SHOULD LARGE URBAN CENTRES DECIDE HOW BEST TO USE HEALTH CARE SERVICES?

by

Suzanne Kathleen Clarke

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
February 2014

For my mum

and

those who unwaveringly support their loved
ones through the ebbs and flows of this life that we lead

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

We assessed how estimates of need-expected inpatient hospital use differ depending on whether need-expected use was estimated for a population of all Canadians, Canadian health regions, or a subpopulation of higher income Canadians, who likely had minimal healthcare access problems. Data came from the 2009/2010 Canadian Community Health Survey, a national cross-sectional survey. Using zero-inflated negative binomial regression, we modeled inpatient hospital use separately based on the three aforementioned choices of population. We adjusted for demographic, health behaviour, health status, socioeconomic, and health care supply factors. We then estimated need-expected inpatient hospital use and compared the estimates across individuals and by income and province. The three choices of the population we used in this study had similar results. Our estimates of the average need-expected use by province or income group were not sensitive to the choice of population used to estimate need-expected use.

(143 words)

LIST OF ABBREVIATIONS USED

BRR	Balanced repeated replication
CCHS	Canadian Community Health Survey
PUMF	Public use microdata file
ZINB	Zero inflated negative binomial regression

ACKNOWLEDGEMENTS

First and foremost, I would like to sincerely thank Dr. Yukiko Asada and Dr. George Kephart. As my co-supervisors, committee members, teachers and academic support system, they have given so much of both themselves to this project. I will forever be grateful for the expansive knowledge they shared with me, the skills they helped me to master and their infinite patience and support throughout this process.

I would also like to thank the Department of Community Health and Epidemiology staff and faculty, in particular Dr. Kathleen MacPherson and Tina Bowdridge, for the knowledge, experience, support, and at times humour, that they shared with me throughout my time in their department.

Finally, I would like to acknowledge and thank my family and friends who have unconditionally believed, supported and loved me for who I am. I could not have finished this project without their unwavering confidence in both this process and myself. Not to mention phone calls, comedic relief and cups of tea. Finishing this project has come to symbolize much more to me than the completion of this master's alone. That said, I am so grateful that my mother, the Gilberts, and most of my loved ones are still here to see it happen.

The past five years have taught me so much and indeed been the most humbling experience. A quote my father often used regarding one's perspectives and experiences in life was, "I cried because I had no shoes, 'til I met a man who had no feet." How true it is ~ not to mention, a little dirt on your feet means you've actually walked somewhere, learned along the way and really, lived this life.

CHAPTER 1 INTRODUCTION

1.1 INTRODUCTION

Researchers widely advocate for the use of a need-based approach when assessing some of the key health care policy issues in various populations and jurisdictions around the world (1-6). Areas in health care systems where a need-based approach is incorporated include resource allocation and forecasting, the assessment of access to health care, human resource planning and health system performance indicators (2-5,7-10). As researchers encourage the use of a need-based approach, it is important to clarify various methodological choices and examine their validity.

The need-based approach, regardless of the context of its application, addresses two methodological questions. The first question is how need should be measured. To answer this question, researchers decide what the best indicators of need for health care use are, how to measure them and how to account for the influence of non-need indicators on health care use. *Need indicators* are those factors that either affect individuals' health care utilization for reasons of need or are proxies for an unmeasured need (e.g., age is routinely used as a proxy for need for health care associated with mortality and morbidity) (2). *Non-need indicators* are those factors unrelated to need, such as measures of access to services (e.g., income or region of residence). This first methodological question addressing how to measure need has received considerable attention in the literature (11).

The second question for the need-based approach is how a *standard* level of health care resource use given need should be estimated (12). Researchers apply this standard to estimate *need-expected use*, which can serve as the basis for need-based allocation or equitable health care use (13). There are different methods for estimating a standard (12). If clinical guidelines exist for determining need for health care, they can serve as an evidence-informed standard for expected resource use given levels of need. Similarly, if accepted benchmarks are available for the various levels of need, one can establish standards using those benchmarks.

However, in a majority of scenarios where the need-based approach is used, a clinical guideline or benchmark is not available due to lack of data or evidence. In such cases, conventionally, researchers have used average health care use based on need in a population to define the standard.

Concerning this second question of the standard, there has been extensive attention in the literature to modeling strategies, including treatment for endogeneity of supply and adjustment for unmeasured variables (14). The choice of population represented in the standard, however, has not received due attention in the literature. Researchers can use different populations to establish the standard. For example, when using a national survey to develop a standard, researchers typically estimate the average health care use based on need in the whole survey population and routinely reflect the target population by weighting. This practice results in metropolitan areas with larger populations having a stronger representation in defining a standard (e.g., in Canada, primarily Toronto, Vancouver and Montreal, where 35 percent of the population reside (15)). Stronger representation of urban populations may be appropriate but also means that rural and remote areas have minimal impact on deciding how health care resources should be used given need. Additional choices of population represented may include the population with equal regional representation or subpopulations believed to face minimal barriers to access health care services, such as high socioeconomic groups.

With little attention to the choice of population represented in the standard, its implications in estimates of need-expected use are unclear. The goal of this research was to compare the estimates of the standard level of health care use given need using different populations represented. Or, illustratively put, we asked: Should large urban centres decide how best to use health care services?

Using a nationally representative survey, the Canadian Community Health Survey, this thesis expanded on previous methodological understandings of the need-based approach and the implications of population choice on estimates of need-expected health care use. This thesis is organized into five chapters. The second chapter presents a background of the need-based approach and examples of different choices of population. The third chapter describes the methods used in this study. The fourth chapter reports the results. The fifth chapter concludes with discussion.

CHAPTER 2 BACKGROUND AND OBJECTIVES

2.1 NEED-BASED APPROACH

Researchers and policy makers widely use the need-based approach when modeling health care use. Though this approach can vary considerably in detail, it consists of two common steps.[1] This section outlines these steps.

Step 1: Modeling observed health care utilization

The need-based approach usually begins with modeling health care utilization. Ideally, researchers estimate this first step using individual-level data (such as survey data). Often, however, due to unavailability of data, they substitute with ecological or small area level data (16). This step models current observed health care utilization using the form:

$$y_i = \alpha X_i + \beta Z_i + \varepsilon_i \qquad (1)$$

where y_i is the observed health care utilization for individual i; α is the coefficient for the vector of the need indicators X_i; β is the coefficient for the vector of the non-need indicators, and the error term is represented by ε_i. Interactions between need indicators (e.g., age and sex) can also be included (2).

The decision to include certain indicators in a model depends on the question a study is addressing, the data availability, and other indicators included in the model. Disagreement exists regarding how to categorize variables as need (X_i) or non-need (Z_i) indicators. Need indicators can include demographic characteristics (e.g., age and sex), health behaviours (e.g., smoking) and health status (e.g., presence of chronic conditions and self-perceived health). Non-need indicators can include variables for geography and access or supply of health services. Some variables can be either a need or non-need indicator depending on the other variables included in

[1] This approach became common after its use in the United Kingdom by the Resource Allocation Working Party (RAWP) in the late 1970s (3).

the model. For example, education is either a need indicator (a proxy for unmeasured need), if the model does not include direct need indicators, such as health status, or a non-need indicator (a proxy for access to the health system), if the model includes direct need indicators. The purpose of this proposed research is not to address variable categorization, as the literature already does so (11). Rather, this study assesses the impact of the choice of the population used in the regression models (in the form of equation [1]) on the estimated need-expected use (Step 2 below).

Step 2: Estimating need-expected health care use

Using the predicted value from equation (1) by purging the influence of non-need indicators (Z_i) (usually set to their means), the second step estimates need-expected health care use for each individual in the data. Need-expected use for individual *i,* $\hat{y}_i^*$, can be estimated as:

$$\hat{y}_i^* = \hat{\alpha} X_i + \hat{\beta} \overline{Z_i} \qquad (2)$$

where non-need indicators are held constant as represented by $\overline{Z_i}$, and the coefficients $\hat{\alpha}$ and $\hat{\beta}$ are estimations from the need and non-need indicators, respectively. Equation (2) represents a *standard level of health care use based on need*, that is, the amount of health care that should be used given the level of need predicted by the model. For planning or resource allocation, the $\hat{y}_i^*$ are often aggregated and averaged over a region or population group to obtain the average need-expected use in the population.

2.2 CHOICE OF POPULATION IN ESTIMATING A STANDARD

One can choose different populations to estimate a standard (i.e., equation [2]). A common choice of population is a whole population, counting each individual in the population equally (1,2). We thus estimate a standard health care use based on the average health care use given need across all individuals in the population. Methodologically, giving equal weight to each individual in the population using survey data means estimating the model using the sample weights for the survey. Both Hurley et al. and Asada and Kephart opted for this choice of the population. Hurley et al. linked administrative data to the Ontario portion of 1997/97 National Population Health Survey and estimated need-expected homecare use (1). The standard in this study reflects the average homecare use based on need among Ontario citizens. To do this, Hurley et al. weighted the analysis to reflect the actual population of Ontario (1). This weighting was necessary to address the oversampling of smaller regions that occurred through the survey's sample design. Urban areas (such as Toronto's census metropolitan area, population of 5,941,500) are thus under-represented in the data, and without the use of sample weights, urban areas would have been underrepresented in estimating the standard, compared to rural regions (such as Georgian Bay, population of 2,124), which were over-represented in the survey (15,17).

Asada and Kephart used a cross-sectional survey, the Canadian Community Health Survey (CCHS), to estimate standard health care use among Canadians across the ten provinces (2). Similar to Hurley et al., they used the sample weights to account for the CCHS's sample design that oversampled smaller regions so that their estimates represent the Canadian population (1,2).

A second population choice when estimating a standard is a population of areas rather than a population of individuals. This means using the whole population and counting each region in the population equally. Carr-Hill et al. partially used such a method (3,16). To estimate a standard hospital in-patient care service use in England, they used small area (electoral population ward) data rather than individual-level data (3,16). They obtained a population that was equally representative of all the individuals within it (similarly to the previous two examples) by weighting the sample proportionally to the total population of the ward. Alternatively, they could have given equal consideration to all regions rather than individuals. If this had been done, the standard estimated would have had equal representation of each region.

The choice between the population of individuals and the population of areas is akin to the difference in population representation between the House of Commons and the Senate in the Parliament of Canada. When electing Members of Parliament to the House of Commons, the electoral system aims to count each eligible voter equally. Appointment of Senators, on the other hand, aims at equal regional representation, thus, Maritime Provinces' populations are overrepresented in the Senate when compared with the House of Commons, for example. Both represent Canadian populations, but in different ways.

Health care is generally organized on a regional level. For example, the largest tertiary hospitals are located in large metropolitan areas, and the largest metropolitan areas have the greatest concentration of specialists. Outside of the largest metropolitan areas, larger centers are often the home of regional hospitals, with smaller hospitals located in more remote regions. This is likely to be reflected as differences in access, supply and patterns of use between regions. Giving equal weight to individuals means that the organization of healthcare in the largest centers will play a disproportionate weight in defining the standard for need-based models. Giving equal weight to regions will gives more equal weight to different ways of organizing and delivering healthcare.

A third population choice when estimating a standard is to use one subpopulation, which can serve as a benchmark. For example, Schofield et al.'s work in Australia estimated a standard based on the use of general practitioner services in urban populations, with the rationale that urban populations faced minimal access problems (4). They therefore estimated a standard that reflected optimal general practitioner service availability. Methodologically, they used data from urban areas to estimate the standard and applied that standard to all areas in the country to estimate need-expected health care use.

Choice of population represented is one of the many important issues of the need-based approach. Diversity in the need-based approach can also come from the availability of data sources (e.g., survey data, administrative data), the selection of variables in the model, and the choice of the modeling techniques (1,2,3,4). Unlike these other issues, few studies exist that investigate the effect of the choice of population represented on estimates of need-expected use.

2.3 OBJECTIVES

We assessed how estimates of need-expected inpatient hospital use differed by the choice of population represented. We examined the following three choices of population:

1. Canadians: Modeling average health care use based on need among Canadians, counting each individual in the population equally.
2. Regions: Modeling average health care use based on need among regions in Canada, counting each region in the population equally (i.e., each region will be given equal weight)
3. Subpopulation of Canadians who likely have minimum access problems: Modeling average health care use based on need among the above average income groups (from here on referred to as the high income group).

CHAPTER 3 METHODS

3.1 OVERVIEW

Using a large Canadian survey, we examined the effects of three population choices on estimates of need-expected inpatient hospital use (hereafter referred to as need-expected use). To model inpatient hospital use and develop standards, we used zero-inflated negative binomial regression models. The three choices of population we used were: Canadians, regions and a subpopulation of high-income earners presumed to have minimal access issues. We compared estimates of need-expected use based on these three population choices, first, across individuals, and second, by income and province.

