Population-level determinants of obesity in Canada: application of econometric techniques

by

Daniel James Dutton

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN COMMUNITY HEALTH SCIENCES

CALGARY, ALBERTA

JULY, 2014

© Daniel James Dutton 2014
Abstract

Obesity is a significant public health problem in Canada. Increasingly, factors that are beyond the individual are being identified as important drivers of the current obesity epidemic. Drivers like prices, taxes, and government policy are jointly identified as population-level determinants of obesity that have resulted in its rapid growth. While these population-level determinants have been identified as important, their impact has not been explicitly quantified in Canada. The geographic distribution of obesity prevalence in Canada is one issue that could potentially highlight the important role that population-level determinants of obesity play.

This dissertation’s objective is to assess the importance of population-level determinants of obesity in Canada by quantifying their impact on individuals living in different regions of the country. Canada has, roughly, an east to west gradient of obesity, with the Atlantic provinces exhibiting the highest prevalence of obesity. I characterize the difference between the Atlantic provinces and other regions of Canada (Quebec, Ontario, the Prairies, and British Columbia) in two ways: the difference in average body mass index (BMI) and the difference in BMI distributions. To estimate the contribution of the population-level determinants to these differences I apply Blinder-Oaxaca decompositions and quantile regression to national level data from the Canadian Community Health Survey. I show that the population-level determinants are important in describing cross-regional differences in obesity in Canada and their importance becomes larger at high percentiles of the BMI distribution, especially in females. I explain how this is consistent with the ecological model of obesity’s portrayal of the population-level determinants of obesity.

Parallel to meeting the overall objective of this dissertation, I assess the added value of corrected BMI values in obesity research. Correction equations are generally used to adjust self-
reported BMI values so they resemble measured BMI values on aggregate. I assess their usefulness by establishing a new correction equation and comparing that correction to established Canadian correction equations, measured BMI, and self-reported BMI. I determine that corrected BMI is not always superior to self-reported BMI and discuss the settings where corrected BMI is useful.
Acknowledgements

This dissertation would not exist without the selfless contributions of others, the foremost being my supervisor Dr. Lindsay McLaren. Dr. McLaren invested hundreds of hours in meeting with me, discussing methods and results, and proof-reading numerous drafts of each chapter of this dissertation. Dr. McLaren also provided me with financial and scholarly support throughout my studies, without which I would have never succeeded. Dr. J. C. Herbert Emery provided numerous forms of support throughout my studies and recounting them here would fill the space allotted for acknowledgements. Briefly, Dr. Emery is responsible for my interest in pursuing a PhD in health issues, my interest in policy studies in general, and was influential in introducing me to the tools used in this dissertation. Thanks also go to Dr. Melanie Rock for preventing me from making critical errors regarding the scope of my dissertation early on.

Personal support from my family was equally instrumental in finishing this dissertation. My mother and sister, Marlene and Jessica, have been supportive from the beginning regarding moving back to Calgary to complete my PhD and the writing of this document. Finally, my wife Patricia has exhibited more patience and understanding than anyone could rightly expect from their partner. PhD students are terrible marriage market prospects so I thank her not only for her support but her faith in me.
# Table of Contents

Abstract .................................................................................................................. ii
Acknowledgements ................................................................................................ iv
Table of Contents ................................................................................................... v
List of Tables .......................................................................................................... viii
List of Figures and Illustrations ........................................................................... ix
List of Symbols, Abbreviations and Nomenclature ................................................ x

CHAPTER ONE: EXPLANATION AND JUSTIFICATION FOR MANUSCRIPT-BASED
THESIS, AND CONTRIBUTIONS OF AUTHORS .............................................. 1
  1.1 Explanation and justification for manuscript-based thesis ............................ 1
  1.2 Contribution of authors ............................................................................. 1
  1.3 Status of manuscripts .............................................................................. 2
  1.4 References ............................................................................................... 4

CHAPTER TWO: INTRODUCTION ..................................................................... 5
  2.1 Outline ....................................................................................................... 5
  2.2 Rationale for Dissertation Research .......................................................... 5
    2.2.1 Conceptual framework .................................................................... 8
  2.3 Determinants of obesity ........................................................................... 13
    2.3.1 Measurement of obesity ................................................................. 13
    2.3.2 A brief overview of the complicated relationship between the determinants of
        obesity ................................................................................................. 13
    2.3.3 Epidemiology and determinants of obesity in Canada ......................... 16
  2.4 Manuscripts ............................................................................................... 19
  2.5 Methods .................................................................................................... 23
    2.5.1 Data Source .................................................................................... 25
    2.5.2 Data access .................................................................................... 25
    2.5.3 Data management and analysis ....................................................... 28
    2.5.4 Ethics approval .............................................................................. 28
  2.6 References ............................................................................................... 30

CHAPTER THREE: EXPLAINED AND UNEXPLAINED REGIONAL VARIATION IN
CANADIAN OBESITY PREVALENCE ................................................................. 39
  3.1 Abstract .................................................................................................... 39
  3.2 Introduction ............................................................................................... 40
  3.3 Methods and Procedures .......................................................................... 43
    3.3.1 Data Source and Study Design ....................................................... 43
    3.3.2 Statistical Analysis ....................................................................... 44
  3.4 Results ...................................................................................................... 45
    3.4.1 OLS Estimates .............................................................................. 47
    3.4.2 Blinder-Oaxaca Decompositions .................................................... 50
  3.5 Discussion ................................................................................................. 56
  3.6 References ............................................................................................... 64
CHAPTER FOUR: HOW USEFUL IS “CORRECTED” BODY MASS INDEX VS. SELF-REPORTED BODY MASS INDEX? COMPARING THE POPULATION DISTRIBUTIONS, SENSITIVITY, SPECIFICITY, AND PREDICTIVE UTILITY OF THREE CORRECTION EQUATIONS USING CANADIAN POPULATION-BASED DATA .......................................................... 67
4.1 Abstract .................................................................................................................. 67
  4.1.1 Background ......................................................................................................... 67
  4.1.2 Methods ............................................................................................................... 67
  4.1.3 Results ................................................................................................................ 68
  4.1.4 Conclusions ......................................................................................................... 68
4.2 Background ............................................................................................................. 69
4.3 Methods .................................................................................................................. 72
  4.3.1 Data .................................................................................................................... 72
  4.3.2 Procedure ........................................................................................................... 73
  4.3.3 Development and estimation of weight-only correction equation ...................... 75
  4.3.4 Evaluation of correction equations: comparison of BMI distributions; estimation of sensitivity and specificity for weight categories; and prediction of health outcomes ....................................................................................... 79
4.4 Results .................................................................................................................... 81
  4.4.1 Distribution of BMI ........................................................................................... 81
  4.4.2 Sensitivity and specificity of corrected and self-reported BMI ......................... 81
  4.4.3 Predictive utility of corrected BMI in health condition models ....................... 85
4.5 Discussion ............................................................................................................... 89
  4.5.1 Distribution of BMI ........................................................................................... 89
  4.5.2 Sensitivity and specificity .................................................................................... 89
  4.5.3 Predictive utility of corrected BMI in health condition models ....................... 90
4.6 Limitations ............................................................................................................. 92
4.7 Conclusions ........................................................................................................... 92
4.8 References ............................................................................................................. 94

CHAPTER FIVE: HOW IMPORTANT ARE DETERMINANTS OF OBESITY MEASURED AT THE INDIVIDUAL-LEVEL FOR EXPLAINING GEOGRAPHIC VARIATION IN BODY MASS INDEX DISTRIBUTIONS? EVIDENCE FROM CANADA. .................................................................................................................... 97
5.1 Abstract .................................................................................................................. 97
  5.1.1 Background ......................................................................................................... 97
  5.1.2 Methods ............................................................................................................... 97
  5.1.3 Results ................................................................................................................ 97
  5.1.4 Discussion ........................................................................................................... 98
5.2 Background ........................................................................................................... 99
  5.2.1 Determinants of obesity measured at the individual level ............................... 99
  5.2.2 Considering the full BMI distribution: quantile regression ............................ 99
  5.2.3 Considering differential impact: decomposition techniques ......................... 102
  5.2.4 The present study ............................................................................................. 103
5.3 Methods ................................................................................................................ 104
  5.3.1 Data .................................................................................................................... 104
  5.3.2 Analytic Procedure ............................................................................................ 105
5.3.2.1 OLS regression .............................................................105
5.3.2.2 Unconditional (marginal) quantile regression ..................105
5.3.2.3 Decomposition of quantile regression ..............................106
5.4 Results ..................................................................................110
5.4.1 OLS Regression ..................................................................110
5.4.2 Unconditional Quantile Regression Models .........................110
5.4.2.1 Females .........................................................................119
5.4.2.2 Males .............................................................................120
Unconditional Quantile Regression Decompositions (Figure 5.6 to Figure 5.9) ...122
5.4.2.3 Males versus female decomposition trends .........................122
5.4.2.4 The explained share of overall regional differences for females 123
5.4.2.5 The explained share of overall regional differences for males 123
5.5 Discussion .............................................................................129
5.6 References .............................................................................132

CHAPTER SIX: CONCLUSION .................................................135
6.1 Manuscript-specific findings .....................................................135
6.1.1 Explained and Unexplained Regional Variation in Canadian Obesity Prevalence ..............................................................135
6.1.2 How useful is “corrected” body mass index vs. self-reported body mass index? Comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data. 136
6.1.3 How important are determinants of obesity measured at the individual-level for explaining geographic variation in body mass index distributions? Evidence from Canada .............................................................................136
6.1.4 Overall conclusions from this dissertation .............................137
6.2 Strengths and Limitations ........................................................137
6.3 Overall implications of the findings: future research ..................142
6.4 References .............................................................................143

CHAPTER SEVEN: WORKS CITED ........................................145

APPENDICES .............................................................................165
Appendix A: Supplemental material to Chapter 3 ..........................165
Appendix B: Ethics approval for this dissertation ...........................167
Appendix C: Permissions for Chapter 3 from journal and co-author ........................................168
Appendix D: Permissions for Chapter 4 from journal and co-author 176
Appendix E: Permission for Chapter 5 from co-author .....................178
Appendix F: Permission to reproduce Figure 2.1 from journal ..........179
Appendix G: Permission to reproduce Figure 2.2 from journal ..........184
List of Tables

Table 3.1: Summary statistics for observations with non-missing variables.............................. 46

Table 3.2: OLS Regression Estimates for the Five Regions, Disaggregated by Region. Dependent variable: measured BMI ................................................................. 48

Table 3.3: Blinder-Oaxaca decompositions using the Atlantic provinces as the reference group. Dependent variable: measured BMI ................................................................. 52

Table 3.4: Disaggregated results of select Blinder-Oaxaca decompositions ............................. 53

Table 4.1: Sensitivity and specificity estimates using self-reported BMI and three correction equations ....................................................................................................................... 84

Table 4.2: Odds ratios from regressions for BMI category and selected health conditions for males, controlling for age. (n=821) ................................................................. 86

Table 4.3: Odds ratios from regressions for BMI category and selected health conditions for females, controlling for age. (n=942) ................................................................. 87

Table 5.1: Summary statistics for all study variables ................................................................. 107

Table 5.2: BMI percentiles by sex and region .................................................................................. 109

Table 5.3: OLS estimates for females by region, dependent variable: self-reported BMI .......... 111

Table 5.4: OLS estimates for males by region, dependent variable: self-reported BMI .......... 113
List of Figures and Illustrations

Figure 2.1: Evans and Stoddart model of the determinants of health. Reproduced with permission. ................................................................. 10

Figure 2.2: The ecological model of obesity showing variables relative to individual-level energy intake and output. Reproduced with permission. ......................................................... 12

Figure 3.1: Average personal income and CPI by region. .......................................................... 60

Figure 4.1: Quantile-quantile plots comparing measured BMI with BMI constructed from combinations of measured and self-reported height and weight. ........................................... 74

Figure 4.2: The distribution of BMI by different correction equations and self-report compared to measured BMI in males. ............................................................... 82

Figure 4.3: The distribution of BMI by different correction equations and self-report compared to measured BMI in females ............................................................ 83

Figure 5.1: Kernel densities of BMI across regions for females and males ................................ 100

Figure 5.2: Quantile regression results for income dummy variables amongst Atlantic females ......................................................................................................................... 116

Figure 5.3: Quantile regression results for income dummy variables amongst Quebec females. .......................................................................................................................... 117

Figure 5.4: Quantile regression results for those females who hold a post-secondary degree by region. ................................................................................................................... 118

Figure 5.5: Quantile regression results for those males who hold a post-secondary degree by region. ................................................................................................................... 121

Figure 5.6: Unconditional quantile decompositions between the Atlantic provinces and Quebec. ......................................................................................................................... 125

Figure 5.7: Unconditional quantile decompositions between the Atlantic provinces and Ontario. ..................................................................................................................... 126

Figure 5.8: Unconditional quantile decompositions between the Atlantic provinces and Prairies. .................................................................................................................... 127

Figure 5.9: Unconditional quantile decompositions between the Atlantic provinces and B.C. . 128
### List of Symbols, Abbreviations and Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>Estimated regression coefficient</td>
</tr>
<tr>
<td>BC</td>
<td>British Columbia</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>CANSIM</td>
<td>Canadian Socio-economic Information Management System</td>
</tr>
<tr>
<td>CCHS</td>
<td>Canadian Community Health Survey</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>CRDCN</td>
<td>Canadian Research Data Centre Network</td>
</tr>
<tr>
<td>DLI</td>
<td>Data Liberation Initiative</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Random disturbance</td>
</tr>
<tr>
<td>$e$</td>
<td>Exponential function</td>
</tr>
<tr>
<td>$h_M^2$</td>
<td>Measured height (in metres) squared</td>
</tr>
<tr>
<td>$h_{3R}^2$</td>
<td>Self-reported height (in metres) squared</td>
</tr>
<tr>
<td>kg</td>
<td>Kilograms</td>
</tr>
<tr>
<td>ln(x)</td>
<td>The natural logarithm of “x”</td>
</tr>
<tr>
<td>m</td>
<td>Metres</td>
</tr>
<tr>
<td>n</td>
<td>Number of observations in the sample</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary least squares</td>
</tr>
<tr>
<td>OR</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>PRRDC</td>
<td>Prairie Regional Research Data Centre</td>
</tr>
<tr>
<td>PUMF</td>
<td>Public use microdata file</td>
</tr>
<tr>
<td>RDC</td>
<td>Research Data Centre</td>
</tr>
<tr>
<td>Sens</td>
<td>Sensitivity</td>
</tr>
<tr>
<td>SES</td>
<td>Socio-economic status</td>
</tr>
<tr>
<td>Spec</td>
<td>Specificity</td>
</tr>
<tr>
<td>$W_M$</td>
<td>Measured weight (kilograms)</td>
</tr>
<tr>
<td>$W_{3R}$</td>
<td>Self-reported weight (kilograms)</td>
</tr>
<tr>
<td>$X'$</td>
<td>A vector of the averages of a number of variables</td>
</tr>
</tbody>
</table>
Chapter One: Explanation and Justification for Manuscript-Based Thesis, and Contributions of Authors

1.1 Explanation and justification for manuscript-based thesis

A manuscript-based thesis is a collection of papers that make up a cohesive program of research (1). It differs from the traditional thesis in that the chapters are written separately with the intention to be submitted for peer-review. I opted to pursue this option because I had no previous experience with the peer-review process and wanted to pursue a career in research. This dissertation includes three manuscripts: the first and second are published and the third is prepared for peer-review submission (described in more detail below).

1.2 Contribution of authors

Daniel J. Dutton is the lead author on each manuscript. He conceptualized each study, identified the appropriate data source(s) and analytic techniques, conducted all data management and analysis, led the interpretation of the results, and led the writing of the manuscript.

Lindsay McLaren is listed as co-author on each study. In her role as supervisor, she contributed to the interpretation of the results, contributed to the writing, and provided feedback and guidance throughout the writing of this dissertation and its component manuscripts.

Designation of both DJD and LM as authors on each manuscript is supported by guidelines produced by the International Committee of Medical Journal Editors (2), which stipulates that to be considered an author of a manuscript one must meet all of the following four criteria:

1. The author should have made substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work;
2. The author should have been involved in drafting the work or revising it critically for important intellectual content;

3. The author should have given final approval of the version to be published;

4. The author should be in agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

DJD and LM acknowledge J. C. Herbert Emery for his comments provided during the development of each of these manuscripts.

1.3 Status of manuscripts

The first manuscript, titled Explained and Unexplained Regional Variation in Canadian Obesity Prevalence was published in the journal *Obesity*, vol. 19 no. 7, July 2011 (3). This manuscript involved two rounds of significant revision between June 2010 and December 2010, which required new analysis to be performed and addition of substantial written material to the Results and Discussion sections. DJD conducted the new analysis for those revisions; both DJD and LM wrote the additional material required for the Results and Discussion sections.

The second manuscript, titled How useful is “corrected” body mass index vs. self-reported body mass index? Comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data, has been published by *BMC Public Health* in May 2014 (4). This document was revised twice since initial submission in January 2013. The first revision request, received June 2013, required a substantial rewrite and the second revision request, received March 2014, required new statistical analyses and additional information in all sections. DJD conducted the additional statistical analyses
required for both revisions; both DJD and LM conceptualized and wrote the additional material required for both revisions.

The third manuscript, titled How important are determinants of obesity measured at the individual-level for explaining geographic variation in body mass index distributions? Evidence from Canada is intended for submission to American Journal of Epidemiology.
1.4 References

(1) Department of Community Health Sciences, Faculty of Medicine - University of Calgary. Guidelines for Manuscript-Based Thesis. 2014.


(4) Dutton DJ, McLaren L. The usefulness of "corrected" body mass index vs. self-reported body mass index: comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data. BMC Public Health 2014;14:430.
Chapter Two: Introduction

2.1 Outline

This dissertation is composed of three sequentially developed manuscripts which are anchored in the overall dissertation purpose: to assess the importance of population-level determinants of obesity in Canada by quantifying their impact on individuals living in different regions of the country. In this second chapter, I provide a rationale for the research, an overview of the health issue of focus, obesity, and how population health and economics contribute to understanding obesity. I then describe in greater detail the manuscripts that make up the dissertation and, finally, describe the data and methods used throughout this dissertation. Chapters three, four, and five contain the first, second, and third manuscripts that make up this dissertation. Each of those chapters contains its own specific background and method sections in addition to the material in this chapter. Chapter six concludes the dissertation with a summary of the main findings of the three manuscripts, overall findings of this dissertation and what it contributes to our current understanding of obesity in Canada, overall strengths and limitations of my dissertation research, and overall conclusions.

2.2 Rationale for Dissertation Research

Obesity is a public health problem in Canada. Obesity, often operationalized as a body mass index (BMI) $\geq 30 \text{ kg/m}^2$, is a major public health concern that has been growing in prevalence in Canada, particularly during recent decades (1-3). Between 1985 and 1998 the prevalence of obesity calculated from measured height and weight among adults in Canada more than doubled, from 5.6% to 14.8% (4), and further increased to 25.4% by 2008 (5). It is possible that contemporary obesity prevalence in Canada is not growing as quickly as it has historically, the 2007-09 Canadian Health Measures Survey reports 24.2% of males and 23.2% of females are
obese (6). While not directly comparable to objectively-measured obesity prevalence, obesity prevalence calculated from self-reported height and weight is reported to have increased between 2008 and 2012 from 17.2% to 18.4% (7). Obesity has been associated with various chronic diseases such as cancer (8), type 2 diabetes (9), arthritis (10), hypertension (11), disability related to cardiovascular disease (12) and can lead to premature death (13,14).

It is important to understand the determinants of obesity so that effective prevention strategies can be developed. A superficial way of thinking about obesity is that it is caused in individuals by energy intake being higher than energy output (13). This basic relationship hides the fact that obesity is in fact multifactorial and jointly determined by individual-level variables like behaviours and genetics; more proximal variables like one's immediate physical environment; and more distal variables like public policy or cultural practices (15). Thus, changes in obesity prevalence are best considered as the end product of a complicated ecological model allowing for the interaction of variables across levels (16). Since obesity is determined by factors beyond the individual it follows that effective prevention strategies would address proximal or distal causes of obesity. To determine the feasibility of any such solutions, the relative importance of determinants beyond the individual level need to be established for Canada.

Techniques from economics are useful for studying population-level determinants of obesity. Economics is the study of how people make decisions about resource allocation in the presence of scarcity (17). Individuals face constrained decisions in their everyday lives based on the amount of money or time they have available, thus individuals need to make trade-offs based on the incentives they face. Even though economics has a primary focus on individuals and the incentives they face, economists have developed unique statistical tools that allow for broader
application to other topics. I use these tools to study obesity from a population health framework (which will be explained in more detail below). The theoretical and applied study of statistics within the discipline of economics is called “econometrics” (18), and I use econometric techniques in this dissertation.

As an example of the unique tools developed studying trade-offs consider the topic of returns to human capital: investing one year of youth in education (and forgoing one year of earnings) should lead to higher income later in life on average. Concerned with potential differences between male and female returns to human capital, economists developed a statistical tool called a Blinder-Oaxaca decomposition. This tool split the difference between male and female earnings into a share attributable to observed determinants of earnings (e.g., years in school, seniority, experience, etc.) and a remaining share (19,20). That remaining share, which resulted in males having higher wages than females, was considered discriminatory since economic theory dictates that earnings should be determined by observable determinants excluding sex. Males and females are two populations which can be substituted for any other populations of interest, and I adopt this tool to study differences in BMI across regions of Canada and whether differences remain after controlling for observable determinants of BMI measured at the individual level.

The study of regional differences in Canadian BMI is a natural fit for econometric techniques. Canadians from the Atlantic provinces (i.e., residents of Prince Edward Island, New Brunswick, Nova Scotia, and Newfoundland and Labrador) have been reported to have a higher average BMI than individuals in Ontario, Quebec, the Prairies (i.e., residents of Manitoba, Saskatchewan, Alberta), and British Columbia (21-23). These regions can be conceptualized as different populations, and as in the human capital example above, differences in BMI across
regions can be split into a share explained by observable determinants of obesity and an unexplained share. In this case the unexplained share would represent population-level factors that influence BMI differences.

This concept might be best explained as a counterfactual. If population-level factors did not influence BMI, that is, if the ecological model of obesity were incorrect and only individual-level factors mattered, then cross-regional differences in BMI would be wholly attributable to cross-regional differences in individual-level determinants of obesity. So the difference between the Atlantic provinces and other regions would be wholly attributable to a less desirable distribution of the individual-level determinants in the Atlantic provinces. None of the difference between the regions would remain unexplained. Thus, by applying econometric techniques to cross-regional BMI differences in Canada the relative importance of population-level determinants of obesity can be established for regional differences.

2.2.1 Conceptual framework

My area of specialization is population and public health; public health in Canada is defined by the Public Health Agency of Canada as “the organized efforts of society to keep people healthy and prevent injury, illness and premature death” through “a combination of programs, services and policies that protect and promote the health of all Canadians.” (24) Population health is outlined by the same organization as an approach that “focuses on the interrelated conditions and factors that influence the health of populations over the life course, identifies systematic variations in their patterns of occurrence, and applies the resulting knowledge to develop and implement policies and actions to improve the health and well-being of those populations.” (25) Defining of the term “population health” is challenging since it can refer to different things in different contexts. For example, population health is a perspective
(i.e., a general way of thinking about health in line with the previous definition), a type of research (i.e., the generation of empirical findings describing population-level health trends or determinants), a framework (i.e., an integration of research findings into an understanding of the determinants of health), and an approach (i.e., using a population health framework to address public health policy) (26).

This dissertation uses a population health framework to study obesity in the Canadian population. Population health in Canada is a young discipline and the formalization of population health studies in Canada is sometimes portrayed as a series of key events (e.g., the Lalonde Report in 1974 was an example of explicit government acknowledgment within a major published report of the importance of determinants of health beyond the individual (27).) One such key event was the publication of a model of the determinants of health by Evans and Stoddart in 1990 (28). The model they proposed explicitly identified health as malleable by manipulation of upstream (distal) variables, that is, variables beyond the individual-level, and criticized the idea that publicly funded health care is the only determinant of health that governments can address. While their model was not universally accepted (29-31) nor the first instance of studying determinants of health beyond the individual-level (shortened to “social determinants of health”) (32) it is a useful tool for describing how the population health discipline views health outcomes: health is determined by multiple factors, some environmental, and understanding their relationship is key to improving the health of populations. The model by Evans and Stoddart is reproduced in Figure 2.1.
Population health research studies the determinants of health at a population-level. That means population health research is particularly interested in trends that exist amongst groups of people and might not be identifiable when considering individual cases (33). The population health interest in groups of people lends itself to explicit consideration of distributions of health variables, the distributions of determinants of health, and the interaction of different types of factors (e.g., biological, economic, social) that can cause those distributions to be different across populations (34). When there are health differences in the aforementioned distributions between groups, population health differentiates between differences that represent inequality between groups (i.e., the groups are not equal) and differences that represent inequities between groups (i.e., the groups are not equal for avoidable and unfair reasons) (35). Since health inequalities
that are equitable are deemed fair or just (e.g., natural variation in biological factors), population health research is particularly concerned with inequitable differences between groups (35). Thus a population health framework allows for the simultaneous consideration of the social determinants of health; the distribution of those determinants and their health consequences across populations; and the way those determinants and health consequences are distributed.

The population health framework can accommodate an ecological model of obesity (15,16,36,37). The ecological model of obesity explicitly identifies individual-level behaviours like energy intake and output as influenced by the physical, economic, and socio-cultural environment. The model defines categories of variables by their position relative to the individual: more proximal variables such as the immediate physical environment operate directly on the individual, whereas more distal variables like taxes and subsidies change in the impact of proximal variables on the individual (15). The factors depicted at these levels are not separate; the model allows for interactions between factors at different conceptual levels. The ecological model of obesity is presented in Figure 2.2.
The ecological model of obesity is similar to the model of the social determinants of health presented in Figure 2.1: the factors beyond the individual level have impacts on individual-level health status. In the ecological model of obesity (Figure 2.2) the authors differentiate between systemic drivers of obesity and environmental drivers of obesity. When the authors use the term “driver” they are referring to what has been called a “fundamental cause” elsewhere (38): drivers are factors that brought about the currently high prevalence of obesity and measures to address obesity need to address these drivers. More proximal to the individual are “environmental moderators”, which the authors distinguish from drivers to emphasize that they interact with the drivers to change individual-level behaviour. Throughout this dissertation I will make no distinction between systemic and environmental drivers, I will refer to them both
collectively as “population-level factors” or “population-level determinants” to emphasize that they are distal from the individual.

