weights (due to their relatively small populations) in this model and the influence of their relatively large distances on the significance of this regression coefficient was reduced. The full model was re-run among urban neighbourhoods only and this variable was no longer significant, regardless of the weighting scheme ($n=80$). Therefore, this suggests that the accessibility of primary care services may have an impact on the rates of preventable hospitalisation in rural neighbourhoods. Among urban neighbourhoods only, the accessibility of primary care – measured as the distance to care – would not be a contributing factor to the rates of neighbourhood preventable hospitalisation.

Within urban environments, primary care clinics are often placed in neighbourhoods to improve accessibility (Baumgardner, Halsmer, Steber, Shah, & Mundt, 2006). But, the association between proximity and the rates of ACS hospitalisation was not studied in an urban context (to the author's knowledge). Yet, there is some evidence of associations between the utilisation of primary care services and the proximity of services in urban neighbourhoods. For example, the utilisation of preventative services such as blood pressure tests was less likely in urban areas in New Zealand if travel times were long (Hiscock, Pearce, Blakely, & Witten, 2008), but this association was not found in a similar study in the US (Baumgardner et al., 2006). Law et al. (2005, p.373) observed that, while no significant neighbourhood effects were observed in Hamilton, "individuals reporting a total travel/visit time of more than 20 minutes were less likely to have reported use of general practitioners in the past 2 weeks". Also, Steward et al. (2005) highlighted that the decision to utilise health services among the poorest Canadians is affected by the location of the facility and the cost associated with getting to the facility. However, there is no evidence that the association between primary care utilisation and primary care accessibility play a role on ACS hospitalisation rates in urban areas. The results of the final models among the urban neighbourhoods only and the above literature suggest that primary care accessibility, in terms of neighbourhood distance to primary care facilities, is not a concern in the urban population of Ottawa.

However, distance to primary care facilities appears to be a contextual factor explaining differences in ACS hospitalisation rates between urban and rural neighbourhoods in Ottawa. The concept of distance decay may explain this relationship. Distance decay refers to people's decreasing tendency to interact with service facilities as their travelling distances increase due to added travelling time, cost and effort and/or reduced knowledge and familiarity with services opportunities (Cromley and McLafferty, 2002, p.235). In the literature, differences between ACS hospitalisation rates have been consistently demonstrated between urban and rural regions in Canada (Fransoo, 2009; Shah et al., 2003;). Yet, the role of distance to care at explaining this gap was studied in only one research conducted by Ansari et al. (2006) in Australia. This study demonstrated that an accessibility and remoteness

index based on road distances to centers providing goods and services was significantly associated with ACS hospitalisation rates, just as among the neighbourhoods of Ottawa. Solutions to increase the accessibility of primary care services in rural neighbourhoods would be to provide telephone consultations with a physician, to provide travel subsidies to get to primary care facilities for people with low mobility or low socio-economic status (e.g. no private or public transportation, disability), to locate community health centers according to the population distributed within their catchment areas, to ensure that facilities have evening and week-end opening hours, to rely on mobile clinics, or to implement compensation fees for primary care physician working in rural areas. Ensuring that territorial justice, meaning that the health care needs of a population are met within pre-defined geographical spaces, is achieved could be a strategy to enhance the accessibility of primary care services in Ottawa (Meade & Erickson, 2000).

On the other side, interpreting distance to health care variables is challenging in health research. In fact, primary care services are provided at fixed sites, yet they are serving a dispersed population where people and services typically cluster in urban areas (Cromley and McLafferty, 2002). In light of this, the frequency distribution of distances is often highly skewed with a large proportion of the population (or neighbourhoods) being close to services and a significant minority being quite distant from services (Cromley and McLafferty, 2002, p.240). In the current investigation, the distribution of the distance variable – where relatively large distances were recorded for the rural neighbourhoods only – followed the distribution described above and may have mimic the behaviour of a dichotomous variable identifying rural neighbourhoods in the regression models.¹²⁶ This would mean that other aspects of places associated with rurality, but not necessarily with primary care accessibility, may explain differences in ACS hospitalisation rates between urban and rural areas in Ottawa. In the literature, authors often explain the gap in ACS hospitalisation rates between rural and urban areas either by differences in

¹²⁶ A dichotomous variable would identify the rural neighbourhoods by 1 and the urban neighbourhoods by 0. The gap generated by the large distances for the rural neighbourhoods could be similar, or even greater, to the gap between 0 and 1 in dichotomous variables.

primary care availability or accessibility (Laditka et al., 2005; Schreiber & Zielinski, 1997; Shah et al., 2003). Studying the impact of a self-assessed measure of primary care accessibility, as well as other aspect of rurality, on ACS hospitalisation rates would provide additional nuances to the results of the current investigation.

5.4.3. Interpretation of Lack of Significance of Other Compositional and Contextual Neighbourhood Aspects

All other compositional and contextual variables in the final models, both for the full final and sensitivity models, were not significantly associated with the rates of preventable hospitalisation in Ottawa neighbourhoods. The lack of significance of these other factors may be linked to the practical and theoretical limitations of the regression models listed in Section 5.2 or due to the lack of a relationship.

The relationships between 'percent of women who never had Pap test' as well as 'percent who did not visit primary care physician', although non-significant, were in the opposite direction than expected. These results may be the consequences of spurious relationships, which generated a milder form of the phenomenon of negative suppression as described in Section 5.3.2.

On the other hand, the variable of primary care physician supply, also non-significant, had a positive regression coefficient; therefore suggesting that the rates of ACS hospitalisation may increase as the supply of primary care physicians increases. Similar findings were observed in two out of 12 studies revised by Ansari (2007). In separate analyses, it was demonstrated that neighbourhoods of higher deprivation status had greater supply of primary care physicians in Ottawa, which may be due to historical investment in the development of health care facilities.¹²⁷ Schreiber and Zielinski (1997, p.281) suggested that "the presence of primary care physicians in an area does not necessarily mean that they will see all who need care, (...), there could be underserved in areas with apparently adequate physician capacity". This may explain the direction of the regression coefficient for the primary care physician supply variable in the final models of the current investigation.

¹²⁷ These analyses were conducted as part of a project for the class Geographies of Health and Health Care (GEG5105) and a journal-like article is available upon request.

Further research on any potential inequities in service distribution, such as contrasts between need and demand for services, among the neighbourhoods of Ottawa would provide more insight on the existence on underserved areas.

Chapter 6. Conclusion

As a conclusion to this current investigation, the research conclusions and contributions are presented in Section 6.1. And, two additional research designs stemming from the current investigation are proposed in the sub-section on future research (Section 6.2).

6.1. Research Contributions

In the literature, the rates of preventable hospitalisation at the group level such as neighbourhood is considered as an indicator of primary care access, yet its validity has been questioned – especially for countries or regions like Canada and Ottawa where universal access to primary care is available. Even so, CIHI and Statistics Canada commonly report the rates of ACS hospitalisation at different geographical levels to assist health care service delivery and decision-making (CIHI 2008a; CIHI 2009a, CIHI, 2009b; CIHI, 2010a; Statistics Canada, 2006a). A better understanding of the factors associated with this group-level indicator was identified as an important research need by CIHI (2008a), and the ecological design of the current investigation seek to address it. Findings of the current investigation provided additional evidence to the presence of a social gradient in preventable hospitalisation rates among the neighbourhoods of Ottawa. Therefore, from a health equity perspective where the goal is the "absence of unfair and avoidable or remediable differences in health among social groups", understanding the factors associated with the rates of ACS hospitalisation is even more important (Solar & Irwin, 2010, p.4).

However, the usefulness of ACS hospitalisation rates as an indicator of primary care access among the neighbourhoods of Ottawa could not be demonstrated in the current research. Further research on the role of primary care services' affordability, accommodation and acceptability is recommended to provide additional insights on the validity of ACS hospitalisations as an indicator of primary care access in predominantly urban areas with universal access to health care. Also, the literature-based hypothesis suggesting that differential primary care utilisation across

neighbourhoods may be associated with the rates of preventable hospitalisation was not supported in the current investigation, especially as it relates to preventive health services utilisation. This may be due to the reduced reliability of the variables and the spurious relationship identified with a variable representing this hypothesis in the analysis models. Further research on the role of differences primary care quality across neighbourhoods, in areas with universal health care is available, is also recommended in order to provide further evidence on the lack or existence of a relationship between primary care utilisation and preventable hospitalisations. Such research may be difficult to carry considering the challenges associated with defining satisfactory measures of health care utilisation, due to its interrelationship with the notions of health need and health care efficacy (Goddard & Smith, 2001).

Two neighbourhood compositional characteristics were strongly associated with ACS hospitalisation rates in Ottawa neighbourhoods: the prevalence of healthy behaviour and the socio-economic index. A weak evidence of a relationship between preventable hospitalisation and prevalence of poor/fair health status was shown for the neighbourhoods of Ottawa.

Contextual differences between rural and urban neighbourhoods in Ottawa, such as distance to primary care facilities, may also be associated with the rates of preventable hospitalisation. However, it could not be determined if contextual differences in primary care access, or other aspect of rurality, is the factor associated with differences in ACS hospitalisation rates. Further research to determine the nature of the contextual aspects of rural neighbourhoods associated with greater rates of preventable is recommended.

It is hoped that the findings of the current research will assist health services planning and other types of local planning in an effort to match expenditure to the notion of need in Ottawa. This expectation is particularly relevant considering the strong association between preventable hospitalisation rates and socio-economic level at the neighbourhood level in Ottawa. Equity in health is concerned with "creating equal opportunities for health and with bringing health differentials down to the lower level possible" (Whitehead, 1992, p.7). To bridge the socio-economic gap

in preventable hospitalisation rates, opportunities within the health care system realms should tackle both horizontal equity – which is "equal treatment is provided for equivalent needs" – and vertical equity – which is "preferential treatment is given to these with greater health needs" (Dahrouge et al., 2011, p.1302).

In light of the results of the current investigation, interventions to reduce the burden of ACS hospitalisations among all neighbourhoods of Ottawa should target healthy behaviour promotion and include smoking prevention and cessation components. To reduce the socio-economic gap in ACS hospitalisations, focussing these interventions on enabling residents of the most deprived neighbourhoods at adopting healthy behaviour, for example quit smoking, would address vertical equity. Some evidence from the current research also suggest that neighbourhoods with greater need for primary health care services, identified indirectly as the neighbourhoods with poorer health status, may not receive the preferential primary care services which could reduce their rates of ACS hospitalisation. Further evidence suggests that rural neighbourhoods in Ottawa may not have access to the full range of primary health services for equal need due to the greater inaccessibility of services. The extent at which avenues within the health care system, either via preventive or curative services or via services' accessibility, may impact the socio-economic gap in preventable hospitalisation could not be demonstrated by the current investigation. Considering the importance of the association between ACS hospitalisations and SES, multidisciplinary policy and program solutions to reduce the SES gap among the neighbourhoods of Ottawa would also address equity in health. This would ensure that people of different neighbourhoods are provided with equal opportunities to avoid preventable hospitalisation. This calls for the integration of different municipal and provincial governments to work together in order to address the social determinants of health.

In all, the current research contributed to the body of knowledge on health inequities, the determinants of health and health care utilisation as well as the meaning of the indicator of ACS hospitalisation rates in a region with a predominantly urban population covered by universal access to health care. A key

strength of this investigation rests on the utilisation of natural neighbourhoods, ensuring that the results are based on meaningful areas which properly reflect processes in which health is created, maintained or damaged. Also, the research outputs provide ways to visualise and understand neighbourhood ACS hospitalisations in the context of Ottawa. And, the interpretations of the results offer avenues for policy makers and local health planners for the delivery of primary health care services as well as for achieving health equity in preventable hospitalisations in Ottawa.

6.2. Future Research

To enrich the results of the current investigation, structural equation modelling with latent variables is a promising alternative analysis model to identify which aspects of neighbourhood are more likely to address the socio-economic gap in ACS hospitalisations. In this type of model, unobservable or complex constructs (latent variables) such as primary care access can be represented by a number of observable variables and a model of dependence between the constructs – rather than among the observed variables themselves – can be developed. These models are represented by two sets of equations: "the measurement equations (describing the relationships between the observed variables and the latent variables) and the structural equations (describing the dependence relationships among the latent variables)" (Lattin et al., 2003, p.381).

For the study of neighbourhood ACS hospitalisation rates, the concepts of interest (e.g. primary care access) could be represented by latent variables and they would be estimated from observable measures (e.g. the accessibility and availability variables). The pattern of dependence between the latent variables could be estimated based on the structural equations, where the pattern of covariance and dependence between the latent variables would be particularly relevant to determine which factors may influence the gap in ACS hospitalisations by neighbourhood socio-economic status. Such model of interdependence could untangle or provide nuances to the relationships between neighbourhood SES, health status, health behaviour as well as primary care access and utilisation. Also, their relative

importance on the rates of ACS hospitalisation could be assessed. This could contribute to identifying neighbourhood aspects to tackle in order to close the SES gap in ACS hospitalisation rates observed in Ottawa neighbourhoods. From the findings of such model, decision makers would be better informed on how to address vertical health equity in ACS hospitalisations in the region.

Stemming from further reflection based on the current investigation, answering the following research question would be of interest for decision makers and primary care providers:

What are the individual level causal factors influencing the risk of ACS hospitalisations, how do they explain neighbourhood level variations in ACS hospitalisations in Ottawa and to what extent are neighbourhood contextual characteristics affecting these variations?

This research question could be tackled in a subsequent research by the Ottawa Neighbourhood Study (ONS) using multi-level modeling – in the event that hospitalisation data is made available at the individual level and that linkages with other individual data sources such as the census and the CCHS is possible. The same information would be needed for individuals who were not hospitalised for ACS conditions. Multilevel analyses could provide insights on the interactions between factors at the individual and at the neighbourhood levels on ACS hospitalisations. In fact, such model allows for the examination of inter-group and inter-individual variability in the outcome variable measured at the individual level as well as for the identification of individual-level and group-level constructs associated with variability in the outcome variable at both level. Therefore, multi-level analysis can be used to draw inferences "regarding causes of inter-individual variation and the extent to which it is explained by individual-level or group-level variables" as well as to examine "between-neighbourhood and within-neighbourhood variability in the outcome and the degree to which between-neighbourhood variability is accounted for by neighbourhood-level or individual-level variables" (Diez Roux, 2003, p.54).

As described above, a multi-level analysis research design would help specifying the processes through which neighbourhood-level and individual-level factors influence the health outcome of ACS hospitalisation as well as test their relative importance. Concepts most relevant at the individual level could be operationalised at that level, rather than as compositional variables, and the relevant neighbourhood variables could be incorporated at the group-level or contextual level (Diez Roux, 2003). For example, the concept of socio-economic status could be incorporated at the individual level and the added effect of neighbourhood income disparity could be assessed. The impact of neighbourhood access to healthy food sources – identified as a potential contributor to ACS hospitalisation by Sanmartin and Khan (2011) – could also be tested. Similarly, other contextual differences between neighbourhoods of rural and urban status could be examined. By answering this research question, policy makers and health care practitioners aiming to redress health and health care inequities in Ottawa, as they relate to the indicator of ACS hospitalisations, would be provided with richer information allowing them to identify at which level and in which regard change in health care delivery could lead to the closure of the socio-economic gap in ACS hospitalisations experienced among the neighbourhoods of Ottawa.