3.2 DATA AND MEASURES

3.2.1 Data

The data for this study came from the 2009/10 Canadian Community Health Survey (CCHS) public use microdata file (PUMF) conducted by Statistics Canada (18). This publically available survey was suited for this study because it provided a large national sample and included information on a variety of need and non-need indicators (19). This survey annually collected self-reported information about health status, health determinants and health care utilization from Canadians twelve years of age or older living in all territories and provinces in selected private dwelling households. Excluded from the sampling frame were Canadians who are full-time members of the Canadian Forces or live in institutions, on Indian Reserves, in certain remote areas or on Crown Lands (2% of the Canadian population aged 12 years or older). The CCHS used a multi-stage stratified cluster sample design. The sampling design enabled the combination of two years of survey data to produce a single population cross-section, thus increasing sample size. For this study, we used the two-year file for 2009/2010. The response rate was 72.3% over the two-year period, resulting in the sample size of 124,188 (18,19).

For this study, we excluded individuals under age 20 (n=13,476) and individuals residing in the territories (n=3082), as modeling their health care use required considerations beyond the scope of this research. We also excluded individuals spending 31 or more days in a hospital, a nursing home or a convalescent home (n=607) as it was likely that these individuals were using these facilities as a substitute for long term care services rather than acute care. In addition, we excluded individuals who did not answer questions regarding the inpatient hospital use (n=193) as we used these questions to create the dependent variable in our analysis. After these exclusions, the sample size available for our analysis was 106,363. Once the descriptive statistics were calculated, those who responded to the self-perceived health question (n=122) or the education question (n=2,836) with "not applicable", "don't know", "refusal" or "not stated", were case-wise deleted. This left 103,417 individuals in the analytic sample.

3.2.2 Variables

Dependent variable: As a measure of health care use, we used inpatient hospital use, more specifically, self-reported number of nights as a patient in a hospital, nursing home or convalescent home in the past 12 months (20). Inpatient hospital use was used because on a per capita level, larger variation in inter-provincial differences have been observed in use and intensity of use when compared with other types of utilization such as general practitioners or specialists (11).

Independent variables: Table 1 lists all independent variables we used for this study. The need indicators included demographic variables (age and sex) and health status (self-perceived health and presence of 12 separate chronic conditions). For those who responded "not applicable", "don't know", "refusal" and "not stated" for questions regarding the presence of chronic conditions (about 1.7%), we regarded them as having no chronic conditions.

The variables included as potential non-need indicators were income and education, as well as regions as a proxy for health care system factors not measured in the

CCHS. If higher income and education were associated with higher use, after adjustment for direct measures of need such as health status, it would indicate inequitable access. Then, we would treat income and education as non-need variables. Conversely, if higher income and education were associated with lower use, after adjustment, the effects of these variables would likely reflect unmeasured need, and we would treat them as need variables. Dummy variables for 98 CCHS health regions were used as a proxy for system variables reflecting access and supply (14). Each dummy variable roughly corresponds to the different administrative health regions that organize and fund healthcare in each province. For income, we created a separate "missing" category for those responding "not applicable", "don't know", "refusal" and "not stated" (17.2%), as these individuals represented a large proportion of the sample and unlikely were random (21).

3.3 ANALYSIS

We first modeled inpatient hospital use to develop standard resource use by need as per equation (1) described in Chapter 2. We developed three separate standards using the three choices of population. Based on each standard, we estimated need-expected use, as per equation (2). We then examined these need-expected use estimates across individuals and compared the average need-expected use by income and province.

3.3.1 Modeling Inpatient Hospital Use to Develop the Standards

The distribution of inpatient hospital use was highly skewed with a large proportion of individuals having zero overnight stays. To handle this distribution, we used zero-inflated negative binomial regression models. These models employed a two-part estimation, with a logistic "inflation" component to estimate the dichotomous binary yes/no overnight hospital use, and a negative binomial component to estimate the number of overnight stays for those who stayed overnight (22,23). The logistic "inflation" component identified zeros, meaning that it estimates the likelihood that an individual would have zero overnight stays (no inpatient nights) compared to those not having zero overnight stays (one or more inpatient nights). The resulting estimations were negative when there was a likelihood of stay and positive when there was a likelihood of no stays.

Because a large proportion of the sample had a dependent variable registering as zero hospital stays, ordinary count models such as Poisson or negative binomial regression were not appropriate (24). We utilized the vuong test to assess whether the zero-inflated negative binomial regression model was more appropriate to use than the standard negative binomial regression model. This test was estimated using unweighted data. A significant vuong test suggested that the zero-inflated model was more appropriate than a negative binomial regression (24). Also, because over-dispersion of the variable was likely, meaning that the variance and mean were different, the option to use zero-inflated Poisson regression was not suitable (24). More specifically, the zero-inflated component of this model addressed the over-dispersion by splitting the dependent variable into zero or non- zero categories. A statistical test was used to assess if the zero-inflated negative binomial regression model was indeed over-dispersed and therefore the preferred method to the zero-inflated Poisson regression model (24). This test used a significant likelihood test, where if alpha equaled zero, the zero-inflated model was more appropriate. Furthermore, since the independent variables affected the two aspects (use and intensity) of the dependent variable in different ways, use of zero-inflated negative binomial regression models was beneficial (24). In sum, by using this method for modeling inpatient hospital use and intensity, we addressed the over-dispersed count outcomes observed in the sample participants while accounting for large numbers of participants with zero hospital stays.

Different indicators and interactions could be included in each part of the model. For this study, the main effects and the interaction between age and sex were included in both portions of the three models, while the interaction between age and self-perceived health was included in only the count component. The age and sex interaction was included to account for differences observed between the sexes in health care utilization due to child bearing and age differences in the onset of chronic conditions. The interaction between age and self-perceived health was included to capture the subjectivity individuals have around their health status. As people age, they tend to become more optimistic about assessment of their health relative to the severity of their symptoms and therefore, rate their health status more highly (25).

All regression models included the same independent variables so that differences in need-expected use obtained from the regression models reflect the choice of populations represented, not the choice of the independent variables. We included all independent variables in Table 1. However, the choice of variables included in the models could affect the results. Specifically, estimates of need-expected use might be more sensitive to the choice of population when fewer need indicators are used. To assess this, we repeated the analysis using only age, sex, self-perceived health, and age-sex interactions.

3.3.2 Specifying the Choice of Populations

We estimated three separate models to develop three separate standards that reflect the following three choices of population.

Choice 1

All Canadians, with each individual represented equally, is the first population choice. To implement, we used data for all subjects in our sample, and to represent all Canadians equally, we applied CCHS survey weights in the estimation. These survey weights reflect the number of people in the Canadian population that each subject represents (19).

Choice 2

For the second population choice, we again used all Canadians, but with each health region represented equally. We weighted the analysis differently from the first choice by giving all the CCHS health regions equal weight. The design of the CCHS facilitates estimation of 117 health regions, which roughly correspond to the sub-provincial regions in use for planning and health care delivery (26). There were 98 CCHS health regions in the public use version, with 97 of them being located in the provinces. Each region contains over 69,999 people, and in certain cases, combines or excludes small sub-provincial health regions (26).

We rescaled the CCHS sample weights to give each region equal representation in the analysis. Letting w_{ir} represent the CCHS sample weight for individual *i* in region *r*, the rescaled sample weights, w_{ir}* are:

$$w_{ir}{*}= w_{ir} / \sum_{\forall i \in r} w_{ir} \qquad (3)$$

Note that $\sum_{\forall i \in r} w_{ir}{*} = 1$, that is, weights within each region sum up to one, and each region has equal weight.

Choice 3

The third population choice is a subpopulation of Canadians who likely have minimal access problems. We defined this subpopulation as those sample participants with above average income (the upper middle or high income groups, referred to as high income). To estimate a model to develop the standard using this choice of the population, we only used data from this subpopulation and gave equal representation of each individual in this population by using the CCHS sample weights.

3.3.3 Estimating Need-Expected Inpatient Hospital Use

We generated three estimates of need-expected use for all CCHS individuals ($\hat{y}_i^*$), using the three regression models estimated from the three choices of population as described above as per equation (2). We held the non-need indicators at their mean.

3.3.4 Comparing Need-Expected Inpatient Hospital Use

3.3.4.1 Bland-Altman Plots

Bland-Altman plots can be used to succinctly illustrate how estimates from two methods differ when compared to the average of the estimates. We used Bland-Altman plots to compare absolute and proportional differences for individuals estimated need-expected use ($\hat{y}_i^*$) between pairs of models (see Figure 2) (27). For all plots, the horizontal axis of the Bland-Altman plots is the mean of the three estimates of need-expected use for each individual in our data as generated by the three choices of population ($_{average}\,\hat{y}_i^*$):

$$_{\mathrm{average}}\,\hat{\mathrm{y}}_{\mathrm{i}}^* = \frac{_{\mathrm{C}}\hat{\mathrm{y}}_{\mathrm{i}}^* + {_{\mathrm{R}}\hat{\mathrm{y}}_{\mathrm{i}}^*} + {_{\mathrm{S}}\hat{\mathrm{y}}_{\mathrm{i}}^*}}{3} \qquad (4)$$

where $_C\hat{y}_i^*$ is the estimate for individual *i* using Canadians, $_{\mathrm{R}}\hat{\mathrm{y}}_{\mathrm{i}}^*$ is the estimate for individual *i* using equal regional representation and $_S\hat{y}_i^*$ is the estimate for individual *i* using the high income subpopulation. Each data point showed results for a respondent in the sample. The vertical axis plotted either absolute or proportionate (i.e., relative) differences between two estimates for each individual, for example, $_C\hat{y}_i^* - {_R\hat{y}_i^*}\ or\ ({_C\hat{y}_i^*} - {_R\hat{y}_i^*})/_{\mathrm{average}}\,\hat{\mathrm{y}}_{\mathrm{i}}^*$, respectively. The plots also include horizontal lines showing the “limits of agreement”, specifying the boundaries within which 95% of the differences lie. These assist with identifying how large the differences are, and with identifying systematic patterns of outliers.

For both the absolute and proportionate differences, we compared the estimates in the following three ways: (1) comparing Canadians against regions, (2) comparing regions against the high income subpopulation, and (3) comparing Canadians against the high income subpopulation. These plots helped us examine the difference between the three estimates of need-expected use, the magnitude (absolute and proportionate) of these differences and any trends that might exist by level of need.

3.3.4.2 Comparing Average Need-Expected Inpatient Hospital Use by Province and Income Group

We computed the average need-expected use for each province using the estimates representing each choice of population, and compared them to assess the impact of choice of population. In a similar manner, we computed the average need-expected use for each income group. We used six income groups, the five income groups and one missing groups described in Table 1. We used CCHS sample weights to calculate the averages.

We used Stata 12 for all analyses (28). Because of lack of information about the sample design in the public use version of the CCHS, our estimates of variance did not account for the complex sample design. Statistics Canada recommends the use of the balanced repeated replication (BRR) methods with BRR replicate weights (this process is referred to as "bootstrapping with bootstrap weights" by Statistics Canada). These replication weights are only available for the master file of the CCHS and not available for the public file, which we used for this study. The alternative method of linearized/robust variance estimation is not feasible because data identifying the sampling units and strata are not provided (29,30). To compensate for this limitation we used a conservative level of statistical significance of $p<0.01$. It is unlikely that this limitation affected main results of our analysis, as the sample size was large and estimates had high precision.

CHAPTER 4 RESULTS

4.1 CHARACTERISTICS OF THE SAMPLE WITH INPATIENT OVERNIGHT STAYS

Table 2 presents the descriptive statistics for the variables included in this study. Almost 5 % of the sample population visited the hospital as an inpatient overnight at least once with a majority staying for a brief time (Figure 2). Estimates are consistent with expectations. Individuals who were older, had lower self-perceived health, low income and one or more chronic conditions were overall most likely to stay overnight in the hospital. Individuals who reported having any particular chronic condition always had a higher mean number of stays than those without one, with the exception of individuals who experienced migraines. When compared with the other provinces, residents of Newfoundland had the lowest proportion of people who stayed hospital overnight (8.36%), while also having the highest mean number of nights spent among those who stayed hospital overnight (6.43).

4.2 THE THREE MODELS FOR INPATIENT OVERNIGHT STAYS

Zero-inflated negative binomial regression fitted significantly better than negative binomial regression and zero-inflated Poisson regression. We confirmed the appropriateness of using zero inflated negative binomial regression by using two different statistical tests on unweighted models. To confirm zero-inflated negative binomial regression was more appropriate than negative binomial regression, we used the Vuong test (Canadian and region models: $v=27.27$, $P<0.001$, High income model: $v=15.45$, $P<0.000$). Similarly, we confirmed that zero-inflated negative binomial regression provided a better fit than zero-inflated Poisson regression (Canadian, region and high income models: $P<0.001$)

Overall, parameter estimates from the three models with difference choices of population were similar, with the high income model being the most different (Table 3). The statistical significance of independent variables was very similar between the models. However, a few variables significant in the Canadian and region models were not significant in the high income model, which was estimated on a much smaller sample. In all three models, age, self-perceived health and region were statistically significant across the count and inflate components. The interaction between age and sex was statistically significant for the inflate component and not significant for the count component. The interaction between age and self-perceived health was only significant for the count component. The number of chronic conditions an individual had was significantly associated with the number of nights they spent as an inpatient in all three models, but was not associated with the likelihood of them having zero overnight stays in two of the three models. The number of statistically significant chronic conditions was largest in the region model. We observed some variance in the effect size of estimates between the models. Most notably, the effect sizes of age, self-perceived health, cancer and heart disease in the count component of the high income model were larger than those of the Canadian and region models.