2.3 Determinants of obesity

2.3.1 Measurement of obesity

Obesity is operationalized using BMI in this dissertation, but adiposity can be measured in multiple ways. BMI is not capable of distinguishing between weight from fat and weight from muscle, therefore individuals with high volumes of muscle mass might incorrectly be labeled as obese (39). While other, more accurate, measures of adiposity exist (e.g., (40,41)) BMI is the measure which World Health Organization recommendations are based on (41,42). BMI is useful for measuring obesity status in populations due to the relative ease of collecting comparable measures on a large scale, the low proportion of individuals likely to be misclassified as obese by muscle instead of fat, and the correlation between BMI and health outcomes (41).

2.3.2 A brief overview of the complicated relationship between the determinants of obesity

Evidence from developed countries has demonstrated how multiple factors play a role in terms of obesity in individuals. At the individual-level, genetic and biological factors play a role in whether an individual becomes obese or not: genetic predispositions to gaining weight or the particular makeup of an individual’s brain can facilitate weight gain (43). However, the rapid increase in obesity prevalence over time makes population-wide changes in biological factors seem unlikely (44).

Diet and physical activity are the two behavioural determinants of energy intake and output in individuals that have been most widely studied. The consumption of certain macronutrients (e.g., carbohydrates (45)) or types of food (e.g., fruits and vegetables (46)) have been linked to obesity status in individuals. Low levels of physical activity are associated with a
higher relative risk of death, even amongst obese individuals (47), and low levels of physical activity have been linked to obesity in many studies. For example, studies have shown a negative correlation between steps-per-day and BMI (48); that high-intensity exercise is associated with a lower odds of obesity (49); and individuals with higher leisure time or occupational physical activity exhibited lower odds of obesity than comparatively inactive groups (50). However, in an ecological model of obesity, these behaviours are considered outcomes of a complicated process that is driven by forces beyond an individual’s control (15,16).

One factor that can affect an individual’s behaviour is their immediate physical environment. Ready access to areas conducive to physical activity is not evenly distributed by neighbourhoods: those with lower socio-economic status (SES) or higher proportions of minorities have been found to have fewer physical activity areas (e.g., parks, pools) than higher SES, low minority neighbourhoods (51). Not only are food options limited by the physical environment, with lower income urban areas being less likely to offer variety in healthy food choices (52-54), but food consumption is affected by neighbourhood characteristics. For example, one study found consumption of fruits and vegetables varies by neighbourhood SES status and that relationship was itself different for different ethnicities (whites, blacks, and Mexican-Americans) (55). The differential relationship between neighbourhood SES status and food consumption by different ethnicities is one way that physical environments can interact with higher level social factors.

Social factors can be local or cultural and operate on other determinants of obesity in complicated ways (15,16). For instance, different ethnic groups might have different ideas or norms about appropriate diet and body size for children (56). While those practices might not have traditionally presented health problems, with corresponding changes to the physical
environment (e.g., cities designed to accommodate car ownership) those practices can result in higher obesity prevalence where it was once rare (57). Since obese children are at least twice as likely to become obese adults as normal weight children (58), increases in childhood obesity prevalence linked to social factors could persist throughout life.

In an ecological model of obesity population-level factors, those causes of obesity furthest from the individual, are ultimately responsible for the current obesity epidemic (15). These population-level factors must be recent and widespread in their scope to be responsible for recent increases in obesity. One proposed population-level factor is food supply practices (15). Consider, as one factor in the larger issue of “food supply”, the case of the cost of food. Food costs are acknowledged in the literature as a determinant of food consumed, and energy dense foods are estimated to be cheaper on a cost-per-unit-of-energy scale in several studies (59-62). Proponents of the theory that sugar-sweetened soft drinks contribute to growing obesity prevalence point to the price of such drinks, which has not increased in real terms as fast as the price of healthy food choices (63). A review examined the role of prices on eating patterns (64) and found that both subsidies and taxes can change food consumption patterns. Subsidies to healthy foods were found to increase the purchasing of healthy foods, yet taxes were only found to have a detectable effect on food consumption patterns when they were “large” (an increase in price at minimum greater than 20%, but some studies showed no effect even at that level.) At a national level, the rise in the quantity of food available corresponding with the decrease in price has been shown to correlate with increases in average weight (65,66).

The price of food is one example of a population-level determinant of obesity, that could be targeted by public policy. While the population-level determinants of obesity are clearly important, the impact of any one particular determinant is quantified when it is explicitly under
study, generally to the exclusion to all other population-level determinants. Continuing with the price of food example, while some claim that the price of food can wholly explain cross-time increases in obesity prevalence (65), other population-level determinants (e.g., trends in infrastructure development conducive to driving) are assumed constant in order to simplify analysis. One contribution of my work in this dissertation is the quantification of population-level determinants in Canada without focusing on any one determinant in particular.

2.3.3 Epidemiology and determinants of obesity in Canada

Because I focus on the Canadian population in this dissertation I will briefly summarize the Canadian evidence for an ecological model of obesity here. Canadian obesity prevalence, as has been mentioned already, has been increasing markedly since at least 1985 (4,5). That increase has not been uniform across the country: maps tracking the prevalence of obesity using self-reported height and weight from 1985 to 1998 were made using data from five national surveys: the 1985 and 1990 Health Promotion Surveys and the 1994, 1996, and 1998 cycles of the National Population Health Survey (4). These maps show that over the whole time period Quebec and British Columbia did not reach a prevalence ≥ 15% by 1998 like the rest of the country, nor did those provinces experience as fast of an increase in obesity prevalence as the rest of the country. Those same maps were updated using data from 2000 to 2011 and show all provinces were above 15% obesity prevalence by 2000 (23). In 2011 Quebec and British Columbia continued to have the lowest obesity prevalence of all provinces, between 20 and 24%. The Atlantic provinces (Newfoundland, Nova Scotia, New Brunswick, and Prince Edward Island) all had obesity prevalence between 30 and 34%. Ontario and the Prairie provinces (Alberta, Saskatchewan, and Manitoba) were between 25% and 29% prevalence. The high prevalence of obesity in the Atlantic provinces relative to the rest of Canada was observed as
early as 1990 (4). Using multiple national datasets\(^1\) to estimate province-specific BMI statuses (normal weight \((18.5 \leq \text{BMI} < 25)\) overweight \((25 \leq \text{BMI} < 30)\), obese\((\text{BMI} \geq 30))\), then fitting a line through those BMI status estimates and extrapolating forwards in time, Twells et al. predicted that by 2019 more than half of the population in each province in Canada will be overweight or obese (22). That increase is estimated to be largest in the Atlantic provinces and driven mainly by increases in obesity prevalence (22). Cross-regional differences in obesity prevalence are persistent; the ecological model suggests such differences would be ultimately caused by region-specific population-level factors.

Canadian evidence on individual-level behavioural causes of obesity is consistent with the international evidence. The consumption of certain foods like fruits and vegetables is correlated with lower BMI cross-sectionally (67,68) and physical activity is associated with lower BMI (69). One exhaustive Canadian report published by the Canadian Institute for Health Information identified that obesity prevalence varies by smoking status, marital status, language, province of residence, employment status, and visible minority status (70). In short, the variation in obesity status amongst individuals is correlated with many determinants measured at the individual level.

Neighbourhood-level, or community-level, associations with obesity have been studied in Canada as potential sources of individual-level behaviour change. Operationally, the physical aspects of community-level variables (distance, geographic boundaries) are sometimes studied alongside the social aspects of community-level variables (neighbourhood-level SES). In the Canadian literature both are found to contribute meaningfully to individual-level variation in

\(^1\) The datasets used in the linear prediction model are the 2001, 2003, 2005, 2007-08, 2009-10, and 2011 Canadian Community Health Surveys.
BMI: low average value of a household in a neighbourhood (71), low proportion of individuals in a neighbourhood with high school education (71), high levels of unemployment (72), low income (73), and close proximity of an individual’s home to fast food outlets (74) are all associated with increased likelihood of obesity. The datasets used to study these proximal causes of obesity are generally national-level and there is little research that examines both neighbourhood effects and allows for provincial or regional differences. Evidence exists to suggest that, while some built environmental factors such as access to green space may be correlated with lower obesity prevalence in general, it is not the case that the effects of environmental factors are homogeneous across different communities (75). This in turn suggests that Canadian regional differences in obesity prevalence could be driven by more distal population-level determinants that interact with more proximal causes such as physical and social environments.

Canadian research on the relative importance of the population-level determinants of obesity and individual-level variables is lacking. While obesity is acknowledged to be the outcome of a complex system (15,16,76), there is a disconnect between how the obesity problem is theoretically understood to be complicated, multifactorial, and driven by factors beyond the individual, and how regional differences in obesity prevalence are sometimes presented in quantitative studies. In Canada, the aforementioned differences in obesity prevalence by geography have been implied to be caused by regional variation in one or more determinants of obesity measured at the individual level (22,23,77,78) such as education, income, physical activity, diet, and others. While this is technically likely true, I have presented evidence showing the implications of this message is too simple. Quantitative research tools that explicitly model the relative importance of population- and individual-level factors are rarely utilized. When they
are utilized (e.g., (79)) they are often used to study individual-level relationships with obesity that change over populations, rather than the effect of population-level factors themselves. My dissertation addresses this key gap in the literature by using methods from economics, a discipline that is itself interested in the theory behind systemic drivers of behaviour.

2.3.4 Obesity in the economics literature

Health economics in a public health context is often wholly characterized as the evaluation of new health technology or estimating costs (80). In reality, evaluation is a part of the wider field of economics, which is concerned with how individuals make decisions in the face of limited resources (17). Economists have developed theories regarding how people might react to incentives and those theories are tested on population data. From that point of view, the study of obesity is within the purview of economics as the study of any health-related behaviour could be: Individuals are making decisions that are resulting in increasing obesity prevalence, how does economic theory explain such a trend?

In trying to answer that question economists consider what changes have occurred that might cause increases in obesity prevalence, and the proposed mechanisms are mostly what would be considered systemic drivers in the ecological model of obesity. For example, “technological change” (44,81-83) is a broad term used by economists to describe the shift towards sedentary productive jobs (thereby shifting physical activity from work time to leisure time) and the simultaneous increase in food production technology that decreased the price paid for different types of food (83,84). Technological change has altered the incentives individuals face, and thereby changed their behaviour, which resulted in an increase in obesity prevalence. Technological change is a systemic driver of the obesity epidemic, but economists also study more proximal causes of obesity. The overall decrease in food prices may be driving up
consumption of calories overall but calorie dense foods are becoming cheaper at a faster rate, which may lead to substitution of those foods for relatively more expensive, healthier options (60,85). The general increase in income in developed countries, a result of transition to a knowledge-based economy, is likely associated with obesity (86). Individual-level behaviours that might be affected by technological change are considered fixed in the economics literature (and known as “preferences”), so it is expected that not all individuals will change their behaviour equally in response to a stimulus (17). This treatment of individual-level behaviours is one of two key differences between the approaches to obesity studies taken by economics and population health.

The two key differences in obesity studies between economics and population health relevant here are the normative recommendations of the analyses and the methodological approaches of the two disciplines. Normative recommendations, or what should be done regarding obesity, tend to be absent from economics and that is a consequence of the discipline’s belief in individual-level decision making ability (87). Population health is concerned with health as an outcome (33), so obesity is considered bad through its association with poor health, thus most obesity studies in population health begin from the premise that decreasing obesity prevalence is important.

This difference becomes evident through an example: consider that obesity might simply be a product of individuals lacking knowledge about what makes one obese. If that lack of information was rectified yet individuals remained obese, the two disciplines may draw different inferences based on normative recommendations. Economists would consider the market failure (lack of information) corrected, therefore if individuals remain obese they must be behaving optimally and simply do not value a lower BMI (87). In that case, no further intervention is
required. Population health researchers might consider this result as partly a failure of knowledge translation, as the desired goal (decreased obesity prevalence) was not achieved (88). In that case, it is not clear from a population health perspective that no further action is necessary\(^2\). Thus, while economists and population health researchers study obesity with similar frameworks (individuals responding to changing incentives determined at the population level vs. the ecological model) the studies are not similar in their conclusions. So while both disciplines consider the obesity problem to be rooted in population-level factors they do not necessarily overlap in their recommendations as to whether any intervention is necessary.

The second key difference between economics and population health is the former’s dedication to specific types of quantitative methodology. While methods used widely in econometrics are becoming known in quantitative population health (e.g., instrumental variables, a common tool for addressing reverse causation in observational models, are covered in graduate-level epidemiology textbooks (89)), population health is multidisciplinary, inclusive of qualitative methods and, as such, is necessarily broad in its methodological treatment of health questions. Thus, there are opportunities to broaden the practice of quantitative population health regarding obesity by borrowing established statistical tools from econometrics.

Cross-application of statistical techniques from econometrics is appropriate for the research I have conducted in this dissertation. Obesity is unique in that (a) economics and population health both consider increasing prevalence of obesity to be a function of determinants beyond the individual, and (b) statistical tools exist that are adaptable to answering questions

\(^2\) Clearly, uptake of information is an intervention at the agentic end of the intervention spectrum (107) and universal uptake would not necessarily be expected anyway by population health researchers. This example was chosen because lack of information by the agent is a market failure in economics and knowledge translation is a strong research topic in health research.
regarding obesity in a population health framework (more on these specific tools below.) I acknowledge throughout this dissertation that I view lower levels of obesity prevalence as desirable and my application of econometric techniques in the manuscripts is within that normative framework.
2.4 Manuscripts

This dissertation contains three manuscripts that were developed sequentially to establish the importance of population-level determinants of obesity in Canada by quantifying their impact on individuals living in different regions of the country.

The objective of the first manuscript was to examine potential sources of geographic variation of obesity in Canada, by analyzing BMI among Canadian working-age men and women in relation to various socio-demographic and behavioural variables across different regions of Canada. This was accomplished by quantifying the relative importance of determinants measured at the individual level and population-level determinants through application of a decomposition technique to Canadian population-level measured height and weight data. The results of the decomposition attribute a share of the difference in average BMI to population-level determinants of obesity, which have not been quantified before as a whole.

Stemming from the first manuscript, it made sense to then consider the entire BMI distribution, because determinants of obesity measured at the individual level have been shown to have differential effects across the BMI distribution (90-93). However, using the decomposition technique across the whole BMI distribution requires many more observations than were available for the first manuscript. This was a major problem for my analysis because Statistics Canada only gathers measured height and weight data occasionally. However, self-reported BMI is gathered more frequently and it is possible to use self-reported BMI in place of measured height and weight if a statistical correction is applied. In Canada one such correction equation is used by researchers working with self-reported height and weight (94). While that correction equation is widely used, I wanted to test its appropriateness for modeling the entire BMI distribution and determinants of the distribution. For ease of presentation, I refer to BMI
computed from measured height and weight as “measured BMI”, BMI computed from self-reported height and weight as “self-reported BMI”, and BMI computed using corrected values of any kind “corrected BMI”.

The purpose of the second manuscript was thus to evaluate the usefulness of existing and new correction equations for BMI in population-based research. To accomplish this, three objectives were pursued: 1) compare the self-reported and corrected BMI distributions; 2) compare self-reported and corrected BMI to measured BMI based on sensitivity and specificity of measured obesity; and 3) compare self-reported and corrected BMI to measured BMI in regression models of various health conditions, in terms of statistical significance, coefficient magnitude, and direction of the coefficient (above or below the measured coefficient). Existing Canadian correction equations (6,94) neglect that the divergence between misreported weight and actual weight in the population grows over actual weight. That is, heavier people as a group misreport to a greater degree than lighter people, and at an increasing rate across weight. Thus, the second manuscript proposes and tests an alternative correction against an existing correction and self-reported BMI. The results of this manuscript show that there is no superior equation for my intended purposes, therefore I could begin the decomposition exercise across the entire distribution using the data available.

The objective of the third manuscript was to examine whether equalizing the identified determinants of BMI across geographic regions of Canada could be reasonably expected to reduce differences in BMI distributions across those regions. This was accomplished by quantifying the relative importance of determinants measured at the individual level and population-level determinants for the entire BMI distribution across regions by using a decomposition strategy similar to the first manuscript. Using self-reported BMI from multiple
cycles of the CCHS, this study was able to compile a large enough sample to utilize the
decomposition tools for the entire BMI distribution. Replacing self-reported BMI with a
corrected BMI did not qualitatively change the results.

2.5 Methods

2.5.1 Data Source

All three manuscripts in this dissertation use national level data collected by Statistics
Canada. The opportunity for me to use Statistics Canada data for my dissertation research
reflects important and relatively recent changes to the national data landscape. In 1996 the
federal government announced that Statistics Canada data would be disseminated to Canadian
academics through university libraries under the Data Liberation Initiative (DLI) (95). Prior to
1996, access to these data was provided on a cost-recovery model: academics were required to
pay for access to microdata. The DLI is a national initiative that aims to increase general use of
Statistics Canada’s proprietary datasets (96). Broadly, the DLI provides universities or research
institutions with Statistics Canada data for a fee; those data can in turn be accessed by members
of those institutions for research purposes. Examples of the types of data made available through
the DLI at the University of Calgary include census data, social survey data (e.g., General Social
Survey), economic data (e.g., the Labour Force Survey), data on special topics of interest (e.g.,
the Ethnic Diversity Survey) and health data such as the Canadian Community Health Survey
(97).

The data for this dissertation come from the Canadian Community Health Survey
(CCHS), which was administered by Statistics Canada bi-annually from 2001 to 2007, and on an
annual, ongoing basis since 2007 (98). The CCHS is a cross-sectional, nationally representative
survey with two primary objectives: to “provide timely, reliable, cross-sectional estimates of
health determinants, health status and health system utilization across Canada; and gather data at the sub-provincial levels of geography (99).” The sub-provincial levels of geography are defined as health regions. To that end, the CCHS collects data on approximately 130,000 individuals, ages 12 and up, every two years (approximately 65,000 annually since 2007). The CCHS excludes some sub-groups, namely, individuals living on Aboriginal reserves, in the Canadian Forces, living in institutions, or selected remote areas; these excluded sub-groups represent less than 3% of the Canadian population.

The CCHS samples individuals living in private dwellings through one of three sampling frames: an area frame, a phone number list frame, and random digit dialing (99). The area frame uses a multistage stratified cluster design wherein the country is broken down into major urban centres, cities, and rural regions. Those regions are further subdivided into geographic and socio-economic strata based on the Labour Force Survey. Within those strata are clusters of households which consist of 250 to 300 households each. Clusters are sampled from the strata with a sampling probability in proportion to their population, then households are randomly sampled from the chosen clusters. Approximately 49% of the CCHS sample is generated by the area frame, the remainder comes from the phone number list frame (50%) and random digit dialing (1%). The phone number list frame maps a directory of Canadian phone numbers to postal codes to complement the area frame at the health region level. Phone numbers are randomly sampled from each health region list. Random digit dialing involves compiling possible phone numbers for a health region.

The administration of the CCHS is conducted using Computer Assisted Interviewing, either in person or via telephone, which involves trained professionals using computers to facilitate the data gathering process (99). The data gathering period in each cluster is split across
11 months to avoid seasonal trends in variables gathered. The CCHS interview takes approximately 40 minutes by phone or in person. There are a number of quality-assurance mechanisms in place; for example, the questions are automatically updated based on previous answers provided and the computer survey program prompts the interviewer so that the interviews are standardized in procedure and content. The central planners of the CCHS can also periodically audit the interviewers to ensure quality and consistency. Interviewees are sent introductory letters in advance of the interviewer arriving to explain the survey and its importance. The interviewers themselves are trained on non-response minimization techniques to encourage interviewees to participate; for instance, interviewers are able to answer questions regarding data use and privacy. If the interviewee is unable to complete the survey multiple attempts are made to accommodate them by phone and in person. This results in high response rates for each cycle of the CCHS (e.g., 79% for the CCHS cycle 3.1) (99).

The CCHS is available in two forms: the public use microdata file (PUMF) and the confidential microdata file, or the master file (100). The master file is composed of most of the original information gathered from the survey by Statistics Canada and a set of variables that can be used as weights for calculating population-level estimates and variances. The master file requires formal application for access (more detail below). The PUMF is a version of the master file that has been censored through aggregation of some variables (e.g., continuous age to age groups), left or right censoring other variables (e.g., BMI), and deleting others (e.g., first three characters of postal code).

For this dissertation I used multiple cycles of the CCHS. I used the PUMF of CCHS Cycle 2.2 for the first manuscript. The PUMF was appropriate for that first manuscript because it contained measured height and weight, but after realizing the relative limitations of the PUMF
(for example, BMI is censored at approximately 56, age is available in grouped categories instead of by year) I decided to switch to master files for both the second and third manuscripts. The access procedure is explained next.

2.5.2 Data access

One specific aspect of the DLI has been to establish Research Data Centres (RDCs) across Canada to facilitate working with Statistics Canada’s master files. The second and third manuscripts in this dissertation required the CCHS master files, access to which was arranged through the Prairie Regional Research Data Centre (PRRDC), the RDC based at the University of Calgary.

In order to access the PRRDC one must complete a two-step process: a research proposal must be submitted for peer review to Statistics Canada analysts, and a security screening, including an oath of secrecy, must be completed to become a “deemed employee” of Statistics Canada (101). After the application process for this dissertation was approved, all work on master file data relevant to this dissertation had to take place within the physical PRRDC at the University of Calgary. Results were then vetted by a Statistics Canada analyst based at the RDC for residual disclosure risk before the results could be taken out of the RDC.

2.5.3 Data management and analysis

This dissertation’s contribution to the literature is the statistical techniques applied to understanding population level determinants of obesity. In particular, the tools used are Blinder-Oaxaca decomposition (19,20), unconditional quantile regression (102-104), and Blinder-Oaxaca decomposition of quantile regression results (102-104). A full description of each of those tools is provided in the chapters where they are used.
Before these analyses could be conducted, significant effort had to be expended managing the data provided by Statistics Canada. Though significant processing of RDC data takes place at Statistics Canada prior to the data’s release to the RDC sites, there is generally still a large amount of data management required to prepare the data for a particular project. Management involves cleaning variables (e.g., replacing values for missing observations coded as “999” with a value the software will recognize as missing, such as “.”; dichotomizing variables into 0/1 from other values; applying labels for analytical ease; removal of subgroups such as pregnant or breastfeeding women; etc.) and merging datasets.

Throughout the dissertation, summary statistics and results are reported separately for males and females. This was done to allow for the relationship between covariates and BMI to vary for males and females. Thus, the analysis allows for effect measure modification by sex which in turn implies there might exist sex-specific responses to region-level effects or policies.

Further, some of my analyses required non-standard statistical software code made available by other researchers, such as the Blinder-Oaxaca decomposition (19,20) command developed by Jann (102), and procedure code for both the unconditional quantile regression and decomposition of those quantile regressions (103-105) developed by Fortin (106). Other analysis required the development of unique code to solve problems for which no built-in macro exists in the software (e.g., solving for the correction equations in the second manuscript numerically). The remaining analyses employed in this dissertation (OLS regression, probit regression, cross-tabulations, etc.) use standardized macro commands included in the software package used. All analyses were done using Stata 10 and Stata 11. All other relevant information on the specific statistical methods used in each manuscript is provided within the manuscript itself.
2.5.4 Ethics approval

Ethics approval for this dissertation was received from the University of Calgary’s Conjoint Health Research Ethics Board. The ethics ID for this dissertation is E-23704.
2.6 References


(5) Canadian Institute for Health Information, Public Health Agency of Canada. Obesity in Canada: A joint report from the Public Health Agency of Canada and the Canadian Institute for Health Information. 2011;HP5-107/2011E.


(7) Statistics Canada. Table 105-0501 - Health indicator profile, annual estimates, by age group and sex, Canada, provinces, territories, health regions (2012 boundaries) and peer groups occasional. CANSIM (database) Accessed 27-05-2014.


(65) Swinburn B, Sacks G, Ravussin E. Increased food energy supply is more than sufficient to explain the US epidemic of obesity. Am J Clin Nutr 2009 Dec;90(6):1453-1456.


(73) Oliver LN, Hayes MV. Effects of neighbourhood income on reported body mass index: an eight year longitudinal study of Canadian children. BMC Public Health 2008 Jan 14;8:16-2458-8-16.


(91) Williams PT. Evidence that obesity risk factor potencies are weight dependent, a phenomenon that may explain accelerated weight gain in western societies. PLoS One 2011;6(11):e27657.


(102) Jann B. A Stata implementation of the Blinder-Oaxaca decomposition. ETH Zurich, Chair of Sociology, ETH Zurich Sociology Working Papers 2008(5).


3.1 Abstract

The objective of our study was to examine socio-demographic and behavioural variables underlying the geographic variation of obesity in Canada. We aimed to quantify the share of regional variation in average body mass index (BMI) attributable to commonly cited determinants of obesity and the remaining share, which is attributable to the idiosyncrasies of the regional environment ("regional effects"). Using data from the Canadian Community Health Survey (2004), OLS regression, and Blinder-Oaxaca decomposition to decompose the difference in mean BMI between regions, we quantify two parts of the difference: a share explained by different levels of the covariates and a share explained by those covariates having different effects on BMI in the different regions, using the Atlantic provinces as the reference group. We observed that some differences (e.g., average BMI for males in Quebec compared to the Atlantic provinces) are mostly explained by the different levels of socio-demographic and behavioural covariates, while others (e.g., average BMI for females in Quebec compared to the Atlantic provinces) are mostly explained by the different effects of the covariates on BMI. In the latter scenario, even if covariates were made to be identical in the different regions, the difference in average BMI would persist. Thus, targeting covariates in different regions through plans like physical activity or nutrition policy, income equalization, or education subsidies will have ambiguous effects for addressing disparate obesity levels, being plausible policy options in some regions but less so in others. Future research and policy would benefit from identifying these region-specific attributes that have local implications for BMI.
3.2 Introduction

Obesity (often defined as a body mass index (BMI) $\geq 30$ kg/m$^2$) is a major public health concern that has been growing more prevalent in Canada and elsewhere since the 1950's but particularly during recent decades (1-3). Between 1985 and 1998 (4) the prevalence of obesity among adults more than doubled, from 5.6% to 14.8%, and further increased to 23.1% by 2004 (5). From a population health perspective (6), obesity's association with various chronic diseases (e.g., cancer (7), leukemia (8), type 2 diabetes (9), dementia (10), arthritis (11), hypertension (12) and disability related to cardiovascular disease (13)) draws attention to the determinants of obesity: individual-level risk factors (e.g., physiology, behaviours) interacting with population-level regional determinants of health (e.g., built environment, societal norms), which may be driving the increase in obesity prevalence by affecting the population as a whole.

In Canada, there exists regional variation in obesity (4,14-17) such that Canadians from the Atlantic provinces (i.e., residents of Prince Edward Island, New Brunswick, Nova Scotia, and Newfoundland and Labrador) have been reported to have a higher average BMI than Western Canadians (i.e., residents of Manitoba, Saskatchewan, Alberta, and British Columbia.) More specifically, Shields and Tjepkema (16) showed that the proportion of obese people was lowest in British Columbia (19% obese), higher in the Prairies (the average of Manitoba, Saskatchewan, and Alberta being 28% obese), then dropped in Ontario and Quebec (Ontario 23% obese, Quebec 22% obese), and increased again in the Atlantic provinces (the average being 28.5%.) The overall average for obesity reported in Canada was 23%.