References

- Aday, L., & Andersen, R. (1974). A framework for the study of access to medical care. *Health Services Research*, 9(3), 208-220.
- Agha, M.M., Glazier, R., & Guttman, A. (2007). Ambulatory care-sensitive hospitalizations persist for up to 9 years among children born in a major Canadian urban center. *Ambulatory Pediatrics*, 7, 258-262.
- Ahern, J., Galea, S., Hubbard, A. & Syme, L. (2009). Neighborhood smoking norms modify the relation between collective efficacy and smoking behavior. *Drug Alcohol Dependence*, 100(1-2), 138-145.
- Ahmad, O.B., Boshi-Pinto, C., Lopez, A.D., Murray, C.J., Lozano, R., & Inoue, M. (2001). Age standardisation of rates: A new WHO standard. *GPE Discussion Paper Series: No.31*, World Health Organization.
- Almirall, J., Bolívar, I., Balanzó, X. & González, C.A. (1999). Risk factors for community-acquired pneumonia in adults: a population-based case-control study. *European Respiratory Journal*, 13(2), 349-355.
- American College of Physicians. (2008). *How is a shortage of primary care physicians affecting the quality and cost of medical care? A comprehensive evidence review*. Philadelphia, PA: American College of Physicians.
- Andersen, R. (1968). *A behavioral model of families' use of health services*. Chicago: University of Chicago, Center for Health Administration Studies.
- Andersen, R. (1995). Revisiting the behavioral model and access to medical care: Does it matter? *Journal of Health and Social Behavior*, 36(1), 1-10.
- Andersen, P.K., & Keiding, N. (2006). Standardization Methods. In *Survival and event history analysis* (pp.428-440). Chichester : Wiley.
- Anderson, G. (1996). Common conditions considered sensitive to ambulatory care. In V. Goel, J. Williams, G. B.-H. Anderson, C. Fooks & C. Naylor (Eds.), *Patterns of health care in Ontario: The ICES practice atlas* (2nd ed.) (pp. 104-110). Ottawa: Canadian Medical Association.
- Ansari, Z. (2007). The concept and usefulness of ambulatory care sensitive conditions as indicators of quality and access to primary health care. *Australian Journal of Primary Health*, 13(3), 91-110.
- Ansari, Z., Laditka, J., & Laditka, S. (2006). Access to health care and hospitalization for ambulatory care sensitive conditions. *Medical Care Research and Review*, 63(6), 719-741.
- Anselin, L., & Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis* 23(2), 112-131.

- Asanin, J., & Wilson, K. (2008). "I spent nine years looking for a doctor": Exploring access to health care among immigrants in Mississauga, Ontario, Canada. *Social Science & Medicine*, 66(6), 1271-1283.
- Bailey, T.C., & Gatrell, A.C. (1995). *Interactive spatial data analysis*. Essex, England: Longman Scientific & Technical.
- Baumgardner, D.J., Halsmer, S.E., Steber, D.L., Shah, D., & Mundt, M.P. (2006). Does proximity to clinic affect immunisation rates and blood pressure? *International Journal of Psychiatry in Medicine*, 36(2), 199-209.
- Bhopal, R.S. (2002). *Concepts of epidemiology: An integrated introduction to the ideas, theories, principles, and methods of epidemiology*. Oxford: Oxford University Press.
- Billings, J., Anderson, G., & Newman, L. (1996). Recent findings on preventable hospitalizations. *Health Affairs*, 15(3), 239-249.
- Billings, J., Zeitel, L., Lukomnik, J., Carey, T., Blank, A., & Newman, L. (1993). Impact of socioeconomic status on hospital use in New York City. *Health Affairs*, 12(1), 163-173.
- Bindman, A., Grumbach, K., Osmond, D., Koramony, M., Vranizan, K., Lurie, N., Billings, J., & Stewart, A. (1995). Preventable hospitalizations and access to health care. *Journal of the American Medical Association*, 274(4), 305-311.
- Bland, J.M., & Altman, D.G. (1996). Statistics notes: Measurement error and correlation coefficients. *British Medical Journal*, 313, 41. doi:10.1136/bmj.313.7048.41.
- Brown, A., Goldacre, M., Hicks, N., Rourke, J., McMurtry, R., Brown, J., & Anderson, G. (2001). Hospitalization for ambulatory care-sensitive conditions: A method for comparative access and quality studies using routinely collected statistics. *Canadian Journal of Public Health*, 92(2), 155-159.
- Casanova, C., Colomer, C., & Starfield, B. (1996). Pediatric hospitalization due to ambulatory care-sensitive conditions in Valencia (Spain). *International Journal for Quality in Health Care*, 8(1), 51-59.
- CIHI. (2004a). *Data quality documentation, Discharge abstract database, 2002–2003*. Ottawa: CIHI.
- CIHI. (2004b). *Data quality documentation, Discharge abstract database, 2003–2004*. Ottawa: CIHI.
- CIHI. (2004c). *Data quality document – NACRS (FY 2002-2003)*. Ottawa: CIHI.
- CIHI. (2004d). *Data quality of the discharge abstract database following the first-year implementation of ICD-10-CA/CCI – Final report*. Ottawa: CIHI.

- CIHI. (2005a). *Data quality documentation, Discharge abstract database, 2004–2005*. Ottawa: CIHI.
- CIHI. (2005b). *Database background documentation: National ambulatory care reporting system (NACRS), FY 2003-2004*. Ottawa: CIHI.
- CIHI. (2005c). *Database background and general data limitations documentation, National ambulatory care reporting system, FY 2004-2005 – Executive summary*. Ottawa: CIHI.
- CIHI. (2006a). *Data quality documentation, Discharge abstract database, 2005–2006—Executive summary*. Ottawa: CIHI.
- CIHI. (2006b). *Database background and general data limitations documentation, National ambulatory care reporting system, FY 2005-2006 – Executive summary*. Ottawa: CIHI.
- CIHI. (2007a). *Health system performance: Effectiveness, technical notes*. Retrieved on September 18, 2010 from http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=tech_notes_e
- CIHI. (2007b). *Data quality documentation, Discharge abstract database, 2006–2007—Executive summary*. Ottawa: CIHI.
- CIHI. (2007c). *Database background and general data limitations documentation, National ambulatory care reporting system, 2006-2007 – Executive summary*. Ottawa: CIHI.
- CIHI. (2008a). *Reducing gaps in health: A focus on socio-economic status in urban Canada*. Ottawa: CIHI.
- CIHI. (2008b). *The cost of acute care hospital stays by medical condition in Canada, 2004-2005*. Ottawa: CIHI.
- CIHI. (2008c). *Data quality documentation, Discharge abstract database, 2007–2008—Executive summary*. Ottawa: CIHI.
- CIHI. (2008d). *Database background and general data limitations documentation, National ambulatory care reporting system (NACRS), 2007–2008—Executive summary*. Ottawa: CIHI.
- CIHI. (2009a). *Health care in Canada 2009: A decade in review*. Ottawa: CIHI.
- CIHI. (2009b). *Health indicators 2009*. Ottawa: CIHI.
- CIHI. (2009c). *Health system performance: Effectiveness, technical notes*. Retrieved on September 18 2010 from http://secure.cihi.ca/cihiweb/dispPage.jsp?cw_page=tech_notes_2009_e
- CIHI. (2009d). *Data quality documentation, Discharge abstract database, 2008–2009—Executive summary*. Ottawa: CIHI.

- CIHI. (2009e). *Database background and general data limitations documentation, National ambulatory care reporting system, 2008-2009 – Executive summary*. Ottawa: CIHI.
- CIHI. (2010a). *Health indicators 2010: Definitions, data sources and rationale*. Retrieved on October 18, 2011 from http://www.cihi.ca/CIHI-ext-portal/pdf/internet/DEFINITIONS_052010_EN
- CIHI. (2010b). *Data quality documentation, National ambulatory care reporting system, 2009-2010, Executive summary* (revised March 2011). Ottawa: CIHI.
- CIHI. (n.d.a) National Ambulatory Care Reporting System. Retrieved on January 15, 2011 from http://www.cihi.ca/CIHI-ext-portal/internet/en/document/types+of+care/hospital+care/emergency+care/services_nacrs
- CIHI. (2012). Analysis in brief: Disparities in primary health care experiences among Canadians with ambulatory care sensitive conditions. Ottawa: CIHI.CIHI. (n.d.b). ICD-10-CA. Retrieved on January 20, 2011 from http://www.cihi.ca/CIHI-ext-portal/internet/en/document/standards+and+data+submission/standards/classification+and+coding/codingclass_icd10
- City of Ottawa. (n.d.). *Ottawa population*. Retrieved on June 2, 2011 from http://ottawa.ca/en/city_hall/statisticsdata/statistics/census/ottawa/index.html
- Cliff, A.D. & Ord, J.K. (1981). *Spatial Processes: Models & Applications*. London: Pion Limited.
- Coker, R., McKee, M., Atun, R., Dimitrova, B., Dodonova, E., Kuznetsov, S., & Drobniowski, F. (2006). Risk factors for pulmonary tuberculosis in Russia: case-control study. *British Medical Journal*, 332, 85-87.
- Conger, A.J. (1974). A revised definition for suppressor variables: A guide to their identification and interpretation. *Educational and Psychological Measurement*, 34(1), 35-46.
- Coultas, D.B. (1998). Passive smoking and risk of adults asthma and COPD: an update. *Thorax*, 53(5), 381-387.
- Cresswell, T. (2004). *Place: A Short Introduction*. Oxford: Blackwell Publishing.
- Cromley, E., & McLafferty, S. (2002). Chapter 9: Analyzing access to health services. In *GIS and public health* (pp. 233-258). New York: The Guilford Press.
- Crouse, D., Ross, N., & Goldberg, M. (2009). Double burden of deprivation and high concentrations of ambient air pollution at the neighbourhood scale in Montreal, Canada. *Social Science & Medicine*, 69(6), 971-981.

- Curtis, S.E. (2004). Landscape of consumption: care and commodifications. In *Health and Inequality* (pp. 119-150). London: Sage Publications.
doi:10.4135/9781849209243
- Dahrouger, S., Hogg, W., Tuna, M., Russell, G., Devlin, R.A., Tugwell, P., & Kristjansson, E. (2011). Age equity in different models of primary care practices in Ontario. *Canadian Family Physician*, 57, 1300-1309.
- de Vaus, D. (2002). *Analyzing social science data - 50 key problems in data analysis*. London: Sage Publications.
- DeMaris, A. (2004). *Regression with social data: Modeling continuous and limited response variables*. Hoboken, NJ: Wiley.
- Demeter, S., Reed, M., Lix, L., MacWilliam, L., & Leslie, W.D. (2005). Socioeconomic status and the utilization of diagnostic imaging in an urban setting. *Canadian Medical Association Journal*, 173(10), 1173-1177.
- Diez Roux, A. (2003). The examination of neighborhood effects on health: Conceptual and methodological issues related to the presence of multiple levels of organization. In I. Kawachi & L. Berkman (Eds.), *Neighborhoods and Health* (pp. 48-50). New York: Oxford University Press.
- Disano, J., Goulet, J., Muhajarine, N., Neudorf, C., & Harvey, J. (2010). Socio-economic status and rates for hospital admission for chronic disease in urban Canada. *Canadian Nurse*, 106(1), 24-29.
- Dorofeev, S., & Grant, P. (2006). *Statistics for real-life sample surveys: Non-simple-random samples and weighted data*. Cambridge, UK: Cambridge University Press.
- Dunlop, S., Coyte, P., & McIsaac, W. (2000). Socio-economic status and the utilisation of physicians' services: Results from the Canadian national population health survey. *Social Science & Medicine*, 51(1), 123-133.
- Dunn, J.R., Schaub, P., & Ross, N.A. (2007). Unpacking income inequality and population health. *Canadian Journal of Public Health*, 97 (Suppl 1), S10-S17.
- Elliott, S., Cole, D., Krueger, P., Voorberg, N., & Wakefield, S. (1999). The power of perception: Health risk attributed to air pollution in an urban industrial neighborhood. *Risk Analysis*, 19(4), 621-634.
- Evans, R.G. & Stoddart, G.L. (1990). Producing health, consuming health care. *Social Science & Medicine*, 31(12), 1347-1363.
- Finkelstein, M.M., Jerrett, M., DeLuca, P., Finkelstein, N., Verma, D., Chapman, K., & Sears, R. (2003). Relation between income, air pollution and mortality: A cohort study. *Canadian Medical Association Journal*, 169(5), 397-402.

- Forthofer, R.N., Lee, E.S., & Hernandez, M. (2007). *Biostatistics: A guide to design, analyses, and discovery* (2nd ed.). Boston: Elsevier Academic Press.
- Fransoo, R.M. (2009). *Manitoba RHA indicators atlas 2009*. Winnipeg: Manitoba Centre for Health Policy.
- Frohlich, K., Potvin, L., & Chabot, P. (2002). Youth smoking initiation: Disentangling context from composition. *Health and Place*, 8(3), 155-166.
- Gale Encyclopedia of Public Health. (2012). *Standardization*. Retrieved on January 22, 2012 from <http://www.answers.com/topic/standardization>
- Galster, G.C. (2001). On the nature of neighbourhood. *Urban Studies* 38, 2111-2124.
- Garson, D. (2009). *Univariate GLM, ANOVA and MANCOVA* (Last modified on June 2 2009). Retrieved on April 16, 2010 from <http://faculty.chass.ncsu.edu/garson/PA765/anova.htm#ctests>
- Gatrell, A., & Elliott, S. (2008). Chapter 3: Methods and techniques in the geography of health. In *Geographies of health: An introduction* (pp. 49-84). Oxford: Blackwell Publishing.
- Gauvin, L., Robitaille, E., Riva, M., McLaren, L., Dassa, C., & Potvin, L. (2007). Conceptualizing and operationalizing neighbourhoods: The conundrum of identifying territorial units. *Canadian Journal of Public Health*, 97(Suppl 1), S18-S26.
- Gee, G., & Payne-Sturges, D. (2004). Environmental health disparities: A framework integrating psychosocial and environmental concepts. *Environmental Health Perspectives*, 112(17), 1645-1653.
- Gibson, D., Richards, H., & Chapman, A. (2008). The national ambulatory care reporting system: Factors that affect the quality of its emergency data. *International Journal of Information Quality*, 2(2), 97-114.
- Glazier, R., Klein-Geltink, J., Kopp, A., & Sibley, L. (2009). Capitation and enhanced fee-for-services models for primary care reform: A population-based evaluation. *Canadian Medical Association Journal*, 180(11), E72-E81.
- Goddard, M. & Smith, P. (2001). Equity of access to health care services: Theory and evidence from the UK. *Social Science & Medicine*, 53, 1149-1162.
- Goddard, M., & Smith, P. (2001). Equity of access to health care services: Theory and evidence from the UK. *Social Science & Medicine*, 53(9), 1149-1162.
- Green, S.B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26(3), 499-510.