Estimates for income and education suggested that they were not acting as non-need indicators. Education was not statistically significant in any of the three models. Income showed a statistically significant negative association with the number of overnight hospital stays, which indicated that income picked up residual unmeasured need and behaved as a need indicator.

4.3 NEED-EXPECTED INPATIENT HOSPITAL USE ACROSS CHOICES OF POPULATION STANDARDS

4.3.1 Differences in Need-Expected Use at the Individual Level

Differences in need-expected use on an individual level between the three choices of population are shown both absolutely and proportionally in the Bland-Altman plot Figures 3 - 8. The mean differences between estimates using regions and Canadians as the choice of population were close to zero (Figure 3-4). The mean absolute difference between the two choices of population estimates (Regions-Canadians) was - 0.01 of a visit (95% limits of agreement: +/- 0.242). The corresponding proportional mean difference was slightly above zero (0.8%) with 95% of the population falling within +/- 26% of this mean difference. The Canadians and regions estimates for need-expected use tended to be smaller when compared with estimates from the high income subpopulation, having larger magnitude negative mean differences. There was also less agreement, with broader 95% limits of agreement (Figure 5-8).

Overall, the absolute and proportionate differences between estimates of the need-expected use showed opposite patterns in where variance was observed. When the average need-expected use across the three models was small, absolute differences between estimates were on average small and their variances were also small (Figures 3,5,7), where as proportionate differences were on average large and their variances were also large (Figures 4,6,8).

4.3.2 Average Need-Expected Inpatient Hospital Use by Province and Income Group

The three models using different choices of population, either weighted or unweighted, yielded similar average estimates of need-expected use (Table 4).

The choice of population resulted in small differences in the estimates of average need-expected use by province or income group, as shown in Figure 9 and 10. Overall, differences were larger in the income comparison than the provincial comparison. Differences in the estimates of average need-expected use based on the three choices of population were the most pronounced among the low income group. Across the ten provinces, differences due to the three choices of population were largest for Quebec.

The choice of variables included in the models might have affected our results. To assess this a sensitivity analysis was run with a simple model that included only age, sex, self-perceived health, and age-sex interactions, as the choice of variables predictors. Study results and conclusions were similar between the full and simple models (Table 5 - 8). Contrary to what was observed in the full models, the simple models with region as the choice of population had the highest average need-expected use within each province and income group (Figure 9 - 12). This led to greater variation in estimates of the average need-expected inpatient hospital use within each province and most income groups across the choices of population.

CHAPTER 5 DISCUSSION AND CONCLUSIONS

This study examined how estimates of need-expected inpatient hospital use differed by the choice of population represented. There is little attention to, or even awareness of, this issue in the need-based literature, despite potentially important implications.

Our results suggest that the need-based models overall were not greatly affected by the choice of population represented. The three choices of the population we used in this study had similar results. All Bland-Altman plots indicated that absolute and proportionate differences in the estimates of need-expected use based on the three choices of population were small; although, estimating need-expected use using the high income subpopulation resulted in higher estimates of need-expected use than the other two choice of population. Our estimates of the average need-expected use by province or income group also did not show great variation across the choices of population. To sum, the choice of population did not make a significant difference overall.

The choice of population becomes an important consideration in need-based resource allocation if it alters winners and losers in allocation decisions. Our results suggest this is not the case. While choice of population did result in some small differences in estimates of average need-expected hospital inpatient use across provinces, and some changes in ranking, these differences were not substantial. In fact, the estimated differences in meeting expected use resulting from choice of population were small relative to differences in need expected use by province. Though within the low income group we observed a noticeable difference by choice of population, this was to some degree expected because our high income model set the standard level of resource use using those individuals with lowest need.

Our choice of variables included in the models might have affected our results. Though estimates were similar when a smaller set of predictor variables were used in the models (Table 5 - 8), larger variations across the choices of population were observed when simpler models were used (Figure 9-12). While fundamental conclusions about income ranking were not altered by the choice of variables included, the conclusions about provincial rankings were affected. This variation in rankings is most obvious for the high income choice of population where only Ontario's ranking (7^{th}) is not altered between the simple and full models and Prince Edward Island's ranking drops from third to eighth (Table 5 and 7). These observed variations are likely explained by the exclusion of region in the simple model. Differences such as these highlight the possibility for resource allocation ramifications when variable inclusions are oversimplified.

The small impact of the choice of population on need-expected inpatient use may not be generalizable to other types of healthcare utilization. General practitioner and specialists services may still be influenced by the choice of population. When compared with inpatient hospital use, these types of healthcare utilization are affected more by practice variation across different populations and contexts. Observed patterns of use based on need may therefore be significantly affected by the choice of population. Future research should explore this possibility.

APPENDIX I: TABLES AND FIGURES

TABLES

Table 1: Independent variables in the models of inpatient hospital use

Independent Variable	Coding of Variables
Need Indicators	
Age [1]	1 = 20-29, 2 = 30-39, 3 = 40-49, 4 = 50-59, 5 = 60-69, 6 = 70-79, 7 = 80+
Sex [1]	0 = Male, 1 = Female
Self-perceived health	1 = excellent, 2 = very good, 3 = good, 4 = fair or poor [4]
Arthritis [2]	0 = no, 1 = yes [5]
Asthma or chronic respiratory disease[3]	0 = no, 1 = yes [5]
Back problems [2]	0 = no, 1 = yes [5]
Bowel disorder, Crohn's disease or colitis	0 = no, 1 = yes [5]
Cancer	0 = no, 1 = yes [5]
Diabetes	0 = no, 1 = yes [5]
Heart disease or stroke	0 = no, 1 = yes [5]
High blood pressure	0 = no, 1 = yes [5]
Mental illness	0 = no, 1 = yes (anxiety or mood disorder) [5]
Migraine	0 = no, 1 = yes [5]
Stomach or intestinal ulcer	0 = no, 1 = yes [5]
Urinary incontinence	0 = no, 1 = yes [5]
Number of major chronic conditions	0 = none, 1 = 1, 2 = 2, 3 = 3, 4 = 4+
Potential Non-need Indicators	
Health region	1 – 97 = Corresponding to each health region or health region grouping [6]
Province of residence	1 = NL, 2 = PEI, 3= NS, 4 = NB, 5 = QC, 6 = ON, 7 = MN, 8 = SK, 9 = AB, 10 = BC [7]
Education	1 = < secondary education, 2 = secondary completed, 3 = some post-secondary, 4 = post-secondary graduate
Household income	1 = low, 2 = lower middle, 3 = middle, 4 = upper middle, 5 = high, 6=missing [8]

(1) We also included the interaction between age and sex
(2) Arthritis and back pain exclude fibromyalgia
(3) Diagnosed with chronic bronchitis, emphysema or chronic obstructive pulmonary disease
(4) "Good" includes responses of "good", "not applicable" and "don't know"
(5) "Yes" suggests diagnosis with the condition. "No" includes responses of "no", "not stated", "not applicable", "don't know" and "refused"
(6) The Canadian Community Health Survey (CCHS) master file has 117 health regions. The CCHS public use file has 98 health regions or health region groupings that contain over 69,999 people. As some of the original 117 small sub-provincial health regions were not large enough, they were either combined into health region groupings or excluded all together from the public use file.
(7) NL: Newfoundland and Labrador; PEI: Prince Edward Island; NS: Nova Scotia; NB: New Brunswick; QC: Quebec; ON: Ontario; MN: Manitoba; SK: Saskatchewan; AB: Alberta; BC: British Columbia
(8) The financial amount of each household income level is: Low = 0 to $19,999, Low middle=$20,000 to $39,999, Middle = $40,000 to $59,999, Upper middle = $60,000 to $79,999 High = $80,000 or more, Missing = unknown

Table 2: Description of independent variables included in the study

Independent Variable	Study Population	Proportion of population with overnight inpatient hospital stays		Number of nights spent for the subjects who answered yes				
	(N=106,363)	No (%) (N=96,206)	Yes (%) (N=10,157)	Mean	50th percentile	75th percentile	90th percentile	95th percentile
Need Indicators								
Age								
20-29	14,377	90.06	9.94	3.22	2	3	6	10
30-39	15,857	91.26	8.74	3.53	2	4	7	12
40-49	15,754	93.76	6.24	4.81	3	6	13	20
50-59	20,275	92.73	7.27	5.15	3	7	13	18
60-69	19,322	90.72	9.28	5.58	4	7	14	18
70-79	13,071	86.77	13.23	6.45	4	8	14	21
80+	7,707	82.3	17.7	7.83	5	11	20	25
Sex								
Male	47,437	91.87	8.13	5.76	4	7	14	20
Female	58,926	89.31	10.69	4.99	3	6	12	18
Self-perceived health								
Excellent	20,093	94.37	5.63	3.26	2	4	6	10
Very good	38,353	93.4	6.6	3.87	2	4	8	14
Good	31,925	90.19	9.81	5.06	3	6	12	16
Fair or poor	15,992	78.98	21.02	7.23	5	10	18	23
Missing	122	87.7	12.3	5.27	5	7	10	14
Arthritis								
No	81,409	91.72	8.28	4.78	3	5	12	17
Yes	24,954	86.29	13.71	6.27	4	8	14	21

Asthma or chronic respiratory disease								
No	94,581	91.16	8.84	5.07	3	6	13	18
Yes	11,782	84.76	15.24	6.28	4	8	15	21
Back problems								
No	81,986	91.3	8.7	5.08	3	6	13	19
Yes	24,377	87.59	12.41	5.75	3	7	14	21
Bowel disorder, Crohn's disease or colitis								
No	100,192	90.93	9.07	5.14	3	6	13	19
Yes	6,171	82.6	17.4	6.45	4	8	15	21
Cancer								
No	96,410	91.26	8.74	4.95	3	6	12	18
Yes	9,953	82.59	17.41	6.88	4	9	16	22
Diabetes								
No	96,965	91.09	8.91	4.98	3	6	13	18
Yes	9,398	83.81	16.19	9.97	5	10	16	22
Heart disease or stroke								
No	97,060	91.8	8.2	4.71	3	5	11.5	15.5
Yes	9,303	76.38	23.62	7.34	5	10	18	22
High blood pressure								
No	100,362	90.66	9.34	5.21	3	6	14	20
Yes	6,001	86.87	13.13	6.14	4	8	14	21
Mental illness								
No	94,652	91.22	8.78	5.11	3	6	13	18
Yes	11,711	84.26	15.74	6.07	3	7	14	21

Migraine								
No	95,870	90.8	9.2	5.3	3	7	14	20
Yes	10,493	87.23	12.77	5.15	3	6	14	20
Stomach or intestinal ulcer								
No	102,476	90.75	9.25	5.22	3	6	14	20
Yes	3,887	82.56	17.44	6.16	4	8	14	21
Urinary incontinence								
No	100,828	90.97	9.03	5.07	3	6	13	18
Yes	5,535	80.9	19.1	7.08	4	10	18	25
Number of major chronic conditions								
None	41,985	94.42	5.58	3.44	2	4	7	11
1	28,942	91.63	8.37	4.81	3	6	12	15
2	17,513	88.24	11.76	5.55	3	7	14	20
3	9,523	84.88	15.12	6.32	4	8	15	21
4 or more	8,400	77.5	22.5	7.08	5	10	17	23
Non-need indicators								
Health region	106,363	96,206	10,157	5.28	3	7	14	20
Province of residence								
Newfoundland	3,312	91.64	8.36	6.43	4	7	16	21
Prince Edward Island	1,658	89.38	10.62	6.21	3	7	15	24
Nova Scotia	4,156	90.33	9.67	5.79	4	5	15	19
New Brunswick	4,302	88.35	11.65	5.45	4	7	14	18
Quebec	19,948	89.92	10.08	4.96	3	6	13	20
Ontario	37,136	91.23	8.77	5.33	3	7	14	20
Manitoba	5,981	89.95	10.05	5.06	3	7	12	16

Saskatchewan	6,489	88.64	11.36	5.55	4	7	14	20
Alberta	10,110	90.17	9.83	4.93	3	6	12	18
British Columbia	13,271	90.95	9.05	5.28	3	7	13	20
Education								
< Secondary education	19,957	87.29	12.71	6.29	4	8	14	21
Secondary completed	17,279	90.81	9.19	5.22	3	6.5	14	19
Some post-secondary	7,247	90.06	9.94	5.34	3	7	14	20
Post-secondary graduate	59,044	91.51	8.49	4.74	3	5	12	16
Missing	2,836	89.18	10.82	5.99	3	7	15	21
Household income (1)								
Low	11,267	84.96	15.04	6.58	4	9	16	21
Low Middle	20,267	88.76	11.24	6.01	4	7	14	21
Middle	17,090	91.09	8.91	4.91	3	6	12	16
Upper Middle	13,417	91.55	8.45	4.47	3	5	10	14
High	27,471	93.08	6.92	3.74	2	4	7	13
Missing	16,851	90.34	9.66	5.62	3	7	14	21