Geographical or spatial analysis of prevalence, while an important tool for public health surveillance (15) and a key starting point for studying obesity, needs to be followed by an investigation of the possible reasons underlying geographic variation in health outcomes such as
obesity. The difference in prevalence between provinces insinuates that there could be possible determinants of obesity so widespread (e.g., applying to entire provinces) in society that they fail to register as risks to individuals (6). These determinants might include the socio-cultural environment, the physical environment or economic policy in a region (18) and are by definition widespread if acting on a regional level. Therefore, estimates of causality or correlation within regions could either miss these important regional risks or over-state minor but heterogeneous ones.

Several factors may contribute to regional variation in obesity. For example, in an attempt to explain the regional variation in obesity across Canada described above, Vanasse et al. (14) estimated fruit and vegetable consumption and physical activity patterns by province. These authors detected substantial regional variation, reporting that 33.5% of British Columbians had low levels of leisure-time physical activity (i.e., energy expenditure of less than 1.5 kcal per kilogram of body weight per day) compared to 71% of Newfoundlanders, and 46.3% of British Columbians ate fewer than 5 servings of fruits and vegetables per day compared to 79.8% of Newfoundlanders. Although these patterns correspond with geographic variation in obesity, it is unlikely that these two variables are sufficient to fully explain geographic variation in obesity; other variables are likely at play. Further, the findings by Vanasse et al. could be interpreted to mean that interventions should focus on facilitating behaviour change, when in fact the impact of even highly effective behaviour change interventions might differ between regions.

Differences in BMI between geographic regions may also be explained, in part, by differences in socio-economic variables across the regions. Research exists to show that, on average, there is an increased risk of obesity associated with lower education and different levels of income (19-23). Thus, a relatively lower obesity rate in one region may reflect different
educational and income attainment profiles in that region if education and income have a homogenous effect across regions. On the other hand, education and income may have heterogeneous effects across regions, in which case region-specific attributes may have a role in the BMI differential by interacting with individual characteristics like household income and educational attainment to produce different outcomes. Region-specific variables may be conceptualized as "ecological" variables to which everyone in a region is exposed; examples include regional variables from the economic domain (e.g. childcare policy and food prices), socio-cultural domain (e.g. attitudes towards physical activity), and physical domain (e.g. the built environment and access to healthy food vendors.)

The objective of this paper was to examine potential sources of geographic variation of obesity in Canada, by analyzing BMI among Canadian working-age men and women in relation to various socio-demographic and behavioural variables across different regions of Canada. Policymakers may find it valuable to know whether the variation reflects a differential distribution of variables (e.g., more people are getting a higher education in one region than another) or whether something else is going on that causes the effect of a variable to manifest differently (e.g., having a higher education has a different [in direction or magnitude] impact on BMI in one region than in another [this phenomenon will henceforth referred to as a higher return to education, for BMI]). Quantifying the contribution of these two separate yet simultaneously-occurring effects on BMI will help predict the magnitude of impact of possible interventions. To separate these two sources of regional differences in BMI we use a Blinder-Oaxaca decomposition (24,25), which partitions the BMI differences between regions into "a share attributed to endowments" (i.e., variation in BMI that is a function of variation in the attributes of residents in the different regions), and "a share attributed to coefficients" (i.e.,
variation in BMI that is a function of variation in the returns to the attributes considered). This approach enables identification of potential policy impacts.

3.3 Methods and Procedures

3.3.1 Data Source and Study Design

This study used a cross-sectional design with data from Statistics Canada's Canadian Community Health Survey (CCHS), Cycle 2.2 (public use version). Details are available at www.statcan.gc.ca. The CCHS was designed to collect health and demographic information from people of all ages living in Canada, with the exclusion of people living on Crown lands or Aboriginal People's Reserves, institutional residents, full-time members of the Canadian forces, and people living in the Territories or remote areas of the provinces. The cycle attained a national response rate of 76.5% (26). Our subsample of interest was adults of working age, which we defined as 20 to 64 years.

We used the following variables: BMI (kg/m$^2$ based on height and weight which were measured as part of a home visit), sex, age in groups of five years, marital status (single, including divorced and never married; or married, including married and common law), highest level of education attained (less than secondary, graduated secondary school, some post-secondary, or graduated post-secondary), employment status (employed or not), total household income levels ($0 to $14,999, $15,000 to $29,999, $30,000 to $49,999, $50,000 to $79,999, $80,000 and up), smoking status (non-smoker, occasional smoker, daily smoker), alcohol intake (non-drinker, occasional drinker, daily drinker), physical activity levels (active, moderate, or inactive based on an index constructed by Statistics Canada using daily energy expenditure), fruit and vegetable consumption (number per day), and whether or not the individual was born in Canada. Canada was split into five regions for this analysis: British Columbia, the Prairies
(Alberta, Saskatchewan, and Manitoba), Ontario, Quebec, and the Atlantic provinces (Nova Scotia, New Brunswick, Prince Edward Island, and Newfoundland). These regions are familiar regional breakdowns of Canada and provinces within the grouped regions are similar in their obesity and overweight profiles (16).

3.3.2 Statistical Analysis

First, we analyzed the conditional mean of BMI using ordinary least squares (OLS) regression with Huber-White standard errors. More specifically, we regressed BMI on the "socio-demographic" variables of whether or not the individual was born in Canada, employment status, marital status, level of education, level of income, and age group, along with the "behavioural" variables of smoking, alcohol intake, physical activity, and fruit and vegetable consumption, for each geographic region and sex separately.

Next, we decomposed the differences in mean BMI between geographic regions to explore the underlying differences between regions. By using a Blinder-Oaxaca decomposition (24,25), we were able to break down the difference between regions into i) the difference in the covariates by region (e.g., does Ontario have higher levels of income than the Atlantic provinces, which in turn contributes to the difference in BMI between the two regions?) and ii) the difference in the returns to those covariates (e.g., does income have a larger effect on BMI in Ontario than in the Atlantic provinces?). We used the Atlantic provinces as a reference group because that region had the highest average BMI. Thus, it provides a compelling counterfactual comparison: by looking at the region with the highest BMI we can see how all the other regions would look if they had the Atlantic region's properties. For example, if Ontario and the Atlantic provinces were made to have the same proportions of people with identical levels of income, would the average difference in BMI between these regions still exist? This can be referred to as
the difference in BMI attributed to the different endowments of income in the Ontario and Atlantic regions. If a difference still exists after the difference in endowments has been accounted for, then it would be considered a difference attributable to different returns to the endowments in the two regions, or "a difference in coefficients".

Finally, for regional comparisons with statistically significant differences in coefficients (i.e., returns to covariates) we investigated which variables most influenced the overall coefficients effect by examining the disaggregated Blinder-Oaxaca output. Regional comparisons with statistically significant differences in the levels of the covariates were also investigated.

All data were analysed using Stata SE Version 10.1. Standard Stata analytic commands were used, with the exception of the Blinder-Oaxaca mean decomposition, for which we used a Stata command by Jann (27). All analyses incorporated the appropriate sampling weight as directed by Statistics Canada.

3.4 Results

Analyses are based on a sample of individuals with complete data on all study variables (n=7,494). Of those individuals age 20-64 with complete data on BMI (n=8,519), 12.03% (n=1,025) were excluded due to missing data on covariates. Relative to included participants, excluded participants were (at a significance level of p<.05) younger, had lower BMI, were less often employed, earned less, were less likely to smoke daily, and were less likely to be daily drinkers. Table 3.1 shows the average BMI and obesity rates for each province of Canada and the whole country in aggregate.
Table 3.1: Summary statistics for observations with non-missing variables

<table>
<thead>
<tr>
<th>Region</th>
<th>Measured BMI</th>
<th>% of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std Deviation</td>
</tr>
<tr>
<td>Full Sample</td>
<td>27.5</td>
<td>5.87</td>
</tr>
<tr>
<td>Males Only</td>
<td>27.67</td>
<td>5.03</td>
</tr>
<tr>
<td>Females Only</td>
<td>27.37</td>
<td>6.46</td>
</tr>
<tr>
<td>Full Sample</td>
<td>26.92</td>
<td>5.73</td>
</tr>
<tr>
<td>Males Only</td>
<td>26.87</td>
<td>4.6</td>
</tr>
<tr>
<td>Females Only</td>
<td>26.96</td>
<td>6.42</td>
</tr>
<tr>
<td>Full Sample</td>
<td>27.38</td>
<td>5.93</td>
</tr>
<tr>
<td>Males Only</td>
<td>27.75</td>
<td>5.3</td>
</tr>
<tr>
<td>Females Only</td>
<td>27.06</td>
<td>6.42</td>
</tr>
<tr>
<td>Full Sample</td>
<td>28.17</td>
<td>5.94</td>
</tr>
<tr>
<td>Males Only</td>
<td>28.04</td>
<td>5.19</td>
</tr>
<tr>
<td>Females Only</td>
<td>27.97</td>
<td>6.56</td>
</tr>
<tr>
<td>Full Sample</td>
<td>27.24</td>
<td>5.82</td>
</tr>
<tr>
<td>Males Only</td>
<td>27.56</td>
<td>5.12</td>
</tr>
<tr>
<td>Females Only</td>
<td>26.98</td>
<td>6.34</td>
</tr>
<tr>
<td>Full Sample</td>
<td>26.56</td>
<td>5.52</td>
</tr>
<tr>
<td>Males Only</td>
<td>26.8</td>
<td>4.71</td>
</tr>
<tr>
<td>Females Only</td>
<td>26.36</td>
<td>6.13</td>
</tr>
<tr>
<td>Full Sample</td>
<td>28.46</td>
<td>6.12</td>
</tr>
<tr>
<td>Males Only</td>
<td>28.14</td>
<td>4.76</td>
</tr>
<tr>
<td>Females Only</td>
<td>28.66</td>
<td>6.82</td>
</tr>
<tr>
<td>Full Sample</td>
<td>27.92</td>
<td>6.02</td>
</tr>
<tr>
<td>Males Only</td>
<td>27.48</td>
<td>4.86</td>
</tr>
<tr>
<td>Females Only</td>
<td>28.21</td>
<td>6.67</td>
</tr>
<tr>
<td>Full Sample</td>
<td>28.06</td>
<td>5.63</td>
</tr>
<tr>
<td>Males Only</td>
<td>28.54</td>
<td>4.79</td>
</tr>
<tr>
<td>Females Only</td>
<td>27.74</td>
<td>6.12</td>
</tr>
<tr>
<td>Full Sample</td>
<td>28.76</td>
<td>5.97</td>
</tr>
<tr>
<td>Males Only</td>
<td>29.12</td>
<td>5.21</td>
</tr>
<tr>
<td>Females Only</td>
<td>28.47</td>
<td>6.53</td>
</tr>
</tbody>
</table>

Matching letters indicates statistical similarity from an ANOVA test with a Scheffé correction
3.4.1 OLS Estimates

The OLS output for the five regions is presented in Table 3.2. Several statistically significant effects were observed in several regions. For example, for males in Quebec, the Prairies, and Ontario, and for females in the Atlantic region, British Columbia, and the Prairies, being born outside of Canada was inversely associated with BMI. Income had a positive association with BMI for men and women in the Atlantic provinces, and for women in the Prairies. In general, a positive association between age and BMI was observed for women. Education was negatively associated with BMI for women in Quebec and British Columbia, and for men in British Columbia and Ontario. These varying estimates of the covariate effects indicate how different the determinants of BMI can be by region. Such differences set the stage for the next step (the Blinder-Oaxaca decompositions) wherein we further investigate these different effects.
Table 3.2: OLS Regression Estimates for the Five Regions, Disaggregated by Region. Dependent variable: measured BMI

<table>
<thead>
<tr>
<th>Region</th>
<th>Atlantic Males</th>
<th>Atlantic Females</th>
<th>Quebec Males</th>
<th>Quebec Females</th>
<th>Ontario Males</th>
<th>Ontario Females</th>
<th>Prairies Males</th>
<th>Prairies Females</th>
<th>British Columbia Males</th>
<th>British Columbia Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Born Outside of Canada</td>
<td>-1.22</td>
<td>-2.47*</td>
<td>-3.86*</td>
<td>4.61</td>
<td>-1.89*</td>
<td>-1.47</td>
<td>-1.33*</td>
<td>-3.81*</td>
<td>-1.69</td>
<td>-3.65*</td>
</tr>
<tr>
<td></td>
<td>(1.06)</td>
<td>(1.25)</td>
<td>(1.35)</td>
<td>(3.38)</td>
<td>(0.74)</td>
<td>(0.77)</td>
<td>(0.64)</td>
<td>(0.84)</td>
<td>(0.99)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>Employed</td>
<td>-1.18</td>
<td>-1.24</td>
<td>-1.49</td>
<td>-0.51</td>
<td>-0.57</td>
<td>-1.66</td>
<td>-1.35</td>
<td>-1.38</td>
<td>-0.031</td>
<td>1.57</td>
</tr>
<tr>
<td></td>
<td>(1.45)</td>
<td>(0.77)</td>
<td>(0.90)</td>
<td>(0.88)</td>
<td>(0.65)</td>
<td>(1.10)</td>
<td>(1.19)</td>
<td>(0.91)</td>
<td>(0.94)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>Married</td>
<td>-0.60</td>
<td>-0.87</td>
<td>0.54</td>
<td>-0.44</td>
<td>0.17</td>
<td>-0.71</td>
<td>-0.12</td>
<td>-1.30</td>
<td>0.18</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>(0.82)</td>
<td>(0.68)</td>
<td>(0.67)</td>
<td>(0.74)</td>
<td>(0.85)</td>
<td>(0.74)</td>
<td>(0.62)</td>
<td>(0.94)</td>
<td>(0.70)</td>
<td>(0.83)</td>
</tr>
<tr>
<td>High Education</td>
<td>0.68</td>
<td>-1.75</td>
<td>-0.97</td>
<td>-2.69*</td>
<td>-1.33*</td>
<td>-0.72</td>
<td>-0.40</td>
<td>1.00</td>
<td>-2.40*</td>
<td>-1.94*</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(1.00)</td>
<td>(0.72)</td>
<td>(1.12)</td>
<td>(0.63)</td>
<td>(0.78)</td>
<td>(0.62)</td>
<td>(0.94)</td>
<td>(0.70)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>Income Level 0 - 14,999</td>
<td>-2.22*</td>
<td>-0.30</td>
<td>-1.69</td>
<td>-0.31</td>
<td>-0.019</td>
<td>-1.37</td>
<td>-1.29</td>
<td>-1.65*</td>
<td>-0.096</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td>(1.01)</td>
<td>(1.09)</td>
<td>(1.21)</td>
<td>(1.11)</td>
<td>(1.37)</td>
<td>(1.14)</td>
<td>(0.75)</td>
<td>(1.33)</td>
<td>(1.35)</td>
</tr>
<tr>
<td>Income Level 15,000 - 29,999</td>
<td>-0.34</td>
<td>1.37</td>
<td>0.25</td>
<td>-1.10</td>
<td>0.087</td>
<td>0.56</td>
<td>0.82</td>
<td>1.15</td>
<td>-1.30</td>
<td>-0.86</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td>(0.94)</td>
<td>(0.77)</td>
<td>(1.07)</td>
<td>(1.04)</td>
<td>(1.27)</td>
<td>(1.16)</td>
<td>(0.85)</td>
<td>(1.11)</td>
<td>(1.30)</td>
</tr>
<tr>
<td>Income Level 50,000 - 79,999</td>
<td>0.70</td>
<td>2.30*</td>
<td>1.55*</td>
<td>1.01</td>
<td>0.67</td>
<td>-1.49</td>
<td>-0.17</td>
<td>0.86</td>
<td>0.98</td>
<td>-0.88</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(1.05)</td>
<td>(0.78)</td>
<td>(1.10)</td>
<td>(0.76)</td>
<td>(0.96)</td>
<td>(0.70)</td>
<td>(0.85)</td>
<td>(0.89)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>Income Level 80,000 and up</td>
<td>-0.37</td>
<td>2.10*</td>
<td>1.27</td>
<td>1.15</td>
<td>0.81</td>
<td>-1.05</td>
<td>1.28</td>
<td>-1.00</td>
<td>-0.016</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>(0.75)</td>
<td>(0.93)</td>
<td>(0.95)</td>
<td>(1.22)</td>
<td>(0.74)</td>
<td>(1.11)</td>
<td>(0.75)</td>
<td>(0.82)</td>
<td>(0.95)</td>
<td>(1.06)</td>
</tr>
<tr>
<td>Age Group 19 to 24</td>
<td>-1.72</td>
<td>-2.38</td>
<td>-5.39*</td>
<td>-0.33</td>
<td>-1.70</td>
<td>-2.58</td>
<td>-0.56</td>
<td>-2.01</td>
<td>-1.36</td>
<td>-1.19</td>
</tr>
<tr>
<td></td>
<td>(1.29)</td>
<td>(1.42)</td>
<td>(1.36)</td>
<td>(1.21)</td>
<td>(1.33)</td>
<td>(1.38)</td>
<td>(0.95)</td>
<td>(1.30)</td>
<td>(1.54)</td>
<td>(1.77)</td>
</tr>
<tr>
<td>Age Group 25 to 30</td>
<td>-1.16</td>
<td>-1.64</td>
<td>-1.90</td>
<td>0.89</td>
<td>0.60</td>
<td>-1.45</td>
<td>1.63</td>
<td>-0.059</td>
<td>-2.33</td>
<td>-1.85</td>
</tr>
<tr>
<td></td>
<td>(1.33)</td>
<td>(1.18)</td>
<td>(1.32)</td>
<td>(1.65)</td>
<td>(0.94)</td>
<td>(1.40)</td>
<td>(1.19)</td>
<td>(1.21)</td>
<td>(1.30)</td>
<td>(1.82)</td>
</tr>
<tr>
<td>Age Group 36 to 40</td>
<td>-0.11</td>
<td>-3.26*</td>
<td>-0.79</td>
<td>2.89*</td>
<td>0.98</td>
<td>1.04</td>
<td>1.22</td>
<td>2.81</td>
<td>0.91</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(1.26)</td>
<td>(1.48)</td>
<td>(1.49)</td>
<td>(1.40)</td>
<td>(1.25)</td>
<td>(1.34)</td>
<td>(0.93)</td>
<td>(1.76)</td>
<td>(2.02)</td>
<td>(1.96)</td>
</tr>
<tr>
<td>Age Group 41 to 45</td>
<td>0.34</td>
<td>1.08</td>
<td>-2.05</td>
<td>1.78</td>
<td>2.29*</td>
<td>-0.76</td>
<td>1.66</td>
<td>2.48</td>
<td>0.60</td>
<td>-0.73</td>
</tr>
<tr>
<td></td>
<td>(1.38)</td>
<td>(1.72)</td>
<td>(1.60)</td>
<td>(1.28)</td>
<td>(0.95)</td>
<td>(1.35)</td>
<td>(0.92)</td>
<td>(1.38)</td>
<td>(1.62)</td>
<td>(1.83)</td>
</tr>
<tr>
<td>Age Group 46 to 50</td>
<td>1.49</td>
<td>-1.23</td>
<td>-2.83*</td>
<td>4.81*</td>
<td>1.34</td>
<td>1.24</td>
<td>1.01</td>
<td>5.24*</td>
<td>-0.62</td>
<td>0.29</td>
</tr>
<tr>
<td>-------------------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(1.25)</td>
<td>(1.22)</td>
<td>(1.91)</td>
<td>(1.09)</td>
<td>(1.04)</td>
<td>(0.94)</td>
<td>(1.64)</td>
<td>(1.63)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>Age Group 51 to 55</td>
<td>2.25</td>
<td>-0.45</td>
<td>0.44</td>
<td>2.24</td>
<td>0.95</td>
<td>2.91*</td>
<td>2.04*</td>
<td>2.29</td>
<td>0.33</td>
<td>1.30</td>
</tr>
<tr>
<td></td>
<td>(1.39)</td>
<td>(1.24)</td>
<td>(1.19)</td>
<td>(1.29)</td>
<td>(1.14)</td>
<td>(1.12)</td>
<td>(0.85)</td>
<td>(1.20)</td>
<td>(1.81)</td>
<td>(1.97)</td>
</tr>
<tr>
<td>Age Group 56 to 60</td>
<td>1.15</td>
<td>-0.23</td>
<td>-0.70</td>
<td>3.11*</td>
<td>1.39</td>
<td>1.86</td>
<td>1.99*</td>
<td>1.84</td>
<td>-0.11</td>
<td>2.54</td>
</tr>
<tr>
<td></td>
<td>(1.63)</td>
<td>(1.21)</td>
<td>(1.30)</td>
<td>(1.40)</td>
<td>(0.97)</td>
<td>(1.11)</td>
<td>(0.98)</td>
<td>(1.20)</td>
<td>(1.25)</td>
<td>(2.13)</td>
</tr>
<tr>
<td>Age Group 61 to 65</td>
<td>0.11</td>
<td>-0.18</td>
<td>-2.12</td>
<td>1.88</td>
<td>2.02*</td>
<td>0.39</td>
<td>2.47</td>
<td>1.25</td>
<td>0.90</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(1.26)</td>
<td>(1.44)</td>
<td>(1.45)</td>
<td>(0.96)</td>
<td>(1.14)</td>
<td>(1.44)</td>
<td>(1.27)</td>
<td>(1.47)</td>
<td>(2.14)</td>
</tr>
<tr>
<td>Smokes Occasionally</td>
<td>-0.12</td>
<td>0.26</td>
<td>-1.42</td>
<td>-1.20</td>
<td>0.36</td>
<td>-0.91</td>
<td>-0.65</td>
<td>3.27*</td>
<td>-1.02</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>(1.57)</td>
<td>(1.69)</td>
<td>(1.23)</td>
<td>(1.01)</td>
<td>(1.17)</td>
<td>(1.52)</td>
<td>(1.32)</td>
<td>(1.32)</td>
<td>(1.99)</td>
<td></td>
</tr>
<tr>
<td>Smokes Daily</td>
<td>-1.80*</td>
<td>-1.26</td>
<td>-0.62</td>
<td>-2.04*</td>
<td>-1.44*</td>
<td>-0.46</td>
<td>-2.71*</td>
<td>0.38</td>
<td>0.95</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.86)</td>
<td>(0.79)</td>
<td>(0.96)</td>
<td>(0.66)</td>
<td>(0.87)</td>
<td>(0.58)</td>
<td>(0.65)</td>
<td>(1.27)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>Drinks at least once a week</td>
<td>0.11</td>
<td>0.0036</td>
<td>0.87</td>
<td>2.87*</td>
<td>1.44</td>
<td>-0.49</td>
<td>-0.65</td>
<td>0.68</td>
<td>-1.75</td>
<td>-1.73</td>
</tr>
<tr>
<td></td>
<td>(1.26)</td>
<td>(0.87)</td>
<td>(0.94)</td>
<td>(1.33)</td>
<td>(0.80)</td>
<td>(0.99)</td>
<td>(0.84)</td>
<td>(0.64)</td>
<td>(1.45)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>Drinks at least once a month</td>
<td>-0.84</td>
<td>-2.67*</td>
<td>0.16</td>
<td>0.10</td>
<td>0.90</td>
<td>-1.35</td>
<td>-1.11</td>
<td>0.54</td>
<td>-1.42</td>
<td>-2.60*</td>
</tr>
<tr>
<td></td>
<td>(1.28)</td>
<td>(0.87)</td>
<td>(0.85)</td>
<td>(1.15)</td>
<td>(0.58)</td>
<td>(0.94)</td>
<td>(0.78)</td>
<td>(0.74)</td>
<td>(1.26)</td>
<td>(1.22)</td>
</tr>
<tr>
<td>Moderate Level of Physical Activity</td>
<td>-1.70*</td>
<td>-0.90</td>
<td>-0.34</td>
<td>-0.94</td>
<td>-1.77*</td>
<td>-0.47</td>
<td>-1.26*</td>
<td>-1.73*</td>
<td>-0.24</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.74)</td>
<td>(0.63)</td>
<td>(0.85)</td>
<td>(0.61)</td>
<td>(0.85)</td>
<td>(0.60)</td>
<td>(0.70)</td>
<td>(0.97)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>Active Level of Physical Activity</td>
<td>-2.25*</td>
<td>-2.92*</td>
<td>0.22</td>
<td>-3.62*</td>
<td>-1.18</td>
<td>-0.44</td>
<td>-1.91*</td>
<td>-2.27*</td>
<td>-0.61</td>
<td>-1.42</td>
</tr>
<tr>
<td></td>
<td>(0.68)</td>
<td>(0.93)</td>
<td>(0.59)</td>
<td>(0.82)</td>
<td>(0.82)</td>
<td>(0.68)</td>
<td>(0.65)</td>
<td>(0.77)</td>
<td>(0.76)</td>
<td>(0.89)</td>
</tr>
<tr>
<td>Fruit and Vegetable Consumption/Day</td>
<td>-0.15</td>
<td>0.40</td>
<td>-0.30*</td>
<td>-0.39*</td>
<td>-0.29*</td>
<td>-0.24</td>
<td>-0.22</td>
<td>0.076</td>
<td>0.20</td>
<td>0.0066</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.24)</td>
<td>(0.10)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.16)</td>
<td>(0.13)</td>
<td>(0.15)</td>
<td>(0.18)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Constant</td>
<td>31.1*</td>
<td>30.6*</td>
<td>31.1*</td>
<td>28.4*</td>
<td>29.1*</td>
<td>32.0*</td>
<td>31.6*</td>
<td>27.2*</td>
<td>30.2*</td>
<td>29.9*</td>
</tr>
<tr>
<td></td>
<td>(2.39)</td>
<td>(1.68)</td>
<td>(1.21)</td>
<td>(1.87)</td>
<td>(1.53)</td>
<td>(1.93)</td>
<td>(1.73)</td>
<td>(1.66)</td>
<td>(2.21)</td>
<td>(2.25)</td>
</tr>
<tr>
<td>Observations</td>
<td>718</td>
<td>1022</td>
<td>574</td>
<td>635</td>
<td>781</td>
<td>960</td>
<td>942</td>
<td>1120</td>
<td>302</td>
<td>440</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.18</td>
<td>0.15</td>
<td>0.27</td>
<td>0.24</td>
<td>0.17</td>
<td>0.10</td>
<td>0.14</td>
<td>0.17</td>
<td>0.22</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

* p<0.05
### 3.4.2 Blinder-Oaxaca Decompositions

The results of the Blinder-Oaxaca decompositions are presented in Table 3.3. The values in Table 3.3 refer to the region in question being compared to the Atlantic provinces. These estimates are reported from the region in question's point of view (e.g., "If British Columbia had the same level of endowments as the Atlantic provinces then BMI in British Columbia would change by this value" or "If British Columbia had the same returns to the variables as the Atlantic provinces then BMI in British Columbia would change by this value.")