- Gulliford, M.C., Sedgwick, J.E.C., & Pearce, A.J. (2003). Cigarette smoking, health status, socio-economic status and access to health care in diabetes mellitus: a cross-sectional survey. *BMC Health Services Research*, 3(4). doi:10.1186/1472-6963-3-4
- Gupta, S., Roos, L.L., Walld, R., Traverse, D., & Dahl, M. (2003). Delivering equitable care: Comparing preventive services in Manitoba. *American Journal of Public Health*, 92(12), 2086-2092.
- Guttmann, A., Shipman, S.A., Lam, K., Goodman, D.C. & Stukel, T.A. (2010). Primary care physician supply and children's health care use, access and outcomes: Findings from Canada. *Pediatrics*, 125(6), 1119-1126.
- Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., & Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science & Medicine*, 67(12), 2007-2016.
- Hawe, P. & Potvin, L. (2009). What is population health intervention research?. *Canadian Journal of Public Health*, 100(1), 1-8.
- Health Canada. (2003). *Health care system – 2003 First Ministers' Accord on Health Care Renewal* (Last modified on May 8, 2006). Retrieved on January 12, 2012 from <http://www.hc-sc.gc.ca/hcs-sss/delivery-prestation/fptcollab/2003accord/index-eng.php>
- Health Canada. (2004a). *Health care system*. Retrieved on July 31, 2010 from <http://www.hc-sc.gc.ca/hcs-sss/prim/index-eng.php>
- Health Canada. (2004b). *Health care system – First Ministers' Meeting on the Future of Health Care 2004 – A 10-year plan to strengthen health care* (Last modified on May 9, 2006). Retrieved on January 12, 2012 from <http://www.hc-sc.gc.ca/hcs-sss/delivery-prestation/fptcollab/2004-fmm-rpm/index-eng.php>
- Health Canada. (2006). *Health care system. About primary care*. Retrieved on February 2, 2010 from <http://www.hc-sc.gc.ca/hcs-sss/prim/about-apropos-eng.php>
- Health Council of Canada. (2010). *How do Canadians rate the health care system? Results from the 2010 Commonwealth Fund International Health Policy Survey*. Canadian Health Care Matters, Bulletin 4. Toronto: Health Council of Canada.
- Health Council of Canada. (2012). *Self-management support for Canadians with chronic health conditions: A focus for primary health care*. Ottawa: Health Council of Canada.
- Hill, A.B. (1965). The environment and disease: Association or causation. *Proceedings of the Royal Society of Medicine* 58, 295-300.

- Hiscock, R., Pearce, J., Blakely, T., & Witten, K. (2008). Is neighbourhood access to health care provision associated with individual-level utilisation and satisfaction?. *Health Services Research*, 43(6), 2183-2200.
- Hogg, W., Rowan, M., Russell, G., Geneu, R., & Muldoon, L. (2008). Framework for primary care organizations: The importance of a structural domain. *International Journal of Quality in Health Care*, 20(5), 308-313.
- Institute of Medicine. (1993). *Access to health care in America*. Washington: National Academy Press.
- Jepson, R.G., Harris, F.M., Platt, S., & Tannahill, C. (2010). The effectiveness of interventions to change six health behaviours: A review of reviews. *BMC Public Health*, 10, 538-554. doi:10.1186/1471-2458-10-538
- Kerr, A.W., Hall, H.K., & Kezurb, S.A. (2002). *Doing statistics with SPSS*. London: Sage Publications.
- Kleinbaum, D., Kupper, L., Nizam, A., & Muller, K.E. (2008). *Applied regression analysis and other multivariable methods*. Belmont: Thomson, Brooks and Cole.
- Kosteniuk, J., & Dickinson, H.D. (2003). Tracing the social gradient in the health of Canadians: primary and secondary determinants. *Social Science & Medicine*, 57(2), 263-276.
- Kristjansson, E., Calhoun, M., Leclair, S., Russell, K., Bilette, J.M., Sawada, M., He, J., McKinnon, J., Prince-Ware, S., & Hole, D. (2010). Neighbourhood and individual characteristics that predict healthy eating and overweight/obese. *Why Place Still Matters: Community Forum* (Ron Kolbus Centre, Ottawa, ON, March 9, 2010). Retrieved on August 16, 2010 from http://137.122.133.36/neighbourhoodstudy.ca/pubs/Forum_FoodEnvironmentbkj.pdf
- Krueger, P. M., & Chang, V. W. (2008). Being poor and coping with stress: Health behaviors and the risk of death. *American Journal of Public Health*, 98(5), 889-896.
- Künzli, N., McConnell, R., Bates, D., Bastain, T., Hricko, A., Lurmann, F., Avol, E., Gilliland, F., & Peters, J. (2003). Breathless in Los Angeles: The exhausting search for clean air. *American Journal of Public Health*. 93(9), 1494–1499.
- Kurtosis. (2010). *How to draw a funnel plot*. Retrieved on January 22, 2012 from <http://www.kurtosis.co.uk/technique/funnel/>
- Laditka, J., Laditka, S., & Probst, J. (2005). More may be better: Evidence of a negative relationship between physician supply and hospitalization for ambulatory care sensitive conditions. *Health Services Research*, 40(4), 1148-1166. doi:10.1111/j.1475-6773.2005.00403.x

- Lasser, K., Himmelstein, D., & Woolhandler, S. (2006). Access to care, health status and health disparities in the US and Canada: Results of a cross-national population-based survey. *American Journal of Public Health*, 96(7), 1300-1307.
- Lattin, J., Carroll, J., & Green, P. (2003). *Analyzing multivariate data*. Belmont, California: Brooks/Cole Cengage Learning.
- Lavoie, J.G., Forget, E.L., Prakash, T., Dahl, M., Martens, P. & O'Neil, J.D. (2010). Have investments in on-reserve health services and initiatives promoting community control improved First Nations' health in Manitoba? *Social Science & Medicine*, 71(4), 717-724.
- Law, M., Wilson, K., Eyles, J., Elliott, S., Jerrett, M., Moffat, T., & Luginaah, I. (2005). Meeting health need, accessing health care: The role of neighbourhood. *Health & Place*, 11(4), 367-377.
- Lawlor, D. A., Frankel, S., Shaw, M., Ebrahim, S., & Smith, G. (2003). Smoking and ill health: Does lay epidemiology explain the failure of smoking cessation programs among deprived populations?. *American Journal of Public Health*, 93(2), 266-270.
- Le, C.T. (2003). *Introductory biostatistics*. Hoboken, NJ: Wiley.
- LeBlanc, P. (2003). *Étude sur les formes d'organisation des services de santé en français dans l'Est de l'Ontario*. Ottawa: Praxis. Retrieved on July 23, 2012 from <http://www.rssfes.on.ca/files/uploads/rssfes-files/formeorganisationoct03reseauquestont.pdf>
- Lee, E., & Forthofer, R. (2006). *Analyzing complex survey data* (2nd ed.). London: Sage Publications.
- Lemstra, M., Neudorf, C., & Opondo, J. (2006). Health disparities by neighbourhood income. *Canadian Journal of Public Health*, 97(6), 435-439.
- Levy, P.S., & Lemeshow, S. (1999). *Sampling of populations, methods and applications*. New York: Wiley.
- Lin, M., Stieb, D., & Chen, Y. (2005). Coarse particulate matter and hospitalization for respiratory infections in children younger than 15 years in Toronto: A case-crossover analysis. *Pediatrics*, 116(2), e235-e240.
- Lui, C.K., & Wallace S.P. (2011). A common denominator: Calculating hospitalization rates for ambulatory care-sensitive conditions in California. *Preventing Chronic Disease*, 8(5), A103. Retrieved on May 18, 2012 from www.cdc.gov/pcd/issues/2011/sep/11_0013.htm
- Macintyre, S., & Ellaway, A. (2003). Neighbourhoods and health: An overview. In I. Kawachi & L. Berkman (Eds.), *Neighbourhoods and health* (pp. 20-42). New York: Oxford University Press.

- Macintyre, S., Macdonald, L., & Ellaway, E. (2008). Do poorer people have poorer access to local resources and facilities? The distribution of local resources by area deprivation in Glasgow, Scotland. *Social Science & Medicine*, 67(6), 900-914.
- Manitoba Centre for Health Policy. (2009). *Concept: Ambulatory care sensitive conditions*. Retrieved on May 18, 2010 from <http://mchp-appserv.cpe.umanitoba.ca/viewConcept.php?conceptID=1023>
- Marshall, M., Klazinga, N., Leatherman, S., Hardy, C., Bergmann, E., Pisco, L., Mattke, S. & Mainz, J. (2006). OECD health care quality indicator project: The expert panel on primary care prevention and health promotion. *International Journal for Quality in Health Care*, 18(1), 21-25.
- Matthews, S.A. (2008). The salience of neighbourhood; some lessons from sociology. *American Journal of Preventive Medicine* 34, 257-259.
- McCusker, J., Roberge, D., Lévesque, J.-F., Ciampi, A., Vadeboncoeur, A., Larouche, D., & Sanche, S. (2010). Emergency department visits and primary care among adults with chronic conditions. *Medical Care*, 48(11), 972-980.
- McDonald, J.T., & Kennedy, S. (2004). Insights into the Health Immigrant Effect: Health Incidence and Health Service Use of Immigrants to Canada. *Social Science & Medicine*, 19(8), 1613-1627.
- McIsaac, W., Goel, V., & Naylor, D. (1997). Socio-economic status and visits to physicians by adults in Ontario, Canada. *Journal of Health Services Research and Policy*, 2(2), 94-102.
- McKenna, H. (1994). The Delphi technique: A worthwhile research approach for nursing? *Journal of Advanced Nursing*, 19(6), 1221-1225.
- Meade, M., & Erickson, R. (2000). Chapter 10: Distributions of health care resources. In *Medical Geography* (2nd ed.) (pp. 343-375). New York: Guildford Press.
- Menec, V.H., Sirski, M., & Attawar, D. (2005). Does continuity of care matter in a universally insured population? *Health Services Research*, 40(2), 389-400.
- Montgomery, D.C., Peck, E.A., & Vining, G.G. (2006). *Introduction to linear regression analyses*. Hoboken, NJ: Wiley.
- National Capital Commission. (2009). *History and culture*. Retrieved on September 10, 2011 from http://www.canadacapital.gc.ca/bins/ncc_web_content_page.asp?cid=16297-16299-9735-9742&lang=1
- Ottawa Neighbourhood Study. (2010). *Ottawa Neighbourhood Study*. Retrieved on July 16, 2012 from <http://www.neighbourhoodstudy.ca/main.php>

- Ottawa Neighbourhood Study. (n.d.a). *Neighbourhood profile: Carleton Heights – Rideauview*. Retrieved on March 25, 2012 from http://www.neighbourhoodstudy.ca/profiles/index.php?page=Carleton_Heights_Rideauview
- Ottawa Neighbourhood Study. (n.d.b). *Neighbourhood profile: Byward Market*. Retrieved on March 25, 2012 from http://137.122.133.36/neighbourhoodstudy.ca/profiles/index.php?page=Byward_Market
- Pampalon, R., Hamel, D., Gamache, P., & Raymond, G. (2009). A deprivation index for health planning in Canada. *Chronic Diseases in Canada*, 29(4), 178-191.
- Pandey, S. & Elliott, W. (2010). Suppressor variables in social work research: Ways to identify in multiple regression models. *Journal of the Society for Social Work and Research*, 1(1), 28-40.
- Pannell, D.J. (1997). Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics*, 16(2), 139-152.
- Parenteau, M., Sawada, M., Kristjansson, E., Calhoun, M., Leclair, S., Labonté, R., Runnels, V., Musiol, A., & Herold, S. (2008). Development of neighbourhoods to measure spatial indicators of health. *URISA Journal*, 20(2), 43-55.
- Parenteau, M., & Sawada, M. (2011). The modifiable areal unit problem (MAUP) in the relationship between exposure to NO₂ and respiratory health. *International Journal of Health Geographies* 10: 58. doi:10.1186/1476-072X-10-58
- Pearce, J., Witten, K., & Bartie, P. (2006). Neighbourhoods and health: A GIS approach to measuring community resource accessibility. *Journal of Epidemiology and Community Health*, 60(5), 389-395.
- Pouliou, T., & Elliott, S. (2010). Individual and socio-environmental determinants of overweight and obesity in urban Canada. *Health and Place*, 16(2), 389-398.
- Romanow, R. (2002). *Building on values: The future of health care in Canada*. Commission on the Future of Health Care in Canada. Ottawa, Canada: National Library of Canada.
- Proctor, J.L., & Badzinski, D.M. (2002). *Introductory statistics for criminal justice and criminology*. Upper Saddle River, New Jersey: Prentice Hall.
- Roos, L., Walld, R., Uhanova, J., & Bond, R. (2005). Physician visits, hospitalizations and socioeconomic status: Ambulatory care sensitive conditions in a Canadian setting. *Health Services Research*, 40(4), 1167-1185.
- Roos, L.L., & Walld, R. (2007). Neighbourhood, family and health care. *Canadian Journal of Public Health*, 98 (Suppl. 1), S54-S61.

- Roos, N., & Mustard, C. (1997). Variation in health and health care use by socioeconomic status in Winnipeg, Canada: Does the system work well? Yes and no. *Milbank Quarterly*, 75(1), 89-11.
- Russell, K., Ali, A., Kristjansson, E., Billette, J.M., & Sawada, M. (2010). Preventative health care, socio-economic status and individual health outcomes in Ottawa Neighbourhoods. *Why Place Still Matters: Community Forum* (Ron Kolbus Centre, Ottawa, ON, March 9, 2010). Retrieved on August 16, 2010 from <http://www.neighbourhoodstudy.ca/pubs/CommunityForum2March%202010A CSHCutilization.pdf>
- Sanderson, C., & Dixon, J. (2000). Conditions for which onset of hospital admission is potentially preventable by timely and effective ambulatory care. *Journal of Health Services Research*, 5(4), 222-230.
- Sanmartin, C., & Khan, S. (2011). *Hospitalizations for ambulatory care sensitive conditions (ACSC): The factors that matter*. Ottawa: Statistics Canada Health Research Working Paper Series.
- Schmidt, F.L. (1971). The relative efficiency of regression in simple unit predictor weights in applied differential psychology. *Educational and Psychological Measurement*, 31, 699-714.
- Schoenbach, V.J., & Rosamond, W.D. (1999). Chapter 6 Standardization of rates and ratios. In *Understanding the fundamentals of epidemiology - an evolving text* (pp.129-151). Retrieved on June 1, 2011 from <http://www.epidemiolog.net/evolving/Standardization.pdf>
- Schreiber, S., & Zielinski, T. (1997). The meaning of ambulatory care sensitive admissions: Urban and rural perspectives. *The Journal of Rural Health*, 13(4), 276-284.
- Shah, B., Gunraj, N., & Hux, J. (2003). Markers of access to and quality of primary care for aboriginal people in Ontario, Canada. *American Journal of Public Health*, 93(5), 798-802.
- Solar, O, & Irwin A. (2010). *A conceptual framework for action on the social determinants of health. Social Determinants of Health Discussion Paper 2 (Policy and Practice)*. Geneva: WHO, Commission on the Social Determinants of Health.
- Starfield, B. (1998). *Primary care: Balancing health needs, services and technology*. New York: Oxford University Press.
- Statistics Canada. (2002a). *CCHS - Health Regions*. Retrieved on June 15, 2010 from <http://www.statcan.gc.ca/concepts/health-sante/regions-eng.htm>
- Statistics Canada. (2002b). *CCHS Cycle 1.1 (2000-2001), Public use microdata file documentation*. Ottawa: Statistics Canada.