(1) The financial amount of each household income level is: Low = 0 to $19,999, Low middle=$20,000 to $39,999, Middle = $40,000 to $59,999, Upper middle = $60,000 to $79,999 High = $80,000 or more, Missing = unknown

Table 3: Zero-inflated negative binomial regression models of number of overnight stays for the population choices of Canadians, regions and high income

Independent Variable	Canadian N=103417				Region N=103417				High income N=40793			
	COUNT		INFLATE (1)		COUNT		INFLATE (1)		COUNT		INFLATE (1)	
	IRR	p-value	OR	p-value	IRR	p-value	OR	p-value	IRR	p-value	OR	p-value
Age		0.000		0.000		0.000		0.000		0.007		0.000
20-29	1.48	0.050	-1.50	0.000	1.69	0.004	-1.50	0.000	2.21	0.004	-1.61	0.000
30-39	1.48	0.065	-1.43	0.000	1.45	0.038	-1.17	0.000	2.40	0.003	-1.66	0.000
40-49	1.68	0.044	-0.40	0.001	1.63	0.027	-0.40	0.001	2.61	0.004	-0.22	0.255
50-59	1.00		1.00		1.00		1.00		1.00		1.00	
60-69	2.57	0.000	-0.14	0.161	2.40	0.001	0.00	0.957	3.67	0.002	-0.15	0.428
70-79	2.65	0.000	-0.36	0.001	2.07	0.001	-0.24	0.013	3.43	0.001	-0.18	0.446
80+	1.63	0.082	-0.55	0.000	2.60	0.000	-0.49	0.000	2.55	0.278	-0.54	0.060
Sex: Male	1.10	0.428	-0.24	0.042	1.29	0.008	-0.13	0.204	1.05	0.831	-.035	0.070
Age*Sex(2)		0.522		0.000		0.131		0.000		0.505		0.000
Self-perceived health		0.000		0.000		0.000		0.000		0.000		0.000
Excellent	1.00		1.00		1.00		1.00		1.00		1.00	
Very good	1.62	0.019	-0.14	0.075	1.54	0.019	-0.08	0.279	2.18	0.011	-0.10	0.322
Good	2.38	0.000	-0.37	0.000	2.38	0.000	-0.28	0.000	2.79	0.000	-0.37	0.002
Fair or poor	3.08	0.000	-0.97	0.000	3.26	0.000	-0.86	0.000	4.17	0.000	-0.80	0.000
Age*SPH (2,3)		0.000				0.001				0.000		

	0.88	0.076	-0.14	0.114	0.86	0.024	-0.07	0.339	0.94	0.659	-0.46	0.016
Asthma and chronic respiratory disease	0.94	0.457	-0.25	0.006	0.91	0.167	-0.15	0.071	0.94	0.696	-0.31	0.126
Back problems	0.86	0.017	-0.06	0.471	0.86	0.008	-0.01	0.931	0.96	0.786	-0.23	0.176
Bowel disorder, Crohn's disease or colitis	1.17	0.045	-0.36	0.001	1.18	0.032	-0.35	0.000	1.33	0.076	-0.65	0.002
Cancer	1.13	0.049	-0.57	0.000	1.19	0.001	-0.50	0.000	1.36	0.013	-0.69	0.000
Diabetes	0.95	0.411	-0.25	0.162	0.99	0.836	-0.15	0.117	1.12	0.551	-0.28	0.151
Heart disease or stroke	1.06	0.006	-0.13	0.000	1.12	0.059	-0.79	0.000	1.45	0.017	-0.94	0.000
High blood pressure	1.20	0.172	-0.80	0.226	1.13	0.103	-0.01	0.908	1.25	0.212	-0.46	0.022
Mental illness	1.14	0.071	-0.12	0.001	1.07	0.304	-0.28	0.001	0.98	0.932	-0.43	0.024
Migraine	0.85	0.033	-0.31	0.030	0.87	0.031	-0.11	0.241	0.76	0.110	-0.40	0.037
Stomach or intestinal ulcer	1.03	0.752	-0.20	0.054	0.92	0.322	-0.27	0.008	1.18	0.427	-0.34	0.117
Urinary incontinence	0.99	0.887	-0.29	0.004	1.02	0.730	-0.32	0.000	0.95	0.771	-0.49	0.020
Number of major chronic conditions		0.000		0.058		0.000		0.001		0.010		0.360
None	1.00		1.00		1.00		1.00		1.00		1.00	
1	1.36	0.000	0.01	0.932	1.27	0.000	-0.10	0.229	1.46	0.012	0.23	0.177
2	1.28	0.018	-0.15	0.268	1.20	0.049	-0.32	0.015	1.22	0.441	0.29	0.327
3	1.64	0.000	0.06	0.730	1.64	0.000	-0.11	0.537	1.73	0.144	0.72	0.097
4 or more	1.73	0.006	0.10	0.713	1.72	0.003	-0.16	0.554	1.64	0.360	0.92	0.148
Region		0.008		0.001		0.001		0.000		0.001		0.012
Education		0.954		0.350		0.200		0.752		0.446	-0.08	0.088
< Secondary education	1.00	0.962	0.01	0.915	0.95	0.280	-0.06	0.355	1.03	0.847	-0.28	0.570
Secondary completed	1.03	0.673	-0.11	0.105	1.07	0.205	-0.03	0.588	0.87	0.203	-0.21	0.002
Some post-secondary	0.98	0.765	0.02	0.814	1.07	0.432	0.08	0.777	0.88	0.276	1.00	0.139
Post-secondary grad	1.00		1.00		1.00		1.00		1.00		1.00	

Household income		0.006		0.489		0.000		0.688	x [4]	x	x	x
Low	1.35	0.000	-0.08	0.339	1.38	0.000	-0.05	0.492	x	x	x	x
Low Middle	1.12	0.126	-0.07	0.371	1.13	0.096	-0.06	0.405	x	x	x	x
Middle	1.05	0.553	0.05	0.547	1.03	0.630	0.01	0.913	x	x	x	x
Upper Middle	1.10	0.196	-0.07	0.396	1.04	0.573	-0.06	0.446	1.16	0.054	-0.04	0.618
High	1.00		1.00		1.00		1.00		1.00		1.00	
Missing	1.02	0.883	0.02	0.851	1.06	0.425	0.04	0.623	x	x	x	x

(1) Show the likelihood of covariate being an always zero (rather than likelihood of spending a night)
(2) The interactions between both age and sex and age and SPH (self-perceived health) were included in the model. Interaction terms for each category within them are not included here
(3) The interaction term between age and SPH (self-perceived health) was included in only in the count component of the model because it was significant
(4) X represents the categories not included in estimation

Table 4: Estimated average need-expected use for the survey sample and Canadian population based on the Canadian, region and higher income models

	Sample (Unweighted)				Canadian Population (Weighted)			
	Mean	Std. err	95%CI	95%CI	Mean	Std. err	95%CI(1)	95%CI
Canadian	0.496	0.0022485	0.491	0.500	0.383	0.0026082	0.377	0.388
Region	0.483	0.0021318	0.479	0.488	0.376	0.0024993	0.371	0.381
Higher income	0.486	0.0029461	0.480	0.491	0.370	0.0031637	0.364	0.376

(1) The confidence intervals for the weighted estimations are underestimated because the standard errors for these measures are not adjusted for the complex survey design

Table 5: **Observed average inpatient use and average need-expected inpatient hospital use based on the three choices of population by province**

Province	Observed average inpatient use[(1)]	Average need-expected inpatient use		
		Canadian	Region	High income
NL	0.463	0.422	0.413	0.423
PEI	0.537	0.415	0.408	0.424
NS	0.435	0.462	0.453	0.478
NB	0.505	0.480	0.466	0.473
QC	0.387	0.368	0.363	0.339
ON	0.321	0.383	0.374	0.374
MN	0.352	0.409	0.402	0.395
SK	0.49	0.405	0.398	0.402
AB	0.381	0.351	0.347	0.349
BC	0.381	0.380	0.374	0.368
Canada	0.368	0.383	0.376	0.370

(1) Observed average inpatient use was calculated by applying the Canadian Community Health Survey sample weights to the sample's reported inpatient hospital use

(2) NL: Newfoundland and Labrador; PEI: Prince Edward Island; NS: Nova Scotia; NB: New Brunswick; QC: Quebec; ON: Ontario; MN: Manitoba; SK: Saskatchewan; AB: Alberta; BC: British Columbia

Table 6: Observed average inpatient use and average need-expected inpatient hospital use based on the three choices of population by income group

Income category	Observed average inpatient use[1]	Average need-expected inpatient use		
		Canadian	Region	High income
Low	0.755	0.846	0.808	0.648
Low Middle	0.559	0.575	0.557	0.535
Middle	0.352	0.354	0.357	0.368
Upper Middle	0.323	0.335	0.313	0.345
High	0.221	0.234	0.237	0.234
Missing	0.379	0.379	0.380	0.406
Canada	0.368	0.383	0.376	0.370

(1) Observed average inpatient use was calculated by applying the Canadian Community Health Survey sample weights to the sample's reported inpatient hospital use

Table 7: Average need-expected inpatient hospital use by province based on the Canadian, region and high income, simple models (1)

Province (2)	Average need-expected inpatient use		
	Canadian	Region	High income
NL	0.380	0.423	0.368
PEI	0.370	0.404	0.362
NS	0.399	0.459	0.392
NB	0.417	0.457	0.409
QC	0.350	0.392	0.344
ON	0.369	0.416	0.365
MN	0.378	0.406	0.371
SK	0.393	0.446	0.397
AB	0.336	0.385	0.327
BC	0.371	0.410	0.367
Canada	0.365	0.416	0.339

(1) The simple model included the independent variables age, sex, self-perceived health, and an interaction term between age and sex.

(2) NL: Newfoundland and Labrador; PEI: Prince Edward Island; NS: Nova Scotia; NB: New Brunswick; QC: Quebec; ON: Ontario; MN: Manitoba; SK: Saskatchewan; AB: Alberta; BC: British Columbia

Table 8: **Average need-expected inpatient hospital use by income group based on the Canadian, region and high income, simple models** (1)

Income category	Average need-expected use		
	Canadian	Region	Higher income
Low	0.587	0.668	0.584
Low Middle	0.512	0.572	0.522
Middle	0.372	0.413	0.368
Upper Middle	0.306	0.346	0.293
High	0.251	0.278	0.238
Missing	0.420	0.478	0.421
Canada	0.365	0.416	0.359

(1) The simple model included the independent variables age, sex, self-perceived health, and an interaction term between age and sex.

FIGURES

Figure 1: An example of Bland –Altman plot

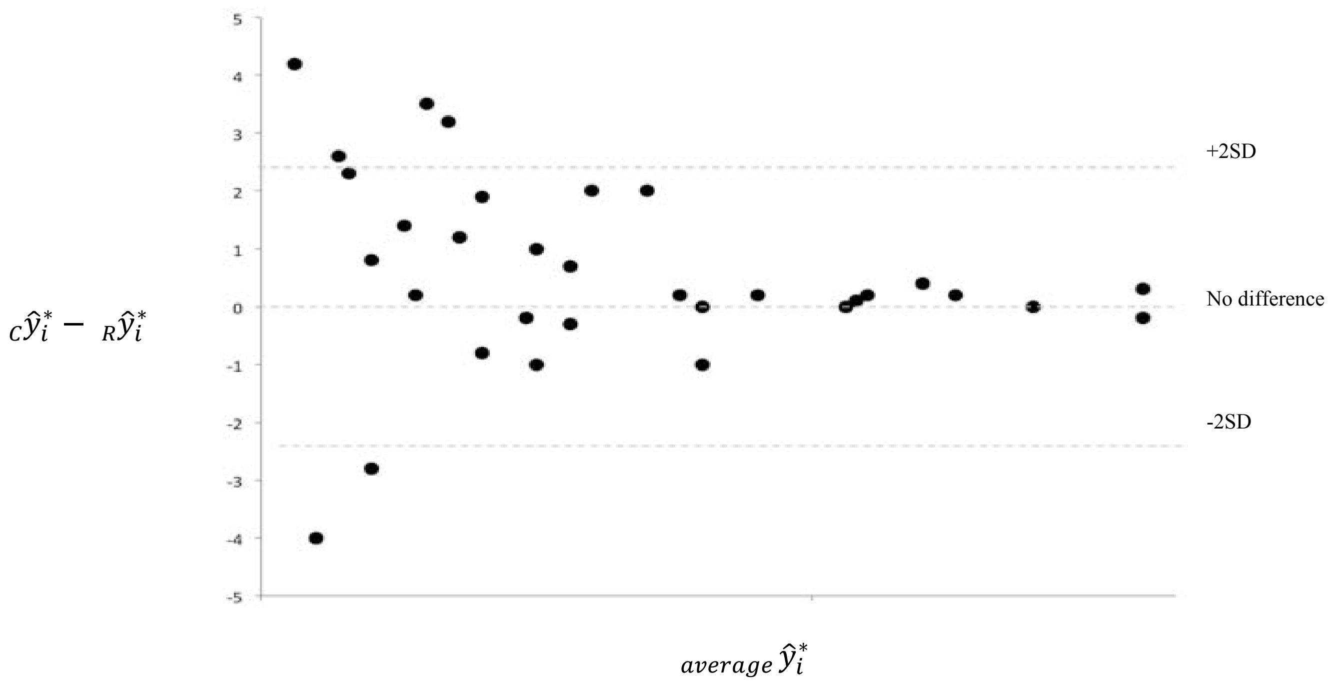


Figure 1 is an illustrative example of a comparison between the mean of the three estimates of need-expected inpatient hospital use for each individual in the data ($_{average}\,\hat{y}_i^*$) against the difference between two choice of populations' individual estimates, the estimate for individual *i* using Canadians and the estimate for individual *i* using equal regional representation. The above example shows that as the $_{average}\,\hat{y}_i^*$ increases, the differences between the two chosen populations' estimations decreases. This suggests that each paired estimate for the top third of the sample is quite similar, with the $_C\hat{y}_i^*$ tending to be slightly larger than $_R\hat{y}_i^*$. A majority of plots fall within the range of the limits of agreement (+/- 2 S.D.). Also, the plots that are in the bottom half of the $_{average}\,\hat{y}_i^*$, have much larger differences observed between the two choice populations' paired estimates. The lowest values of $_{average}\,\hat{y}_i^*$ range over +/- 2SD from each other, suggesting their estimates differ greatly, with$_C\hat{y}_i^*$ tending (though not consistently) to be larger.