For each two-way comparison, the following information is reported: difference (the predicted difference in mean measured BMI between the two regions given the covariates specified in the model [Atlantic provinces - Region X (the region indicated in the row)]), endowments (how would BMI in Region X change if Region X had the Atlantic provinces' level of the covariates in the model), coefficients (how would BMI in Region X change if they had the Atlantic provinces' returns to the covariates in the model), interaction (the term that is used to balance the equation of the predicted difference; this is often referred to as the "unexplained difference" between the Atlantic provinces and Region X, since the correction of both coefficients and endowments levels need not reduce the difference to zero), and the number of observations on which the two-way comparison is based.

Statistically significant findings for the males appeared only in the Quebec - Atlantic provinces comparison. If males in Quebec had the same endowments as males in the Atlantic provinces, Quebec's expected average BMI would be 0.76 units higher than its current value. The overall difference between the regions is 1.06 BMI units, so the statistically significant estimate of the effects of the endowments means that the overall difference between the regions can be mostly explained by different levels of the covariates included in the model. The share of the
difference attributable to the different returns to the covariates between regions was not statistically significant for any regional comparison.

For females, most regions exhibit statistically significant differences when compared to the Atlantic provinces. Females in British Columbia have an overall average BMI that is 1.60 units less than the average predicted BMI for women in the Atlantic provinces, and it is almost completely explained by the endowments (different levels of covariates) between the two regions, which account for 1.46 units of the difference. Females in Quebec have a statistically significant overall difference in BMI from women in the Atlantic provinces, with the females in Quebec having a predicted average BMI 1.50 units lighter than females in the Atlantic provinces. Unlike the British Columbia - Atlantic comparison, the Quebec - Atlantic difference for females is entirely explained by the share of the difference attributable to the difference in the coefficients (i.e. different returns to the covariates.) Finally, the females in Ontario have a statistically significant difference in predicted BMI from the females in the Atlantic provinces, with Ontario being 1.49 BMI units lighter. The overall difference is attributable to both the difference in the endowments (levels of the covariates) (0.91 BMI units) and the coefficients (returns to the covariates) (1.27 BMI units).
Table 3.3: Blinder-Oaxaca decompositions using the Atlantic provinces as the reference group. Dependent variable: measured BMI

| Region         | Difference | Endowments | Coefficients | Interaction | N   | Difference | Endowments | Coefficients | Interaction | N   |
|----------------|------------|------------|--------------|-------------|-----|------------|------------|--------------|-------------|-----|-----|
| 1  British Columbia | 0.67       | 0.85       | 0.15         | -0.33       | 1020| 1.60*      | 1.46*      | 1.15         | -1.01       | 1462|
| 2  Prairies     | -0.28      | 0.29       | -0.35        | -0.21       | 1660| 0.92       | 0.68*      | 0.63         | -0.38       | 2142|
| 3  Ontario      | 0.49       | 0.68*      | 0.032        | -0.22       | 1499| 1.49*      | 0.91*      | 1.27*        | -0.70       | 1982|
| 4  Quebec       | 1.06*      | 0.76*      | 0.96         | -0.66       | 1292| 1.50*      | 0.64       | 1.50*        | -0.63       | 1657|

These estimates should be interpreted as the effect on the measured BMI of the region in question if it had the same endowments/coefficients/interaction as the Atlantic provinces.

Robust standard errors in parentheses

* p<0.05
<table>
<thead>
<tr>
<th></th>
<th>ON Females vs. ATL Females</th>
<th>QC Females vs. ATL Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endowments</td>
<td>Coefficients</td>
</tr>
<tr>
<td>Born Outside of Canada</td>
<td>0.27</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Employed</td>
<td>0.046</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(1.10)</td>
</tr>
<tr>
<td>Married</td>
<td>0.043</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.72)</td>
</tr>
<tr>
<td>High Education</td>
<td>0.018</td>
<td>-0.70</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.86)</td>
</tr>
<tr>
<td>Income Level 0 - 14,999</td>
<td>-0.065</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Income Level 15,000 - 29,999</td>
<td>0.060</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Income Level 50,000 - 79,999</td>
<td>0.052</td>
<td>1.16*</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Income Level 80,000 and up</td>
<td>0.22</td>
<td>1.26*</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Age Group 19 to 24</td>
<td>0.031</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Age Group 25 to 30</td>
<td>-0.0083</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Age Group 36 to 40</td>
<td>-0.0088</td>
<td>-0.53*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Age Group 41 to 45</td>
<td>-0.0090</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Age Group 46 to 50</td>
<td>0.0045</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Age Group 51 to 55</td>
<td>-0.080</td>
<td>-0.44</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Age Group 56 to 60</td>
<td>0.11</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Age Group 61 to 65</td>
<td>0.00041</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Smokes Occassionally</td>
<td>0.00063</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Smokes Daily</td>
<td>-0.016</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>Drinks at least once a week</td>
<td>-0.041</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.62)</td>
</tr>
<tr>
<td>Category</td>
<td>Mean Difference</td>
<td>p-value</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Drinks at least once a month</td>
<td>0.096</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.43)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.26)</td>
</tr>
<tr>
<td>Moderate Level of Physical Activity</td>
<td>-0.0041</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>Active Level of Physical Activity</td>
<td>0.018</td>
<td>-0.49</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.24)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.11)</td>
</tr>
<tr>
<td>Fruit and Vegetable Consumption/Day</td>
<td>0.18</td>
<td>2.91*</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(1.33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.39*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.16)</td>
</tr>
<tr>
<td>Total</td>
<td>0.91*</td>
<td>1.27*</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.40)</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
These results pertain to Blinder-Oaxaca decompositions displayed in Table 3.3
For females, in both the Quebec-Atlantic comparison and the Ontario-Atlantic comparison, the coefficients (returns to covariates) dominate as determinants of BMI differences between the regions. For these two comparisons, we investigated which variables most influenced the overall coefficients effect by looking at the disaggregated Blinder-Oaxaca output, as shown in Table 3.4. For females in Ontario, higher income has a different return to earning high income in the Atlantic provinces, with Ontario females in the $50,000 to $79,000 range being 1.16 BMI units lighter than their Atlantic province counter-parts. Similarly, those Ontario women in the $80,000 and up income category were 1.26 BMI units lighter than their Atlantic province counter-parts. Eating fruits and vegetables has a statistically significant and large effect on the coefficients estimate for Ontario women compared to Atlantic women, with the same fruit and vegetable designation resulting in the average BMI in Ontario being 2.91 units lower than in the Atlantic provinces. Finally, Ontario women who had an "active" level of physical activity were 0.49 BMI units heavier than women in the Atlantic provinces that had the same level of physical activity. For females in Quebec, the estimated contribution of fruit and vegetable consumption was statistically significant and large such that comparable levels of fruit and vegetable consumption in the two regions results in an average BMI 3.8 units lower in Quebec than in the Atlantic provinces.

We also compared the disaggregated Blinder-Oaxaca output for those comparisons with statistically significant contributions from the different endowments (not shown). For females in British Columbia, the majority of the statistically significant difference with females from the Atlantic provinces was attributed to the different proportions of those born outside of Canada. The average BMI of females in British Columbia would be 1.08 BMI units higher if the mix of people born in Canada were the same for both regions. For males in Quebec, the only statistically
significant difference with males from the Atlantic provinces was attributed to the different levels of consumption of fruits and vegetables. Males in Quebec would be 0.27 BMI units heavier on average if fruit and vegetable consumption among males was the same in both regions. For the Ontario-Atlantic comparison for women, none of the covariates was found to contribute to the statistically significant endowments effect, in the disaggregated analysis.

Finally, due to the right-skewed nature of the BMI variable we considered models with the dependent variable log-transformed. We re-ran the Blinder-Oaxaca analysis with the natural log of the measured BMI as the dependent variable (not shown) and the results were qualitatively similar. The only differences were (i) the overall difference in mean BMI for females in the Prairies versus the Atlantic provinces became statistically significant, and (ii) the endowments (levels of covariates) became a statistically significant contributor to the difference in mean BMI between females in Quebec and the Atlantic provinces, yet the magnitude of the effect was approximately half that of the coefficients, which remained the same. Due to the more natural interpretation of BMI units (versus the natural log of BMI units) coupled with the substantively similar findings between the two formats, we opted to present the findings based on the untransformed BMI variable.

3.5 Discussion

The OLS regressions showed that the different regions of Canada exhibit different associations between BMI and covariates but we receive additional information about regional differences by applying a Blinder-Oaxaca decomposition. According to the Blinder-Oaxaca decompositions on the same model estimated for the OLS regression, the relative contributions of both levels of covariates and returns to covariates differ between regions.
The difference between males in Quebec and the Atlantic provinces is due to the statistically significant endowments (i.e., levels of covariates.) This indicates that correction of the different levels of covariates in the Atlantic provinces may be a viable policy option towards obesity prevention for males, if bringing the Atlantic provinces' average male BMI in line with those of Quebec males is the goal of policy. For females in other regions, such as Ontario or Quebec compared to the Atlantic provinces, correcting levels of covariates is not necessarily the most effective policy target, because it is the different returns to the covariates - rather than the levels - that are important. The scenarios in which the differences in average BMI are primarily attributable to different returns to covariates draw attention to factors that may be present in one province but not in another, and impact BMI outcomes in combination with personal characteristics. An example would be females in Ontario versus the Atlantic provinces, a comparison for which the different levels of the covariates are not able to wholly explain the difference in average BMI. The significant contribution of the coefficients (returns to the covariates) in this example indicates that there is some sort of region-specific effect that differs between the regions and causes the covariates to have different effects in the two regions, resulting in a higher average BMI in the Atlantic provinces.

Regional effects are unmeasured in our analysis and therefore we cannot state with certainty what these might be. However, we can look to other sources for insight. Taking the comparison between females in Quebec and females in the Atlantic provinces as an example, one plausible contributor is regional differences in government social policy. In 1997, Quebec instituted a subsidy that made the cost of daycare $5/day (no Atlantic province had funding available in this fashion) and this subsidy was available to a large portion of the population (28). Given that formal child care arrangements from age 3 to 5 have been linked with a reduced risk
of obesity among children at age 6 to 12 (29) and considering the evidence linking childhood obesity with obesity in adulthood (30,31), this is one example of an age-related social policy that could be contributing to regional differences. If children in Quebec had a better starting position than children in the Atlantic provinces through policy, then we might expect those advantages to carry through to adulthood, and thus mitigate the effect of factors like physical activity on BMI in the Atlantic provinces compared to Quebec.

Food prices are another factor that could affect the cross-regional obesity rates and manifest as different returns to behavioural and socio-economic variables. Residents of the different provinces are subject to different food prices, and are also subject to different provincial levels of income. There is evidence that in at least one Canadian province it would not be possible for a family earning minimum wage or seniors on social assistance to afford a nutritious basket of food monthly (32,33), and through such a mechanism, the relationship between socioeconomic circumstances and BMI could differ between regions. To explore, post hoc, the plausibility of regional differences in food price in contributing to regional differences in BMI we compiled a simple graph (Figure 3.1) that shows the average personal income after tax by region and consumer price indices (CPIs) for food for the year 2004 (to correspond with the CCHS data used in this paper). The data for Figure 3.1 comes from Statistics Canada's database of national statistics CANSIM (Canadian Socio-economic Information Management System) and more information can be found at http://cansim2.statcan.gc.ca/. One can see that there is substantial variation between the regions, reflecting a difference in the ability of regional incomes to match the regional prices of food. The Atlantic provinces have the lowest purchasing power for fruits and vegetables and food bought in stores, and relatively speaking, food bought in a restaurant is the cheapest alternative. Such a price differential, and the associated incentive
that it gives the regional population, could be a regional effect as the other regions may be able to afford both more money towards food and more money towards healthy food options, therefore potentially relying less on cheaper processed food (34) or restaurant foods that are generally served in larger portions and encourage overeating (35). Different food environments could potentially impact the protective effects of other covariates on BMI, such as income, as the population in these regions simply cannot make the same choices as they could in a different food environment (e.g., income in the Atlantic provinces has a different meaning, relative to BMI, than income in other regions, such that it places constraints on food purchase to a larger extent in the Atlantic region.)
Figure 3.1: Average personal income and CPI by region.

The results of the disaggregated Blinder-Oaxaca decomposition analysis for different coefficients provides some support for these proposed causes of the contribution of the coefficients to some regional differences in women. The income effect observed when comparing Ontario and the Atlantic provinces means that despite a similar and relatively high level of personal financial resources in the two regions, there are different circumstances that define how the women in the Atlantic provinces can use this income compared to those in Ontario, which in turn have implications for BMI. Thus, the food environment (or environment in general) could be quite different between regions such that a protective effect of higher income for BMI is region-specific: stronger in Ontario than in the Atlantic provinces.
The case of fruits and vegetables having a different effect in both Ontario and Quebec when compared to the Atlantic provinces is less straightforward. One possibility is that the quality of fruits and vegetables consumed differs by region. For example, fruit servings in the Atlantic provinces may consist of more canned fruits (due to a climate less conducive to growing a variety of fruits or geographical distance affecting the efficiency of transporting perishables to the region), which are often packed in syrup, adding to the calorie count, compared to fresh fruits. Another plausible explanation is that consumption of other foods varies regionally, and differentially offsets the impact of fruit and vegetable consumption. For example, if high levels of consumption of fruits and vegetables in the Atlantic region are associated with higher consumption of food overall (including less healthy foods), perhaps reflecting dietary social norms, then we would observe different returns to the consumption of fruits and vegetables. Finally, it is possible that the variable "Fruit and Vegetable consumption" is not actually capturing just that, and regionally there is variation in what it is capturing. However, this measure has been shown to be a valid indicator of diet quality (36). The difference in the coefficients attributable to fruit and vegetable consumption is intriguing and further research should interrogate what fruit and vegetable consumption entails in different regions, and investigate what factors are associated with fruit and vegetable consumption in the Atlantic provinces that are causing these less favourable returns.

Taken altogether, our results indicate that for females, there is interesting regional variation in BMI which manifests among those living in Ontario and Quebec compared with the Atlantic region. For males, on the other hand, there is less average regional variation in BMI and what does exist can be readily explained by different levels of covariates between the different
regions. These gender differences raise the interesting possibility, for pursuit in future research, that males may be more resilient than females to regional effects.

Our study has some limitations. One limitation is that the inclusion of behavioural variables may introduce statistical bias into the study. Specifically, inclusion of these variables introduces endogeneity into the interpretation of the model (possible reverse causation from the dependent variable to the independent variables). Endogeneity contributes to the impossibility of distinguishing causality, for example, do people who exercise have a lower BMI through exercise, or are people who have a lower BMI more likely to participate in exercise, because they find it easier or more enjoyable? Notwithstanding this problem, we judged the inclusion of behavioural variables in our models to be important, considering the demonstrated association of these variables with obesity (37).

A second limitation is that any omitted variables that have an effect on BMI will have their share of the difference erroneously attributed to differences in coefficients (i.e. the returns to the covariates). Variables such as race/ethnicity could be considered important omitted variables, since people of East Asian descent are known to have a lower BMI than the general population of Canada, and there are regional differences in the proportion of the population of East Asian descent (e.g., British Columbia has a larger immigrant proportion from East Asia than the other regions in our analysis) (5,38,39). Our analysis identified situations in which the endowments (i.e. covariates) were statistically significant even in the face of this limitation, so it suggests that sizable contributions to BMI were captured in our choice of variables. A third limitation is that the explanatory power of the models overall was low in some cases, with values of $R^2$ ranging from .06 to .24 for the OLS models. Also, it is important to note that the population distribution of BMI is not Gaussian but, rather, right-skewed (1). Although we
ascertained that the results presented here are not unduly affected by the right skew, we acknowledge that our approach (which focuses on the mean) neglects the potentially policy-relevant possibility that effects and returns are different at higher percentiles of the distribution (amongst the heaviest individuals). Future research would need to accommodate the right-skew by using techniques such as quantile regression to explicitly model such issues (23) in the decomposition context.
3.6 References


(5) Tjepkema M. Adult obesity in Canada: Measured height and weight. 2005;Statistics Canada - Cat. No. 82-620-MWE.


(27) Jann B. A Stata implementation of the Blinder-Oaxaca decomposition. ETH Zurich, Chair of Sociology, ETH Zurich Sociology Working Papers 2008(5).


(31) Dietz WH. Periods of risk in childhood for the development of adult obesity--what do we need to learn? J Nutr 1997 Sep;127(9):1884S-1886S.


Chapter Four: How useful is “corrected” body mass index vs. self-reported body mass index? Comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data.

4.1 Abstract

4.1.1 Background

National data on body mass index (BMI), computed from self-reported height and weight, is readily available for many populations including the Canadian population. Because self-reported weight is found to be systematically under-reported, it has been proposed that the bias in self-reported BMI can be corrected using equations derived from data sets which include both self-reported and measured height and weight. Such correction equations have been developed and adopted. We aim to evaluate the usefulness (i.e., distributional similarity; sensitivity and specificity; and predictive utility vis-à-vis disease outcomes) of existing and new correction equations in population-based research.

4.1.2 Methods

The Canadian Community Health Surveys from 2005 and 2008 include both measured and self-reported values of height and weight, which allows for construction and evaluation of correction equations. We focused on adults age 18-65, and compared three correction equations (two correcting weight only, and one correcting BMI) against self-reported and measured BMI. We first compared population distributions of BMI. Second, we compared the sensitivity and specificity of self-reported BMI and corrected BMI against measured BMI. Third, we compared the self-reported and corrected BMI in terms of association with health outcomes using logistic regression.
4.1.3 Results

All corrections outperformed self-report when estimating the full BMI distribution; the weight-only correction outperformed the BMI-only correction for females in the 23-28 kg/m² BMI range. In terms of sensitivity/specificity, when estimating obesity prevalence, corrected values of BMI (from any equation) were superior to self-report. In terms of modelling BMI-disease outcome associations, findings were mixed, with no correction proving consistently superior to self-report.

4.1.4 Conclusions

If researchers are interested in modelling the full population distribution of BMI, or estimating the prevalence of obesity in a population, then a correction of any kind included in this study is recommended. If the researcher is interested in using BMI as a predictor variable for modelling disease, then both self-reported and corrected BMI result in biased estimates of association.
4.2 Background

Obesity’s rise in prevalence over the past 30 years (1), coupled with knowledge of its public health burden (2-6), has opened debate over the best way to measure adiposity in populations. The body mass index - BMI [weight (kg) / height\(^2\) (m\(^2\))] - is a common measure in population-based surveys: it is relatively inexpensive, simple, and non-intrusive. Notwithstanding the limitations of BMI as a measure of adiposity (7-10), it continues to be recommended by the World Health Organization (10) as the appropriate criteria to assess obesity status in populations due to the high correlation of high BMI with excess body fat and poor health outcomes (1).

The practice of gathering self-reported height and weight data from survey respondents has raised concerns about the inaccuracy of self-reported data. Where comparison to measured weight is possible, studies have demonstrated misreporting (both under- and over-reporting) of weight, which varies by sex, age, race/ethnicity, and BMI (11-19). The consequence of this misreporting is that the potential exists for large bias in estimates of prevalence and measures of association in studies that use self-reported BMI (7).

By using population-based datasets which contain both measured and self-reported values of BMI for the same individuals, it is possible to develop statistical adjustments that bring self-reported values of BMI closer to measured values. These correction methods can then be applied to datasets that only contain self-reported height and weight, resulting in a corrected height and weight and thus improving the usability of those datasets. Attempts to develop correction methods for use with Canadian data have resulted in two important papers. Connor Gorber et al. (20) used data from the 2005 Canadian Community Health Survey (CCHS) and, after considering many potential covariates, concluded that the most parsimonious and effective
correction equations for men and women rely on one variable: self-reported BMI. Shields et al. (21) used data from the 2007-2009 Canadian Health Measures Survey to develop a correction equation, which was then used to correct self-reported BMI measures from the 2008 CCHS. This second paper likewise concluded that a correction equation using only self-reported BMI is appropriate for correcting BMI values. In both studies, additional covariates did not add enough predictive power to the models to justify the added complexity of including them in the correction equations.

The practice of correcting based on BMI alone (i.e., no other covariates) is appealing in its simplicity. Other authors have included covariates in an effort to increase the accuracy of their correction equations; including age (22), leisure time physical activity, self-reported health (14), education level (23), and ethnicity (24), among others. These studies find that including additional covariates in a correction equation can increase the correction equation’s accuracy when adjusting BMI or categorizing individuals by obesity status (e.g., BMI ≥ 30). In this paper we did not include additional covariates to maximize comparability with the existing Canadian correction equations.

A separate issue is whether corrections should be based on BMI, or weight alone. Studies from the United States (11,13), as well as from Sweden (14), and France (15) verify that average misreporting increases as measured weight and/or BMI increases, suggesting that correcting on either BMI or weight may be acceptable. Correcting on weight only raises the question of whether or how to deal with height. It is well-documented that height is subject to bias in the elderly, i.e., those over age 60 (11,22) and that bias in height has been shown to be substantial (25,26). Furthermore, international evidence shows that height bias also exists among those under age 60, is strongest for the shortest males and for females, and that bias might be changing
over time (23,24,27). Correction equations developed by other authors have directly corrected for either BMI as a whole (14,20,21,23,26), or height and weight separately and calculated a corrected BMI from those values (12,23,24) to incorporate both weight and height. However, there is evidence to suggest that average bias in self-reported height among adults as a whole is quite small: based on five national surveys conducted in Canada and the United States on males and females aged 18 to 74, the range of average bias in self-reported height ranged from 0.2 cm to 1.4 cm (28). Because the bias in working age adults is almost wholly from bias in self-reported weight (11-17,26), it is worth considering the value of correcting on weight only rather than overall BMI for the general population. This paper will consider such a correction.

The purpose of this study is to evaluate the usefulness of existing and new correction equations for BMI in population-based research. To accomplish this, we have three objectives: 1) compare the self-reported and corrected BMI distributions; 2) compare self-reported and corrected BMI to measured BMI based on sensitivity and specificity of measured obesity; and 3) compare self-reported and corrected BMI to measured BMI in regression models of various health conditions, in terms of statistical significance, coefficient magnitude, and direction of the coefficient (above or below the measured coefficient). We compare three correction equations: first, an existing Canadian correction equation (20) (a correction that used self-reported BMI, so will be referred to as the “BMI-only” correction); second, a new correction equation developed here which corrects values of weight only; and third, another correction equation developed here which is a computationally simpler version of the weight-only correction. The term “weight-only” means that we use a corrected value for weight but self-reported height to correct overall BMI.
4.3 Methods

4.3.1 Data

We used data from two cycles (2005 and 2008) of Statistics Canada’s Canadian Community Health Survey (CCHS). The CCHS is a repeated cross-sectional survey that provides socio-demographic and health information for individuals living in the ten Canadian provinces. The CCHS uses a multi-stage cluster sampling procedure to derive a sample that is representative of the Canadian population, excluding those who live in institutions, on First Nations reserves, on Canadian Forces bases, and in certain remote areas; the CCHS is representative of approximately 98% of the Canadian population over age 12 (29). The overall response rates for households were 87.0% for the 2005 CCHS and 85% for the 2008 CCHS (21).

In the 2005 and 2008 iterations of the survey, a random sub-sample of individuals was asked to self-report their height and weight; those values were subsequently measured by the interviewer. The respondents were not told they would be measured when they self-reported their height and weight. The response rate for the sub-sample was 64.2% for 2005 CCHS and 59.7% for the 2008 CCHS; no information is available for reason of refusal to be measured (21). We focused on individuals aged 18 to 65 for whom both self-reported and measured BMI data were available. We excluded adults over age 65, because of observed over-reporting of height in the over 65 age group (11,25,26).

We used the master file versions of the CCHS, accessed through the Research Data Centres (secure data laboratories) program in Canada. Access was granted by Statistics Canada via the Canadian Research Data Centre Network (CRDCN) through a standardized application process. All analyses incorporated sampling weights as directed by Statistics Canada and were
Ethics approval for this project was obtained from the University of Calgary’s Conjoint Health Ethics Research Board (Ethics ID: E-23704).

4.3.2 Procedure

Below, we first describe the development of the weight-only correction equations. Then, we describe the procedure for achieving our three objectives that compare self-reported and corrected BMI to measured BMI. All analyses, including those involved in developing the correction equations, are conducted for males and females separately.

The justification for modelling misreporting based on weight only is best shown graphically. Figure 4.1 shows three quantile-quantile plots comparing measured BMI to three other BMI measures: self-reported BMI (graph a); a BMI constructed from self-reported weight and measured height (graph b); and a BMI constructed from measured weight and self-reported height (graph c), for males and females. The graphs show the average BMI for each percentile of measured BMI against the average BMI for each percentile of the BMI measures containing at least one self-reported value. Note that the quantile-quantile plots in graphs a and b look similar, that is, there is very little improvement in modelling measured BMI by replacing self-reported height with measured height. Graph c shows that there is a large improvement in modelling measured BMI using only measured weight, which indicates that the majority of the measurement error of the distribution of self-reported BMI comes from self-reported weight, not height, in our sample. Thus, the weight-only correction should be addressing the main source of measurement error in the sample of working age individuals.
Figure 4.1: Quantile-quantile plots comparing measured BMI with BMI constructed from combinations of measured and self-reported height and weight.
4.3.3 Development and estimation of weight-only correction equation

We can model the self-reported value of BMI as being a function of an individual’s measured (true) BMI multiplied by a misreporting term:

$$\frac{W_{SR}}{h_{SR}^2} = \frac{W_M}{h_M^2} \cdot e^{\varepsilon + W_M}$$  \hspace{1cm} 4-1

That is, an individual’s self-reported BMI (which is their self-reported weight ($W_{SR}$) over their self-reported height squared ($h_{SR}^2$)) is equal to their measured BMI and a misreporting term that is made up of random noise, $\varepsilon$, and measured weight $W_M$. Equation 4-1 was chosen because, with the right parameters\(^3\), we can mimic the nonlinear relationship between measured and self-reported BMI described in the literature, whereby the discrepancy increases across measured BMI at an increasing rate due to increases in weight. By including weight in the exponential term, we allow the difference between self-reported BMI and measured BMI to grow at an increasing rate as measured weight increases, a relationship that is supported by published literature (11-17). A linear error term would not accurately capture that association.