- Statistics Canada. (2003). *CCHS - Health Regions*. Retrieved on June 15, 2010 from http://www.statcan.gc.ca/concepts/health-sante/cycle2_1/regions-eng.htm
- Statistics Canada. (2005a). *CCHS - Health Regions*. Retrieved on June 15, 2010 from http://www.statcan.gc.ca/concepts/health-sante/cycle3_1/regions-eng.htm
- Statistics Canada. (2005b). *Canadian community health survey 2003 - User guide for the public use microdata file*. Ottawa: Statistics Canada.
- Statistics Canada. (2006a). *Comparable health indicators: 12-PC hospitalization rate for ambulatory care sensitive conditions* (Last modified December 21, 2006). Retrieved on October 11, 2011 from <http://www.statcan.gc.ca/pub/82-401-x/2006000/considerations/pc/4064357-eng.htm>
- Statistics Canada. (2006b). *Canadian community health survey (CCHS) Cycle 3.1 (2005) - Public use microdata file (PUMF) user guide*. Ottawa: Statistics Canada.
- Statistics Canada. (2008). *Canadian Community Health Survey – Annual Components (CCHS)* (Last modified June 5, 2008). Retrieved on Nov 24, 2011 from <http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SurvId=3226&SurvVer=1&Instald=15282&InstaVer=4&SDDS=3226&lang=en&db=imdb&adm=8&dis=2>
- Statistics Canada. (2007). *2006 Community Profiles. 2006 Census. Statistics Canada Catalogue no. 92-591-XWE*. Ottawa. Retrieved on July 16, 2012 from <http://www12.statcan.ca/census-recensement/2006/dp-pd/prof/92-591/index.cfm?Lang=E>
- Statistics Canada. (2009). *Canadian community health survey (CCHS) - Annual component - User guide 2007-2008* (microdata files). Ottawa: Statistics Canada.
- Statistics Canada. (2010). *Canadian community health survey (CCHS) - Annual component - User guide 2009* (microdata files). Ottawa: Statistics Canada.
- Statistics Canada. (2011). *Application process and guidelines* (Last modified on March 22, 2011). Retrieved on January 22, 2012 from <http://www.statcan.gc.ca/rdc-cdr/process-eng.htm>
- Steiner, J., Braun, P., Melinkovich, P., Glazner, J., Chandramouli, V.L., & Davidson, A. (2003). Primary-care visits and hospitalizations for ambulatory-care-sensitive conditions in an inner-city health care system. *Ambulatory Pediatrics*, 3(6), 324-328.
- Stewart, M., Reutter, L., Makwarimba, E., Rootman, I., Williamson, D., Raine, K., Wilson, D., Fast, J., ..., Rutakumwa, W. (2005). Determinants of health-

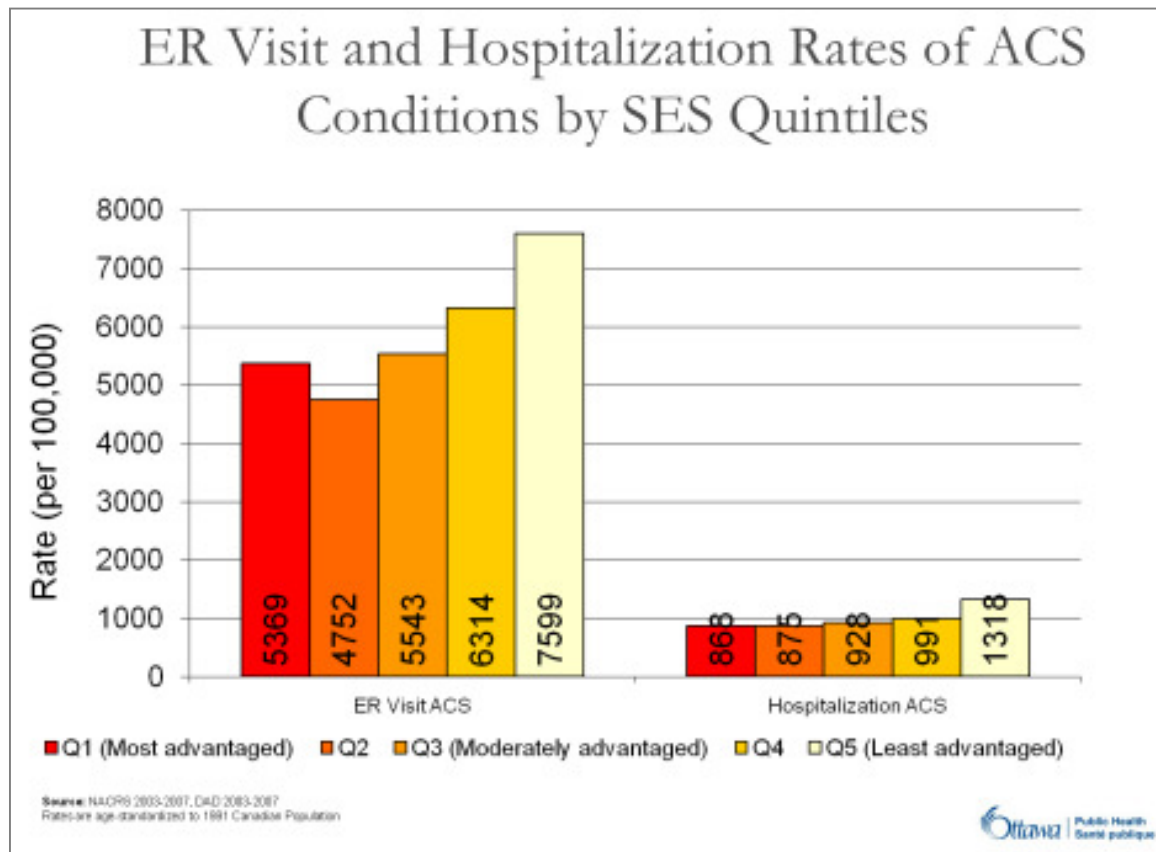
- service use by low-income people. *Canadian Journal of Nursing Research*, 37(3), 104-131.
- Sturgis, P. (2004). *Analysing complex survey data: Clustering, stratification and weights*. Social Research Update Issue 43. Surrey: University of Surrey, Department of Sociology.
- Tabachnick, B.G., & Fidell, L.S. (1989). *Using multivariate statistics* (2nd ed.). Cambridge, MA: Harper & Row.
- Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.
- Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson Education / Allyn & Bacon.
- Thomas, S., Saraffin, C., & Simard, M. (2007). Review of the weighting strategy for the Canadian Community Health Survey. *SSC Annual meeting, June 2007. Proceedings of the survey methods section*. Ottawa: Statistics Canada. Retrieved on November 9, 2010 from http://www.ssc.ca/survey/documents/SSC2007_C_Saraffin.pdf
- Thomas, S., & Tremblay, S. (2010). *Interpreting estimates from the redesigned Canadian Community Health Survey (CCHS)* (Last modified on June 14, 2010). Ottawa: Statistics Canada. Retrieved on November 7, 2011 from http://www.statcan.gc.ca/imdb-bmdi/document/3226_D64_T9_V1-eng.htm.
- Thomas, S., & Wannell, B. (2009). Combining cycles of the Canadian Community Health Survey. *Health Reports*, 20(1), 53-58.
- Townsend, P. (1987). Deprivation. *Journal of Social Politics*, 16(1), 125-146.
- Trochim, W.M., & Donnelly, J.P. (2008). *The research methods knowledge base* (3rd ed). Mason, OH: Atomic Dog.
- Trosset, M.W. (2006). *An introduction to statistical inference and its applications*. Retrieved on November 12, 2011 from http://www.astrohandbook.com/ch17/statistical_inference.pdf
- Tu, Y.K., Gunnell, D., & Gilthorpe, M.S. (2008). Simpson's Paradox, Lord's Paradox and Suppression Effects are the same phenomenon – the reversal paradox. *Emerging Themes in Epidemiology*, 5(2), 1-9.
- Valery, L., Anke, O., Inge, K.K. & Johannes, D. (2008). Effectiveness of smoking cessation interventions among adults: a systematic review of reviews. *European Journal of Cancer Prevention* 17, 535-544.
- van Gageldonk-Lafeber, A.B., van der Sande, M.A., Heijnen, M.L.A., Peeters, M.F., Bartelds, I.M. & Wilbrink, B. (2007). Risk factors for acute respiratory tract

- infections in general practitioner patients in The Netherlands: a case-control study. *BMC Infectious Diseases* 7, 35. doi:10.1186/1471-2334-7-35
- Vyas, S., & Kumaranayake, L. (2006). Constructing socioeconomic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459-468.
- Wakefield, S., & McMullan, C. (2005). Healing in places of decline: (re)imagining everyday landscapes in Hamilton, Ontario. *Health and Place*, 11(4), 299-312.
- Walker, J., Teare, G.F., Hogan, D.B., Lewis, S., & Maxwell, C. (2009). Identifying potentially avoidable hospital admissions from Canadian long-term care facilities. *Medical Care*, 47(2), 250-254.
- Ward, M.D., & Gleditsch, K.S. (2007). *An Introduction to Spatial Regression Models in the Social Sciences*. Colchester, England: University of Essex. Retrieved on July 16, 2012 from <http://www.duke.edu/web/methods/pdfs/SRMbook.pdf>
- Warner, R.M. (2008). *Applied statistics: From bivariate through multivariate techniques*. London: Sage Publications.
- Weisberg, S. (2005). *Applied linear regression (3rd ed.)*. Hoboken, NJ: Wiley.
- Whitehead, M. (1992). The concepts and principles of equity and health. *International Journal of Health Services*, 22(4), 429-445.
- Wilkins, K., Berthelot, J.-M., & Ng, E. (2002). Trends in mortality by neighbourhood income in urban Canada from 1971 to 1996. *Supplement to Health Reports*, 13, 1-28.
- World Health Organisation. (2002). The world health report 2002: Reducing risks, promoting healthy life. Geneva: WHO.
- Worthington, C., McLeish, K., & Fuller-Thomson, E. (2012). Adherence over time to cervical cancer screening guidelines: insights from the Canadian National Population Health Survey. *Journal of Women's Health*, 21(2), 199-208.

Appendix 1

ER visits and hospitalisations rates for 26 ACS conditions from 2003 to 2007 for the 91 Ottawa neighbourhoods sorted by SES quintiles (Previous Study)

Figure A-1: ACS ER Visit and Hospitalisation Rates by SES Quintiles



(Russell, Ali, Kristjansson, Billette, & Sawada, 2010)

Appendix 2

Ottawa Neighbourhoods Population, SES Quintile and ACS Hospitalisation and ER Rates

Table A-1: List of Ottawa Neighbourhoods and Selected Characteristics

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
5	Barrhaven	17,902	2	490.8 (442.2, 539.4)	0.82	3013.2 (2915.3, 3111.0)	0.86
6	Bayshore	7,878	5	920.9 (830.9, 1010.9)	1.53	4717.7 (4554.4, 4881.0)	1.34
7	Beacon Hill South - Cardinal Heights	6,903	4	575.3 (511.2, 639.4)	0.96	3380.1 (3233.8, 3526.4)	0.96
8	Beaverbrook	5,237	3	468.4 (400.7, 536.1)	0.78	3166.3 (2994.4, 3338.1)	0.90
9	Bells Corners East	4,510	3	724.5 (641.8, 807.2)	1.21	4514.9 (4303.2, 4726.6)	1.28
10	Bells Corners West	4,392	4	807.3 (694.9, 919.7)	1.34	4502.1 (4297.5, 4706.7)	1.28
11	Billings Bridge - Alta Vista	11,807	4	473.0 (427.0, 519.1)	0.79	3246.8 (3136.9, 3356.7)	0.92
85	Blackburn Hamlet	8,493	3	515.5 (457.0, 574.1)	0.86	2864.0 (2734.7, 2993.3)	0.81
13	Borden Farm - Stewart Farm - Parkwood Hills - Fisher Glen	10,178	4	479.1 (430.1, 528.2)	0.80	3066.2 (2946.7, 3185.7)	0.87
14	Braemar Park - Bel Air Heights - Copeland Park	7,393	3	776.5 (705.3, 847.8)	1.29	3894.4 (3744.7, 4044.1)	1.11
15	Briar Green - Leslie Park	4,935	3	629.0 (551.6, 706.4)	1.05	4096.8 (3914.0, 4279.7)	1.16
82	Bridlewood - Emerald Meadows	19,129	1	408.1 (366.7, 449.4)	0.68	2664.0 (2573.2, 2754.8)	0.76
94	Britannia Village	7,806	5	920.9 (841.6, 1000.3)	1.53	5348.9 (5190.1, 5507.8)	1.52
95	ByWard Market	4,525	3	999.4 (892.3, 1106.5)	1.66	5378.6 (5193.3, 5564.0)	1.53

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
61	CFB Rockcliffe-NRC	5,392	5	891.5 (803.3, 979.7)	1.48	5722.2 (5527.0, 5917.5)	1.63
16	Carleton Heights – Rideauview	6,353	5	415.8 (353.0, 478.5)	0.69	3464.5 (3305.6, 3623.4)	0.98
18	Carlington	10,107	5	1148.4 (1066.6, 1230.1)	1.91	6068.7 (5924.8, 6212.5)	1.72
20	Carlingwood West - Glabar Park - McKellar Heights	5,298	3	445.6 (380.7, 510.4)	0.74	3179.3 (3008.6, 3350.0)	0.90
75	Carp - Hardwood Plains*	8,981	1	499.2 (440.7, 557.7)	0.83	4244.7 (4092.8, 4396.6)	1.21
21	Carson Grove - Carson Meadows	8,031	4	712.9 (648.9, 777.0)	1.19	3874.6 (3734.0, 4015.2)	1.10
23	Centrepoinste	7,300	2	457.7 (396.3, 519.1)	0.76	2651.2 (2512.4, 2790.0)	0.75
24	Centretown	24,173	4	588.0 (550.0, 626.0)	0.98	3337.2 (3261.6, 3412.9)	0.95
90	Chapman Mills - Rideau Crest - Davidson Heights	14,750	1	365.4 (312.4, 418.3)	0.61	2761.0 (2645.5, 2876.4)	0.78
26	Cityview - Skyline - Fisher Heights	6,277	4	460.4 (396.5, 524.3)	0.77	3362.3 (3205.8, 3518.8)	0.96
27	Civic Hospital-Central Park	9,920	3	430.8 (378.6, 482.9)	0.72	2923.2 (2805.3, 3041.0)	0.83
28	Crestview – Meadowlands	8,521	4	606.1 (547.7, 664.4)	1.01	3858.5 (3720.2, 3996.8)	1.10
93	Crystal Bay - Lakeview Park	4,146	2	487.5 (414.3, 560.7)	0.81	3368.7 (3164.7, 3572.7)	0.96
1	Cumberland*	4,503	1	778.8 (679.3, 878.2)	1.30	3124.4 (2945.7, 3303.0)	0.89
71	Cummings	8,650	5	954.6 (873.5, 1035.8)	1.59	4887.2 (4737.0, 5037.4)	1.39
69	East Industrial	8,257	4	256.0 (213.6, 298.5)	0.43	1447.6 (1350.3, 1545.0)	0.41
70	Elmvale - Eastway - Riverview - Riverview Park West	15,178	4	797.3 (745.3, 849.3)	1.33	4565.3 (4455.8, 4674.9)	1.30

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
29	Emerald Woods - Sawmill Creek	5,133	5	619.6 (533.8, 705.3)	1.03	3214.9 (3041.2, 3388.5)	0.91
2	Fitzroy Harbour - West Carleton*	11,371	2	604.8 (550.5, 659.1)	1.01	6077.5 (5938.1, 6216.9)	1.73
30	Glebe - Dows Lake	10,920	1	346.1 (302.2, 390.0)	0.58	2326.6 (2225.7, 2427.6)	0.66
77	Glen Cairn - Kanata South Business Park	7,875	3	943.1 (863.3, 1022.9)	1.57	4578.3 (4429.4, 4727.3)	1.30
88	Greely*	11,681	1	497.5 (445.6, 549.3)	0.83	2838.9 (2725.9, 2951.9)	0.81
92	Greenbelt	5,667	2	539.5 (465.9, 613.0)	0.90	3204.1 (3039.8, 3368.4)	0.91
31	Greenboro East	9,561	4	491.3 (425.6, 557.0)	0.82	2956.8 (2825.7, 3088.0)	0.84
32	Hawthorne Meadows - Sheffield Glen	6,465	5	768.6 (685.5, 851.7)	1.28	4290.6 (4125.3, 4455.9)	1.22
33	Hintonburg – Mechanicsville	9,542	5	867.9 (797.8, 937.9)	1.44	4700.6 (4568.4, 4832.9)	1.34
34	Hunt Club - Ottawa Airport	4,385	4	334.4 (257.6, 411.3)	0.56	2066.5 (1906.6, 2226.5)	0.59
35	Hunt Club East - Western Community	8,563	4	475.0 (415.0, 535.1)	0.79	2918.1 (2786.9, 3049.2)	0.83
68	Hunt Club Park	7,113	3	568.4 (486.4, 650.3)	0.95	3531.9 (3370.2, 3693.6)	1.00
12	Hunt Club Upper -Blossom Park – Timbermill	7,222	3	568.9 (494.0, 643.7)	0.95	3081.8 (2933.4, 3230.1)	0.88
80	Hunt Club Woods - Quintarra – Revelstoke	5,649	1	380.5 (323.9, 437.2)	0.63	2330.9 (2185.1, 2476.8)	0.66
36	Iris	6,816	5	868.3 (785.6, 951.1)	1.45	4673.1 (4509.4, 4836.8)	1.33
37	Island Park	5,165	1	470.5 (394.6, 546.4)	0.78	2609.1 (2449.7, 2768.4)	0.74
79	Kanata Lakes - Marchwood Lakeside - Morgan's Grant - Kanata North Business Park	18,214	1	343.2 (304.0, 382.3)	0.57	2397.2 (2308.0, 2486.5)	0.68