Figure 2:

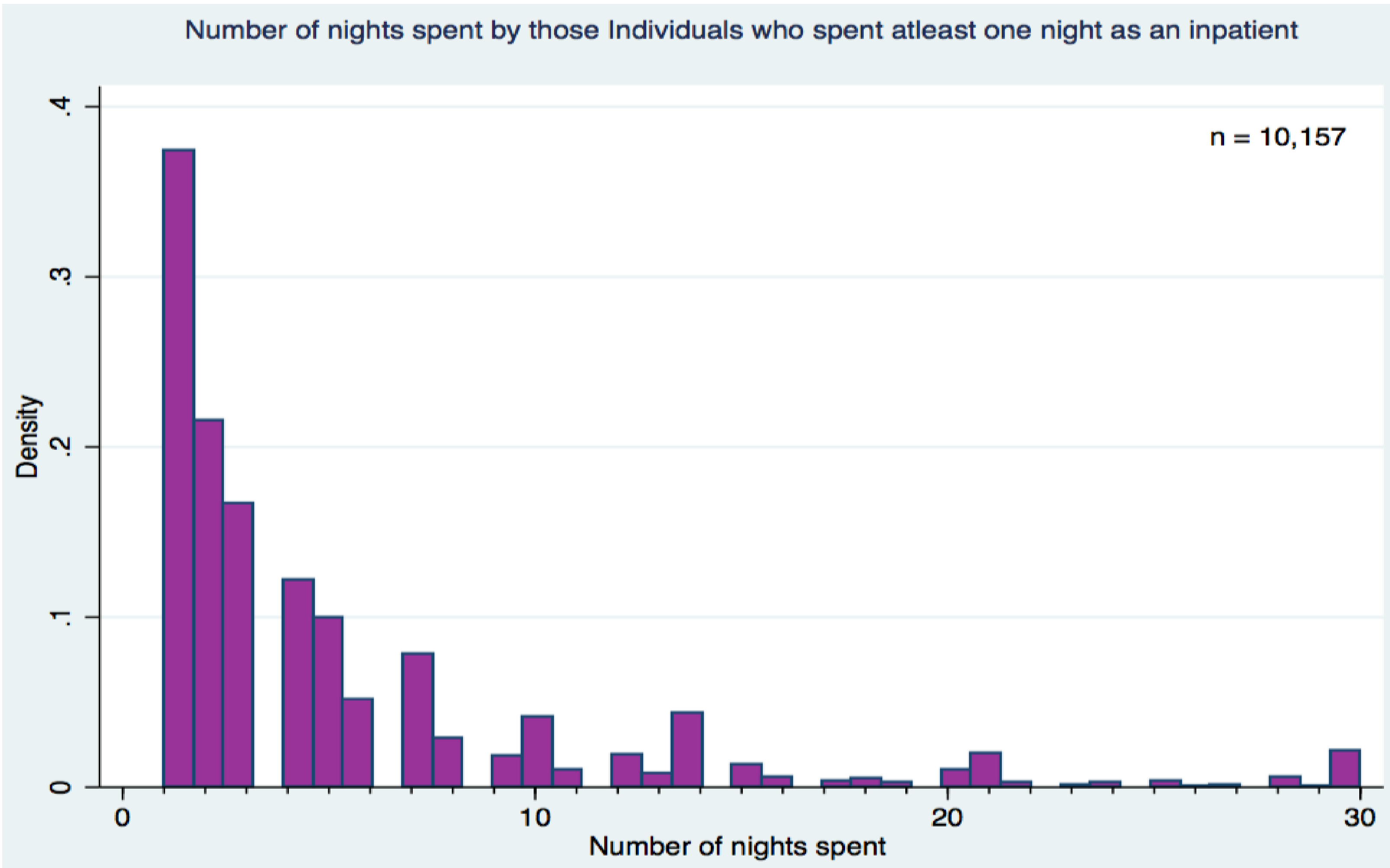
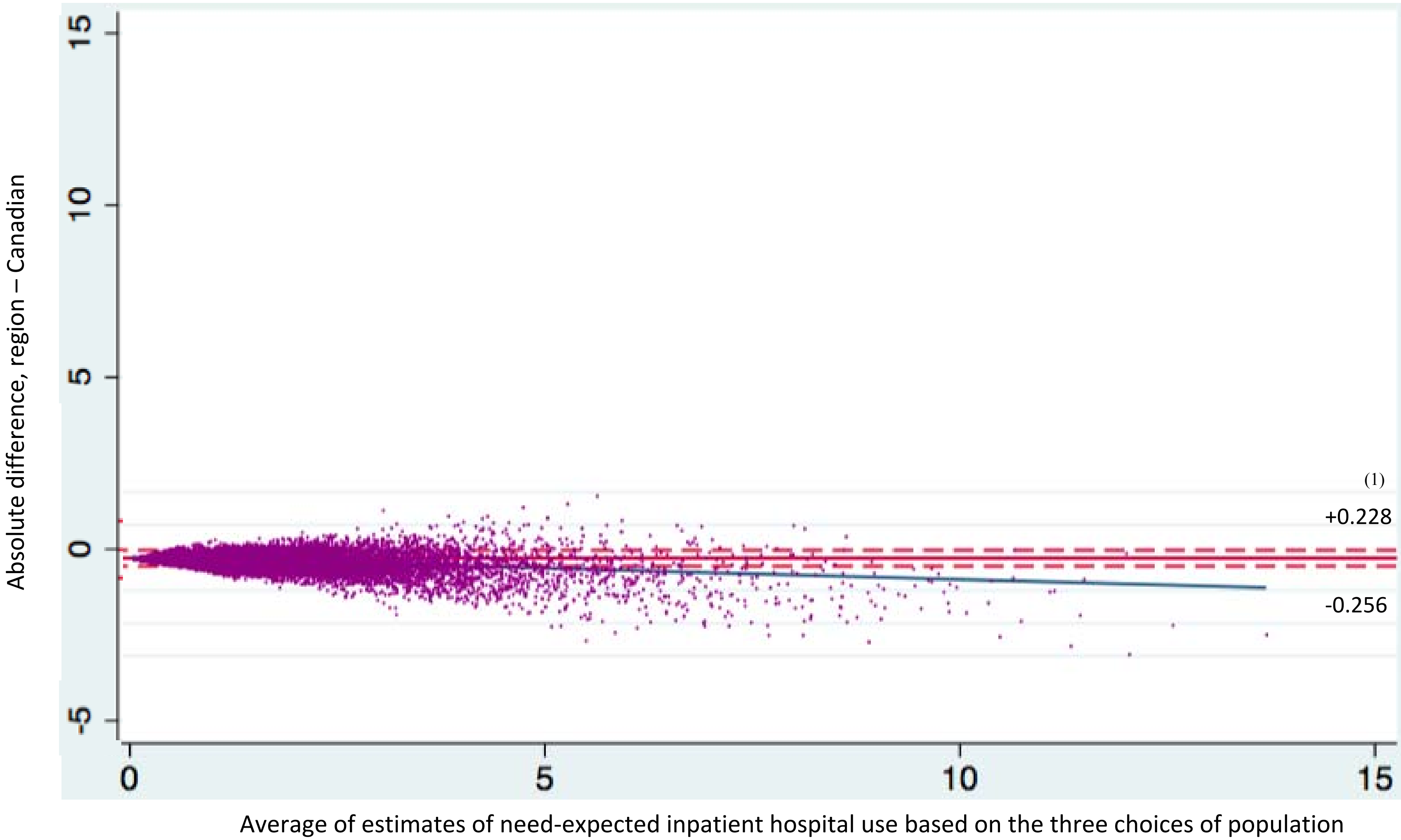
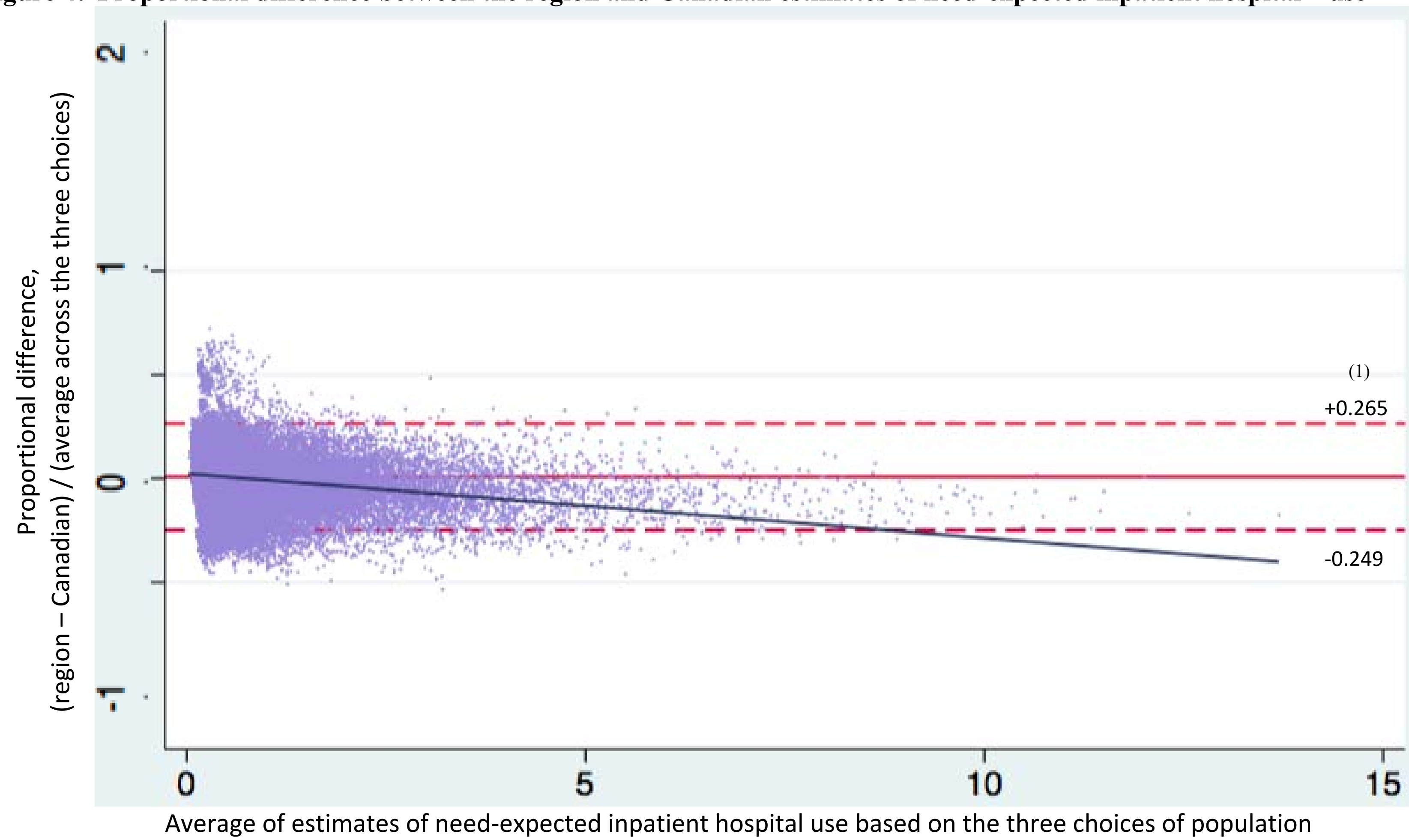

Number of nights spent by those Individuals who spent atleast one night as an inpatient
n = 10,157
Density
0
.1
.2
.3
.4
0
10
20
30
Number of nights spent

Figure 3: Absolute difference between the region and Canadian estimates of need-expected inpatient hospital use