Next, we can take the natural logarithm of both sides of the equation to reduce the BMI relationship into its constituent parts, which can then be rearranged into the following equation:

$$\ln(W_{SR}) = \ln(W_M) + W_M + 2 \ln\left(\frac{h_{SR}}{h_M}\right) + \varepsilon$$

where $\ln(W_{SR})$, the natural logarithm of an individual’s weight, is a function of measured weight and the ratio of self-reported to measured height (the relative misreport in height). This is an

\(^3\)“The right parameters” implies that the correct equation would not have a coefficient of 1 on every term, as it stands in equation 4-1. For instance, a data generating process with the coefficients 0.01 on the measured weight term in the error and 0.01 on the noise term would generate an exaggerated version of the relationship we observe in the literature. So it is just a matter of finding the right parameters to match the data.
equation for which we can estimate regression parameters, using a sample of individuals with both measured and self-reported height. The estimated regression equation is:

$$\ln(W_{SR,i}) = \hat{\beta}_0 + \hat{\beta}_1 \cdot \ln(W_{M,i}) + \hat{\beta}_2 \cdot W_{M,i} + \hat{\beta}_3 \cdot 2 \ln\left(\frac{h_{SR,i}}{h_{M,i}}\right)$$  \hspace{1cm} (4-2)

Where \( i \) denotes individual values. If the ratio of self-reported to measured height is not related to self-reported weight on average (which we would expect from the literature on non-seniors), then \( \hat{\beta}_3 \) would be statistically equal to zero, leaving us with the equation:

$$\ln(W_{SR,i}) = \hat{\beta}_0 + \hat{\beta}_1 \cdot \ln(W_{M,i}) + \hat{\beta}_2 \cdot W_{M,i}$$  \hspace{1cm} (4-3)

The restriction necessary for equation 4-3 to be an appropriate step, that \( \hat{\beta}_3 \) is statistically equal to 0, was tested using our dataset during the model building exercise and is not an assumption. Specifically, using a t-test, we could not reject the null hypothesis that the coefficient \( \hat{\beta}_3 \) was equal to 0 (p-value of 0.609 in males and 0.559 in females). Removal of the ratio of self-reported to measured height did not impact the values of the other coefficients in the model.

Equation 4-3 can be rearranged to put measured weight in terms of self-reported weight and estimated parameters (constants):

$$\ln(W_{M,i}) + \frac{\hat{\beta}_2}{\hat{\beta}_1} \cdot W_{M,i} = \frac{\ln(W_{SR,i}) - \hat{\beta}_0}{\hat{\beta}_1}$$  \hspace{1cm} (4-4)

One can solve this equation numerically using a dataset that contains both measured and self-reported weight. Using self-reported values, we back-solved for measured weight by iteratively substituting in values for measured weight until the equality held at a predetermined tolerance (number of decimal places, in our case 0.001) to solve for a corrected weight in place of measured weight.
The parameter associated with the measured weight term in misreporting, \( \hat{\beta}_2 \) in equation 4-2, could be quite small in practice. \( \hat{\beta}_2 \) could be quite small because it is the coefficient of a variable that is measured in kilograms while the other terms in the equation are measured in the natural log of kilograms. Thus, measured weight, the variable that \( \hat{\beta}_2 \) is associated with, need only have a small effect to make a large impact on the natural log of self-reported weight. It might be the case that, for the ranges of BMI exhibited by the majority of the population, \( \hat{\beta}_2 \) is essentially zero. To accommodate this possibility, we also consider an alternative form of equation 4-1:

\[
\frac{W_{SR}}{h_{SR}^2} = \frac{W_M}{h_M^2} \cdot e^{\hat{\epsilon}} \tag{4-5}
\]

Equation 4-5 assumes that natural log of self-reported weight depends only the natural log of measured weight and on a constant term that captures the average effect of unobserved variables. This equation leads to a much simpler regression equation:

\[
\ln(W_{SR,i}) = \hat{\beta}_0 + \hat{\beta}_1 \cdot \ln(W_{M,i}) \tag{4-6}
\]

and correction equation if we follow the same steps as outlined for equation 4-1:

\[
W_{M,i} = e^{\frac{\ln(W_{SR,i}) - \hat{\beta}_0}{\hat{\beta}_1}} \tag{4-7}
\]

For the remainder of this paper, equation 4-4 will be referred to as the “weight-only correction”, and equation 4-7 as the “simple weight-only correction”. Equation 4-6 is similar to the equation developed by Connor Gorber et al. (20) in the sense that it does not include other covariates other than the measured and self-reported versions of the variable being corrected. The evaluation process below aims to test correction equations that would be widely usable due to their simplicity. Connor Gorber et al. (20) showed that the inclusion of other covariates such
as age, perception of one’s own weight, life dissatisfaction, ethnicity, and activity limitations did not importantly improve the accuracy of their models. Thus, we do not include any other variables in the correction equations to facilitate comparison with the existing recommended Canadian correction equations.

We first defined the full sample and identified outliers. The full sample consists of the pooled cross-section of the 2005 and 2008 CCHS respondents who provided measured and self-reported height and weight (n=6294: 3208 female, 3086 male), restricted to working-age individuals (age 18 to 65) and non-breastfeeding, non-pregnant women. Outliers (n=145) were defined as those for whom the discrepancy between self-reported and measured height or weight exceeded three standard deviations from the sex- and cycle-specific mean discrepancy. The identification and removal of outliers served to remove their potentially undue influence during correction generation.

The full sample of non-outliers was split randomly in half. One half, randomly selected, was used to calibrate the weight-only correction model parameters (the “model generating group”, n=3084). The other half (“test group”, n=3210) was used to test the model parameters to see how well the adjusted BMI values compared to measured BMI. The previously excluded outliers were included in the test group to simulate a real dataset where outliers may appear to have reasonable values of self-reported height or weight and may be impossible to identify and exclude. The regression equations of interest, 4-3 and 4-6, were run on the model generating group for males and females separately to obtain the necessary parameters for the correction equations. Male- and female-specific correction equations were developed separately to allow for known different trends in misreporting weight for males and females (12,13), and to match the
convention used in other Canadian corrections (20,21) and international corrections (14,23,24). This stratification by sex was maintained for all analyses.

The model-generating group consisted of only working age adults (i.e., 18-65), but the correction equations are appropriate to apply to adults over age 65. This was confirmed by a Chow test for equation 4-3 which was run separately for males and females. The mutually exclusive groups under consideration were 1) working age adults and 2) adults over age 65. The null hypothesis that the models have the same coefficients for both working age adults and adults over age 65 could not be rejected (for males the p-value was 0.167 for females the p-value was 0.292).

After obtaining estimates for the parameters, we applied correction equations 4-4 and 4-7 to the test group. This model solves for corrected weight, so to make a corrected value of BMI with the weight-only model we used corrected weight with self-reported height. For the test group, we report weight-only as well as simple weight-only corrections, along with BMI-only corrections using the Connor Gorber et al. model (20).

4.3.4 Evaluation of correction equations: comparison of BMI distributions; estimation of sensitivity and specificity for weight categories; and prediction of health outcomes

The different weight-only correction equations (regular and simple), which are calibrated versions of equations 4-4 and 4-7, for males and females, are given below:

Males (Weight-only Correction): \[ \ln(W_M) + \frac{.002}{.788} W_M = \frac{\ln(W_{SR}) - .710}{.788} \]

Females (Weight-only Correction): \[ \ln(W_M) + \frac{.001}{.811} W_M = \frac{\ln(W_{SR}) - .697}{.811} \]
Males (Simple weight-only Correction): \[ W_M = e^{\left( \ln(W_{SR}) - 0.304 \right) / 0.926} \]

Females (Simple weight-only Correction): \[ W_M = e^{\left( \ln(W_{SR}) - 0.209 \right) / 0.942} \]

We evaluate the correction equations for usefulness in three ways: first, we compare self-reported, corrected, and measured BMI distributions. Second we compare self-reported and corrected BMI to measured BMI by estimating the sensitivity/specificity of the equations vis-à-vis BMI categories of normal weight (18.5 ≤ BMI < 25), overweight (25 ≤ BMI < 30), and obese (BMI ≥ 30). Lastly, we compare self-reported and corrected BMI to measured BMI by association with selected health outcomes in regression models. We compare the regression results in terms of statistical significance, coefficient magnitude, and direction of the coefficient (above or below the measured coefficient). Assessment of association with health outcomes entailed comparing coefficients (statistical significance, magnitude, and direction) across the different corrected BMI values in regression equations modeling arthritis, heart disease, diabetes, high blood pressure, self-reported health, and activity limitation, following Connor Gorber et al. (20). We use the BMI categories of normal weight (18.5 ≤ BMI < 25), overweight (25 ≤ BMI < 30), obese (30 ≤ BMI < 35), and obese class II or higher (BMI ≥ 35) for this part of the analysis. We further restrict this part of the analysis to individuals aged 40 or older, to follow the convention set by Connor Gorber et al. (20); however we also test the disease association models with the full age range (18-64). Throughout our analysis we do not report results for underweight individuals because there were so few underweight individuals by measured BMI.
4.4 Results

4.4.1 Distribution of BMI

Figures 4.2 and Figure 4.3 show the BMI distributions estimated from the measured, self-reported, and corrected BMI values. For males the corrected distributions all trend together, and all are closer to the measured distribution than to the self-reported distribution. For females, the weight-only corrections resemble the measured BMI distribution more closely than the BMI-only correction does, most notably between BMI 23 kg/m$^2$ and 28 kg/m$^2$.

4.4.2 Sensitivity and specificity of corrected and self-reported BMI

Table 4.1 displays the sensitivity and specificity estimates based on the self-reported and corrected values for BMI for males and females. Focusing on those instances with at least a 5 percentage point difference in sensitivity or specificity within sex-BMI category groups, we observed the following patterns: All corrections were similar to one another and superior to self-report in specificity of normal weight among women, sensitivity of overweight among both men and women, and sensitivity of obese among both men and women. There were two instances in which self-report was superior to corrected values: sensitivity of normal weight among both men and women. Overall, any corrected BMI was superior to self-reported BMI in estimating prevalence within weight categories.
Figure 4.2: The distribution of BMI by different correction equations and self-report compared to measured BMI in males.
Figure 4.3: The distribution of BMI by different correction equations and self-report compared to measured BMI in females
Table 4.1: Sensitivity and specificity estimates using self-reported BMI and three correction equations

**MALES - Selected BMI indicators versus measured BMI**

|                  | Normal Weight | | | Overweight | | | Obese | |
|------------------|---------------|---|---|-------------|---|---|-------------|---|---|
|                  | Sens          | Spec          | Sens          | Spec          | Sens          | Spec          | Sens          | Spec          |
| Self-Reported    |               |               |               |               |               |               |               |               |
| BMI Correction   |               |               |               |               |               |               |               |               |
| Weight-only Corr. (Simple) |               |               |               |               |               |               |               |               |
|                  |               |               |               |               |               |               |               |               |

95% confidence intervals in brackets
4.4.3 Predictive utility of corrected BMI in health condition models

Table 4.2 and Table 4.3 show the results of models regressing six health conditions on BMI (measured; self-report; BMI-only correction; weight-only corrections) for men and women, controlling for age. Focusing on differences in statistical significance (i.e., presence/absence) between coefficients in the measured BMI model and coefficients in each of the other models, the following is apparent: For men, of the 18 coefficients in each column, 16 in the self-reported BMI column had the same statistical significance status as measured BMI. The numbers for the BMI-only, weight-only, and simple weight-only corrections were 15/18, 16/18, and 15/18, respectively. For women, of the 18 coefficients in each column, 13 in the self-reported BMI column had the same statistical significance status as measured BMI. The numbers for the BMI-only, weight-only, and simple weight-only corrections were 15/18, 15/18, and 14/18, respectively. From this preliminary consideration of statistical significance, the correction equations and self-report BMI appear to perform similarly for men, while for women the correction equations appear to perform similarly to each other and better than self-reported BMI.
Table 4.2: Odds ratios from regressions for BMI category and selected health conditions for males, controlling for age. (n=821)

<table>
<thead>
<tr>
<th>Health Condition</th>
<th>Measured BMI OR 95% CI</th>
<th>Self-Reported BMI OR 95% CI</th>
<th>BMI Correction OR 95% CI</th>
<th>Weight-only Correction OR 95% CI</th>
<th>Weight-Only Correction (Simple) OR 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Arthritis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.92</td>
<td>0.35 to 2.46</td>
<td>1.81</td>
<td>0.90 to 3.64</td>
<td>1.78</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>0.98</td>
<td>0.34 to 2.81</td>
<td>2</td>
<td>0.86 to 4.67</td>
<td>1.61</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>1.52</td>
<td>0.43 to 5.33</td>
<td>3.79*</td>
<td>1.40 to 10.28</td>
<td>2.54*</td>
</tr>
<tr>
<td><strong>Heart Disease</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.84</td>
<td>0.30 to 2.37</td>
<td>1.23</td>
<td>0.46 to 3.35</td>
<td>1.45</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.36</td>
<td>0.41 to 4.53</td>
<td>1.42</td>
<td>0.48 to 4.22</td>
<td>1.21</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>0.79</td>
<td>0.19 to 3.27</td>
<td>1.82</td>
<td>0.40 to 8.20</td>
<td>2.21</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.99</td>
<td>0.33 to 2.95</td>
<td>0.75</td>
<td>0.29 to 1.93</td>
<td>1.3</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>2.90*</td>
<td>1.02 to 8.22</td>
<td>2.05</td>
<td>0.71 to 5.92</td>
<td>2.28</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>4.81*</td>
<td>1.48 to 15.64</td>
<td>5.89*</td>
<td>2.04 to 17.01</td>
<td>5.40*</td>
</tr>
<tr>
<td><strong>High Blood Pressure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>1.01</td>
<td>0.50 to 2.04</td>
<td>1.56</td>
<td>0.85 to 2.87</td>
<td>1.28</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>2.98*</td>
<td>1.38 to 6.44</td>
<td>3.46*</td>
<td>1.64 to 7.32</td>
<td>3.30*</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>4.56*</td>
<td>1.81 to 11.50</td>
<td>5.33*</td>
<td>2.04 to 13.95</td>
<td>3.00*</td>
</tr>
<tr>
<td><strong>Self-Reported Health (Fair or Poor)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.75</td>
<td>0.35 to 1.63</td>
<td>0.84</td>
<td>0.42 to 1.66</td>
<td>0.61</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.26</td>
<td>0.59 to 2.71</td>
<td>1.65</td>
<td>0.79 to 3.47</td>
<td>1.32</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>1.33</td>
<td>0.48 to 3.71</td>
<td>1.5</td>
<td>0.57 to 3.96</td>
<td>1.55</td>
</tr>
<tr>
<td><strong>Activity Limitation (Often or Sometimes)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overweight</td>
<td>1.13</td>
<td>0.66 to 1.93</td>
<td>1.08</td>
<td>0.65 to 1.78</td>
<td>0.95</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.53</td>
<td>0.79 to 2.96</td>
<td>1.7</td>
<td>0.90 to 3.24</td>
<td>1.18</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>2.55*</td>
<td>1.15 to 5.65</td>
<td>3.53*</td>
<td>1.51 to 8.25</td>
<td>1.88</td>
</tr>
</tbody>
</table>

*statistically significantly different from Normal Weight Odds Ratio (p<0.05)

Models control for age. There were so few observations in the underweight category that those individuals are excluded from this analysis. Regression performed only using individuals age 40 or higher.
Table 4.3: Odds ratios from regressions for BMI category and selected health conditions for females, controlling for age. (n=942)

<table>
<thead>
<tr>
<th>Health Condition</th>
<th>Measured BMI OR</th>
<th>95% CI</th>
<th>Self-Reported BMI OR</th>
<th>95% CI</th>
<th>BMI Correction OR</th>
<th>95% CI</th>
<th>Weight-only Correction OR</th>
<th>95% CI</th>
<th>Weight-Only Correction (Simple) OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arthritis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>1.08</td>
<td>0.61 to 1.90</td>
<td>1.04</td>
<td>0.61 to 1.77</td>
<td>1.21</td>
<td>0.69 to 2.11</td>
<td>1.22</td>
<td>0.70 to 2.13</td>
<td>1.02</td>
<td>0.58 to 1.77</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.24</td>
<td>0.66 to 2.34</td>
<td>2.72*</td>
<td>1.56 to 4.74</td>
<td>1.67</td>
<td>0.94 to 2.97</td>
<td>1.70*</td>
<td>0.96 to 3.00</td>
<td>1.54</td>
<td>0.87 to 2.74</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>4.97*</td>
<td>2.45 to 10.08</td>
<td>4.29*</td>
<td>1.85 to 9.97</td>
<td>5.04*</td>
<td>2.45 to 10.37</td>
<td>4.92*</td>
<td>2.41 to 10.04</td>
<td>4.36*</td>
<td>2.16 to 8.79</td>
</tr>
<tr>
<td>Heart Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>0.47</td>
<td>0.10 to 2.23</td>
<td>2.5</td>
<td>0.69 to 8.96</td>
<td>1.78</td>
<td>0.48 to 6.65</td>
<td>1.9</td>
<td>0.51 to 7.09</td>
<td>1.97</td>
<td>0.53 to 7.33</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.45</td>
<td>0.34 to 6.16</td>
<td>2.64</td>
<td>0.96 to 7.22</td>
<td>1.63</td>
<td>0.55 to 4.84</td>
<td>1.64</td>
<td>0.54 to 4.95</td>
<td>1.69</td>
<td>0.56 to 5.10</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>2.04</td>
<td>0.39 to 10.64</td>
<td>4.54</td>
<td>0.83 to 24.67</td>
<td>3.29</td>
<td>0.82 to 13.14</td>
<td>3.44</td>
<td>0.88 to 13.55</td>
<td>3.43</td>
<td>0.87 to 13.47</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>0.71</td>
<td>0.21 to 2.46</td>
<td>1.43</td>
<td>0.41 to 4.94</td>
<td>6.28*</td>
<td>1.67 to 23.56</td>
<td>4.30*</td>
<td>1.25 to 14.77</td>
<td>4.45*</td>
<td>1.29 to 15.32</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>0.9</td>
<td>0.24 to 3.33</td>
<td>1.85</td>
<td>0.65 to 5.26</td>
<td>3.51*</td>
<td>1.01 to 12.13</td>
<td>2.53</td>
<td>0.83 to 7.71</td>
<td>2.42</td>
<td>0.87 to 7.51</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>3.13</td>
<td>0.78 to 12.44</td>
<td>5.26*</td>
<td>1.23 to 22.36</td>
<td>14.56*</td>
<td>3.79 to 55.93</td>
<td>10.29*</td>
<td>3.01 to 35.25</td>
<td>10.55*</td>
<td>3.12 to 35.60</td>
</tr>
<tr>
<td>High Blood Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>1.45</td>
<td>0.70 to 3.01</td>
<td>2.12*</td>
<td>1.12 to 4.00</td>
<td>1.84</td>
<td>0.92 to 3.67</td>
<td>1.94</td>
<td>0.97 to 3.85</td>
<td>2.00*</td>
<td>1.01 to 3.99</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>2.26*</td>
<td>1.09 to 4.69</td>
<td>3.70*</td>
<td>1.99 to 6.89</td>
<td>3.19*</td>
<td>1.65 to 6.13</td>
<td>3.27*</td>
<td>1.70 to 6.28</td>
<td>3.38*</td>
<td>1.75 to 6.50</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>5.28*</td>
<td>2.20 to 12.68</td>
<td>7.15*</td>
<td>2.57 to 19.90</td>
<td>5.50*</td>
<td>2.27 to 13.27</td>
<td>5.54*</td>
<td>2.32 to 13.26</td>
<td>5.55*</td>
<td>3.12 to 35.60</td>
</tr>
<tr>
<td>Self-Reported Health (Fair or Poor)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>0.7</td>
<td>0.34 to 1.44</td>
<td>1.05</td>
<td>0.51 to 2.17</td>
<td>0.95</td>
<td>0.46 to 1.95</td>
<td>0.98</td>
<td>0.47 to 2.01</td>
<td>1.02</td>
<td>0.49 to 2.10</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.19</td>
<td>0.54 to 2.62</td>
<td>2.25*</td>
<td>1.12 to 4.53</td>
<td>1.31</td>
<td>0.63 to 2.69</td>
<td>1.34</td>
<td>0.65 to 2.75</td>
<td>1.18</td>
<td>0.59 to 2.37</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>2.29</td>
<td>0.84 to 6.22</td>
<td>2.55</td>
<td>0.69 to 9.30</td>
<td>2.37</td>
<td>0.86 to 6.50</td>
<td>2.37</td>
<td>0.87 to 6.46</td>
<td>2.70*</td>
<td>1.03 to 7.12</td>
</tr>
<tr>
<td>Activity Limitation (Often or Sometimes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal Weight</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>1.23</td>
<td>0.75 to 2.02</td>
<td>1.31</td>
<td>0.79 to 2.19</td>
<td>0.96</td>
<td>0.58 to 1.57</td>
<td>0.97</td>
<td>0.59 to 1.58</td>
<td>0.9</td>
<td>0.55 to 1.47</td>
</tr>
<tr>
<td>Obese (Class I)</td>
<td>1.81</td>
<td>0.96 to 3.43</td>
<td>2.03*</td>
<td>1.15 to 3.57</td>
<td>1.74</td>
<td>0.93 to 3.23</td>
<td>1.74</td>
<td>0.94 to 3.24</td>
<td>1.57</td>
<td>0.84 to 2.94</td>
</tr>
<tr>
<td>Obese (Class II+)</td>
<td>3.80*</td>
<td>1.90 to 7.62</td>
<td>3.70*</td>
<td>1.50 to 9.13</td>
<td>3.20*</td>
<td>1.49 to 6.91</td>
<td>3.19*</td>
<td>1.50 to 6.82</td>
<td>3.30*</td>
<td>1.56 to 6.98</td>
</tr>
</tbody>
</table>

*statistically significantly different from Normal Weight Odds Ratio (p<0.05)

Models control for age. There were so few observations in the underweight category that those individuals are excluded from this analysis. Regression performed only using individuals age 40 or higher.
In terms of the magnitude and direction of the coefficients for the relationships between BMI (self-reported and corrected) and disease outcomes (Table 4.2 and Table 4.3), the self-reported and corrected BMI measures did not exhibit a consistent pattern for males or females across diseases. For males, the corrected odds ratios were higher than the measured odds ratios in almost all comparisons; exceptions were heart disease and diabetes (obese class I), high blood pressure (obese class II+), and self-reported health (overweight). All of the corrected odds ratios for activity limitation were lower than the measured estimates. For females, the corrected odds ratios are higher than the measured odds ratios for heart disease, diabetes, high blood pressure, and self-reported health (except for the simple weight-only correction for obese class I). The corrected odds ratios for activity limitation are all lower than the measured odds ratios, and for arthritis the weight-only corrections were lower than the measured odds ratios for overweight (simple correction only) and obese class II+. Thus, the magnitudes and directions of the estimated coefficients in the health condition models do not clearly point to a superior correction equation.

Furthermore with respect to estimates in Table 4.2 and Table 4.3, the corrected odds ratios tended to have wider confidence intervals (suggesting less precision of estimate) than the corresponding measured odds ratios, but not in every case. Considering only obese class II+ as an example, for males the corrected confidence intervals were wider than the measured confidence intervals for arthritis, heart disease, diabetes, and self-reported health. For females the

---

4 The disease models were repeated for the entire sample (age 18 to 65). The results of those models (not reported) are consistent with those from the truncated sample; namely, they do not clearly point to any correction equation being superior.
corrected confidence intervals were wider than the measured confidence intervals in all cases but one (simple weight-only correction for arthritis).

4.5 Discussion

4.5.1 Distribution of BMI

A corrected BMI distribution, regardless of whether BMI-only or weight-only, was found to be more accurate than the self-reported BMI distribution (see Figures Figure 4.2 and Figure 4.3). However, this only refers to the ability of the correction equations to bring the distribution of self-reported BMI into line with the distribution of measured BMI, unconditional on any other variables or restrictions. In terms of which correction is better, for males, it is not obvious that there is a superior corrected distribution. For females, the weight-only corrections follow the measured BMI distribution more closely than the BMI-only correction, making the weight-only corrections the best overall candidate correction for simply constructing the BMI distribution.

4.5.2 Sensitivity and specificity

Findings from the sensitivity and specificity analyses were mixed, with the weight-only and BMI-only performing similarly well in some cases (specificity for normal weight women, sensitivity for overweight men and women, and sensitivity for obese men and women), and self-report best in others (sensitivity for normal weight men and women). The high sensitivity for self-reported BMI in the normal categories for females is consistent with past observations that normal BMI females, on average, under- or over-report weight to a lower degree than other BMI groups (20,21). In the absence of a superior correction equation, researchers must trade-off sensitivity and specificity when choosing a correction equation. The outcome of these trade-offs depend on the situation: for example, if a researcher were interested in studying a sample of predominantly normal weight males, where the weight-only corrections result in losing 10
percentage points in sensitivity in exchange for 6 percentage points in specificity, it is not clear that a weight-only correction is superior to self-reported BMI. Simply stating that one percentage point gain offsets another is inappropriate, especially when specificity may be more important for conditions like obesity, where the majority of the population is not obese (7).

4.5.3 Predictive utility of corrected BMI in health condition models

Taking the population distribution and the sensitivity/specificity findings as a whole, our results suggest that if a researcher is interested in BMI statistics across a population, including estimating prevalence within weight categories, any correction presented here will be preferable to self-reported BMI. However, findings from the analysis of associations with health outcomes tell a different story: namely, findings were mixed, such that no correction was uniformly superior and in some cases the self-reported data outperformed the corrected data. None of the correction equations were able to consistently provide coefficient estimates closer to the measured BMI estimates than those provided by the self-reported estimates. The implication of this finding is that in a disease modelling context correction equations are not necessarily more useful than self-reported BMI.

Our findings differ from other Canadian correction equation studies that have showed the BMI-only correction consistently provides coefficient estimates closer to the measured values (20,21). In statistics, it is recognized that including a variable exhibiting measurement error on the right hand side of the regression equation can provide a biased estimate of effect (30). The magnitude of this bias depends on the variance of the mis-measured variable (in our case, measured BMI) and the variance of the measurement error itself (in our case, the difference between measured and self-reported BMI). While the corrections presented here adequately correct the average BMI at percentiles of interest and can be used to estimate the distribution of
BMI or the prevalence of obesity, they are unable to correct the variance of self-reported BMI. As a consequence, corrected BMI measures provide biased and inconsistent estimates of association when used as regressors, just as any mis-measured variable would. The key issue in our case is whether the magnitude of the bias is substantial (i.e., clinically or socially significant). We suggest that the magnitude of this bias is too large, and the direction of the bias too unpredictable, for corrected BMI variables to be used in this context of modelling health outcomes.