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
3	Kars – Osgoode*	8,057	2	687.9 (617.6, 758.2)	1.15	4501.9 (4341.6, 4662.3)	1.28
38	Katimavik – Hazeldean	14,750	3	544.6 (495.1, 594.0)	0.91	2845.5 (2747.7, 2943.2)	0.81
19	Laurentian	9,517	3	403.6 (357.5, 449.7)	0.67	2431.7 (2315.9, 2547.6)	0.69
40	Ledbury - Heron Gate - Ridgemont – Elmwood	13,648	5	709.9 (648.0, 771.9)	1.18	4368.9 (4247.7, 4490.2)	1.24
41	Lindenlea - New Edinburgh	5,253	1	461.7 (395.8, 527.7)	0.77	2665.3 (2514.2, 2816.5)	0.76
42	Lowertown	8,166	5	1058.1 (974.2, 1142.0)	1.76	5370.1 (5216.3, 5523.9)	1.53
86	Manotick - North Gower*	9,151	1	407.7 (356.3, 459.2)	0.68	3526.6 (3388.3, 3664.9)	1.00
72	Merivale Gardens - Grenfell Glen - Pineglen - Country Place	2,719	1	408.0 (327.9, 488.1)	0.68	2975.9 (2735.2, 3216.7)	0.85
4	Metcalfe*	6,216	1	591.4 (516.1, 666.7)	0.98	4849.5 (4664.9, 5034.0)	1.38
87	Munster Hamlet – Richmond*	11,765	2	547.7 (496.7, 598.7)	0.91	4649.3 (4520.1, 4778.4)	1.32
89	Navan – Vars*	9,747	1	691.8 (622.8, 760.7)	1.15	2425.4 (2311.9, 2538.9)	0.69
78	New Barrhaven – Stonebridge	21,328	2	444.0 (399.8, 488.1)	0.74	2801.6 (2713.2, 2890.1)	0.80
83	Orleans Avalon - Notting Gate - Fallingbrook - Gardenway South	31,058	2	430.7 (391.9, 469.5)	0.72	2149.8 (2080.6, 2219.0)	0.61
67	Orleans Central	3,461	3	902.7 (796.1, 1009.2)	1.50	3322.4 (3121.5, 3523.3)	0.94
25	Orleans Chapel Hill	8,550	2	308.2 (252.5, 364.0)	0.51	1908.8 (1789.8, 2027.8)	0.54
84	Orleans Chapel Hill South	5,561	1	274.0 (210.5, 337.5)	0.46	1880.4 (1730.1, 2030.8)	0.53
74	Orleans Chatelaine Village	4,672	3	544.8 (449.1, 640.5)	0.91	2149.7 (1995.2, 2304.3)	0.61

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
64	Orleans North West	11,141	2	541.3 (484.4, 598.2)	0.90	2536.5 (2428.6, 2644.4)	0.72
48	Orleans Queenswood Heights	13,452	2	493.9 (441.1, 546.8)	0.82	2261.4 (2162.5, 2360.3)	0.64
43	Orleans Village – Chateaufneuf	12,911	2	451.1 (400.5, 501.8)	0.75	2068.8 (1974.2, 2163.4)	0.59
97	Ottawa East	4,878	4	372.1 (297.8, 446.4)	0.62	2762.5 (2596.9, 2928.1)	0.78
96	Ottawa South	8,153	2	369.4 (313.1, 425.7)	0.62	2417.1 (2291.0, 2543.1)	0.69
44	Overbrook – McArthur	11,463	5	1051.6 (983.1, 1120.1)	1.75	5372.1 (5242.1, 5502.0)	1.53
45	Pineview	5,566	4	693.7 (603.8, 783.6)	1.15	3840.4 (3664.6, 4016.2)	1.09
46	Playfair Park - Lynda Park - Guildwood Estates	6,041	2	398.1 (341.6, 454.5)	0.66	2693.0 (2537.2, 2848.8)	0.77
47	Qualicum - Redwood Park	4,318	4	619.2 (529.4, 709.1)	1.03	4656.9 (4444.3, 4869.6)	1.32
49	Riverside Park	4,610	3	456.1 (386.3, 525.9)	0.76	3017.8 (2843.0, 3192.6)	0.86
91	Riverside South – Leirtrim	7,104	1	383.2 (312.1, 454.4)	0.64	2485.9 (2337.6, 2634.2)	0.71
60	Rockcliffe - Manor Park	5,228	2	474.7 (401.6, 547.8)	0.79	3266.4 (3092.8, 3440.1)	0.93
73	Rothwell Heights - Beacon Hill North	10,311	2	589.5 (538.6, 640.5)	0.98	2851.8 (2734.9, 2968.7)	0.81
50	Sandy Hill - Ottawa East	10,879	5	551.6 (492.5, 610.7)	0.92	3311.7 (3189.9, 3433.5)	0.94
52	South Keys - Heron Gate - Greenboro West	4,667	3	279.7 (222.8, 336.6)	0.47	1986.3 (1838.8, 2133.9)	0.56
76	Stittsville	15,470	1	482.8 (438.3, 527.2)	0.80	3472.0 (3360.5, 3583.5)	0.99
53	Tanglewood	4,895	3	511.3 (434.9, 587.6)	0.85	3451.4 (3276.0, 3626.9)	0.98

ID	Neighbourhood	Population	SES Quintile	Hosp. Rate [¥]	Hosp. Ratio	ER Rate [¥]	ER. Ratio
51	Trend-Arlington	3,918	2	450.9 (374.6, 527.2)	0.75	3592.4 (3388.7, 3796.0)	1.02
63	Vanier North	8,272	5	1161.5 (1080.6, 1242.4)	1.93	6174.5 (6023.1, 6325.9)	1.75
54	Vanier South	7,052	5	1002.9 (924.4, 1081.4)	1.67	5267.1 (5108.0, 5426.2)	1.50
55	West Centertown	11,951	5	849.3 (786.7, 911.9)	1.41	4284.1 (4165.9, 4402.3)	1.22
56	Westboro	9,437	3	678.1 (617.6, 738.7)	1.13	3721.1 (3595.4, 3846.7)	1.06
57	Whitehaven - Queensway Terrace North	11,355	5	989.8 (923.9, 1055.7)	1.65	4885.3 (4758.2, 5012.3)	1.39
59	Woodroffe - Lincoln Heights	4,356	4	872.9 (779.7, 966.2)	1.45	4386.2 (4177.4, 4595.0)	1.25
58	Woodvale - Craig Henry - Manordale - Estates of Arlington Woods	8,683	4	594.7 (531.1, 658.3)	0.99	3682.8 (3544.5, 3821.1)	1.05

*Considered rural

[¥] The variance and standard errors of the standardised rates were calculated according to Chiang approach – based on the binomial distribution – presented in Andersen and Keiding (2006). This method assumed that: 1) the standard rates and population were stable and their sampling errors can be ignored, 2) the neighbourhood population were fixed and 3) the events in each age and gender group were independent. Therefore, the age and gender specific counts of events in the neighbourhood population (embedded in rate r_i) were the only random variable considered while deriving the variances and standard errors of the rates. Their variance was calculated following the formula:

$$\frac{r_i(2 - y_i r_i)}{r_i(2 + y_i r_i)}$$

Where:

- r_i was the event rate in the age and gender group i of the neighbourhood
- y_i was the number of years in the age group i

The variance of the 20-79 years old standardised rates themselves was $\sum w_i^2 \text{var}(r_i)$, where the weights $w_i = N_i / N$. Their standard errors were the square roots of the variances.

With regard to estimating the confidence intervals from the standard errors calculated, it was assumed that the rates follow a normal distribution. Andersen and Keiding (2006) proposed the basic method – to add and subtract $1.96 \times \text{standard error}$ from the point estimate – but cautioned the user about the possibility of deriving negative confidence intervals when the rates are small. This was not a concern for the confidence intervals of ACS hospitalisation or the ACS ER visits rates.

Appendix 3

Application Process and Approval – Secured Access to CCHS Master Files - COOL RDC

Secure access to the confidential master files of the CCHS was a necessity for the current investigation. Such access is available at the Carleton, Ottawa, Outaouais Local Research Data Centre (COOL RDC) located in the Morisset Library of the University of Ottawa and provided at no charge to university-based professors and students for research projects approved by the Social Sciences and Humanities Research Council (SSHRC).

A Master student needing access to confidential data at the COOL RDC for a project that is an integral part of his/her thesis can be granted access to the centre. The student must submit an application online on the SSHRC website in collaboration with her/his supervisor. This application includes an online form, a support letter from the supervisor, a 5-page project proposal highlighting the necessity of confidential data access, a document listing the research contributions of the student and the supervisor as well as their online CVs. The application is reviewed by a Statistics Canada analyst based on:

- "the scientific merit and viability of the proposed research;
- the relevance of the methods to be applied and the data to be analyzed;
- the demonstrated need for access to detailed microdata; and
- the expertise and ability of the researchers to carry out the proposed research as illustrated in the CVs and list of contributions" (Statistics Canada, 2011, Step 3).

Following approval of the project, the applicant must complete the security screening process, where Statistics Canada performs an Enhanced Reliability Check on the researchers needing access to the data. Conditional on the approval of the research project and the security status of the applicants, a contract between the researchers and Statistics Canada is signed and access to the microlevel data is granted.

For the current investigation, an application was submitted by Geneviève Prud'homme (M.Sc. Candidate) and Elizabeth Kristjansson (supervisor) to SSHRC on September 30, 2009. The Statistics Canada reviewer pointed some data limitations and requested additional information on the methodology that would be used to address them. A re-submission based on the reviewer's comments was submitted and access to the CCHS master files was granted to the researchers. A contract with Statistics Canada was signed on December 9th 2009, allowing the researchers to have access to the CCHS cycles needed for the current investigation until January 2012. An extension of this contract was requested and obtained, providing access to the CCHS data until December 2012 to the researchers. Additional details on the content of the application and re-submission is available upon request.

Appendix 4

Summary of Variables of Interest and Related Information

Table A- 2: List of Variables of Interest and Related Information

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Neighbourhood Outcome Characteristic							
Preventable hospitalisation	1	ACSC Hospitalisation Rate (Hosp_Rate)	Discharge Abstract Database (DAD) – CIHI	Rate – Continuous	• 19 ACS conditions combined • Patients aged 20-79 years old • Fiscal year 2002-2003 to 2009-2010 timeframe • Age- and gender- standardized using direct method and 10 years age intervals • Standardised using the 2006 Ottawa census population	Yes	Yes
Neighbourhood Compositional Characteristics							
Socio-economic status: Material and social deprivation	2	SES Index (SESindexcont)	Census (2006)	Index score – Continuous	• Variables included in PCA (principal component analysis) (Parenteau et al., 2008) • % with less than high school education • % single-parent families • % unemployed residents • % households below LICO • Average household income	Yes	Yes
	3	SES Quintiles	Census (2006)	Categorical – Ordinal	• Based on the SES index score, the neighbourhoods were divided into quintiles where Q1 is comprised of the most affluent neighbourhoods and Q5 of the most deprived neighbourhoods. • See above for elements of the SES index score.	Yes	Yes

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Health Status	4	Percent with fair or poor self-rated health status (health_status4_5_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7309 respondents - No neighbourhood-level imputation performed	Yes	Yes
	5	Percent with 1 or more ACS conditions (chronic_dis_1_plus_ag_st)	CCHS (all cycles combined)	% – Continuous	- Representing the percent of CCHS respondents who mentioned have been diagnosed with one or more of the following ACS chronic conditions: - Asthma - Heart disease - Diabetes - High blood pressure - Ottawa sample size = 7282 respondents	No	No
	6	Percent with 2 or more ACS conditions (chronic_dis_2_plus_ag_st)	CCHS (all cycles combined)	% – Continuous	- Representing the percent of CCHS respondents who mentioned have been diagnosed with two or more of the following ACS chronic conditions: - Asthma - Heart disease - Diabetes - High blood pressure - Ottawa sample size = 7282 respondents	No	No
Healthy Behaviour	7	Percent of people smoking (smoker_ag_st)	CCHS (all cycles combined)	% – Continuous	- Refers to the respondents smoking daily or occasionally - Item asked to all CCHS respondents - Ottawa sample size = 7285 respondents - No neighbourhood imputation performed	Yes	Yes
	8	Percent of people binge drinking in past year (binge_drink_yr_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7194 respondents - No neighbourhood imputation performed	No	No

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Healthy Behaviour	9	Percent of people binge drinking in past month (binge_drink_mth_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7194 respondents - No neighbourhood imputation performed	Yes	No
Utilisation and quality of health care services	10	Mean number of contacts with a primary care physician in the past year (no_visit_gp_ag_st)	CCHS (all cycles combined)	Number – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7297 respondents - No neighbourhood imputation performed - Neighbourhood mean influenced by extreme values - Variable closely related to concept of health care need or neighbourhood health status	No	No
	11	Percent who did not visit or talk with a primary care physician in past year (didnt_visit_gp_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7297 respondents - No neighbourhood imputation performed - Representing both respondents who did not have an annual physical exam as well as those who were sick and did not seek care – therefore concept closely related to propensity to use health services	Yes	Yes
	12	Mean number of contacts with a nurse the past year (no_visit_nurse_ag_st)	CCHS (all cycles combined)	Number – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7307 respondents - No neighbourhood imputation performed - Neighbourhood mean influenced by extreme values	No	No
	13	Percent who did not visit or talk with a nurse in past year (didnt_visit_nurse_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7307 respondents - No neighbourhood imputation performed	Yes	No

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Utilisation and quality of health care services	14	Mean number of contacts with a medical specialist in the past year (no_visit_otherdoc_ag_st)	CCHS (all cycles combined)	Number – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7306 respondents - No neighbourhood imputation performed - Neighbourhood mean influenced by extreme values - Variable representing both concept of health care need and quality of primary care received.	No	No
	15	Percent who did not visit or talk with a medical specialist in past year (didnt_visit_otherdoc_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7306 respondents - No neighbourhood imputation performed - Variable representing both concept of health care need and quality of primary care received.	Yes	No
	16	Percent who never received a flu shot (flu_ag_st)	CCHS (all cycles combined)	% – Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7228 respondents - No neighbourhood imputation performed	Yes	No
	17	Percent who never received a Pap test (pap_input)	CCHS (not available for Cycle 2009 - optional content)	% – Continuous	- Item asked to women 18+ years old among the CCHS respondents - Ottawa sample size = 3559 respondents 13 neighbourhood imputations performed - This variable was age-standardized using the female population only.	Yes	Yes
	18	Percent who never received a breast examination (breast_input)	CCHS (not available for Cycle 4.1 and Cycle 2009 - optional content)	% – Continuous	- Item asked to women 18+ years old among the CCHS respondents - Ottawa sample size = 2662 respondents 21 neighbourhood imputations performed - This variable was age-standardized using the female population only.	Yes	No