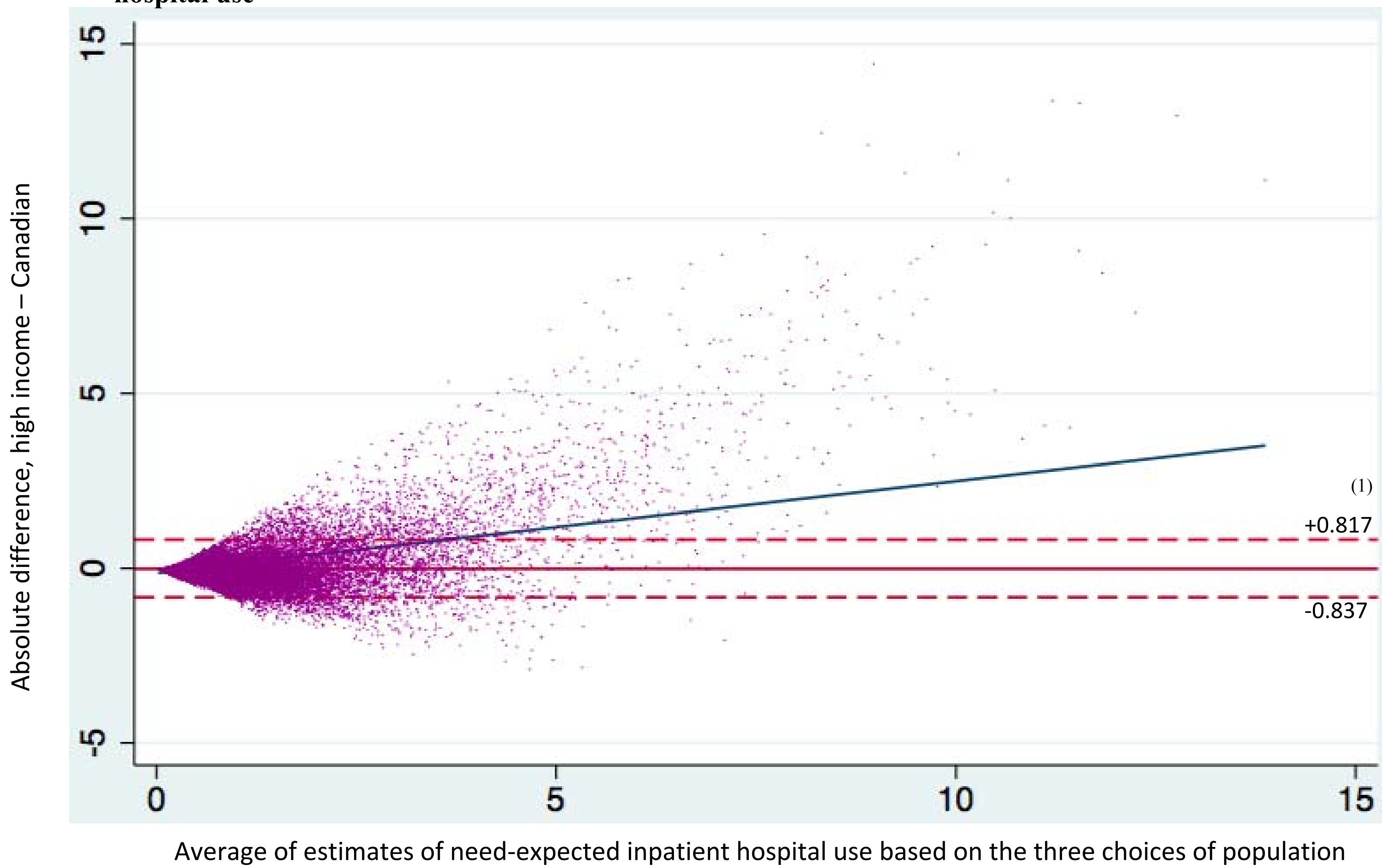
(1) The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the absolute difference)) and the solid red line is the mean of the absolute difference(-0.014)

Figure 4: Proportional difference between the region and Canadian estimates of need-expected inpatient hospital use



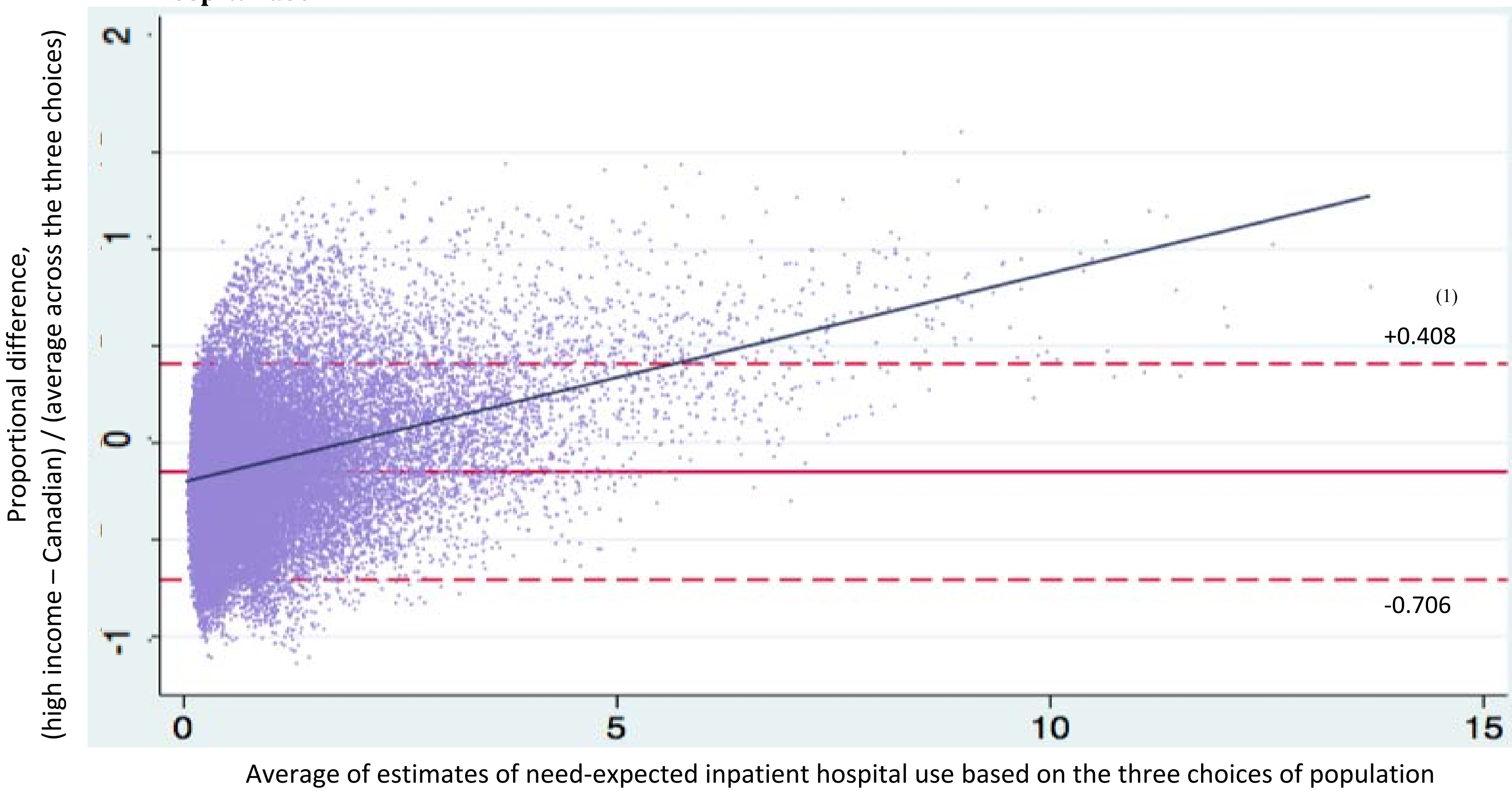
(1) The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the proportional difference)) and the solid red line is the mean of the proportional difference (0.008)

Figure 5: Absolute difference between the high income and Canadian estimates of need-expected inpatient hospital use



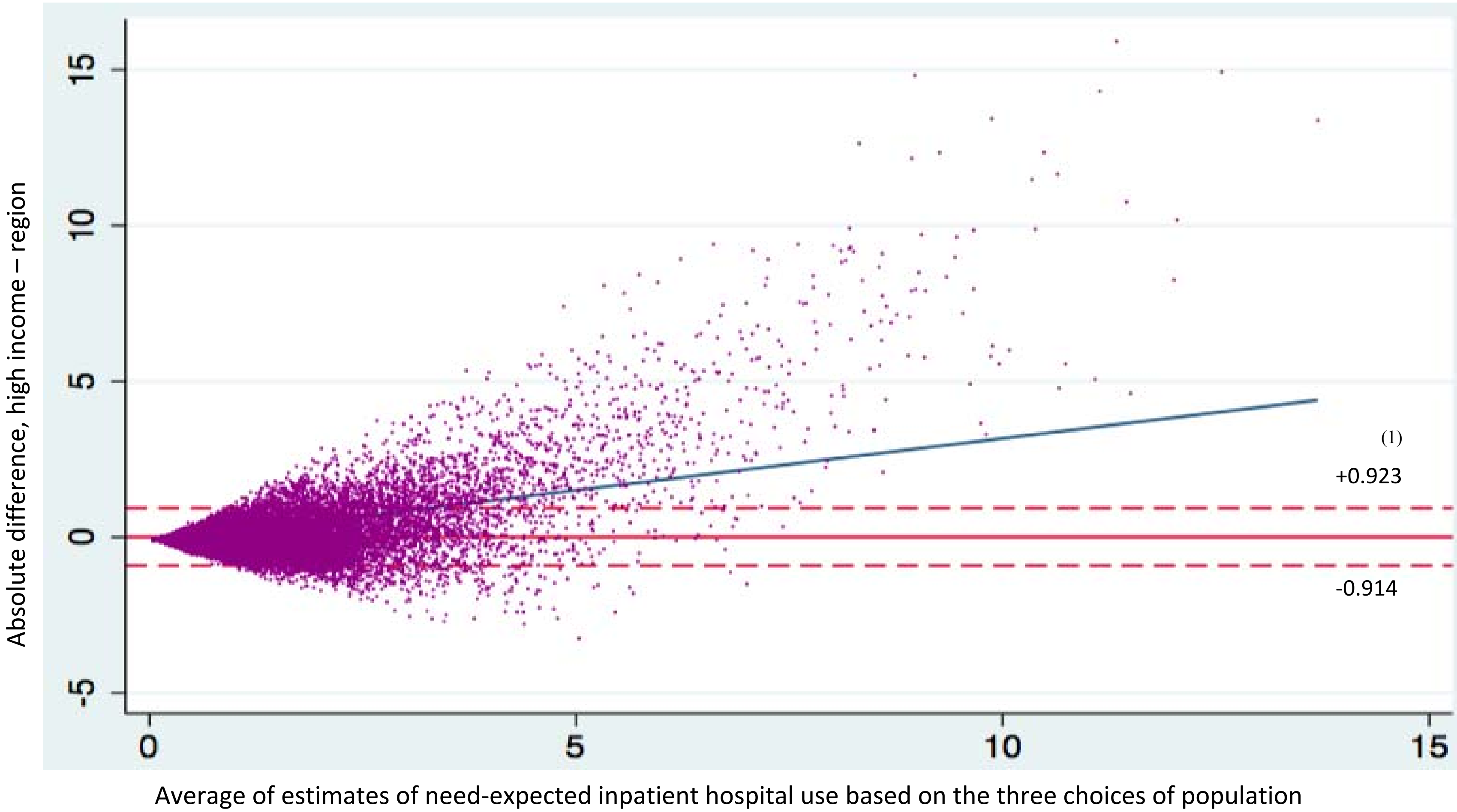
(1)The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the absolute difference)) and the solid red line is the mean of the absolute difference (-0.010)

Figure 6: **Proportional difference between the high income and Canadian estimates of need-expected inpatient hospital use**