The health outcome association analysis shows that, given a disease in a logit model and a categorical BMI variable (a very common modelling convention), researchers should not necessarily use a correction equation in an effort to improve self-reported BMI. Further, if a researcher, seeking a conservative range of estimates, chose to report both a corrected and a self-reported estimate, it is not clear that the difference between the two should be meaningful anyway in this regression framework: when the corrected estimate is larger than the self-reported coefficient, it does not conform to the idea that the self-reported BMI estimate serves as an upper bound; when the corrected estimate is smaller than the self-reported coefficient, it is not a guarantee that the corrected estimate is closer to the measured BMI estimate or even on the correct side of the null value. Resorting back to self-reported BMI is going to provide biased estimates of association between BMI and the dependent variable, as well as the biased and inconsistent estimates of association for all the other regressor variables in the model (30). In short, our results suggest that if a researcher is interested in using BMI as a predictor variable for modelling disease, then both self-reported and corrected BMI result in biased estimates of association.
4.6 Limitations

It is important to acknowledge that, because correction equations are based on particular populations, which change over time and place, the equations themselves are likewise somewhat time and place dependent and should be updated over time. Changing misreporting patterns over time have been shown using data from the United States (27), and Ireland (31) and may reflect, in part, changing social attitudes about obesity. The fact that correction equations can change across time is important for modelling the BMI distribution, but updating a correction equation will not fix the issue of a corrected BMI providing biased and inconsistent regression results unless the update somehow deals with the variance of the misreported variable.

Although the response rate for the CCHS surveys as a whole were reasonably high (87.0% for the 2005 CCHS was and 85% for the 2008 CCHS), another limitation of the study is the lower response rate for the sub-samples among which both self-reported and measured data were available and which constituted the basis for this study. The overall response rate for the sub-sample that provided measured height and weight was 55.9% for 2005 CCHS and 50.7% for the 2008 CCHS, most of the non-response was from refusal to be measured (21). It is unlikely that those who refuse to be measured are randomly distributed across the BMI distribution, so if the individuals who are heavier are refusing to be measured, any correction equation based on this dataset will be inaccurate, including others that have been developed (20,21). Strategies to improve response rate across the population are thus desirable.

4.7 Conclusions

The BMI-only correction has been applied to Canadian data extensively (e.g., Orpana et al. (32), Janssen et al. (33), and Barberio and McLaren (34)), attesting to the value of such corrections in the literature. Our findings support the use of BMI-only corrections if the
researcher is interested in reporting the distribution of BMI, the prevalence of those above and below the obesity threshold of BMI 30 kg/m$^2$, or any other cut-point. On the other hand, if the researcher is interested in estimating the effect of BMI on a health condition, then our findings suggest that corrected BMI, using any of the methods examined here, does not represent an improvement over self-report data.
4.8 References


Chapter Five: How important are determinants of obesity measured at the individual-level for explaining geographic variation in body mass index distributions? Evidence from Canada.

5.1 Abstract

5.1.1 Background

Obesity prevalence varies across geographic regions in Canada. The reasons for this variation are unclear, but likely implicate both individual characteristics and population-level forces. The objective of this study is to examine whether equalizing the identified determinants of body mass index (BMI) across geographic regions of Canada could be reasonably expected to reduce differences in BMI distributions between those regions.

5.1.2 Methods

We analyzed data from four cycles of the Canadian Community Health Survey (CCHS): 2001, 2003, 2005, and 2007, for both males and females. First, we modelled determinants of obesity using OLS regression separately by region. Second, we modelled those same determinants using quantile regression. Third, in an effort to disentangle individual-level and population-level determinants, we compare regions using Blinder-Oaxaca decomposition to determine how much of the cross-regional difference in BMI is attributable to different levels of the measured determinants (“explained share”) and the share that remains unexplained.

5.1.3 Results

Findings from OLS regression analyses were consistent with the literature (e.g., higher income associated with lower BMI in females for most regions). Findings from quantile regression reveal that average effects captured by OLS do not necessarily match effects from across the BMI distribution. For example, post-secondary education (versus high school or less) has a negative effect at the higher end of the BMI distribution for females in the Atlantic region, a relationship that was insignificant in OLS. Findings from the decomposition analysis show that
the determinants account for only a small share of the overall differences between regions. For example, for females in the Atlantic provinces versus Quebec, the explained share makes up less than half of the overall difference between regions.

5.1.4 Discussion

Determinants such as income, education, physical activity level, fruit and vegetable consumption, smoking status, drinking status, and marital status are capable of explaining some individual variation in BMI within geographic regions. Those determinants are not, however, capable of explaining variation in BMI between regions, which implies the importance of population-level factors. Equalization of the identified determinants measured at the individual level across regions cannot be reasonably expected to reduce differences in the BMI distributions across regions.
5.2 Background

5.2.1 Determinants of obesity measured at the individual level

Research from various disciplines has focused on quantifying the effect of determinants of obesity and body mass index (BMI). For example, income has a negative association with BMI in females in developed countries, while findings for men have been mixed, with positive, negative, and statistically insignificant associations being found in the literature (1). Education and occupation are negatively associated with BMI in both men and women. Other important determinants include diet (e.g., fruit and vegetable consumption (2)); physical activity level (3); sex and ethnicity (4); smoking status (5); mental health conditions such as depression (6); age (7); social support (8); and others.

Many quantitative studies on the determinants of BMI measured at the individual level focus on the average effects of the determinants, e.g., using OLS regression to infer how average BMI should respond to changes in average levels of determinants. While such an approach has yielded important contributions to the knowledge base, average effects do not necessarily capture the impact determinants have on those at the high end of the BMI distribution, those with the heaviest BMIs. The BMI distribution is skewed (see Figure 5.1 for examples from Canada); therefore, variables that are associated with average BMI do not necessarily affect higher BMI to the same magnitude or even in the same direction.

5.2.2 Considering the full BMI distribution: quantile regression

A few studies have examined the potentially differential effects of determinants of BMI across the BMI distribution. These studies use quantile regression (9,10), which allows for the effect of determinants to vary depending on the quantile (or percentile) of BMI under consideration.
Figure 5.1: Kernel densities of BMI across regions for females and males

Self-reported BMI by region for females

Self-reported BMI by region for males

Self-reported BMI truncated at 60
Three American studies examined effects of determinants across the distribution of BMI. The first study found that local fast food prices at the county level had the largest effect on the heaviest individuals, those at the 75th and 90th percentiles (11). For example, a $1 decrease in the price of fast food was associated with a 0.1 kg/m² increase in average BMI for males, but a 1.0 kg/m² increase in BMI for those at the 75th and 90th percentiles of BMI. The second study showed that among females, participation in a food stamp program was associated with an increase of 1.6 kg/m² on average, but an increase of 2.9 kg/m² at the 90th percentile of BMI (12). The third study estimated the effect of educational attainment on BMI in females (13). One additional year of education was associated with a 0.165 kg/m² decrease in BMI for females at the median BMI level, and a 0.3 kg/m² decrease in BMI for females at the 90th percentile of BMI. A French study observed that the effect of socio-economic position differed across the BMI distribution and by sex (14). For example, the negative relationship between BMI and socio-economic position was statistically significant for females at the 70th BMI percentile or higher, but not at lower percentiles. For males there was virtually no relationship across quantiles. Finally, analysis of Canadian BMI data showed a negative association between BMI and education for men and women on average, and that association was stronger for women at the 60th percentile and up (15).

Overall, research that has studied the determinants of BMI across the distribution using quantile regression is consistent in its message: average estimates sometimes suppress interesting effects that are unique to certain sections of the BMI distribution (e.g., the higher percentiles). This is an important consideration in population health research because it could be that some determinants serve to increase the spread of the BMI distribution; an effect that would be missed when considering only the average. A variable that increases the spread of the BMI distribution
is important for determining the causes of obesity incidence, as well as potential leverage points for prevention, when comparing across populations (16).

5.2.3 Considering differential impact: decomposition techniques

When determinants of BMI measured at the individual level are identified, they may be suggested as potential targets for policymakers interested in lowering average BMI or shifting the BMI distribution. For instance, studies have implied that targeting income levels (through redistribution via taxes or wages) is one potential solution for addressing obesity inequality that arises through complex pathways related to income (e.g., (17,18)). However, such an approach may not have the desired impact if income has a different effect on BMI in different groups. One important example of such groups in Canada is geographic region: according to recent BMI maps (19) there is an East to West gradient whereby those in the Atlantic provinces (eastern part of Canada) have a higher prevalence of obesity than the rest of Canada, and British Columbia (westernmost province) has a low prevalence of obesity. Thus, if income is lower in the Atlantic provinces, it seems reasonable to suggest that raising income in those provinces could decrease obesity prevalence. However, this might not have the desired impact if income has a different effect on BMI in the Atlantic provinces than elsewhere.

Decomposition techniques may be used to determine the relative importance of 1) the levels and 2) the effects of determinants measured at the individual level across geographic regions. Continuing with the income example, by applying decompositions across geographic regions, one can derive two shares of the overall between-region difference: 1) a share attributable to the levels of income across the two regions (the “explained share”), and 2) a share that remains unexplained. If the share of the overall difference that remains unexplained is large, it would indicate that income (or whatever determinant(s) one includes) does a poor job of
explaining the difference across regions and is not a relevant policy target for reducing between-region differences in BMI.

A few studies have been published that use decomposition techniques to explore geographic differences in BMI across geographic or ethnic groups. For example, a study comparing men and women in Italy (a country with low obesity prevalence) to those in Spain (a country with higher obesity prevalence) showed that determinants (i.e., age; years of education; marital, employment and smoking statuses, all measured at the individual level) could not explain the difference in BMI distributions between countries, for women or men (20). A second study using Malaysian data compared BMI distributions between ethnic Malays and Chinese living in Malaysia (21). The study concluded that the determinants studied (i.e., age; household income; education level; and living in an urban or rural area; all measured at the individual level) accounted for very little of the BMI distribution difference between the Malays (heavier group) and Chinese (lighter group). Some analyses have focused on decomposing geographic differences in BMI in Canada (22,23), including one study by our group (22). Briefly, these studies found that obesity is concentrated among poorer individuals in the Atlantic provinces, a concentration that is not present to the same extent in other regions (and reversed in some provinces like Alberta) (23) and that addressing determinants measured at the individual level will have ambiguous effects for equalizing average BMI across regions (22). However, little effort has been devoted to decomposing differences across the entire BMI distribution, despite the fact that conclusions may differ substantively from decompositions based on averages.

5.2.4 The present study

The objective of this study is to examine whether equalizing the identified determinants of BMI across geographic regions of Canada could be reasonably expected to reduce differences
in BMI distributions across those regions. By using a combination of quantile regression and
decomposition techniques, this study will present an estimate of the difference in regional BMI
distributions that would remain if we were to hypothetically equalize the determinants across the
different regions.

5.3 Methods

5.3.1 Data

Because of the data intensive nature of the analytic procedures, this study uses merged
data from four cycles (from 2001, 2003, 2005, and 2007) of the Canadian Community Health
Survey (CCHS). The CCHS is a repeated, nationally representative cross-sectional survey
conducted by Statistics Canada. The target population is individuals over age 12 living in
Canada’s ten provinces. Those living in institutions, remote areas, the northern territories, on
Crown lands, and military bases are excluded. The survey represents 98% of the target Canadian
population (24). The CCHS uses a multi-stage stratified cluster sampling procedure, and weights
are provided to account for this complex sampling technique.

The variables we used are based on past literature on the determinants of BMI measured
at the individual level (1,2,25-27). Specifically, we used: age; income decile (specifically,
household income adjusted for household size and community size\(^5\)); smoking status (non-
smoker, occasional smoker, daily smoker); alcohol use status (non-drinker, occasional drinker,
former drinker, daily drinker); physical activity level (active, moderate, or inactive based on an
index constructed by Statistics Canada using daily energy expenditure); fruit and vegetable

\(^5\) Income deciles are computed by Statistics Canada in the CCHS at the provincial-level from all respondents who
provided a valid income value. The decile is based on the ratio of household income of the respondent to the low
income cut-off appropriate for that respondent’s household and community size. The deciles are dummy variables,
so the 10\(^{th}\) income decile indicates an individual falls into the lowest 10% of such ratios in their province. The
percentiles are computed for males and females together.
consumption (number per day); whether the respondent had a family doctor (yes/no); marital status (single vs. other); urban vs. rural residence; education (holds a post-secondary degree or higher vs. not); and employment in the past 12 months (was employed vs. not). Separate models were run for males and females in each Canadian region (Atlantic provinces, Quebec, Ontario, Prairie provinces, and British Columbia). Our sample of interest was working age individuals (age 18-65).

BMI was constructed from self-reported height and weight, which is the format in which most Canadian national data is available. We re-ran all analyses replacing self-reported BMI with two different corrected measures of BMI (28), one of which we developed (29). Results were substantively identical in all cases, and those from self-reported data are presented.

5.3.2 Analytic Procedure

We took three steps to address the study objective: OLS regression, unconditional quantile regression, and decomposition of quantile regression.

5.3.2.1 OLS regression

First, we regressed BMI on the determinants listed above, separately by region and sex.

5.3.2.2 Unconditional (marginal) quantile regression

Next, we ran unconditional (or marginal) quantile regression using the same model specification as the OLS regression. We estimated every second percentile from the 2nd to the 98th. At this stage of the analysis, we were particularly interested in whether the average effect

---

6 The term “unconditional”, when used in this context, refers to a way of interpreting the coefficients obtained from a particular estimation strategy. Unconditional quantile regression uses the insight that re-centring the influence function of a statistic of interest (in this case, a quantile), and using that value in a regression, can obtain coefficients with an unconditional (marginal) interpretation by using a least squares estimator. This technique contrasts standard quantile regression (9,10) and is desirable in the case of BMI in particular, where an unconditional quantile regression (30-32) allows for marginal effect interpretations of the determinants measured at the individual level. The convention used by the authors who developed the technique is to use the term “unconditional” over “marginal” to avoid confusing marginal probabilities with marginal effects.
of a variable as estimated by OLS was similar to the effect of that same variable across the BMI distribution.

5.3.2.3 Decomposition of quantile regression

Finally, we ran Blinder-Oaxaca-type decompositions across estimated quantiles using the results from the second step (30-32). We previously applied a standard Blinder-Oaxaca decomposition to BMI data across regions of Canada and found that the unobserved share of the difference played a large role in some cross-regional comparisons (22). That study, however, focused only on the differences in average BMI. It is important to build on that previous study to consider the different distributions of BMI across regions. The reason is, because the BMI distribution has become more right-skewed over time, differences between regions may be associated with differently skewed distributions, as opposed to different averages in similar distributions (33). We use the Atlantic provinces as the comparison region because they have the most skewed BMI distribution and thus the most informative counterfactual. The counterfactual can be stated as, “If we were to equalize determinants across regions, would the BMI distributions become more similar?”
Table 5.1: Summary statistics for all study variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Atlantic Provinces</th>
<th>Quebec</th>
<th>Ontario</th>
<th>Prairies</th>
<th>BC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td>n</td>
<td>8246</td>
<td>9440</td>
<td>13207</td>
<td>14064</td>
<td>30681</td>
</tr>
<tr>
<td>Self-reported BMI (kg/m^2)</td>
<td>27.34</td>
<td>26.57</td>
<td>26.12</td>
<td>24.73</td>
<td>27.07</td>
</tr>
<tr>
<td>Age (Years)</td>
<td>43.5</td>
<td>42.82</td>
<td>42.52</td>
<td>42.94</td>
<td>42.24</td>
</tr>
<tr>
<td>1st Income Decile (%)</td>
<td>9%</td>
<td>14%</td>
<td>6%</td>
<td>10%</td>
<td>6%</td>
</tr>
<tr>
<td>2nd Income Decile (%)</td>
<td>6%</td>
<td>10%</td>
<td>6%</td>
<td>9%</td>
<td>6%</td>
</tr>
<tr>
<td>3rd Income Decile (%)</td>
<td>8%</td>
<td>8%</td>
<td>7%</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>4th Income Decile (%)</td>
<td>8%</td>
<td>9%</td>
<td>8%</td>
<td>9%</td>
<td>8%</td>
</tr>
<tr>
<td>5th Income Decile (%)</td>
<td>10%</td>
<td>9%</td>
<td>10%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td>6th Income Decile (%)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>7th Income Decile (%)</td>
<td>10%</td>
<td>10%</td>
<td>11%</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>8th Income Decile (%)</td>
<td>11%</td>
<td>10%</td>
<td>12%</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>9th Income Decile (%)</td>
<td>12%</td>
<td>10%</td>
<td>13%</td>
<td>10%</td>
<td>13%</td>
</tr>
<tr>
<td>10th Income Decile (%)</td>
<td>15%</td>
<td>10%</td>
<td>14%</td>
<td>10%</td>
<td>15%</td>
</tr>
<tr>
<td>Education beyond high school (%)</td>
<td>59%</td>
<td>61%</td>
<td>62%</td>
<td>62%</td>
<td>66%</td>
</tr>
<tr>
<td>Daily Smoker (%)</td>
<td>29%</td>
<td>26%</td>
<td>29%</td>
<td>27%</td>
<td>26%</td>
</tr>
<tr>
<td>Occasional Smoker (%)</td>
<td>4%</td>
<td>4%</td>
<td>5%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>Occasional Drinker (%)</td>
<td>14%</td>
<td>30%</td>
<td>10%</td>
<td>21%</td>
<td>12%</td>
</tr>
<tr>
<td>Former Drinker (%)</td>
<td>14%</td>
<td>16%</td>
<td>9%</td>
<td>10%</td>
<td>9%</td>
</tr>
<tr>
<td>Never Drinker (%)</td>
<td>3%</td>
<td>6%</td>
<td>2%</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>Physically Active (%)</td>
<td>22%</td>
<td>18%</td>
<td>23%</td>
<td>17%</td>
<td>28%</td>
</tr>
<tr>
<td>Moderately Physically Active (%)</td>
<td>24%</td>
<td>25%</td>
<td>25%</td>
<td>26%</td>
<td>25%</td>
</tr>
<tr>
<td>Fruits &amp; Vegetables Per Day (mean)</td>
<td>3.89</td>
<td>4.51</td>
<td>4.37</td>
<td>5.3</td>
<td>4.31</td>
</tr>
<tr>
<td>No Family Doctor (%)</td>
<td>15%</td>
<td>8%</td>
<td>36%</td>
<td>20%</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>36%</td>
<td>41%</td>
<td>43%</td>
<td>37%</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Single (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural (%)</td>
<td>46%</td>
<td>43%</td>
<td>27%</td>
<td>26%</td>
<td>20%</td>
</tr>
<tr>
<td>Worked in the Past 12 Months (%)</td>
<td>84%</td>
<td>73%</td>
<td>85%</td>
<td>73%</td>
<td>89%</td>
</tr>
</tbody>
</table>

Note: Income deciles are computed by Statistics Canada in the CCHS at the provincial-level from all respondents who provided a valid income value. The decile is based on the ratio of household income of the respondent to the low income cut-off appropriate for that respondent’s household and community size. The deciles are dummy variables, so the 10th income decile indicates an individual falls into the lowest 10% of such ratios in their province. The percentiles are computed for males and females together. Our sample is not evenly split by decile since not all individuals with valid income measures are included in our study (i.e., they did not all provide valid responses for all other variables used).
### Table 5.2: BMI percentiles by sex and region

<table>
<thead>
<tr>
<th>n</th>
<th>Atlantic Provinces</th>
<th></th>
<th>Quebec</th>
<th></th>
<th>Ontario</th>
<th></th>
<th>The Prairies</th>
<th></th>
<th>BC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>10th Percentile</td>
<td>8246</td>
<td>9440</td>
<td>21.4</td>
<td>19.4</td>
<td>22.0</td>
<td>19.8</td>
<td>22.1</td>
<td>20.0</td>
<td>21.6</td>
</tr>
<tr>
<td>20th Percentile</td>
<td>23.7</td>
<td>21.9</td>
<td>22.8</td>
<td>20.6</td>
<td>23.4</td>
<td>21.1</td>
<td>23.5</td>
<td>21.3</td>
<td>23.0</td>
</tr>
<tr>
<td>30th Percentile</td>
<td>24.9</td>
<td>23.1</td>
<td>23.8</td>
<td>21.6</td>
<td>24.2</td>
<td>22.2</td>
<td>24.7</td>
<td>22.5</td>
<td>24.1</td>
</tr>
<tr>
<td>40th Percentile</td>
<td>25.6</td>
<td>24.2</td>
<td>24.9</td>
<td>22.5</td>
<td>25.4</td>
<td>23.3</td>
<td>25.6</td>
<td>23.6</td>
<td>24.9</td>
</tr>
<tr>
<td>50th Percentile</td>
<td>26.7</td>
<td>25.5</td>
<td>25.6</td>
<td>23.6</td>
<td>26.3</td>
<td>24.5</td>
<td>26.4</td>
<td>24.8</td>
<td>25.6</td>
</tr>
<tr>
<td>60th Percentile</td>
<td>27.8</td>
<td>26.9</td>
<td>26.4</td>
<td>24.8</td>
<td>27.2</td>
<td>25.6</td>
<td>27.5</td>
<td>26.3</td>
<td>26.7</td>
</tr>
<tr>
<td>70th Percentile</td>
<td>29.0</td>
<td>28.4</td>
<td>27.7</td>
<td>26.4</td>
<td>28.4</td>
<td>27.2</td>
<td>28.7</td>
<td>28.0</td>
<td>27.8</td>
</tr>
<tr>
<td>80th Percentile</td>
<td>30.6</td>
<td>30.7</td>
<td>29.1</td>
<td>28.1</td>
<td>29.9</td>
<td>29.5</td>
<td>30.2</td>
<td>29.9</td>
<td>29.3</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>33.0</td>
<td>33.9</td>
<td>31.3</td>
<td>31.4</td>
<td>32.2</td>
<td>32.7</td>
<td>32.7</td>
<td>33.2</td>
<td>31.6</td>
</tr>
</tbody>
</table>
5.4 Results

Table 5.1 shows summary statistics for males and females by region. Table 5.2 shows BMI deciles by sex and region.

5.4.1 OLS Regression

For females, income decile was negatively correlated with BMI in the Atlantic provinces, Ontario, and the Prairies (see Table 5.3). Higher education in females was negatively associated with BMI in all regions but the Atlantic provinces. Each determinant was statistically significantly associated with BMI in at least one region. Three variables were statistically correlated with BMI in all regions: Age (positive), occasional drinker (positive), and physical activity (negative).

There are some differences between male (Table 5.4) and female results. Where income effects were present for females (Atlantic, Ontario, the Prairies) they were negative. In contrast, income effects for males (notably in Ontario, the Prairies, and BC) were positive. Also unique to males in most regions were the statistically significant negative effects on BMI of being single and having no family doctor. On the other hand, the effects of higher education (negative), physical activity (positive), and age (positive) on BMI in men were similar to those observed in women.

5.4.2 Unconditional Quantile Regression Models

Quantile regression results are presented graphically. Due to the very large volume of data, and for ease of presentation, we present variables that serve as examples of a particular pattern. We furthermore focus on variables for which findings from quantile regression are substantively different from the OLS regression results. Full graphical results are available from the authors.
Table 5.3: OLS estimates for females by region, dependent variable: self-reported BMI.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Atlantic</th>
<th>Quebec</th>
<th>Ontario</th>
<th>Prairies</th>
<th>B.C.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.04*</td>
<td>0.06*</td>
<td>0.09*</td>
<td>0.08*</td>
<td>0.08*</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>2\textsuperscript{nd} decile of Income</td>
<td>-0.45</td>
<td>0.35</td>
<td>-0.14</td>
<td>-0.30</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.31)</td>
<td>(0.22)</td>
<td>(0.30)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>3\textsuperscript{rd} decile of Income</td>
<td>-0.62</td>
<td>-0.11</td>
<td>-0.08</td>
<td>-0.64*</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.42)</td>
<td>(0.27)</td>
<td>(0.21)</td>
<td>(0.27)</td>
<td>(0.32)</td>
</tr>
<tr>
<td>4\textsuperscript{th} decile of Income</td>
<td>-0.81*</td>
<td>0.11</td>
<td>-0.02</td>
<td>0.19</td>
<td>-0.35</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.29)</td>
<td>(0.24)</td>
<td>(0.28)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>5\textsuperscript{th} decile of Income</td>
<td>-1.35*</td>
<td>0.07</td>
<td>-0.38*</td>
<td>-0.01</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.30)</td>
<td>(0.19)</td>
<td>(0.27)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>6\textsuperscript{th} decile of Income</td>
<td>-1.20*</td>
<td>-0.33</td>
<td>-0.33</td>
<td>-0.00</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.27)</td>
<td>(0.20)</td>
<td>(0.29)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>7\textsuperscript{th} decile of Income</td>
<td>-1.08*</td>
<td>0.03</td>
<td>-0.46*</td>
<td>-0.55*</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.40)</td>
<td>(0.20)</td>
<td>(0.26)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>8\textsuperscript{th} decile of Income</td>
<td>-1.53*</td>
<td>-0.08</td>
<td>-0.63*</td>
<td>-0.62*</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.29)</td>
<td>(0.19)</td>
<td>(0.25)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>9\textsuperscript{th} decile of Income</td>
<td>-1.58*</td>
<td>-0.17</td>
<td>-0.56*</td>
<td>-0.63*</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.29)</td>
<td>(0.20)</td>
<td>(0.27)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>10\textsuperscript{th} decile of Income</td>
<td>-1.92*</td>
<td>-0.46</td>
<td>-0.99*</td>
<td>-0.79*</td>
<td>-0.58*</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.29)</td>
<td>(0.20)</td>
<td>(0.26)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Holds post-secondary degree</td>
<td>-0.34</td>
<td>-0.50*</td>
<td>-0.48*</td>
<td>-0.54*</td>
<td>-0.47*</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.15)</td>
<td>(0.11)</td>
<td>(0.13)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Daily Smoker</td>
<td>-1.49*</td>
<td>-0.77*</td>
<td>-0.26*</td>
<td>-0.72*</td>
<td>-0.22</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.17)</td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Occasional Smoker</td>
<td>-0.82*</td>
<td>-0.83*</td>
<td>-0.35*</td>
<td>0.06</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.21)</td>
<td>(0.18)</td>
<td>(0.24)</td>
<td>(0.34)</td>
</tr>
<tr>
<td>Occasional Drinker</td>
<td>1.26*</td>
<td>1.47*</td>
<td>1.14*</td>
<td>1.31*</td>
<td>0.92*</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.18)</td>
<td>(0.11)</td>
<td>(0.15)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Former Drinker</td>
<td>1.75*</td>
<td>1.32*</td>
<td>0.95*</td>
<td>0.80*</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.23)</td>
<td>(0.17)</td>
<td>(0.19)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Never Drinker</td>
<td>0.65</td>
<td>1.55*</td>
<td>-0.92*</td>
<td>-0.42</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.39)</td>
<td>(0.17)</td>
<td>(0.28)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Physically Active</td>
<td>-1.28*</td>
<td>-0.79*</td>
<td>-1.13*</td>
<td>-1.26*</td>
<td>-1.02*</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.18)</td>
<td>(0.12)</td>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Moderately Active</td>
<td>-0.77*</td>
<td>-0.51*</td>
<td>-0.47*</td>
<td>-0.31*</td>
<td>-0.33*</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.15)</td>
<td>(0.10)</td>
<td>(0.14)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Fruits &amp; Veggies Per Day</td>
<td>-0.08*</td>
<td>-0.05*</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>No Family Doctor</td>
<td>-0.36</td>
<td>-0.44*</td>
<td>-0.22</td>
<td>-0.30</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.18)</td>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Category</td>
<td>Estimate</td>
<td>SE</td>
<td>Estimate</td>
<td>SE</td>
<td>Estimate</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------</td>
<td>-----</td>
<td>-----------</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>Single</td>
<td>-0.42*</td>
<td>(0.19)</td>
<td>-0.12</td>
<td>(0.16)</td>
<td>-0.07</td>
</tr>
<tr>
<td>Rural</td>
<td>0.14</td>
<td>(0.17)</td>
<td>0.19</td>
<td>(0.14)</td>
<td>0.49*</td>
</tr>
<tr>
<td>Worked in the past 12 months</td>
<td>-0.15</td>
<td>(0.22)</td>
<td>-0.37*</td>
<td>(0.17)</td>
<td>0.00</td>
</tr>
<tr>
<td>Constant</td>
<td>27.01*</td>
<td>(0.59)</td>
<td>23.06*</td>
<td>(0.41)</td>
<td>22.39*</td>
</tr>
<tr>
<td>Observations</td>
<td>9,440</td>
<td></td>
<td>14,064</td>
<td></td>
<td>31,681</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.07</td>
<td></td>
<td>0.08</td>
<td></td>
<td>0.08</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses, * p<0.05, weighted estimates
Table 5.4: OLS estimates for males by region, dependent variable: self-reported BMI