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Utilisation and quality of health care services	19	Percent who never received a mammography (mammo_input)	CCHS (all cycles combined)	% – Continuous	- Item asked to women 35+ years old among the CCHS respondents - Ottawa sample size = 2575 respondents 23 neighbourhood imputations performed - This variable was age-standardized using the female population above 30 years old only.	Yes	No
	20	Percent who rated the quality of the care provided by physician as being fair or poor (quality_doc_input)	CCHS (not available for Cycle 3.1 and respondent-s of year 2008 in Cycle 4.1)	% – Continuous	- Item asked only to CCHS respondents who received care by a physician (either a primary care physician or a medical specialist) as their last contact with health care system - Ottawa sample size = 2448 respondents 31 neighbourhood imputations performed - Concept representing patient satisfaction of the quality of physician care received, regardless of type of services received.	Yes	No
	21	Rate of ER visits for ACSC (ER_rate)	National Ambulatory Care Reporting System (NACRS) – CIHI	Rate – Continuous	19 ACS conditions combined - Patients aged 20-79 years old - Excluding patients who were subsequently admitted to the hospital (based on quality of data in NACRS) - Fiscal years 2002-2003 to 2009-2010 - Age- and gender- standardized using direct method and 10 years age intervals - Standardised using the 2006 Ottawa census population - Variable potentially representing primary care access issues or reduced service continuity	Yes	No

Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Neighbourhood Contextual Characteristics							
Primary Care Access	22	Average distance to closest four primary care facilities (dist_4_physician)	ONS – C.T. Lamont Primary Health Care Research Centre	Number - Continuous	- Expressed in meters - Calculated from neighbourhood population centroid using network analyses of roads - Representing concept of primary care accessibility	Yes	Yes
	23	Primary care physician-to-population ratio (physician_1000)	ONS – C.T. Lamont Primary Health Care Research Centre	Rate - Continuous	- Expressed by 1000 people - Calculated based on number of physicians within neighbourhood itself - Representing concept of primary care availability	Yes	Yes
	24	Percent without a family physician (fam_doc_ag_st)	CCHS (all cycles combined)	% - Continuous	- Item asked to all CCHS respondents - Ottawa sample size = 7310 respondents - No neighbourhood imputations performed - Representing the concept of primary care availability - Could also be perceived as a compositional characteristic	Yes	Yes
	25	Percent who rated the availability of health care in the community as being fair or poor (availability_com_input)	CCHS (not available for Cycle 1.1 and Cycle 3.1)	% - Continuous	- Item asked to all CCHS respondents aged 15+ - Ottawa sample size = 5470 respondents 9 neighbourhood imputations performed - Representing the concept of availability and accommodation of health care - Item not specific to primary care only, but to the health care as a whole	Yes	No

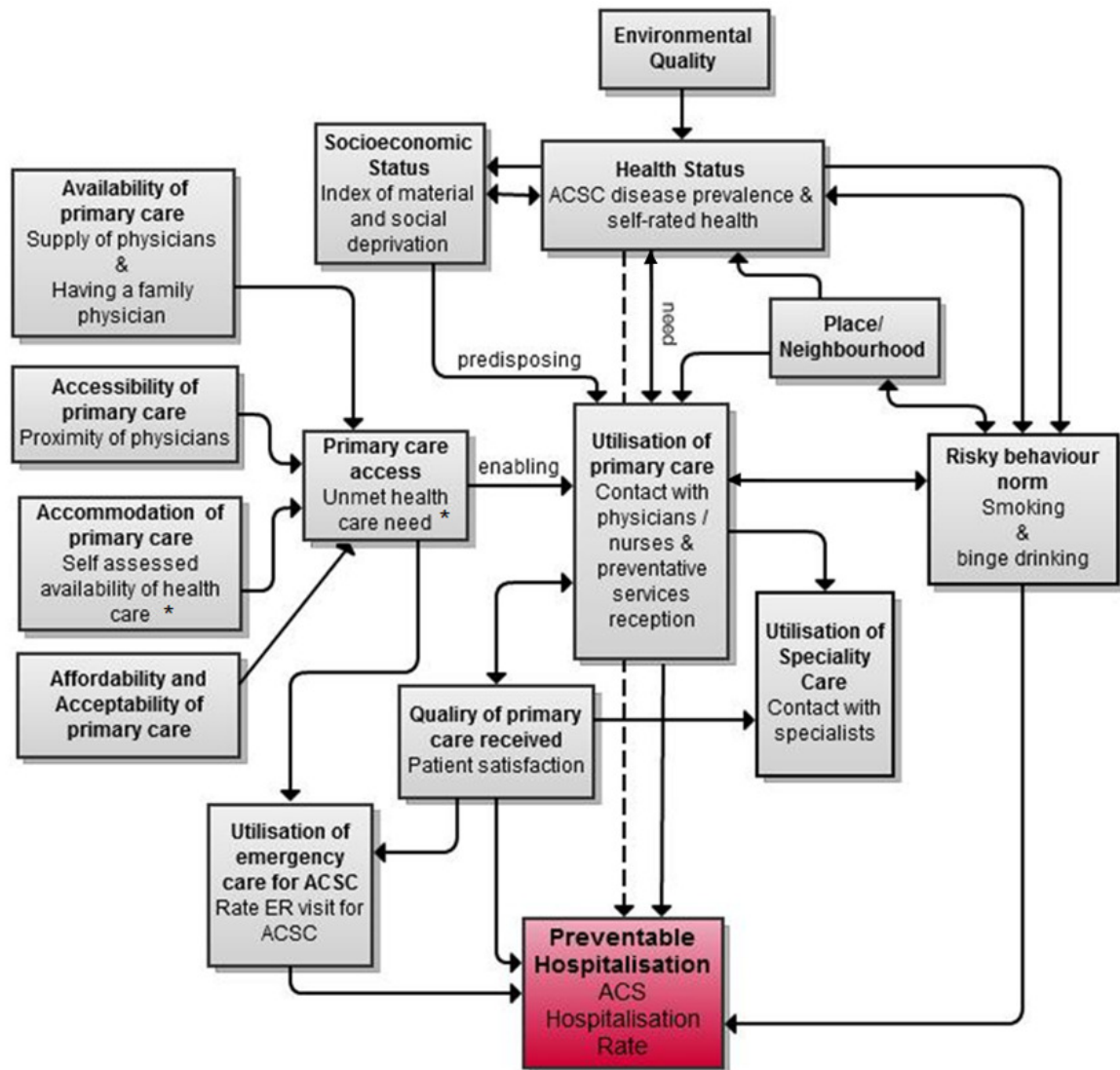
Concept Represented by Variable	#	Variable Name (Label)	Data Source (Coverage)	Variable Type	Variable Elements – Characteristics	Inclusion in Exploratory Models	Inclusion in Final Models
Primary Care Access	26	Percent with unmet health care need in past year (unmet_care_input)	CCHS (not available for Cycle 4.1 and Cycle 2009)	% – Continuous	• Item asked to all CCHS respondents • Ottawa sample size = 4788 respondents • 7 neighbourhood imputations performed • Representing concept of access to health services • Item not specific to primary care only, but to the health care as a whole • Could also be perceived as a compositional characteristic	Yes	No

Appendix 5

Causal Model of Preventable Hospitalisations Including All Variables of Interest

Concepts (in bold) and the Variables (not in bold)

Figure A-2: Causal Model Including All Variables of Interest



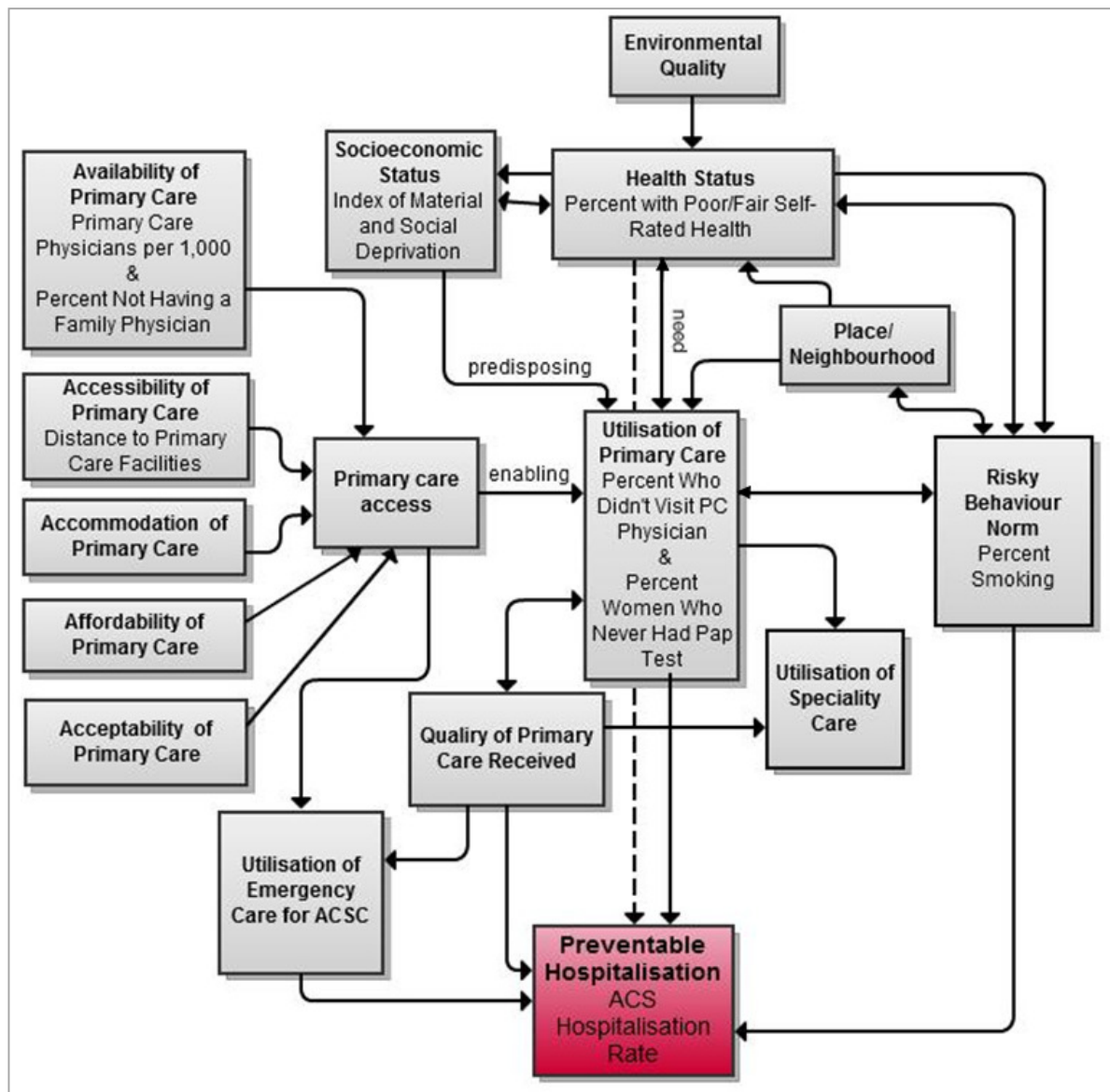
* Health care does not refer to primary care exclusively

Appendix 6

Causal Model of Preventable Hospitalisations Represented by Variable Selected for Final Model

Concepts (in bold) and the Variables (not in bold)

Figure A-3: Causal Model Representing Variables Selected for Final Model



Appendix 7

Information about the Data Quality of the Discharge Abstract Database

The data quality of the DAD is assessed by a systematic program of re-abstraction of the hospitalisation charts included in the database. As part of this program, the hospitalisation discharges of selected institutions are re-coded using the original medical charts and re-captured in the database.¹²⁸ Comparisons in the coding between the first and second abstractions are performed in terms of completeness and correctness, where "completeness represents the proportion of observations 'about the world' that are actually recorded and correctness represents the proportion of observations that reflect the 'true state of the world'" (CIHI, 2006b, p.9).

For the fiscal year 2005-2006, the DAD's quality of coding for ACS conditions was assessed – based on the ICD-10-CA codes of the indicator developed by CIHI. This indicator requires that one of the following conditions is coded as the '*most responsible diagnosis*': COPD, heart failure, pulmonary edema, angina, diabetes, hypertension, epilepsy and asthma. From the data quality report (CIHI, 2006b, p.26), the completeness was assessed as follows: "of all ACSC hospitalizations identified during the chart review, 87% were similarly identified in the database across institutions in Canada". In Ontario, this proportion was 92%, demonstrating that the method to identify ACS hospitalisation provides complete information in the DAD. For data correctness, "of all acute care ACSC hospitalizations in DAD, 86% were confirmed following the chart review" (87% for Ontario) (CIHI, 2006b, p.26). CIHI considers that this proportion provides correct information for the indicator of ACS hospitalisations. However, there may be an issue of over-representation of ACS hospitalisations for angina and hypertension. Overall, this demonstrates that the

¹²⁸"A re-abstraction study involves returning to the original source of data (a hospital chart, for example) in order to compare what is contained in a patient's health record with what exists in CIHI's Discharge Abstract Database. To measure the accuracy of selected administrative and clinical data elements in the DAD, a sample set of records that was submitted to the database in a previous fiscal year is re-abstracted. The re-abstracted data are then compared to the original submissions." (CIHI, 2004d, p.iii)

DAD's coding quality of ACS hospitalisations, based on CIHI indicator, is good in terms of its completeness and correctness in Canada and in Ontario. Similar quality in the identification of the ICD-10-CA codes and categorisation of the '*most responsible diagnostic*' can be expected for the current investigation.

However, the coding quality of diagnoses on the chart may be of lower quality for the earlier fiscal years included in this data extraction. In fact, the ICD-10-CA codes were used for the first time in the fiscal year of 2002-2003 in Ontario¹²⁹. This implementation affected the quality of the diagnostic coding, since the data coders were not accustomed with the differences in the classification. For Ontario, in a re-abstraction study assessing the impact of the new classification, 7% of the diagnostic codes¹³⁰ were different between the first and second abstraction, yet these remained in the same condition. It is important to note that the coding for COPD with acute lower respiratory infection (pneumonia) (J44) is significantly different under the new classification (ICD-10-CA) compared to the older version (ICD-09-CA) (CIHI, 2004b). This led to a significant increase in the number of charts re-coded with a COPD diagnostic in the re-abstraction for Ontario (330%), but this increase is not necessarily associated with the '*most responsible*' diagnosis field used for the extraction in the current investigation (CIHI, 2004d). Based on this information, for the first fiscal years of the current investigation, the number of ACS hospitalisations due to COPD was most certainly underestimated. It is assumed that the impact of this under-estimation was the same for the different neighbourhoods in Ottawa; therefore not precluding comparisons across neighbourhoods.

With regard to the non-clinical field extracted, age and gender, the data is of very high quality. For the fiscal year of 2002-2003 to 2007-2008¹³¹, 100% of the age and gender information were coded accurately based on the re-abstraction studies (CIHI, 2004a; 2004b; 2005a; 2006a; 2007b; 2008c; 2009d).

¹²⁹ In prior years, the ICD-09-CA codes were used to classify the diagnostics.

¹³⁰ These diagnostic codes are not specific to ACS conditions.

¹³¹ The information is not available for the fiscal year of 2008-2009.

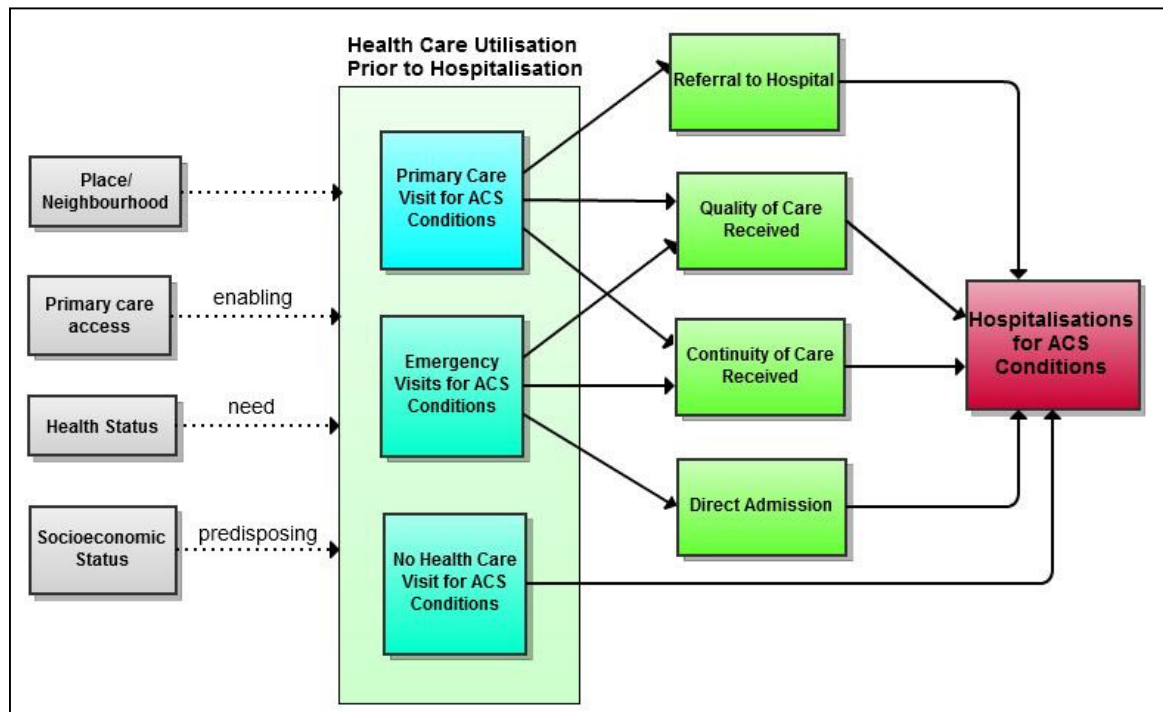
Appendix 8

About the Complexity of Interpreting the Role of Emergency Care Utilisation in the Relationships between Primary Care Access, Health Care Utilisation and Hospitalisations for ACS Conditions

In the causal model of preventable hospitalisations, the ACS ER visits rates are thought to be a proxy of primary care access as well as a proxy of service quality and continuity. The ER visit rate was considered as a proxy for primary care access, considering that reduced neighbourhood access to primary care was hypothesised to cause higher reliance on emergency care. Also, it was hypothesised that ER visit rates may be a factor influencing the rates of ACS hospitalisation, especially since continuity of care tend to be poorer in emergency setting compared to primary care setting. In this context, a significant relationship between ER visits rates and ACS hospitalisation rates in the regression model could be indicative of a lack of primary care access or a reduced continuity of care in emergency setting. Yet, this subtlety in the interpretation of the regression could not be evaluated. The interpretations of the results would be further compromised by the inability of the model at differentiating the effect of ER visits rates due to direct admissions – caused by the potential overlap in the counts – or to the suspected poorer continuity of care received in an emergency setting.

Figure A-4 below illustrates how the relationships between primary care access, health care utilisation and hospitalisations for ACS conditions are further complicated once the ER visits rates are considered in the causal model. This demonstrates the lack of construct validity of this measure as a proxy of primary care access or quality/continuity of care.

Figure A-4: Relationship Diagram: Primary Care Access, Health Care Utilisation, ACS Hospitalisations and ACS ER Visits Rates



Appendix 9

Overview of the CCHS Sampling Strategy and Weight Calculation Method

The CCHS sampling strategy is complex¹³²; it involves three sampling frames¹³³ (area, telephone and random digit) and two sampling stages (household and individual). Different sampling strategies are implemented based on the sampling frame and stage: 1) households are selected either via a stratified random sampling process or via a multistage stratified cluster design where the clusters are selected randomly and the household selected systematic within each cluster, and 2) individuals are selected using varying probabilities taking into account the age structure of the household members¹³⁴. Attempts to contact the selected household or individual are conducted either by phoning or by visiting the residence. Similarly, the interviews are conducted either in person or over the phone at any time of the year. Information about the sampling and data collection phases of the survey is recorded and used for the calculation of the regional sampling weights.

Using this information, Statistics Canada develops a final weight for each respondent of the same health region and cycle¹³⁵ following a 4-step process¹³⁶. The first step involves the calculation of a weight for the area frame respondents and a weight for the telephone frame¹³⁷ respondents, considering their different sampling processes and non-response patterns. The second step is the integration of the two

¹³² The descriptions of the CCHS sampling methodology is based on the User Guides of the Microdata Files (Statistics Canada, 2002b; 2005b; 2006b; 2009; 2010).

¹³³ The area frame is based on clusters marrying demographic and geographic attributes and do not cover the whole region of Ottawa-Carleton. Therefore, some neighbourhoods may not be represented in the area frame. The telephone list frame covers the whole region of Ottawa and yet it contains only listed phone numbers. Overall, the telephone list frame is covering about 70% of the population in Ottawa, independently of the neighbourhoods. (M.-C. Duval, personal communication, June 13, 2011).

¹³⁴ For Cycle 1.1, two interviewees could be selected in a same household if youth were residing in the household (12-19 years old). Yet, the respondents below 20 years old were not included in the sample of the current investigation.

¹³⁵ For the Cycle 4.1 and 2009, the process to calculate the weight was slightly modified to take into consideration the difference in the data collection methodology (Thomas & Tremblay, 2010). For these cycles, the non-response calibrations at the individual level are performed after the weights from the telephone and area frames are integrated (Thomas, Saraffin, & Simard, 2007).

¹³⁶ The description of the weight calculations method is based on the User Guides of the Microdata Files (Statistics Canada, 2002b; 2005b; 2006b; 2009; 2010).

¹³⁷ Since only minor differences differentiate the Random Digit Sampling frame from the telephone frame, these are treated as one sampling frame in the development of the weights.

sets of weights based on the relative importance of each samples in the total sample of respondents in the health region. The third step involves adjusting the weights based on the time of the year at which the interview was conducted. At this stage, any outlier weight is adjusted downward in order to reduce its influence of the regional population estimates. Finally, the fourth step is post-stratification. "Post-stratification is done to ensure that the sum of the final weights corresponds to the population estimates defined at the health region, for all 10 age-sex groups of interest, that is, the five age groups 12-19, 20-29, 30-44, 45-64, 65+, for both males and females" (Statistics Canada, 2006b, p.35).

Appendix 10

Descriptive Statistics – Variables of Interest (The meaning of the variable labels is available in Appendix 4.)

Table A- 3: Descriptive Statistics – Variables of Interest

Concept	#	Variable	N	Min	5th Pctl	Lower Quartile	Median	Upper Quartile	95th Pctl	Max	Mean	IQR	Std Dev	Coeff of Variation	Skewness	Kurtosis	Correlation with Hosp. Rate	Correlation with SES
Preventable Hospitalisations	1	Hosp_Rate	90	256.1	334.5	451.0	543.0	724.5	1003.0	1161.6	599.0	273.5	217.0	36.2	0.792	-0.220	1.00	0.67
SES	2	SESindexcont	90	-1.36	-1.17	-0.79	-0.17	0.50	2.05	2.78	0.01	1.3	0.997	7368.3	0.758	-0.266	0.67	1.00
Health Status	4	health_status4_5_ag_st	90	0.0%	2.2%	7.0%	10.0%	13.2%	20.5%	32.0%	10.3%	6.2%	0.056	54.9	0.857	1.869	0.53	0.58
	5	chronic_dis_1_plus_ag_st	90	10.7%	16.3%	21.4%	26.2%	30.8%	40.5%	46.9%	26.7%	9.4%	0.072	26.9	0.479	0.350	0.30	0.23
	6	chronic_dis_2_plus_ag_st	90	0.0%	0.0%	3.1%	5.3%	7.9%	11.2%	18.6%	5.7%	4.8%	0.034	59.8	0.608	1.145	0.18	0.22
Healthy Behaviour	7	smoker_ag_st	90	0.0%	4.8%	14.4%	18.8%	26.2%	32.6%	42.9%	19.9%	11.8%	0.083	41.6	0.223	-0.032	0.62	0.54
	8	binge_drink_yr_ag_st	90	14.3%	22.5%	30.9%	36.6%	41.2%	49.6%	53.9%	35.9%	10.3%	0.079	21.9	-0.294	0.041	-0.15	-0.25
	9	binge_drink_mth_ag_st	90	3.0%	8.9%	12.3%	15.4%	18.9%	25.4%	31.7%	16.0%	6.6%	0.053	32.8	0.384	0.264	-0.02	-0.12
Utilisation of Health Services	10	no_visit_gp_ag_st	90	1.3	2.1	2.7	3.1	3.5	4.9	7.0	3.2	0.9	0.872	27.3	1.362	3.713	0.39	0.43
	11	didnt_visit_gp_ag_st	90	6.0%	8.7%	15.8%	19.1%	23.0%	30.8%	39.7%	19.5%	7.2%	0.067	34.1	0.438	0.459	0.14	0.22
	12	no_visit_nurse_ag_st	90	0.0	0.1	0.2	0.4	0.8	3.1	8.8	0.8	0.6	1.264	150.7	3.919	18.981	0.29	0.29
	13	didnt_visit_nurse_ag_st	90	59.4%	80.0%	84.9%	88.6%	91.5%	96.4%	100.0%	87.9%	6.6%	0.060	6.8	-1.440	5.655	-0.12	-0.07
	14	no_visit_otherdoc_ag_st	90	0.4	0.6	1.0	1.3	1.8	2.6	4.3	1.4	0.8	0.671	46.7	1.365	3.035	0.12	0.18
	15	didnt_visit_otherdoc_ag_st	90	37.9%	54.0%	61.0%	66.2%	70.3%	76.1%	82.8%	65.4%	9.3%	0.072	11.0	-0.862	2.352	0.09	0.01
Utilisation of Preventative Health Services	16	flu_ag_st	90	21.0%	23.1%	32.0%	36.1%	40.5%	48.4%	57.3%	36.0%	8.4%	0.072	19.9	0.082	0.316	0.02	-0.01
	17	pap_input	90	0.0%	0.0%	3.7%	6.5%	9.8%	15.3%	20.7%	7.0%	6.2%	0.046	65.4	0.549	-0.018	0.25	0.52
	18	breast_input	90	0.0%	3.7%	8.1%	12.9%	16.2%	25.4%	37.6%	13.1%	8.1%	0.067	51.3	0.886	1.481	0.08	0.34
	19	mammo_input	90	0.0%	8.0%	14.8%	17.0%	18.9%	25.7%	31.8%	16.9%	4.1%	0.053	31.5	-0.244	2.161	0.06	-0.03
Quality of Primary Care	20	quality_doc_input	90	0.0%	0.0%	5.8%	7.0%	9.6%	14.7%	22.7%	7.4%	3.8%	0.041	55.0	0.775	1.760	0.00	-0.03
Accessibility of Primary Care	22	dist_4_physician	90	2	257	639	1117	1714	7676	16733	1884	1075	2635	139.9	3.644	15.263	-0.10	-0.37
Availability of Primary Care	23	physician_1000	90	0	0	0.19	0.59	1.47	3.64	14.45	1.14	1.28	1.814	159.5	4.831	32.442	0.25	0.20
	24	fam_doc_ag_st	90	1.6%	4.7%	8.4%	10.9%	16.8%	26.4%	32.8%	13.0%	8.4%	0.066	51.2	0.876	0.236	0.32	0.50
Accommodation of Primary Care	25	availability_com_input	90	19.3%	21.7%	31.8%	36.4%	40.7%	50.8%	75.9%	36.8%	8.9%	0.099	26.8	1.104	3.265	-0.02	-0.09
Primary Care Access	26	unmet_care_input	90	0.0%	7.4%	10.6%	12.9%	16.3%	23.8%	43.5%	13.9%	5.8%	0.058	42.1	1.850	7.142	0.24	0.20

Appendix 11

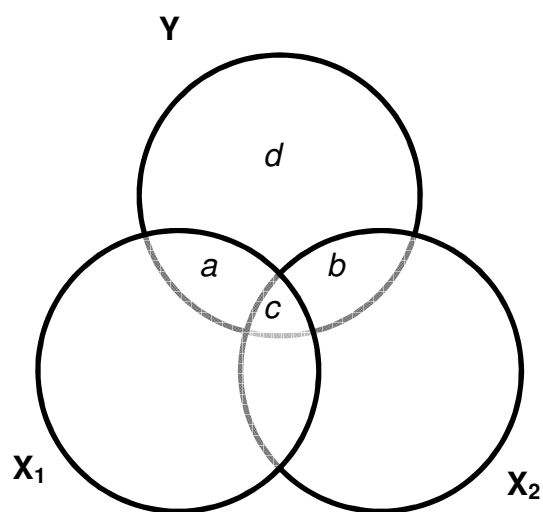
Partition of Variance Explained in Regression – Example of a Two-predictor Model

In Figure A-5 below is a graphical representation of the partition of the variance in the dependent variable explained by two correlated predictor variables using regression. The circle for the dependent variable Y , represented by the sum of the area $a+b+c+d$, is the total variance contained in this variable and equals 1¹³⁸. The circles of X_1 and X_2 represent the variances of the two predictor variables. Their variance is overlapping considering that the predictor variables are correlated. The area represented by the sum of $a+b+c$ is the proportion of the variance in the dependent explained by the two predictor variables. This proportion is also called R^2 . The proportion of the variance in the dependent variable that is predictable uniquely from the variable X_1 , when its correlation with X_2 is removed, is represented by the area a . Similarly, the area b represents the proportion of the variance of the dependent variable explained uniquely by the variable X_2 , when its correlation with X_1 is removed. These proportions are also called the squared semi-partial correlation (sr^2). The area c represents the proportion of the variance in the dependent variable that could be explained either by X_1 or by X_2 , which is also called the shared variance. And, finally, the area d represents the proportion of the variance in Y not explained by the predictor variables included in the model. (Tabachnick & Fidell, 2007, p. 131-132; Warner, 2008, p.429)

By applying a linear regression model on these variables, the partition of the area a , b , c and d is obtained. In standard multiple regression, the shared variability explained by all the variables in the model is not attributed to any predictor and its significance is not tested independently. In sequential regression models, the shared variance is attributed to the variable(s) entered first in the regression. (Tabachnick & Fidell, 2007, p. 138)

¹³⁸ Assuming the dependent variable Y is expressed in z score units.

Figure A-5: Partition of the Variance in Dependent Variable in Multivariate Regression Model



Appendix 12

Summary of Linear Model Assumptions and Diagnostic Methods

Table A- 4: Linear Model Assumptions and Diagnostic Methods

Assumption	Diagnostic Method	Suggested Remedy
Matrix of predictor variable is full rank – no multicollinearity	- Variance Inflated Factor (VIF) < 10 - Condition Index (CI) < 30 - Elevated Variance Proportions for Component with CI < 30	If multicollinearity is identified based on the CI and VIF diagnostic, removing the predictor variables involved in multicollinearity should be performed (based on the assessment of their variance proportions)
Predictors are fixed	Researcher knowledge of the predictor variables: possibility of greater variance in measurement error for neighbourhoods of small population size having an impact on structure of the model errors	Population-weighted regression analyses – see sensitivity analyses
Errors are independent	Durbin-Watson Test $\approx$ approximately 2	Special parameter estimation technique taking into consideration the auto-correlation of the errors
Errors are normally distributed	- Histogram of residual $\rightarrow$ bell-shaped - Scatter plot of studentized residuals and predicted values $\rightarrow$ 95% of residuals within ± 2 STD. - Q-Q plot of residuals $\rightarrow$ no strong departure from straight line	Transformation of the independent variables or on the dependent variable could be improving the distribution of the residuals
Errors have constant variance	- Scatter plot of residuals and predicted values $\rightarrow$ constant vertical spread - Partial regression plots $\rightarrow$ constant vertical spread - The White test $\rightarrow$ p non-significant	Transformations on the variables thought to cause heteroscedasticity or weighted regression
Dependent variable is linearly related to predictor variables	- Partial regression plots $\rightarrow$ no departure from line where residuals = 0 - Scatter plot of dependent and independent variables $\rightarrow$ linear shape	Transformations on the non-linear variables, addition of quadratic term in the regression model or non-linear regression

Summary of Diagnostic Tests for Multivariate Outliers, Leverage and Influence

Table A- 5: Diagnostic Test for Multivariate Outliers, Leverage and Influence

	Diagnostic Test	Cut-off values
Multivariate outliers	Mahalanobis Distance	Chi-square statistic with $p < 0.001$
Leverage	Leverage Values	$2p/n$
Influence on model in general	Cook's Distance	Any values considerably greater than for the majority of cases
Influence on regression coefficient	DFBETAS	2 or $2/\sqrt{n}$
Influence on fitted values	DFFITS	$2/(\sqrt{p/n})$

Note: n = number of cases in model & p = number of predictor variables in model

Appendix 13

Scatter Plots of ACS Hospitalisation Rates by Predictor Variables – Demonstration of Linearity of Relationships

Figure A- 6: Scatter Plot of SES Index vs ACS Hospitalisation Rates

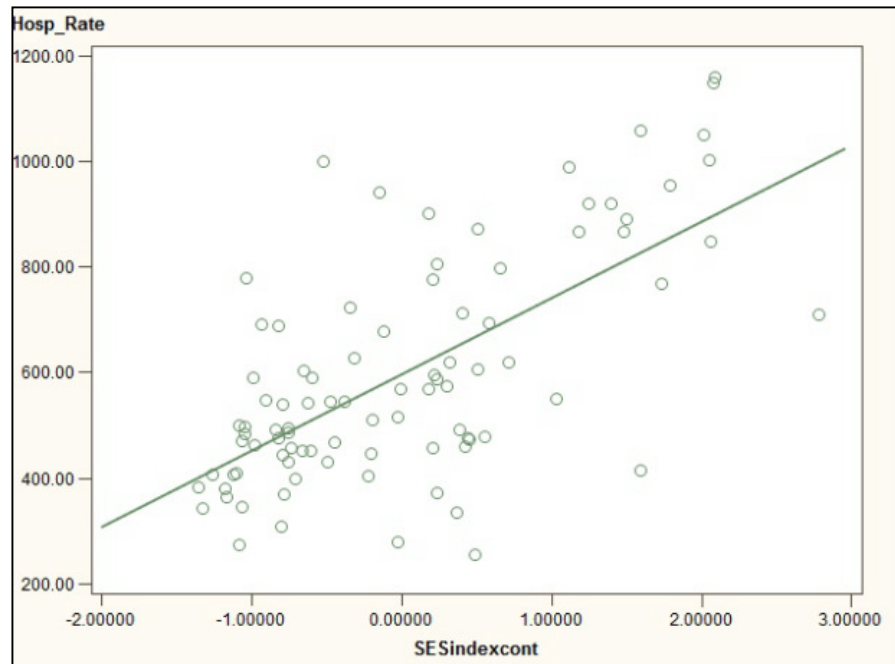


Figure A- 7: Scatter Plot of 'Percent with poor or fair health status' vs ACS Hospitalisation Rates

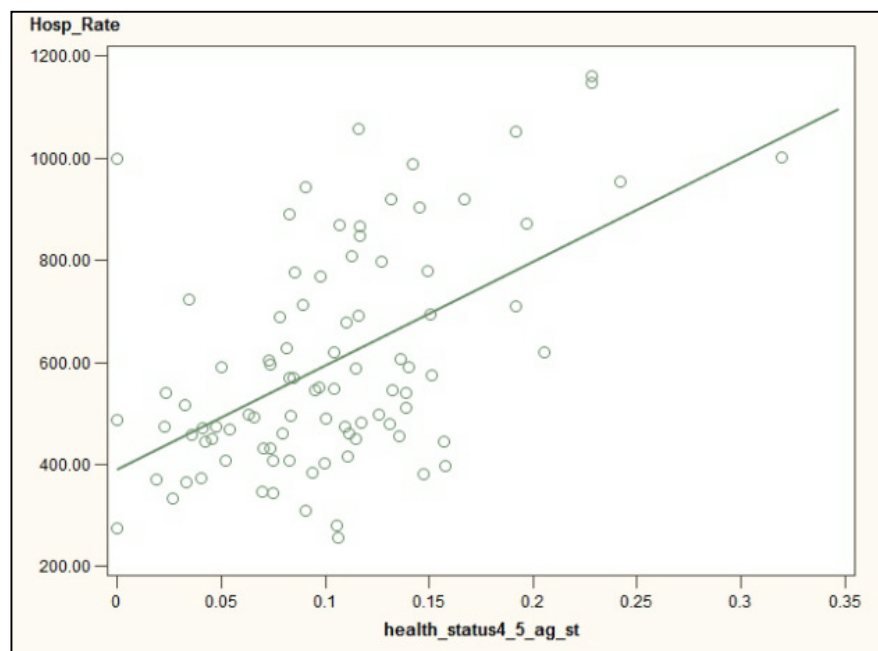


Figure A- 8: Scatter Plot of 'Percent without a family physician' vs ACS Hospitalisation Rates

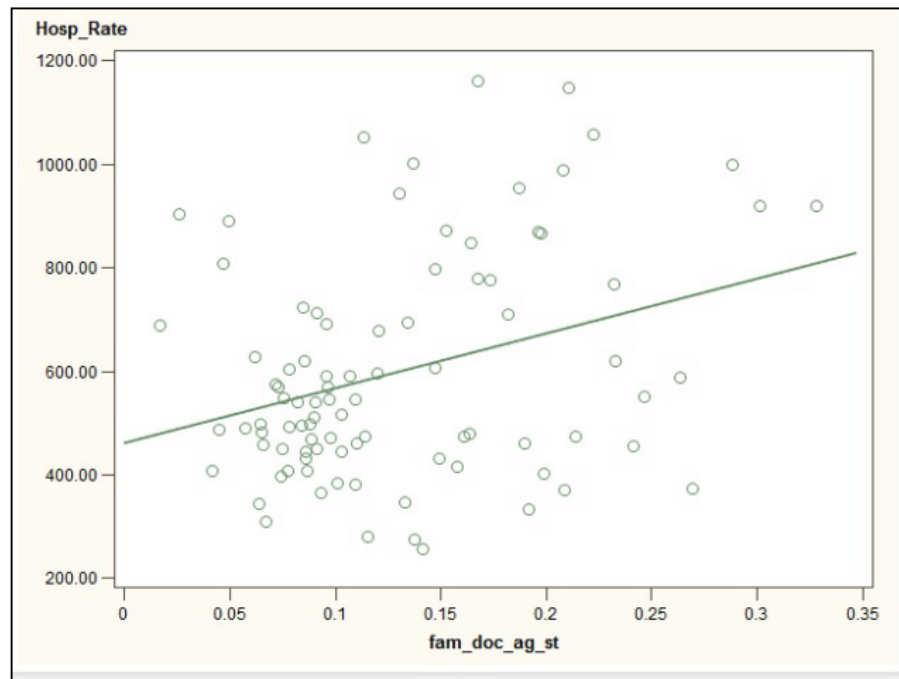
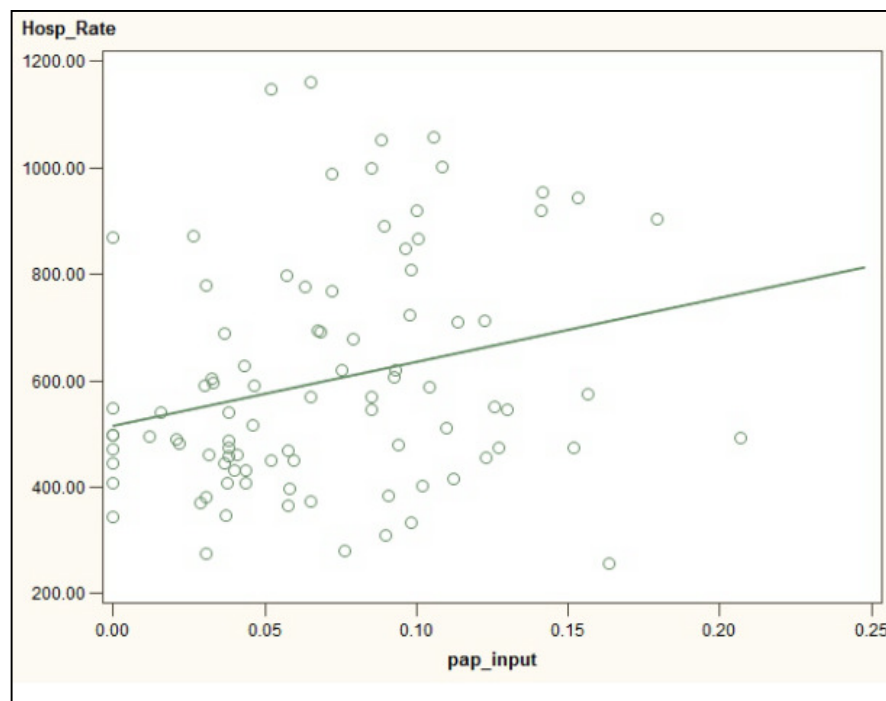


Figure A- 9: Scatter Plot of 'Percent of women who never had a Pap test' vs ACS Hospitalisation Rates



Appendix 14

Univariate and Bivariate Regression Results – All Variables of Interest

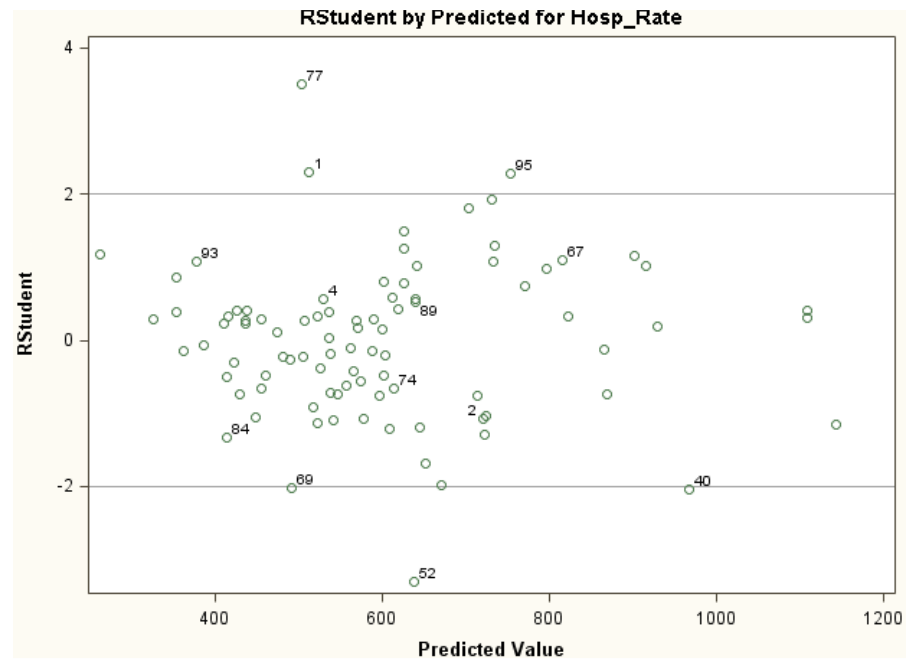
Table A- 6: Univariate & Bivariate Regression Results – Variables of Interest

Dependent	Variable Label (meaning in Appendix 4)	Univariate Models					Models Including SES Index				
		R ²	Coefficient Estimate	Std Err	t value	p value*	R ²	Coefficient Estimate	Std Err	t value	p value*
Hosp_Rate	SESindexcont	0.442	144.7	17.3	8.35	0.00	n/a				
	health_status4_5_ag_st	0.279	2034.3	348.8	5.83	0.00	0.473	831.9	367.2	2.27	0.03
	chronic_dis_1_plus_ag_st	0.088	896.0	307.6	2.91	0.01	0.463	451.5	233.1	1.85	<i>0.07</i>
	chronic_dis_2_plus_ag_st	0.034	1165.8	664.8	1.75	<i>0.08</i>	0.444	235.8	520.6	0.45	0.65
	smoker_ag_st	0.386	1628.6	219.0	7.44	0.00	0.541	973.8	225.7	4.31	0.00
	binge_drink_mth_ag_st	0.000	-63.8	440.5	-0.14	0.89	0.446	270.4	332.1	0.81	0.42
	binge_drink_yr_ag_st	0.022	-405.1	291.0	-1.39	0.17	0.443	60.7	228.2	0.27	0.79
	no_visit_gp_ag_st	0.156	98.3	24.4	4.03	0.00	0.456	32.8	21.8	1.5	0.14
	didnt_visit_gp_ag_st	0.021	466.4	344.1	1.36	0.18	0.442	-19.7	267.9	-0.07	0.94
	no_visit_nurse_ag_st	0.085	50.0	17.5	2.86	0.00	0.453	18.5	14.2	1.3	0.20
	didnt_visit_nurse_ag_st	0.015	-436.9	384.3	-1.14	0.26	0.448	-271.2	290.0	-0.93	0.35
	no_visit_otherdoc_ag_st	0.015	40.0	34.2	1.17	0.25	0.442	1.9	26.3	0.07	0.94
	didnt_visit_otherdoc_ag_st	0.008	273.2	319.1	0.86	0.39	0.449	251.6	239.2	1.05	0.30
	flu_ag_st	0.000	63.2	323.0	0.20	0.85	0.443	86.8	242.5	0.36	0.72
	pap_input	0.064	1201.8	491.5	2.45	0.02	0.453	-589.0	440.8	-1.34	0.18
	breast_input	0.006	250.2	344.2	0.73	0.47	0.466	-532.3	269.3	-1.98	0.05
	mammo_input	0.004	254.0	433.1	0.59	0.56	0.449	340.5	324.1	1.05	0.30
	quality_doc_input	0.000	21.5	566.7	0.04	0.97	0.443	140.1	425.6	0.33	0.74
	dist_4_physician	0.010	0.0	0.0	-0.96	0.34	0.468	0.0	0.0	2.04	0.04
	physician_1000	0.065	30.4	12.3	2.47	0.02	0.458	15.4	9.6	1.60	0.11
	fam_doc_ag_st	0.105	1057.8	329.6	3.21	0.00	0.442	-45.9	302.6	-0.15	0.88
	availability_com_input	0.000	-43.8	234.4	-0.19	0.85	0.444	94.4	176.6	0.53	0.59
	unmet_care_input	0.058	890.6	384.5	2.32	0.02	0.455	424.3	299.9	1.41	0.16

* Highlighted in bold are results significant at p<0.05, highlighted in italics are the results significant at p<0.1.

Appendix 15

Figure A-10: Scatter plot of Studentized Residuals and Predicted Values – Exploratory Model



* The labels are the neighbourhood identification numbers, refer to Appendix 2 for the associated neighbourhood names

Figure A-11: Q-Q Plot of Residuals – Exploratory Model