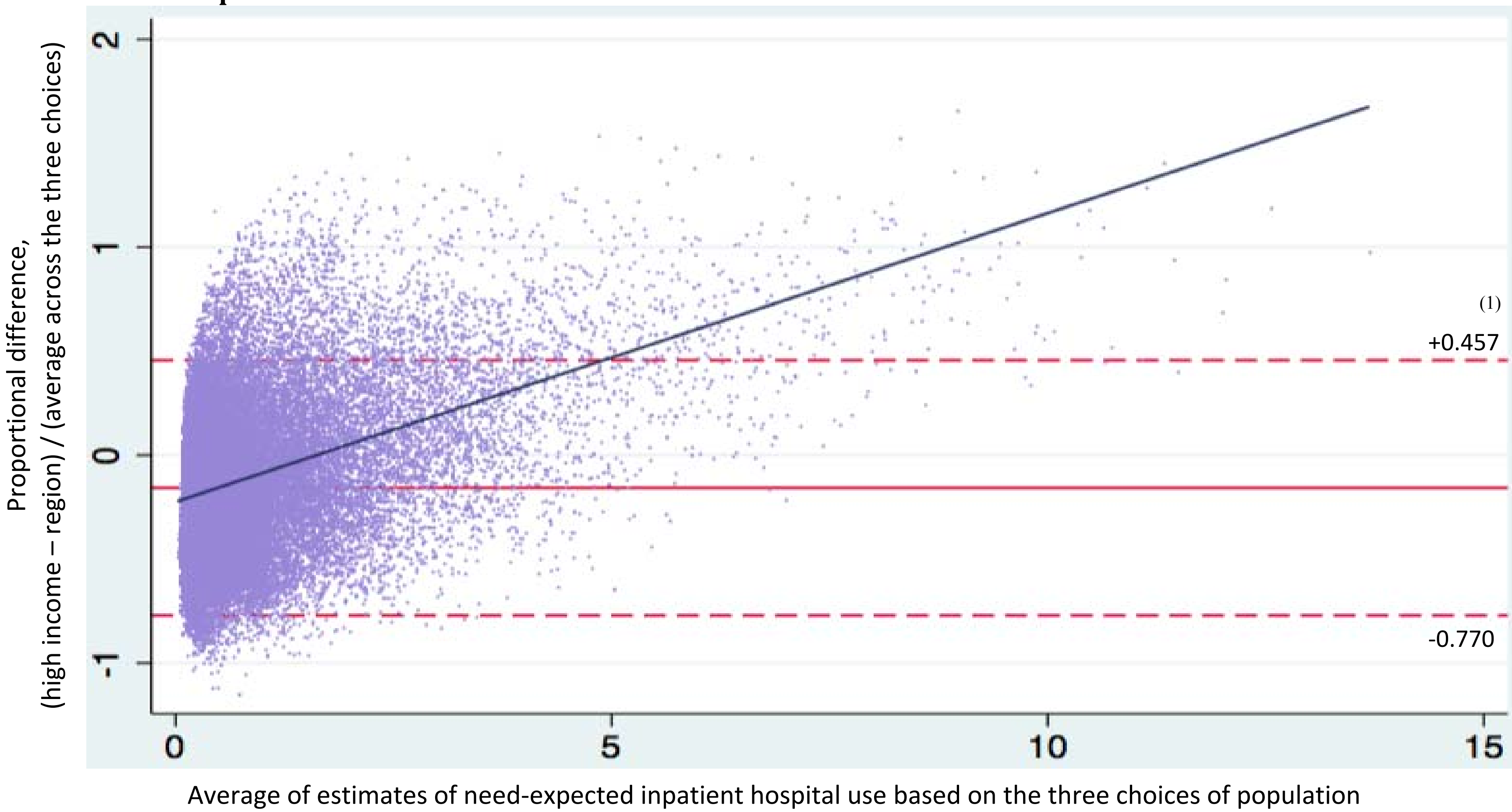
(1) The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the proportional difference)) and the solid red line is the mean of the proportional difference (-0.149)

Figure 7: **Absolute difference between the high income and region estimates of need-expected inpatient hospital use**



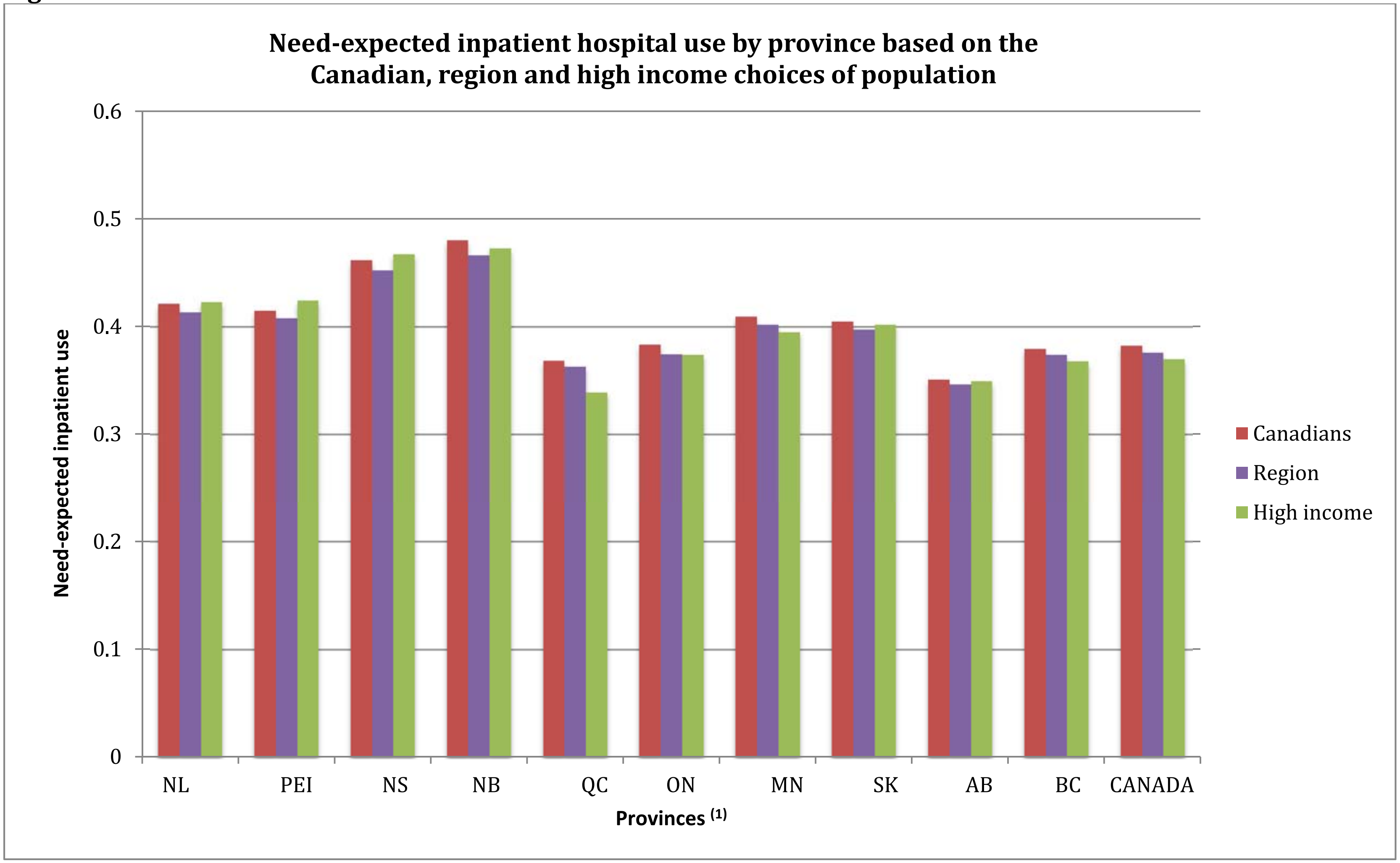
(1)The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the absolute difference)) and the solid red line is the mean of the absolute difference (-0.004)

Figure 8: Proportional difference between the high income and region estimates of need-expected inpatient hospital use



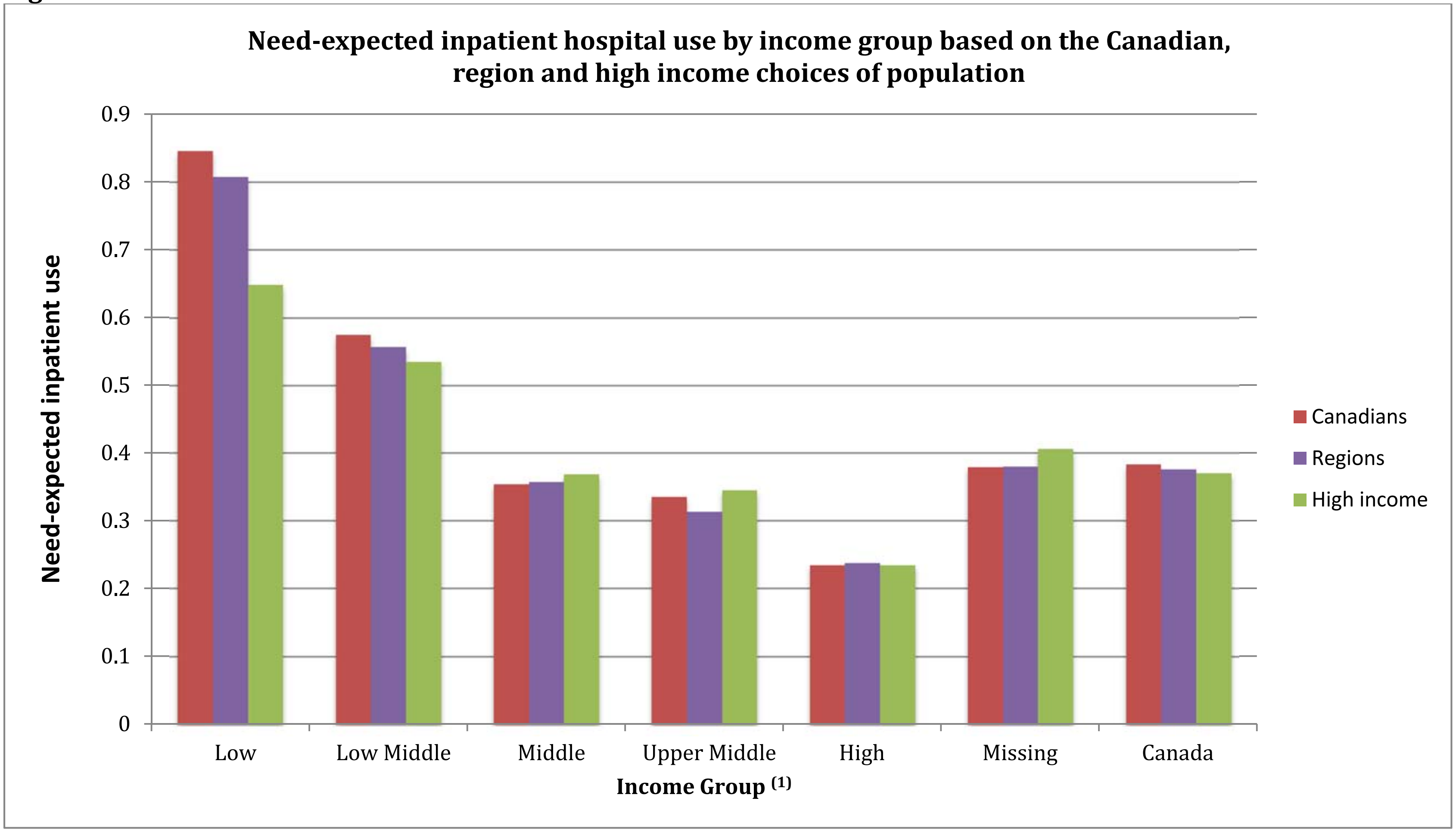
(1) The dashed lines are the 95% limits of agreement (mean +/- (1.96 * standard deviation of the proportional difference)) and the solid red line is the mean of the proportional difference (-0.157)

Figure 9:



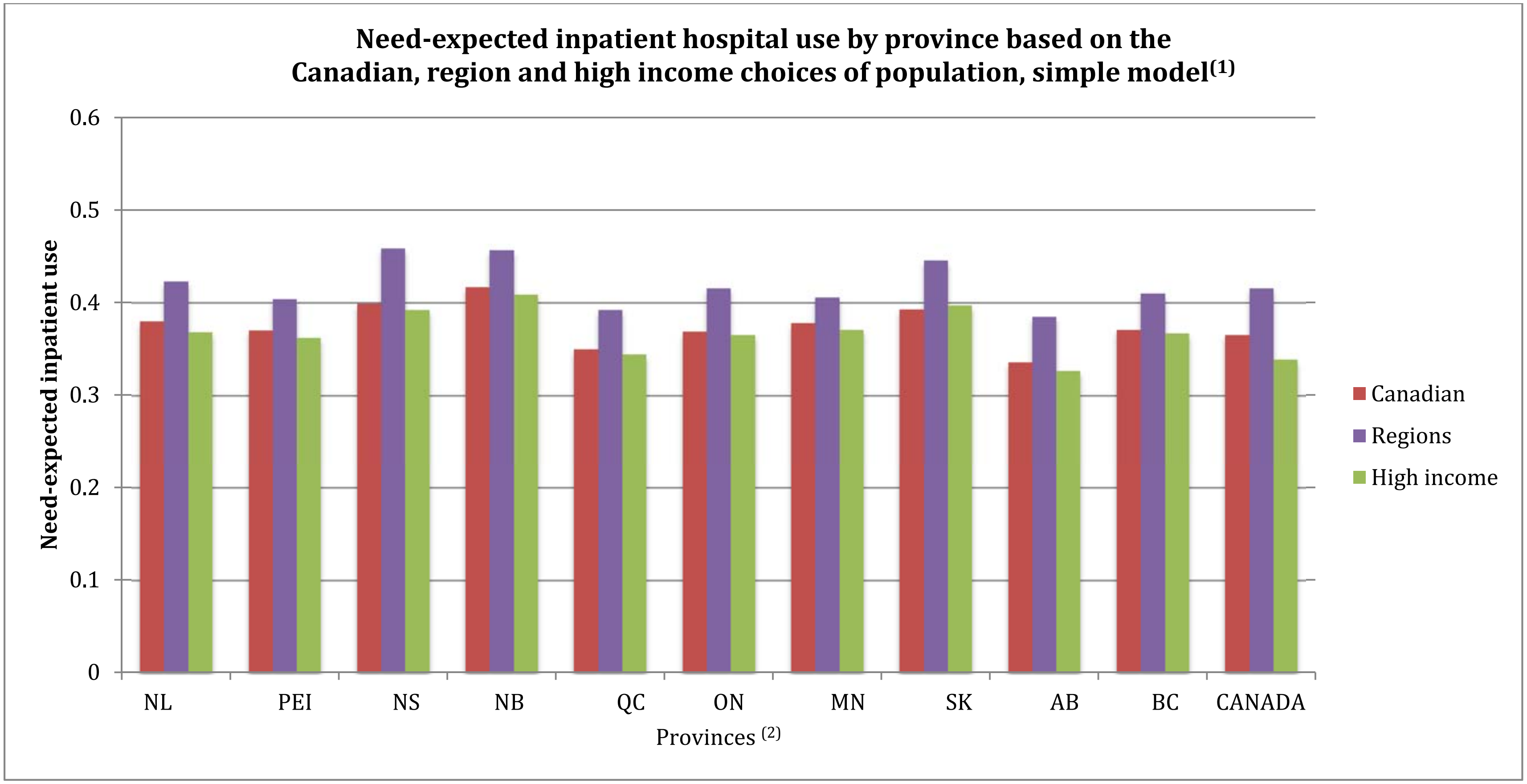
(1) NL: Newfoundland and Labrador; PEI: Prince Edward Island; NS: Nova Scotia; NB: New Brunswick; QC: Quebec; ON: Ontario; MN: Manitoba; SK: Saskatchewan; AB: Alberta; BC: British Columbia

Figure 10:



(1) The financial amount of each household income level is: Low = 0 to $19,999, Low middle=$20,000 to $39,999, Middle = $40,000 to $59,999, Upper middle = $60,000 to $79,999 High = $80,000 or more, Missing = unknown

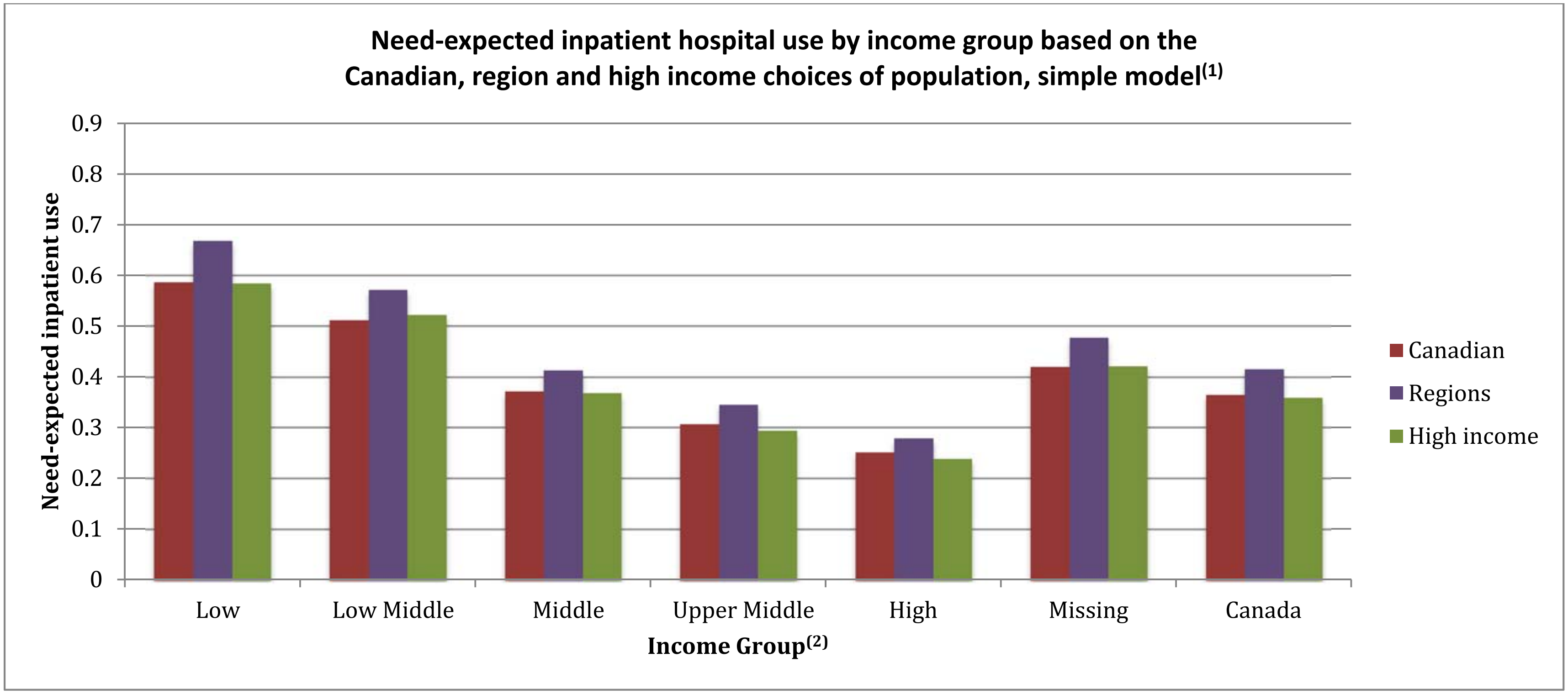
Figure 11:



(1) The simple model included the independent variables age, sex, self-perceived health, and an interaction term between age and sex.

(2) NL: Newfoundland and Labrador; PEI: Prince Edward Island; NS: Nova Scotia; NB: New Brunswick; QC: Quebec; ON: Ontario; MN: Manitoba; SK: Saskatchewan; AB: Alberta; BC: British Columbia

Figure 12:



(1) The simple model included the independent variables age, sex, self-perceived health, and an interaction term between age and sex.

(2) The financial amount of each household income level is: Low = 0 to $19,999, Low middle=$20,000 to $39,999, Middle = $40,000 to $59,999, Upper middle = $60,000 to $79,999 High = $80,000 or more, Missing = unknown

Bibliography

1. Hurley J, Hutchison B, Buckley G, Woodward C. Needs-based funding for home care and community support services in Ontario: A new approach based on linked survey and administrative data. CHEPA Working Paper Series. Ontario: McMaster University; 2003. Available at: http://www.chepa.org/Files/Working%20Papers/04-01.pdf. Accessed 07/23, 2013.

2. Asada Y, Kephart G. Equity in health services use and intensity of use in Canada. BMC Health Serv Res. 2007;7(41). Available at: http://www.biomedcentral.com/1472-6963/7/41. Accessed 07/23, 2013.

3. Carr-Hill R, Hardman G, Martin S, Peacock S, Sheldon T, Smith P. A formula for distributing NHS revenues based on small area use of hospital beds. York, England: University of York; 1994. Available at: http://econpapers.repec.org/scripts/redir.pf?u=http%3A%2F%2Fwww.york.ac.uk%2Fmedia%2Fche%2Fdocuments%2Fpapers%2Foccasionalpapers%2FCHE%2520Occasional%2520Paper%252022.pdf;h=repec:chy:respap:22cheop. Accessed 07/23, 2013.

4. Schofield D, McRae I, Shretha R. Equity, Poverty and access to GP services in Australia. International Medical Workforce Conference; Edinburgh, Scotland; 2008. Available at:http://crmcc.medical.org/publicpolicy/imwc/Equity_%20Australia.pdf. Accessed 07/23, 2013.

5. Carr-Hill R, Jamison J, O'Reilly D, Stevenson M, Reid J, Merriman B. Risk adjustment for hospital use using social security data: cross sectional small area analysis. Brit Med J. 2002;16(324):1-4.

6. Zhou Z, Su Y, Gao J, Campbell B, Zhu Z, Xu LF, et al. Assessing equity of healthcare utilization in rural China: results from nationally representative surveys from 1993 to 2008. Int J Equity Health. 2013;12(34). Available at: http://www.equityhealthj.com/content/12/1/34. Accessed 07/23, 2013.

7. Birch S, Kephart G, Tomblin-Murphy G, O'Brien-Pallas L, Alder R, MacKenzie A. Human Resources Planning and the Production of Health: A Needs-Based Analytical Framework. Can Public Pol. 2007;33(0):S1-S16.

8. Hurley J. Regionalization and the allocation of healthcare resources to meet population health needs. Healthc Pap. 2004;5(1):34-39.

9. Gusen P. Expenditure Need: Equalization's Other Half. Mowat Centre for Policy Innovation: University of Toronto; 2012. Available at: http://www.mowatcentre.ca/pdfs/mowatResearch/47.pdf. Accessed 07/23, 2013.

10. Advisory Committee on Resource Allocation. A brief history of resource allocation in the NHS 1948–98. Resource Allocation Working Paper. 1998;4. Available at: http://www.dh.gov.uk/prod_consum_dh/groups/dh_ digitalassets/@dh/@en/documents/digitalasset/dh_4109403.pdf. Accessed 07/23, 2013.

11. Kephart G, Asada Y. Need-based resource allocation: different need indicators, different results? BMC Health Serv Res. 2009;9(1):122. Available at: http://www.biomedcentral.com/1472-6963/9/122. Accessed 07/23, 2013.

12. Asada Y, Kephart G. Understanding different methodological approaches to measuring inequity in health care. Int J Health Serv. 2011;41(2):195–207.

13. Asada Y, Kephart G. Equity, poverty, and access - Canada. International Medical workforce collaborative conference; September 2008; Edinburgh, Scotland; 2008. Available at: http://rcpsc.medical.org/publicpolicy/ imwc/IMWC-Equity_Poverty_Access-Canada_Aug1_2008_YA.pdf. Accessed 07/23, 2013.

14. Gravelle H, Sutton M, Morris S, Windmeijer F, Leyland A, Dibben C, et al. Modeling supply and demand influences on the use of health care: implications for deriving a needs-based capitation formula. Health Econ. 2003;12(12):985-1004.

15. Statistics Canada. Population of census metropolitan areas. 2013; Available at: http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/ demo05 a-eng.htm. Accessed 07/09, 2013.

16. Carr-Hill RA, Sheldon TA, Smith P, Martin S, Peacock S, Hardman G. Allocating resources to health authorities: development of method for small area analysis of use of inpatient services. BMJ. 1994;309(6961):1046–9.

17. Statistics Canada. Census Profile: Georgian Bay. 2013; Available at: http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=3544065&Geo2=PR&Code2=35&Data=Count&SearchText=Georgian%20Bay&SearchType=Begins&SearchPR=01&B1=All&Custom=. Accessed 07/09, 2013.

18. Statistics Canada. Canadian Community Health Survey 2009-2010 Public Use Microdata File. 2011.

19. Statistics Canada. Canadian Community Health Survey (CCHS) annual component: User guide 2010 and 2009-2010 Microdata files. June 2011. Available at: http://www23.statcan.gc.ca/imdb-bmdi/pub/document/3226 _D7_T9_V8-eng.htm. Accessed 07/23, 2013.

20. Statistics Canada. CCHS 2009-2010: Data dictionary public use microdata file - 24 month. 2011. Available at:http://search1.odesi.ca/documentation/CCHS_2009-2010/cchs-escc2009-2010cbk-eng.pdf. Accessed on 07/23, 2013.

21. Kim S, Egerter S, Cubbin C, Takahashi ER, Braverman P. Potential Implications of Missing Income Data in Population-Based Surveys: An Example from a Postpartum Survey in California. Pub Heal Rep. 2007;122(6):753-763. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1997243/. Accessed 08/18, 2013.

22. Cameron A, Trivedi P. Regression analysis of count data. USA: The press syndicate of the University of Cambridge; 1998.

23. Sarma S, Simpson, W. A microeconometric analysis of Canadian health care utilization. Health Econ. 2006;15:219-239.

24. Greene WH. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. NYU Working Paper. 1994;94(10). Available at: http://papers.ssrn.com/sol3/Delivery.cfm/2451_26263.pdf?abstract id=1293115&mirid=1. Accessed 07/23, 2013.

25. Layes A, Asada A, Kephart G. Whiners and deniers – What does self-rated health measure? Soc Sci Med. 2012; 75(1). Available at: http://www.ncbi.nlm.nih.gov/pubmed/22265085Accessed 08/18.

26. Statistics Canada. Canadian Community Health Survey (CCHS) annual component: Complement to the user guide 2010 and 2009-2010 microdata files. November 2011. Available at: http://www23.statcan.gc.ca/imdb-bmdi/document/3226_D74_T1_V1-eng.htm. Accessed 07/23, 2013.

27. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;i:307-310.

28. StataCorp. Stata Statistical Software: Release 12. 2011;12.

29. Lohr SL. Sampling: Design and analysis. USA: Brooks/Cole publishing company; 1999.

30. Rao J, Wu C. Resampling inference with complex survey data. J Amer Statist Assoc. 1988;88:231-241.