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Atlantic</th>
<th>Quebec</th>
<th>Ontario</th>
<th>Prairies</th>
<th>B.C.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.02*</td>
<td>0.04*</td>
<td>0.04*</td>
<td>0.03*</td>
<td>0.04*</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.00)</td>
<td>(0.01)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>2\textsuperscript{nd} decile of Income</td>
<td>0.16</td>
<td>0.18</td>
<td>-0.07</td>
<td>0.13</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(0.57)</td>
<td>(0.22)</td>
<td>(0.32)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>3\textsuperscript{rd} decile of Income</td>
<td>0.56</td>
<td>0.08</td>
<td>-0.05</td>
<td>0.30</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.34)</td>
<td>(0.23)</td>
<td>(0.30)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>4\textsuperscript{th} decile of Income</td>
<td>0.41</td>
<td>0.05</td>
<td>0.54*</td>
<td>0.25</td>
<td>0.51*</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.33)</td>
<td>(0.20)</td>
<td>(0.39)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>5\textsuperscript{th} decile of Income</td>
<td>0.28</td>
<td>-0.02</td>
<td>0.45*</td>
<td>0.55*</td>
<td>0.51*</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.32)</td>
<td>(0.20)</td>
<td>(0.26)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>6\textsuperscript{th} decile of Income</td>
<td>0.57</td>
<td>0.13</td>
<td>0.57*</td>
<td>0.61*</td>
<td>0.79*</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.32)</td>
<td>(0.19)</td>
<td>(0.28)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>7\textsuperscript{th} decile of Income</td>
<td>0.19</td>
<td>0.43</td>
<td>0.71*</td>
<td>0.69*</td>
<td>0.76*</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.32)</td>
<td>(0.19)</td>
<td>(0.26)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>8\textsuperscript{th} decile of Income</td>
<td>0.82*</td>
<td>0.13</td>
<td>0.70*</td>
<td>0.75*</td>
<td>0.90*</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.33)</td>
<td>(0.19)</td>
<td>(0.26)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>9\textsuperscript{th} decile of Income</td>
<td>0.59</td>
<td>0.32</td>
<td>0.72*</td>
<td>0.79*</td>
<td>1.09*</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.32)</td>
<td>(0.18)</td>
<td>(0.25)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>10\textsuperscript{th} decile of Income</td>
<td>0.58</td>
<td>0.16</td>
<td>0.88*</td>
<td>0.67*</td>
<td>0.98*</td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.32)</td>
<td>(0.21)</td>
<td>(0.26)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Holds post-secondary degree</td>
<td>-0.29</td>
<td>-0.28*</td>
<td>-0.56*</td>
<td>-0.30*</td>
<td>-0.25*</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.14)</td>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Daily Smoker</td>
<td>-1.48*</td>
<td>-1.07*</td>
<td>-0.72*</td>
<td>-1.07*</td>
<td>-0.84*</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.12)</td>
<td>(0.09)</td>
<td>(0.12)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Occasional Smoker</td>
<td>-0.98*</td>
<td>-0.43</td>
<td>-0.14</td>
<td>-0.50*</td>
<td>-0.24</td>
</tr>
<tr>
<td></td>
<td>(0.29)</td>
<td>(0.28)</td>
<td>(0.14)</td>
<td>(0.21)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Occasional Drinker</td>
<td>0.86*</td>
<td>0.59*</td>
<td>0.05</td>
<td>0.22</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.20)</td>
<td>(0.13)</td>
<td>(0.16)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>Former Drinker</td>
<td>0.44</td>
<td>0.54*</td>
<td>0.05</td>
<td>0.10</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.23)</td>
<td>(0.15)</td>
<td>(0.20)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Never Drinker</td>
<td>-0.38</td>
<td>0.75</td>
<td>-1.38*</td>
<td>-1.30*</td>
<td>-0.87*</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(1.06)</td>
<td>(0.22)</td>
<td>(0.34)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>Physically Active</td>
<td>-0.85*</td>
<td>-0.58*</td>
<td>-0.48*</td>
<td>-0.61*</td>
<td>-0.43*</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.15)</td>
<td>(0.10)</td>
<td>(0.12)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Moderately Active</td>
<td>-0.25</td>
<td>-0.06</td>
<td>-0.16</td>
<td>-0.07</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.14)</td>
<td>(0.09)</td>
<td>(0.15)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>Fruits &amp; Vegs Per Day</td>
<td>0.00</td>
<td>-0.00</td>
<td>-0.03*</td>
<td>-0.08*</td>
<td>-0.08*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>No Family Doctor</td>
<td>-0.46*</td>
<td>-0.51*</td>
<td>-0.46*</td>
<td>-0.22</td>
<td>-0.37*</td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.12)</td>
<td>(0.10)</td>
<td>(0.13)</td>
<td>(0.14)</td>
</tr>
<tr>
<td></td>
<td>Single</td>
<td>Rural</td>
<td>Worked in the past 12 months</td>
<td>Constant</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.94*</td>
<td>-0.75*</td>
<td>-0.61*</td>
<td>27.39*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.18)</td>
<td>(0.13)</td>
<td>(0.09)</td>
<td>25.47*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.53*</td>
<td>0.41*</td>
<td>0.35*</td>
<td>25.55*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.49*</td>
<td>0.10</td>
<td>0.43*</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.09)</td>
<td>(0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.14</td>
<td>0.10</td>
<td>(0.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.28</td>
<td>-0.30</td>
<td>-0.30*</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.28)</td>
<td>(0.14)</td>
<td>(0.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.07</td>
<td>-0.07</td>
<td>-0.07</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.25)</td>
<td>(0.25)</td>
<td>(0.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.35*</td>
<td>0.35*</td>
<td>24.91*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.17)</td>
<td>(0.17)</td>
<td>(0.35)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>8,246</td>
<td>13,207</td>
<td>30,681</td>
<td>16,901</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.06</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses, * p<0.05, weighted estimates
The quantile regression results (Figure 5.2 to Figure 5.4) are interpreted as the effect of a change in the level of the variable on the value of each estimated quantile of BMI. For instance, the coefficient of the 5th decile of income for Atlantic females (Figure 5.2) at the 80th percentile of BMI (quantile 0.8) is -2, which means that the 80th percentile of the Atlantic female BMI distribution for those with an income in the 5th decile is estimated to be 2 kg/m$^2$ lower than the 80th percentile of the BMI distribution for those in the 1st decile (the reference group). In marginal terms, the expected shift in the BMI distribution overall would be 2 kg/m$^2$ at the 80th percentile.

The quantile regression graphs can be summarized succinctly by describing what change they imply in the marginal BMI distribution. For example, if the line depicted on the quantile regression graph is downwards sloping and crosses zero (i.e., the line is positive at the low BMI quantiles and negative at the high BMI quantiles) that implies a narrowing of the BMI distribution associated with that determinant. An upward sloping line that crosses zero implies a widening of the associated BMI distribution. If the depicted line is above or below zero for all quantiles of BMI then the BMI distribution associated with that determinant is shifted right or left of the reference group’s distribution (a “location shift” of the distribution with minimal change in spread). If the depicted line is above zero, goes below zero, then rises back up, that indicates a BMI distribution growing more right-skewed.
Figure 5.2: Quantile regression results for income dummy variables amongst Atlantic females
Figure 5.3: Quantile regression results for income dummy variables amongst Quebec females.
Figure 5.4: Quantile regression results for those females who hold a post-secondary degree by region.
5.4.2.1 Females

The heaviest women in the Atlantic provinces were in the lowest (1st and 2nd decile) income groups. Figure 5.2 shows that for individuals in the higher income groups, the heaviest individuals were lighter (i.e., bigger departure from zero in a negative direction) compared to the reference group in the 1st decile. Income decile displayed a similar pattern in Ontario, the Prairie provinces, and BC with minimum values of -3, -3, and -2 BMI at the highest quantile (not shown). Results for Quebec (Figure 5.3) indicated a different pattern: the BMI distribution grew more right-skewed as income rose, with the median shifting up to 1 BMI lower for the 6th and higher income deciles.

As shown in Figure 5.4, higher education for females narrowed the BMI distribution in the Atlantic provinces by lowering the 70th and higher percentiles of BMI to a minimum value of -2 BMI at the highest quantile. The BMI distribution for the Prairie provinces was also narrowed by higher education, but the narrowing was across the entire distribution to an approximate minimum of -0.75 BMI. The results for Quebec, Ontario, and BC indicate a location shift of the entire BMI distribution to the left by approximately 0.5 BMI.

The variables daily smoker, occasional smoker, being physically active, being moderately physically active, and working in the past 12 months active were associated with a narrowing of the BMI distribution for most regions (not shown). Being single was associated with a widening of the BMI distribution for most regions (not shown). The variables age, occasional drinker, former drinker, and living in a rural area were associated with a location shift to the right for most regions (not shown). The variables never drinker, fruit and vegetable consumption, and not having a family doctor were associated with a location shift to the left for most regions (not shown).
5.4.2.2 Males

Across all regions, higher education narrowed the BMI distribution (Figure 5.5). The impact of this effect increased in magnitude across the BMI distribution for each region, from a minimum of approximately -0.7 in Quebec to a minimum of approximately -1.5 in the Prairie provinces. The effect of higher education on the BMI distribution for males by region was qualitatively similar to females in the corresponding region. The 6th income decile and higher were associated with location shifts to the right for the Atlantic provinces of approximately 0.5 BMI; for Ontario, the Prairies, and BC of approximately 1 BMI. For Quebec the 6th income decile was associated with a location shift right by approximately 0.5 BMI and an increasing degree of right-skew (not shown).

The variables daily smoker, occasional smoker, being physically active, and being moderately physically active, not having a family doctor, and working in the past 12 months were associated with a narrowing of the BMI distribution to some extent for most regions (not shown). The variables occasional drinker and former drinker were associated with a widening of the distribution for most regions (not shown). The variables age and living in a rural area were associated with a location shift to the right for most regions (not shown). The variables never drinker, fruit and vegetable consumption, and being single were associated with a location shift to the left for most regions (not shown).
Figure 5.5: Quantile regression results for those males who hold a post-secondary degree by region.
5.4.3 **Unconditional Quantile Regression Decompositions (Figure 5.6 to Figure 5.9)**

Decomposition results are presented graphically. The figures depict each region decomposed against the Atlantic provinces for males and females separately. The y-axis is BMI (measured in kg/m²), so the “overall” line shows the size of the overall BMI difference between the two regions at each BMI percentile (quantile) in kg/m². The “explained” line shows the share of the overall difference at each percentile that is explained by the differential levels of the determinants measured at the individual level between the two regions. The “unexplained” line shows the remaining share. For example, between the Atlantic provinces and Quebec for females (Figure 5.6, left panel), the overall difference at the 40th quantile is approximately -2, which means the 40th quantile of the BMI distribution for Quebec females is 2 BMI lighter than the 40th quantile for Atlantic females. The explained share of that difference is just under -0.5, which means we would expect the overall difference at the 40th quantile between the two regions to equal approximately 1.5 BMI if the regions had identical levels of the determinants measured at the individual level.

5.4.3.1 Males versus female decomposition trends

The overall difference between any region and the Atlantic provinces was larger for females than males at all quantiles. That means that the difference between the heaviest individuals in the Atlantic provinces and those in other regions was greater for women than men. Quebec (Figure 5.6) and B.C. (Figure 5.9) had the lowest BMIs on average, and that difference persisted across the BMI quantiles. Furthermore, the difference between the regions increased more for females than males as BMI quantile increased.

For both males and females the overall difference between regions (green line) increased more than the explained difference (blue line) across the BMI distribution for most regions, so
the ability of the explanatory variables to account for the difference between the regions became weaker for heavier individuals. That weakening of explanatory power was more pronounced for females.

5.4.3.2 The explained share of overall regional differences for females

Overall, differences in female determinants measured at the individual level were not effective at explaining the cross-regional differences in BMI. For instance, when comparing the Atlantic provinces to Quebec (Figure 5.6, left panel), the unexplained share accounted for almost the entire difference between the regions across the whole BMI distribution. The Atlantic provinces versus Ontario comparison (Figure 5.7, left panel) showed an increase in explanatory power of the determinants as BMI quantile increased since the unexplained share (red line) is fairly constant across the BMI distribution while the explained share (blue line) gets larger. Still, this increase in explained difference made up less than half of the overall difference at any quantile. The Atlantic provinces versus Prairies comparison (Figure 5.8, left panel) showed a fairly even share of the overall difference split between the explained and unexplained shares. Finally, the Atlantic provinces versus BC comparison (Figure 5.9, left panel) showed a similar pattern to the Atlantic provinces versus Quebec: the explained share accounted for a greater share of the difference as BMI percentile increased, but never reached the same amount as the unexplained share.

5.4.3.3 The explained share of overall regional differences for males

Males exhibited the same qualitative patterns as females, the magnitude is just much smaller. The Atlantic provinces versus Quebec comparison (Figure 5.6, right panel) showed that the explanatory variables explained nearly none of the overall difference between regions (similar to women). The Atlantic provinces versus Ontario comparison (Figure 5.7, right panel)
was similar in that the explanatory variables explained a small magnitude of the overall
difference, but the overall difference was smaller so the share explained is larger as a proportion
of the total difference. The Atlantic provinces versus Prairies overall difference (Figure 5.8, right
panel) was quite small, so the explained share was large compared to the overall difference
despite being of tiny magnitude. The Atlantic provinces versus BC comparison (Figure 5.9, right
panel) was fully unexplained until approximately the 60th percentile, after which there was a
slight increase in the ability of the explanatory variables to explain the overall difference
between the two regions.
Figure 5.6: Unconditional quantile decompositions between the Atlantic provinces and Quebec.
Figure 5.7: Unconditional quantile decompositions between the Atlantic provinces and Ontario.
Figure 5.8: Unconditional quantile decompositions between the Atlantic provinces and Prairies.
Figure 5.9: Unconditional quantile decompositions between the Atlantic provinces and B.C.
5.5 Discussion

The OLS estimates presented here are consistent with the literature in terms of magnitude and direction. Income has been shown to be negatively associated with BMI in females and have an unclear relationship with BMI in males, education is negatively associated with BMI, and behavioral variables like physical activity, consuming fruits and vegetables, and smoking have the expected signs (1,2,26,27). The quantile regressions show that in some situations the average effect is a poor descriptor of how a determinant measured at the individual level would be expected to impact the BMI distribution. For example, the effect of higher education for females in the Atlantic provinces was a narrowing of the BMI distribution not captured by the corresponding OLS estimate, which itself was statistically insignificant. The narrowing or widening of the BMI distribution is important from the point of view of population-level interventions for obesity. Information about the change in the BMI distribution can be characterized through quantile regression, making this tool valuable for population-level research on obesity and complementary to OLS regression.

Although several of the variables were statistically significantly associated with BMI in the OLS regression, the decomposition results show that the variation in cross-regional BMI distributions is not attributable to different levels of those determinants. The comparison between the Atlantic provinces and the Prairie provinces for males is the only exception, where the overall difference was very small. The determinants measured at the individual level, thus, are capable of explaining some of the within-region variation in BMI (as shown in the OLS and quantile regressions) but are generally unable to explain population-level differences in BMI (as shown by the decomposition). It has been proposed that the relative importance of determinants measured at the individual level in Canada has weakened over time as obesity prevalence has
increased (15). Past decomposition efforts have proposed different reasons for the weak explanatory power of those determinants; examples include population-level factors such as easy availability and low cost of calorie-dense foods (22), and cultural and social practices that overwhelm determinants measured at the individual level (21). It is also possible that the weak explanatory power reflects unmeasured behavioral variables (20). Thus, our study agrees with the existing decomposition literature: our main finding is that the equalization of identified determinants of BMI cannot be reasonably expected to reduce differences in BMI distributions across regions.

Our study has some limitations. One limitation is that the cross-sectional nature of our data means we cannot attribute changes in BMI outcomes to direct causal changes in the regressor variables. Furthermore, we use behavioural variables in our regression which may be a source of endogeneity in our model (34). However, these variables have been cited as important determinants of BMI in the past and excluding them would lead to unmeasured confounding (35,36), so they were ultimately included in the model. A second limitation is that we used self-reported BMI as our outcome measure, which is considered biased compared to measured BMI when measuring obesity prevalence (37). We attempted to address this issue by applying two correction algorithms (28,29). We found that our results change very little upon correction, which increases our confidence in the use of self-reported BMI. A third limitation is that anything unmeasured in our decomposition analysis would appear to be part of the unexplained share of the difference. One example is ethnicity, which we did not include because it is measured very crudely in the CCHS. However, we did ensure inclusion of a broad cross-section of ultimately statistically significant variables in our within-region models in an effort to reduce this unmeasured variable problem. Furthermore, our analysis showed that the explained share...
does show variation across the BMI distribution, so it is likely that we captured significant individual variation with the variables that were included.

When considered as a whole these results suggest regional variation in Canadian obesity prevalence reflects population-level factors, rather than differences in determinants measured at the individual-level between regions. The unexplained shares of the overall differences identified in this study may be understood as population-level factors, or causes of incidence as described by Rose (16). Our findings demonstrate that the difference in BMI distributions between regions in Canada is predominantly rooted in these population-level factors. Our study thus provides quantitative support for the view that efforts aimed at encouraging individual-level behaviour change are insufficient to address obesity on a population scale (38). Our study furthermore builds on that point to show that individual-level behavioral change efforts, even if they were successful for changing behaviors at the individual level, are ineffective for equalizing BMI distributions across regions.
5.6 References


(13) Williams PT. Evidence that obesity risk factor potencies are weight dependent, a phenomenon that may explain accelerated weight gain in western societies. PLoS One 2011;6(11):e27657.


(29) Dutton DJ, McLaren L. The usefulness of "corrected" body mass index vs. self-reported body mass index: comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data. BMC Public Health 2014;14:430.


Chapter Six: Conclusion

The overall purpose of this dissertation was to assess the importance of population-level determinants of obesity in Canada by quantifying their impact on individuals living in different regions of the country. In this chapter I will review the manuscript-specific findings from Chapters 3, 4, and 5; state overall findings of this dissertation and what it contributes to our current understanding of obesity in Canada; overall strengths and limitations; and outline overall implications for future research.

6.1 Manuscript-specific findings

6.1.1 Explained and Unexplained Regional Variation in Canadian Obesity Prevalence

Chapter 3 showed that the differences in average BMI across regions are explained by individual-level covariates, to varying degrees, depending on the sex and region under comparison. For females, in the comparisons between the Atlantic provinces and both Quebec and Ontario, the contribution of the coefficients were statistically significant. The most important variables included in the model for these differences were income and fruits and vegetables consumed per day. This is important because it highlighted the potential magnitude of population-level determinants, how they could be different for males and females, and what observed variables could be contributing to this phenomenon. This manuscript generated questions for later research. In particular, the focus on the mean BMI difference between regions could be expanded by considering the rest of the BMI distribution. Faster increases in obesity prevalence in the Atlantic provinces (1,2) means that the BMI distribution in the Atlantic provinces might be growing differently from the other regions, and cross-regional distributions for males and females could be different. This idea was the catalyst for Chapter 5, but sufficient
measured height and weight data did not exist at the national level to answer this question. In Chapter 4 I addressed the issue of measured height and weight.

6.1.2 How useful is “corrected” body mass index vs. self-reported body mass index? Comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data.

The manuscript presented in Chapter 4 had three major findings with respect to its three objectives, and those findings jointly had the same implication: automatically resorting to corrected BMI measures does not guarantee estimates closer to those produced by measured BMI. These findings have implications for obesity research in general: automatically using a corrected BMI value in place of self-reported BMI does not result in less biased regression coefficients. Relevant to this dissertation, however, was a clear recommendation on how to proceed in Chapter 5. Since we were particularly interested in modeling the BMI distribution this manuscript recommends using a correction, in particular the weight-only correction since we modeled males and females separately. Although it became clear that self-reported and corrected BMI produced virtually identical results in the Chapter 5 analysis, and I eventually used self-reported BMI, this manuscript allowed for the analysis in Chapter 5 to proceed.

6.1.3 How important are determinants of obesity measured at the individual-level for explaining geographic variation in body mass index distributions? Evidence from Canada.

The manuscript presented in Chapter 5 found that almost all cross-regional differences in the BMI distributions under study had large portions unexplained by determinants of BMI measured at the individual level. In arriving at that finding, this manuscript also demonstrated that OLS regression estimates can mask interesting nuances in the relationship between determinants of obesity measured at the individual level and BMI. Similarly, the decompositions across the BMI distribution in Chapter 5 exposed interesting nuances compared to the
decompositions of the average provided in Chapter 3. For instance, the importance of the population-level determinants grew over the BMI distribution for most comparisons, especially in females. I concluded that cross-regional differences in BMI are influenced by population-level determinants to the point where equalization of determinants measured at the individual level would not be expected to equalize BMI distributions across regions.

6.1.4 Overall conclusions from this dissertation

These findings represent some important contributions to Canadian obesity research: namely, the potential value of importing methods from other disciplines into population health and reflections on best modeling practices.

Only after decomposing the BMI distributions was the potential magnitude of the impact of population-level determinants made clear. The fact that population-level determinants are drivers of increasing obesity prevalence (3,4) is supported by the large share of the cross-regional BMI difference left unexplained by the model. These findings are a result unique to the tools I utilized in my analysis. The current literature using the CCHS dataset is reliant on statistical tools that are more classically rooted in an epidemiological framework: Of the literature utilizing regression to analyze CCHS data, approximately 85% of studies relied on logistic regression (5). Logistic regression is used as an extension of stratified analysis, one way of considering confounding in epidemiology (6). The high prevalence of logistic regression in the literature is indicative of a strong epidemiological practice in Canada. Chapters 3 and 5 contain findings that the current emphasis on logistic regression would not otherwise have exposed. The consideration of the entire BMI distribution, even in the absence of decomposition, is important for understanding obesity in Canada and the tools presented in this document allow that.
Econometric research on obesity using the tools presented in this document is rare, especially in Canada, but past literature has focused on the effect of specific individual variables when using a decomposition method. For example, one social dimension of interest when studying obesity in the United States is ethnicity, and researchers have decomposed BMI across black and white males and females (7,8). The conclusions reached in both of these studies stated the existence of unexplained differences between blacks and whites were associated with one or a combination of potential factors ranging from distal to more proximal: culture, genetics, or unobserved (unmeasured) confounders. Outside of the United States, decompositions between ethnic groups (e.g., Chinese and Malays) (9), between countries (e.g., Italy and Spain) (10), across sexes (11), or over time (12) indicate that there can be large unexplained differences over the groups under consideration. The reasons cited for the existence of these unexplained differences include cultural differences (9,10); increased availability of food compared to what was traditionally available, which the different groups differentially take advantage of (11); and unmeasured behavioural differences (9,10,12). The one Canadian manuscript studying provincial decompositions of obesity, which focuses on the concentration of obesity through use of a concentration index, notes that groups where obesity is concentrated are potential targets for intervention (13). The population-level determinants of obesity are rarely mentioned explicitly as such, but they are sometimes presented in conjunction with individual preferences.

As one example, Etilé finds increased levels of education in France lowered BMI, but simultaneously the returns to education became less favourable for those with lower levels of education (12). He describes this phenomenon as “more-educated women are now relatively more efficient than in the past at controlling their body weight. This finding suggests that technical changes in food production, food supply or on-the-job physical exercise have been
partially biased in favour of the well-off.” (12) Food production and food supply are population-level drivers of the obesity epidemic (3,4) but they are framed as devices that allow or disallow women to control their body weight based on their education level. My work is unique in that it explicitly acknowledges the population-level determinants of obesity as what is represented by the unexplained share of the decompositions. This document contains the only Canadian evidence using this approach.

The use of these tools requires appropriate data. Chapter 4 showed that when modeling obesity in Canada at the national level the practice of substituting corrected BMI for self-reported BMI is not a guarantee of more valid results. While this could be interpreted as discouraging for obesity research, it serves to highlight the value that increased measurement of height and weight by Statistics Canada provides (as is being done in newer surveys, like the Canadian Health Measures Survey (14)). Further, only when BMI is used as a regressor is the use of corrected BMI of no clear benefit due to measurement error (15), so corrected BMI still has value. Appropriate justification for correction would take into account the research question. Prior to the publication of Chapter 4 the recommendations for the use of correction equations in Canadian population-level data was less nuanced.

The overall purpose of this dissertation was to assess the importance of population-level determinants of obesity in Canada by quantifying their impact on individuals living in different regions of the country. This was accomplished by establishing the importance of population-level determinants of cross-regional differences in both average BMI and the BMI distribution, and by ensuring appropriate BMI measures were used. The implications for Canadian obesity research outlined above lead to further research questions, which I will cover below.
6.2 Strengths and Limitations

I relied on high quality Statistics Canada data to address the objectives of each manuscript in this dissertation. The amount of effort that goes into data gathering, cleaning, and validation of the CCHS (16) would be infeasible for any one researcher. I am confident that the statistics produced in these manuscripts reflect data that are representative of the Canadian population. Further, through the RDC initiative (17,18), I was able to access master file versions of the CCHS which allowed for analysis of disaggregated and uncensored versions of the variables available in the PUMF. The high quality of these data allows for confidence in the estimates of association presented in this document.

I have been able to make a unique contribution to the obesity literature through the manuscripts in this dissertation because of the statistical tools I applied. I used two statistical tools that are underutilized in population health: Blinder-Oaxaca decomposition (19,20) and unconditional quantile regression (21-23). Both of these tools have been rigorously scrutinized in the econometric literature and are an excellent fit into a population health framework for studying obesity. The increased use of these tools in population health research will lead to an increase in the ability to analyze both distributions and population-level determinants of health.

The manuscripts in this dissertation also have limitations. The cross-sectional nature of the data means I cannot draw causal implications from the majority of the findings in this paper (24). The use of the term “determinant” implies causation, which is itself explicitly stated by the ecological model of obesity, yet causation is not established by this data format. So while the ecological model of obesity would suggest that population-level determinants cause obesity, the data itself precludes making such conclusions outright. However, conclusions regarding the importance of population-level determinants of obesity need not be causal as obesity is the result
of a complex system (25): identifying the magnitude of the unexplained share of the difference between regions in a cross-sectional framework draws attention to their possible importance.

A second limitation of the manuscripts presented in this dissertation is the likely presence of endogeneity in the estimated regression models in Chapters 3 and 5. Endogeneity refers to a situation where the regressors in a model are possibly determined jointly with, or possibly caused by, the regressand. For example, BMI is associated with physical activity (26). However, if we use observational data and model the effect of physical activity on BMI we are not certain that the estimated effect represented by the coefficient is unbiased: people who enjoy physical activity might also enjoy being healthy (BMI and physical activity jointly determined), and people who do not enjoy exercise might find it more difficult (high BMI causes low physical activity). A consequence of endogeneity is biased estimates, but leaving important determinants of BMI out of the models would also result in biased estimates through omitted variables bias (27). Ultimately such variables, like physical activity, were included since they are identified as important determinants of BMI.

Another limitation is that measured BMI is strictly superior to self-reported BMI but Chapter 5 relied on self-reported BMI. If measured BMI were available I would have used it, but the CCHS gathers self-reported BMI in each cycle and measured BMI is comparatively rarer. Chapter 4 established that the measurement error introduced from self-reported BMI can cause biased estimates in particular types of models and corrected BMI might not be a solution (i.e., when BMI is a predictor variable of a health condition), but our models do not use BMI as a regressor. We used self-reported BMI and two different forms of corrected BMI to estimate the models in Chapter 5 and they were qualitatively similar. Thus, the potential for bias in this situation was low.
6.3 Overall implications of the findings: future research

The findings presented in this dissertation jointly imply that further research is necessary to identify why the regional differences in Canadian BMI are so persistent. Now that we understand how large a role the population-level determinants play it is important to identify what they are specifically. For example, it is unknown how important cultural differences are in determining cross-regional differences in BMI as opposed to other population-level determinants, such as food prices. It is possible that the high overall BMI in the Atlantic provinces is driven by a population-level determinant not considered in the ecological model framework. Further, it is unknown whether policy targeting population-level determinants of obesity would interact with more proximal variables in an expected way in the face of co-existing unique population-level determinants. These questions will best be answered using a suite of methods including, but not limited to, application of sophisticated statistical tools to population health survey data.
6.4 References


(8) Sen B. Using the oaxaca-blinder decomposition as an empirical tool to analyze racial disparities in obesity. Obesity (Silver Spring) 2014 Apr 15.


Chapter Seven: Works Cited


Department of Community Health Sciences, Faculty of Medicine - University of Calgary. *Guidelines for Manuscript-Based Thesis* 2014.


Fortin, Nicole M. "Nicole M. Fortin's personal website, data and software section." <http://faculty.arts.ubc.ca/nfortin/datahead.html>.


Sen, B. "Using the Oaxaca-Blinder Decomposition as an Empirical Tool to Analyze Racial Disparities in Obesity." *Obesity (Silver Spring, Md.)* 22.7 (2014): 1750 -5.


---. "Canadian Community Health Survey - Annual Component (CCHS)." 2013b. 


---. "The RDC program: Application process and guidelines." 2014d. 


Tjepkema, Michael. "Adult Obesity in Canada: Measured Height and Weight" Statistics Canada (2005): Cat. No. 82-620-MWE.


Williams, P. T. "Evidence that Obesity Risk Factor Potencies are Weight Dependent, a Phenomenon that may Explain Accelerated Weight Gain in Western Societies." *PloS one* 6.11 (2011): e27657.


Appendices

Appendix A: Supplemental material to Chapter 3

When Chapter 3 was published online, material explaining the Blinder-Oaxaca decomposition with an interaction term was made available as a supplement to the manuscript. That material is reproduced with permission here:

The expected value of the body mass index of region $i$, $\overline{\text{BMI}}_i$, is predicted by the average (or expected) value of the vector of variables in a regression specific to that region, $\overline{X}_i$, multiplied by the vector of estimated regression coefficients for that region $\hat{\beta}_i$. For example, with regions A and B, the average BMI equations, as estimated by OLS regression, look like:

$$\overline{\text{BMI}}_A = \overline{X}_A' \hat{\beta}_A$$
$$\overline{\text{BMI}}_B = \overline{X}_B' \hat{\beta}_B$$

Thus, the difference between the two regions is:

$$\overline{\text{BMI}}_A - \overline{\text{BMI}}_B = \overline{X}_A' \hat{\beta}_A - \overline{X}_B' \hat{\beta}_B$$

We can add and subtract the same term three times from this equation in the form of other $\overline{X}_i$ and $\hat{\beta}_i$ combinations to get:

$$\overline{\text{BMI}}_A - \overline{\text{BMI}}_B = \overline{X}_A' \hat{\beta}_A - \overline{X}_B' \hat{\beta}_B + \overline{X}_A' \hat{\beta}_B - \overline{X}_B' \hat{\beta}_A + \overline{X}_A' \hat{\beta}_A - \overline{X}_A' \hat{\beta}_B + \overline{X}_B' \hat{\beta}_B - \overline{X}_B' \hat{\beta}_B$$

Some rearranging yields the useful equation:

$$\overline{\text{BMI}}_A - \overline{\text{BMI}}_B = (\overline{X}_A' - \overline{X}_B') \hat{\beta}_B + (\hat{\beta}_A - \hat{\beta}_B) \overline{X}_B' + (\overline{X}_A' - \overline{X}_B')(\hat{\beta}_A - \hat{\beta}_B)$$

Thus we see the three terms that make up the three-fold Blinder-Oaxaca decomposition used in this paper. The first summand of the equation:
\[(X'_A - X'_B)\hat{\beta}_B\]

represents the share of the difference in BMI between the regions attributable to the difference in the levels of the variables between the regions. This particular piece indicates how much region B’s estimated BMI would change if region B were endowed with region A’s average level of the \(X_i\) variables.

The second summand:

\[(\hat{\beta}_A - \hat{\beta}_B)X'_B\]

represents the share of the difference in BMI between regions attributable to the difference in the returns to the variables between the regions. This piece indicates how much region B’s estimated BMI would change if region B had the same coefficients as region A.

The third summand:

\[(X'_A - X'_B)(\beta_A - \beta_B)\]

represents the share of the difference in BMI between the two regions that still exists after removing the first two pieces. This is known as the “interaction” term and is the difference in the coefficients and the difference in the endowments considered together. As a mathematical term, it is the “leftover” difference and, as such, we interpret this as being the part of the difference that affects BMI without interacting with the variables we included in the model.

Appendix B: Ethics approval for this dissertation

2011-05-20

Dr. Lindsay McLaren
University of Calgary
Community Health Sciences
3330 Hospital Dr. NW
Calgary, Alberta
T2N 4N1

Dear Dr. McLaren:

RE: An investigation into the population-level determinants of obesity in Canada - Policy, demographics, and socio-cultural factors

Ethics ID: E-23704

Student: Mr. Daniel Dutton

The above-noted proposal including the Protocol and Committee Sign Off has been submitted for Board review and found to be ethically acceptable.

Please note that this approval is subject to the following conditions:
1) access to personal identifiable health information was not requested in this submission;
2) a copy of the informed consent form must have been given to each research subject, if required for this study;
3) a Progress Report must be submitted by May 30, 2012, containing the following information:
   i) the number of subjects recruited;
   ii) a description of any protocol modification;
   iii) any unusual and/or severe complications, adverse events or unanticipated problems involving risks to subjects or others;
   iv) withdrawal of subjects from the research, or complaints about the research;
   v) a summary of any recent literature, finding, or other relevant information, especially information about risks associated with the research;
   vi) a copy of the current informed consent form;
   vii) the expected date of termination of this project.
4) A Final Report must be submitted at the termination of the project.

Please note that you have been named as the principal collaborator on this study because students are not permitted to serve as principal investigators. Please accept the Board's best wishes for success in your research.

Yours sincerely,

[Signature]

Deputy Chair, Conjoint Health Research Ethics Board

SP/eng

cc: Dr. C. Doig (information) Ms. Sharon Van Oot, Research Services – Main Campus Mr. Daniel Dutton (Student)
Office of Information & Privacy Commissioner

CREATING THE FUTURE OF HEALTH An innovative medical school committed to excellence and leadership in education, research and service to society.
Appendix C: Permissions for Chapter 3 from journal and co-author

The following permissions allow for the reproduction of the material in Chapter 3 within this document.
RE: Permission request: Dissertation manuscript 1

Lindsay McLaren

Sent: 30 May 2014 17:45
To: Daniel James Dutton

Permission granted

------------------------------------------
Lindsay McLaren PhD
Associate Professor and Alberta Innovates - Health Solutions Population Health Investigator
Scientific Co-Director, Population Health & Inequities Research Centre, Institute for Public Health
President-Elect, Alberta Public Health Association
Department of Community Health Sciences, University of Calgary

TRW3, 3280 Hospital Dr. NW, Calgary, Alberta, Canada T2N 4Z9
Tel: (403) 210-9424, Email: lmclaren@ucalgary.ca

From: Daniel James Dutton
Sent: Wednesday, May 28, 2014 3:18 PM
To: Lindsay McLaren
Subject: Permission request: Dissertation manuscript 1

Hello Lindsay,

I am preparing my dissertation document and require your permission to include our publication listed below in the document:

*Explained and unexplained regional variation in Canadian obesity prevalence, Obesity (Silver Spring); 19(7), 2011 July.*

In addition to allowing me to include this manuscript in my dissertation, you will also be agreeing to allow the manuscript to be submitted to the University of Calgary Thesis Vault (http://theses.ucalgary.ca) according to the Faculty of Graduate Studies electronic thesis and dissertations program. I will also be allowing Library and Archives Canada permission to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of my dissertation, among other things, by signing a Non-Exclusive License.

The title of my dissertation is: Population-level determinants of obesity in Canada: application of econometric techniques

Please respond with your permission no later than June 4, 2014.

Thank you,
Daniel

Daniel J. Dutton, PhD(e)
Research Associate, Health Policy
This is a License Agreement between Daniel J Dutton ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number 3146810265355
License date May 12, 2013
Licensed content publisher John Wiley and Sons
Licensed content publication Obesity
Licensed content title Explained and Unexplained Regional Variation in Canadian Obesity Prevalence
Licensed copyright line 2011 North American Association for the Study of Obesity (NAASO)
Licensed content author Daniel J. Dutton, Lindsay McLaren
Licensed content date Sep 10, 2012
Start page 1460
End page 1468
Type of use Dissertation/Thesis
Requestor type Author of this Wiley article
Format Print and electronic
Portion Full article
Will you be translating? No
Total 0.00 USD

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively WILEY). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC"s Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at

https://s100.copyright.com/CustomerAdmin/PDF.jsp?ref=bcf0d41e-f904-4433-8466-7d02a0f6f600
Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the journal/book, and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT.
AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

https://s100.copyright.com/CustomerAdminPLF.jsp?fn=bc9d4e-e-fb04-4423-8445-fb09b00d

172
14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC’s Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state’s conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

Wiley publishes Open Access articles in both its Wiley Open Access Journals program [http://www.wileyopenaccess.com/view/index.html] and as Online Open articles in its subscription journals. The majority of Wiley Open Access Journals have adopted the Creative Commons Attribution License (CC BY) which permits the unrestricted use, distribution, reproduction, adaptation and commercial exploitation of the article in any medium. No permission is required to use the article in this way provided that the article is properly cited and other license terms are observed. A small number of Wiley Open Access journals have retained the Creative Commons Attribution Non Commercial License (CC BY-NC), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Online Open articles Authors selecting Online Open are, unless particular exceptions apply, offered a choice of Creative Commons licenses. They may therefore select from the CC BY, the CC BY-NC and the Attribution-NoDerivatives (CC BY-NC-ND). The CC BY-NC-ND is more restrictive than the CC BY-NC as it does not permit adaptations or modifications without rights holder consent.

Wiley Open Access articles are protected by copyright and are posted to repositories and websites in accordance with the terms of the applicable Creative Commons license referenced on the article. At the time of deposit, Wiley Open Access articles include all changes made during peer review, copyediting, and publishing. Repositories and websites that host the article are responsible for incorporating any publisher-supplied amendments or retractions issued subsequently. Wiley Open Access articles are also available without charge on Wiley’s publishing platform, Wiley Online Library or any successor sites.

Conditions applicable to all Wiley Open Access articles:

https://s100.copyright.com/CustomerAdmin/PDF.jsp?rel=bc6d3e1-6b04-44f2-84f4-7d9280f5cb00

173
• The authors' moral rights must not be compromised. These rights include the right of "paternity" (also known as "attribution" - the right for the author to be identified as such) and "integrity" (the right for the author not to have the work altered in such a way that the author's reputation or integrity may be damaged).

• Where content in the article is identified as belonging to a third party, it is the obligation of the user to ensure that any reuse complies with the copyright policies of the owner of that content.

• If article content is copied, downloaded or otherwise reused for research and other purposes as permitted, a link to the appropriate bibliographic citation (authors, journal, article title, volume, issue, page numbers, DOI and the link to the definitive published version on Wiley Online Library) should be maintained. Copyright notices and disclaimers must not be deleted.

Creative Commons licenses are copyright licenses and do not confer any other rights, including but not limited to trademark or patent rights.

• Any translations, for which a prior translation agreement with Wiley has not been agreed, must prominently display the statement: "This is an unofficial translation of an article that appeared in a Wiley publication. The publisher has not endorsed this translation."

Conditions applicable to non-commercial licenses (CC BY-NC and CC BY-NC-ND)

For non-commercial and non-promotional purposes individual non-commercial users may access, download, copy, display and redistribute to colleagues Wiley Open Access articles. In addition, articles adopting the CC BY-NC may be adapted, translated, and text- and data-mined subject to the conditions above.

Use by commercial "for-profit" organizations

Use of non-commercial Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Commercial purposes include:

• Copying or downloading of articles, or linking to such articles for further redistribution, sale or licensing;

• Copying, downloading or posting by a site or service that incorporates advertising with such content;

• The inclusion or incorporation of article content in other works or services (other than normal quotations with an appropriate citation) that is then available for sale or licensing, for a fee (for example, a compilation produced for marketing purposes, inclusion in a sales pack)

https://s100.copyright.com/CustomerAdminPLF.jsp?rep=bc86916-4d04-4b23-8466-7d3280ff7b00
Use of article content (other than normal quotations with appropriate citation) by for-profit organizations for promotional purposes

Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;

Use for the purposes of monetary reward by means of sale, resale, license, loan, transfer or other form of commercial exploitation such as marketing products

Print reprints of Wiley Open Access articles can be purchased from: corporatesales@wiley.com

The modification or adaptation for any purpose of an article referencing the CC BY-NC-ND License requires consent which can be requested from RightsLink@wiley.com.

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

v1.8

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNS01019427. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

https://s100.copyright.com/CustomerAdminPL.jsp?ref=bc006d46-5465-100-4455-8466-7d028007b00
Appendix D: Permissions for Chapter 4 from journal and co-author

To publish in the journal *BMC Public Health* authors must agree to the Creative Commons Attribution License v4.0 (available here: http://creativecommons.org/licenses/by/4.0/), which allows for copying and sharing of the published material for any reason as long as it is properly attributed. Thus no explicit licenses are granted to reproduce material from *BMC Public Health*. 


RE: Permission request: Dissertation manuscript 2
Lindsay McLaren

Sent: 30 May 2014 17:45
To: Daniel James Dutton

Permission granted

Lindsay McLaren PhD
Associate Professor and Alberta Innovates - Health Solutions Population Health Investigator
Scientific Co-Director, Population Health & Inequities Research Centre, Institute for Public Health
President-Elect, Alberta Public Health Association
Department of Community Health Sciences, University of Calgary

TRW3, 3280 Hospital Dr. NW, Calgary, Alberta, Canada T2N 4Z6
Tel: (403) 210-5424, Email: lmclaren@ucalgary.ca

From: Daniel James Dutton
Sent: Wednesday, May 28, 2014 3:18 PM
To: Lindsay McLaren
Subject: Permission request: Dissertation manuscript 2

Hello Lindsay,

I am preparing my dissertation document and require your permission to include our publication listed below in the document:

The usefulness of “corrected” body mass index vs. self-reported body mass index: comparing the population distributions, sensitivity, specificity, and predictive utility of three correction equations using Canadian population-based data. *BMJ Public Health*, 14:430, 2014.

In addition to allowing me to include this manuscript in my dissertation, you will also be agreeing to allow the manuscript to be submitted to the University of Calgary Thesis Vault ([http://theses.ucalgary.ca](http://theses.ucalgary.ca)) according to the Faculty of Graduate Studies electronic thesis and dissertations program. I will also be allowing Library and Archives Canada permission to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of my dissertation, among other things, by signing a Non-Exclusive License.

The title of my dissertation is: Population-level determinants of obesity in Canada: application of econometric techniques

Please respond with your permission no later than June 4, 2014.

Thank you,
Daniel

Daniel J. Dutton, PhD(o)
Appendix E: Permission for Chapter 5 from co-author

The following permission allows for the reproduction of the material in Chapter 5 in this document.

**RE: Permission request: Dissertation manuscript 3**

Lindsay McLaren

Sent: 30 May 2014 17:45
To: Daniel James Dutton

Permission granted

----------------------------------
Lindsay McLaren PhD
Associate Professor and Alberta Innovates - Health Solutions Population Health Investigator
Scientific Co-Director, Population Health & Inequities Research Centre, Institute for Public Health
President-Elect, Alberta Public Health Association
Department of Community Health Sciences, University of Calgary

TRW3, 3280 Hospital Dr. NW, Calgary, Alberta, Canada T2N 4Z5
Tel: (403) 210-9424, Email: lmclaren@ucalgary.ca

From: Daniel James Dutton
Sent: Wednesday, May 28, 2014 3:17 PM
To: Lindsay McLaren

Subject: Permission request: Dissertation manuscript 3

Hello Lindsay,

I am preparing my dissertation document and require your permission to include our manuscript intended for submission listed below in the document:

**How important are determinants of obesity measured at the individual-level for explaining geographic variation in body mass index distributions? Evidence from Canada.**

In addition to allowing me to include this manuscript in my dissertation, you will also be agreeing to allow the manuscript to be submitted to the University of Calgary Thesis Vault [http://theses.ucalgary.ca] according to the Faculty of Graduate Studies electronic thesis and dissertations program. I will also be allowing Library and Archives Canada permission to reproduce, communicate to the public on the Internet, loan, distribute or sell copies of my dissertation, among other things, by signing a Non-Exclusive License.

The title of my dissertation is: Population-level determinants of obesity in Canada: application of econometric techniques

Please respond with your permission no later than June 4, 2014.

Thank you,

Daniel

Daniel J. Dutton, PhD(c)
Research Associate, Health Policy
Appendix F: Permission to reproduce Figure 2.1 from journal

ELSEVIER LICENSE
TERMS AND CONDITIONS
Jun 02, 2014

This is a License Agreement between Daniel J Dutton ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Elsevier Limited</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Boulevard, Langford Lane</td>
<td>Killington, Oxford, OX5 1GB, UK</td>
</tr>
<tr>
<td>Registered Company Number</td>
<td>1982084</td>
</tr>
<tr>
<td>Customer name</td>
<td>Daniel J Dutton</td>
</tr>
<tr>
<td>Customer address</td>
<td>Faculty of Medicine, U Calgary</td>
</tr>
<tr>
<td></td>
<td>Calgary, AB T2N 4Z6</td>
</tr>
<tr>
<td>License number</td>
<td>3396170604024</td>
</tr>
<tr>
<td>License date</td>
<td>May 25, 2014</td>
</tr>
<tr>
<td>Licensed content publisher</td>
<td>Elsevier</td>
</tr>
<tr>
<td>Licensed content title</td>
<td>Producing health, consuming health care</td>
</tr>
<tr>
<td>Licensed content author</td>
<td>Robert G. Evans, Gregory L. Stoddart</td>
</tr>
<tr>
<td>Licensed content date</td>
<td>1990</td>
</tr>
<tr>
<td>Licensed content volume</td>
<td>31</td>
</tr>
<tr>
<td>Licensed content issue</td>
<td>12</td>
</tr>
<tr>
<td>Number of pages</td>
<td>17</td>
</tr>
<tr>
<td>Start Page</td>
<td>1347</td>
</tr>
<tr>
<td>End Page</td>
<td>1363</td>
</tr>
<tr>
<td>Type of Use</td>
<td>reuse in a thesis/dissertation</td>
</tr>
<tr>
<td>Intended publisher of new work</td>
<td>other</td>
</tr>
<tr>
<td>Portion</td>
<td>figures/tables/illustrations</td>
</tr>
<tr>
<td>Number of figures/tables/illustrations</td>
<td>1</td>
</tr>
<tr>
<td>Format</td>
<td>both print and electronic</td>
</tr>
</tbody>
</table>

179
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at [http://myaccount.rightlink.com](http://myaccount.rightlink.com)).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

   "Reprinted from [Publication title], Vol [volume number], Author(s), Title of article / title of chapter, Pages No. Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modyifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC’s Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC’s Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC’s Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher’s written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher’s behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC’s Billing and Payment terms and conditions. These terms and conditions, together with CCC’s Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC’s Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to
LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the website must maintain the copyright information line on the bottom of each image. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/sxxxx or the Elsevier homepage for books at http://www.elsevier.com. This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the website must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The website must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

For journal authors: the following clauses are applicable in addition to the above: Permission granted is limited to the author accepted manuscript version* of your paper.

*Accepted Author Manuscript (AAM) Definition: An accepted author manuscript (AAM) is the author’s version of the manuscript of an article that has been accepted for publication and which may include any author-incorporated changes suggested through the processes of submission, processing, peer review, and editor-author communications. AAMs do not include other publisher value-added contributions such as copy-editing, formatting, technical enhancements and (if relevant) pagination.

You are not allowed to download and post the published journal article (whether PDF or HTML proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/sxxxx. As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier’s online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article’s Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be
included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Posting to a repository: Authors may post their AAM immediately to their employer’s institutional repository for internal use only and may make their manuscript publicly available after the journal-specific embargo period has ended.

Please also refer to Elsevier’s Article Posting Policy for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution’s repository.

20. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Elsevier Open Access Terms and Conditions

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals.

Authors publishing in an Open Access journal or who choose to make their article Open Access in an Elsevier subscription journal select one of the following Creative Commons user licenses, which define how a reader may reuse their work: Creative Commons Attribution License (CC BY), Creative Commons Attribution – Non Commercial - ShareAlike (CC BY NC SA) and Creative Commons Attribution – Non Commercial – No Derivatives (CC BY NC ND)

Terms & Conditions applicable to all Elsevier Open Access articles:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author’s honour or reputation.

The author(s) must be appropriately credited.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:
Appendix G: Permission to reproduce Figure 2.2 from journal
INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking “accept” in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. (“CCC”), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

“Reprinted from Publication title, Vol./edition number, Author(s). Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER].” Also Lancet special credit - “Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier.”

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)
6. If the permission fee for the requested use of our material is waived in this instance, please be
advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of
(i) the license details provided by you and accepted in the course of this licensing transaction, (ii)
these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon
issuance of the license at the end of the licensing process for the transaction, provided that you have
disclosed complete and accurate details of your proposed use, no license is finally effective unless
and until full payment is received from you (either by publisher or by CCC) as provided in CCC's
Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any
license preliminarily granted shall be deemed automatically revoked and shall be void as if never
granted. Further, in the event that you breach any of these terms and conditions or any of CCC's
Billing and Payment terms and conditions, the license is automatically revoked and shall be void as
if never granted. Use of materials as described in a revoked license, as well as any use of the
materials beyond the scope of an unrevoked license, may constitute copyright infringement and
publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed
material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their
respective officers, directors, employees and agents, from and against any and all claims arising out
of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned,
or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed
by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase
order, acknowledgment, check endorsement or other writing prepared by you, which terms are
inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions.
These terms and conditions, together with CCC's Billing and Payment terms and conditions (which
are incorporated herein), comprise the entire agreement between you and publisher (and CCC)
concerning this licensing transaction. In the event of any conflict between your obligations
established by these terms and conditions and those established by CCC's Billing and Payment
terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this
License at their sole discretion, for any reason or no reason, with a full refund payable to you.
Notice of such denial will be made using the contact information provided by you. Failure to
receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright
Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a
result of a denial of your permission request, other than a refund of the amount(s) paid by you to

https://e103-copyright.com/CustomerAdmin/FUF.jsp?ref=a037914f4462-4e47-a091-366df6a18662
3/7

186
LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. Posting licensed content on any Website: The following terms and conditions apply as follows. Licensing material from an Elsevier journal. All content posted to the web site must maintain the copyright information line on the bottom of each image. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxxx or the Elsevier homepage for books at http://www.elsevier.com. Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com. All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable. The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

For journal authors: the following clauses are applicable in addition to the above: Permission granted is limited to the author accepted manuscript version* of your paper.

* Accepted Author Manuscript (AAM) Definition: An accepted author manuscript (AAM) is the author’s version of the manuscript of an article that has been accepted for publication and which may include any author-incorporated changes suggested through the processes of submission, peer review, and editor-author communications. AAMs do not include other publisher value-added contributions such as copy-editing, formatting, technical enhancements and (if relevant) pagination.

You are not allowed to download and post the published journal article (whether PDF or HTML, proof or final version) nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxxx. As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier’s online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article’s Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be...
included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

Posting to a repository: Authors may post their AAM immediately to their employer’s institutional repository for internal use only and may make their manuscript publicly available after the journal-specific embargo period has ended.

Please also refer to Elsevier’s Article Posting Policy for further information.

18. For book authors the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution’s repository.

20. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Elsevier Open Access Terms and Conditions

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals.

Authors publishing in an Open Access journal or who choose to make their article Open Access in an Elsevier subscription journal select one of the following Creative Commons user licenses, which define how a reader may reuse their work: Creative Commons Attribution License (CC BY), Creative Commons Attribution – Non Commercial – ShareAlike (CC BY NC SA) and Creative Commons Attribution – Non Commercial – No Derivatives (CC BY NC ND)

Terms & Conditions applicable to all Elsevier Open Access articles:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author’s honour or reputation.

The author(s) must be appropriately credited.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license: