Health Economics, Policy and Law (2016), **11**, 39–65 © Cambridge University Press 2015
doi:10.1017/S1744133115000274

The determinants of efficiency in the Canadian health care system

SARA ALLIN*
Canadian Institute for Health Information, Toronto, ON, Canada
University of Toronto, Toronto, ON, Canada
MICHEL GRIGNON
McMaster University, Hamilton, ON, Canada
LI WANG
Canadian Institute for Health Information, Toronto, ON, Canada
McMaster University, Hamilton, ON, Canada

Abstract: In spite of the vast number of studies measuring economic efficiency in health care, there has been little take-up of this evidence by policy-makers to date. This study provides an illustration of how a system-level study drawing on best practice in empirical measurement of efficiency may be of practical use to health system decision makers and managers. We make use of the rich data available in Canada to undertake a robust two-stage data envelopment analysis to calculate efficiency at the regional (sub-provincial) level. Decisions about what the health system produces (the outcome to measure efficiency against) and what are the resources it has to produce that outcome were based on interviews and consultation with health system decision makers. Overall, we find large inefficiencies in the Canadian health care system, which could improve outcomes (here, measured as a reduction in treatable causes of death) by between 18 and 35% across our analyses. Also, we find that inefficiencies are the result of three main sets of factors that policy makers could pay attention to: management factors, such as hospital re-admissions; public health factors, such as obesity and smoking rates; and environmental factors such as the population's average income.

Submitted 11 August 2014; revised 1 April 2015; accepted 2 April 2015;
first published online 7 May 2015

1. Introduction

Payers for health care, either private insurers or governments using taxpayers' money, have been looking for decades to make the most of the resources they purchase with insured or taxpayers' contributions. This being the very definition of (output-oriented) technical, or productive, efficiency in economics, one could

*Correspondence to: Sara Allin, Senior Researcher, Canadian Institute for Health Information (CIHI), 4110 Yonge Street, Suite 300, Toronto, ON M2P 2B7, Canada. Email: sallin@cihi.ca

assume economists (and possibly, health economists) have been enroled to help solve the problem and provide decision makers with good evidence-based recommendations on how to reform or steer their health care 'systems'. However, as shown in Hollingsworth (2012), in spite of the close to 400 publications measuring economic efficiency in health care to date, there has been little take-up of this evidence by policy makers. Hollingsworth (2012) suggests that the existing studies are not framed in a way that is relevant to policy makers, and that to help bridge the gap between research and practice, a set of guidelines is needed for efficiency analysis in health care. These guidelines could help improve take-up by policy makers by addressing two main shortcomings: first, poor average quality of the studies (at least in their ability to derive actionable recommendations); and second, the lack of clear and universally accepted criteria that would allow decision makers to assess the quality of such studies.

As much as we agree with the need for guidelines in efficiency analyses of health care systems, we suggest an additional explanation for the limited take-up of empirical evidence on economic efficiency in health care among policy makers: the very notion of systemic economic efficiency is seen as a black box. In contrast to 'bottom-up' studies that focus on clearly identified problems such as waste of resources, or poor outcomes on a single disease with clear practice guidelines, 'top-down', or system-level, studies that link resources to outcomes and seek to identify inefficiencies on the basis of what systems or sub-systems actually do rather than by comparison with a guideline of what ought to be, are often viewed as lacking any actionable recommendation (Häkkinen *et al.*, 2013). The aim of this study is to provide a tentative demonstration that 'top-down' studies can be of help to health system decision makers and managers.

Unlike studies of facilities such as hospitals and nursing homes, studies of economic efficiency of health care systems are actually quite rare (Hollingsworth, 2008). Studies of hospitals or nursing homes can be of interest to economists (e.g. to answer questions such as 'are for-profit hospitals more or less efficient than public ones?') but such questions are not necessarily of much help to decision makers who are more interested in reforms at the intensive margin (e.g. improving efficiency for all hospitals, whatever their ownership status) than in radical ones altering the distribution of hospitals by ownership status. Also, conclusions drawn at sub-system levels (e.g. hospitals or nursing homes) may not be helpful if apparent efficiency is the result of a better ability to select cases or to shift costs to other providers in the system (Felder and Tauchmann, 2013). Policy-makers are responsible for the system as a whole and want to make sure that they get the best out of resources they purchase. In technical terms, the level of Decision Making Unit (DMU) chosen to conduct efficiency analyses matters much and we posit here that studies conducted at the district or country level (aggregating all providers on a given territory contributing to a given health outcome) can be of help to policy-makers.

We briefly indicate here how our study differs from most (if not all) studies published so far in the small literature on efficiency at the health system (rather

than hospitals or nursing homes) level, and how these differences contribute to a step in the direction of efficiency measurement as evidence for policy.

First, we based the main strategic decisions needed to conduct a study of efficiency (outcomes, inputs, DMU level and estimation of the frontier) on a systematic approach rather than on opportunity. The first such strategic decision regards the outcome, or what health systems produce: health systems are complex and do many different things, but what they are supposed to produce is never explicitly discussed. In that sense, a health system is very different from a firm (whose output is often self-obvious, even though, as Newhouse (1994) state, it would be quite impossible to tell what an airline company is actually selling) or even from a hospital, and the first step in any measure of efficiency of the health system is to make sure the analyst takes the right measure of outcome. We find that, in most of the literature at the country or district level, this choice of outcome variable is never made explicit and often driven by data availability. Since life expectancy or health-adjusted life expectancy is often available at the regional or national level, and since it is a 'reasonable' objective for a health system, studies often use such measures of average population health as their outcome measure, without discussing whether this is what policy-makers are truly mandated to maximise (Smith and Street, 2005). Our approach was radically different: we started with an agnostic view of what the outcome of a health system should be and used a series of tools (official documents scan, elite interviews, policy dialogue, all are detailed below) to elicit the answer to such an important question. Once we had that answer, we started looking for good data to represent and measure the outcome of the health system in the population. Since our approach did not yield a unanimous answer to the question of determining the objective of a health system, we compared the findings of our assessment of efficiency using various measures of such an outcome, such as age standardised treatable potential years of life lost (PYLL), survival rate and mortality rate.

Similarly, as far as the second decision (inputs) was concerned, we found that the definition of the sets of inputs and factors explaining efficiency scores was often *ad hoc*, the result of what is available in the data and not of an explicit discussion of what should be considered a resource available to the system and what should be considered an environmental factor to explain efficiency scores of individual DMUs. A crucial question is whether inputs should be accounted for as quantities (number of beds or physicians per 1000 population) or values (dollars spent on beds or physicians incomes). We used the same approach to determine the level of DMU at which to conduct the analysis: in a federal country such as Canada, this could be the province or sub-provincial districts such as health regions. All the above decisions were systematically made, based on a method that we documented, so that others could replicate and amend them. This, arguably, can help contribute to a more standardised and scientific approach to measuring efficiency in a way that health decision makers and managers can use.

Finally, ours is one of the first studies to use a robust two-step approach, following a Simar-Wilson method to account for serial correlation among efficiency

scores (Simar and Wilson, 2007). Other such studies are Felder and Tauchmann (2013) (for districts at the sub-Lander level in Germany) and Wranik (2012) (for countries in the OECD), and we see our study as a contribution to the field in the same spirit as theirs: how can economists measure efficiency scores for a series of DMUs within a larger set (jurisdiction or club of countries such as the OECD) and, rather than providing a league table that might be used for naming and shaming, 'explain' these scores by factors that managers can alter or under which they have to operate. The former (factors that managers can alter) will help local decision makers improve their performance (and decision makers at a higher level assess that performance), while the latter will mainly help decision makers at a higher level (central) allocate resources so that DMUs operating under more challenging environments are compensated.

2. Method

The method consists of two stages. First, data envelopment analysis (DEA) is used to estimate efficiency at the regional level, which allows us to generate the average level of efficiency in Canada as a whole. Second, regression analyses are undertaken in order to identify the factors that help explain variation in inefficiency across regions (we analyse regional efficiency scores and treat regions as observations in this second-stage statistical analysis). The software package Frontier Efficiency Analysis with R (FEAR 1.0) was used to carry out the DEA estimations and Stata 11.0 was used for the second stage regression.

2.1 DEA

DEA and SFA are the two main techniques to measure efficiency. SFA is a parametric technique, which involves econometric estimation of the production frontier and can accommodate random noise. DEA is non-parametric and applies linear programming to calculate the maximum attainable output for every level of input. In this study, we used DEA because it does not rely on *a priori* assumptions on the specification of the production frontiers and the random error distribution, and is less sensitive than SFA to underperforming outliers. Also, DEA can incorporate inputs and outputs, which are measured in different units (Valdmanis *et al.*, 2008), such as dollar values or quantities or activities. With DEA, one region is inefficient to some extent if a linear combination of two other regions observed in the data that would use the same level of inputs can produce more (Jacobs *et al.*, 2006). (The Appendix provides more details on the method.)

2.2 Robustness of DEA

As described earlier, the DEA approach has several advantages; however, there are two main limitations. First, results of DEA are sensitive to high-performing outliers.

If the outliers are undetected, the large random variation affecting a frontier DMU moves the entire frontier, which influences the estimates of all other DMUs. This study employs the method proposed by Wilson (1993) to detect the outliers in DEA. Second, DEA is non-stochastic: it implicitly assumes that the entire deviation from the frontier is caused by the true level of inefficiency of a given region, ignoring the fact that the discrepancy may partially be due to measurement error (in the data) or random error (the region experienced a random shock in the very year measures were taken) (Zere *et al.*, 2001). To address this drawback, we introduced the smoothed bootstrap method developed by Simar and Wilson (1998) to correct the estimates of efficiency for random noise. Details on the statistical outlier detection and bootstrapping methods can be found in the Appendix.

Smith (1997) notes that a well-specified DEA model will always overestimate efficiency; therefore, robust estimates are consistently lower than the point estimates. Because DEA is a descriptive approach that defines the efficiency frontier with actual regions, some of them will have an efficiency score of 1, by construction. It is highly likely, though, that the true frontier is further out and that no unit actually achieves perfect efficiency. The robust estimates reflect that in the sense that, as in SFA, none of the observed DMUs achieves an efficiency of 1 once the bias is corrected.

Finally, in order to ensure the level of variation in the efficiency estimates is sufficient relative to noise, we followed the approach recommended by Badunenko *et al.* (2012). Essentially, the ratio of the variation in efficiency to the variation in noise should be greater than one in order to proceed with efficiency analyses. In our study, the ratio of variation in efficiency to noise was 2.24, which supports our choice to undertake the analysis using DEA.

2.3 Sensitivity analysis

DEA is sensitive to the number of inputs and outputs with respect to the number of DMUs used in the analysis. Efficiency scores are likely to be overestimated if the number of DMUs is small relative to the number of inputs (Mohammad, 1998). Also, a large number of DEA inputs and outputs can result in an excessive number of DMUs lying on the frontier, which reduces the ability to identify inefficient DMUs (McCallion *et al.*, 2000). Hollingsworth and Peacock (2008) recommend that the number of DMUs should be at least three times the total number of input and output variables. In our case, working with more than 80 DMUs, we can enter as many as 25 input and output variables in our model, which is more than we need. This, as a result, is not a binding constraint in our case.

The selection of inputs and outputs may result in potential model mis-specification. This can occur in the form of omitted variables or the inclusion of irrelevant variables. No test exists to assess the suitability of a particular model specification (Smith, 1997). Therefore, we carried out a series of separate sensitivity

analyses to assess the robustness of the results that are obtained from the DEA analysis. We examined whether the efficiency and the rank of health regions were robust to seven model specifications (detailed below). Our preferred model, that we then use as the baseline to conduct the sensitivity analyses, is the one with the lowest 'bias' detected following bootstrapping (Section 2.2). We calculated correlations across analyses between the (robust) efficiency estimates of health regions, and their rankings in terms of the (robust) efficiency estimates.

2.4 Explanatory analysis on the robust estimates of efficiency

Once the robust estimates of efficiency are generated for each region, we conduct a multivariate analysis of the determinants of efficiency scores, using a straightforward log-linear regression. Because some of the variables of interest are likely to be correlated and number of factors entered in this second stage is limited due to DMU size (still the number of regions), we used a backward step-wise regression to select variables to include in the second stage. The logarithm of the robust efficiency score was the dependent variable. Results were compared with a forward step-wise regression and they were largely unchanged. The final regression models included only those variables with statistically significant associations with efficiency estimates using a criterion of a p-value <0.2 (Steyerberg *et al.*, 2000).

Although this type of two-stage analysis is widely used (Simar and Wilson, 2007); it seldom is correct for a likely statistical problem with the dependent variable: because efficiency scores have been estimated in a first stage in which the same observations as in the second stage are used and all contribute to the efficiency scores of all, it is likely that the values of the dependent variable for this second stage are serially correlated and thus violate conventional regression assumptions (Simar and Wilson, 2007). Therefore, conventional tests for possible violations of the regression assumptions were undertaken including scatter plots of residuals to detect the magnitude of serial correlation, tests for homoscedasticity, and variance inflation tests for multicollinearity among the independent variables. These tests showed no violations of the conventional regression assumptions.

3. Data

The data used in this study were collected over the period of 2007–2009 from 89 health regions in Canada through various data resources. Health regions are administrative bodies that are legislated by the provincial ministries of health. They tend to be defined by geographical areas and are generally responsible for maintaining the health of their respective populations and for providing health services to residents. More information about the health regions in each province can be found in Table A1.

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

3.1 Inputs and outputs in DEA models

Decisions on the inputs and outputs of health production for this analysis of efficiency were made following document review and stakeholder consultation. Specifically, a scoping review of official statements of health system objectives was undertaken, and in-depth interviews were conducted with senior health ministry officials primarily to understand their views regarding the objectives of the health system (Abelson and Pasic, 2011). Another set of health sector decision makers and stakeholders participated in a facilitated dialogue on health system inputs and outputs (Lavis, 2011). The results of these consultations suggested that there is a good deal of agreement across provinces that the objective of the health system that we should measure efficiency against is to ensure that Canadians have access to effective care when they are sick or need care [Abelson and Pasic, 2011; Lavis, 2011; Canadian Institute for Health Information (CIHI), 2012a]. To measure this objective of ensuring access to care when needed, we used the reduction of PYLL from causes of death that are considered to be amenable to health system intervention (Nolte *et al.*, 2002; CIHI, 2012b). Some examples of amenable, or treatable, causes of death in the Canadian context include sepsis, pneumonia, colorectal cancer, breast cancer in women, hypertensive diseases, asthma and most other respiratory diseases, renal failure, pregnancy and childbirth (CIHI, 2012b).

In this study, we set the age cut-off for defining premature death at 80 years, since about half of all deaths occur before that age. However, premature deaths are currently reported based on deaths occurring before 75 years in Canada and internationally (Nolte and McKee, 2008). Therefore, we tested the sensitivity of the results to different age cut-offs (ages 75 and 85), and the results remained relatively unchanged (see Section 4 and Table A3 for more details).

Rates of PYLL per 100,000 population for the latest available time period (the average for the period 2007–2009) were age-standardised in order to account for the different age structures across regions (Statistics Canada, 2013a). They were also transformed by taking the inverse of PYLL (1/PYLL × 100,000) in order to ensure that, all other things being equal, increased inputs should reduce efficiency and increased outputs should increase efficiency. These estimates were based on Canada's vital statistics database held at Statistics Canada. The other alternative outcomes included the treatable mortality rate and the survival rate using the formula (8,000,000 – PYLL)/PYLL.

Inputs were selected in order to best represent the factors of production as suggested by the health system decision makers. Participants in the stakeholder dialogue were almost unanimous to recommend using dollar values rather than quantities to measure resources: the idea was that, if a region or province manages to pay its physicians less than another region, for the same quality, it will fairly be deemed more efficient. Over-paying providers was clearly identified as inefficient. For the type of inputs included the dialogue agreed that all major components of

health spending – hospitals, physician services (including both family doctors and specialists), pharmaceuticals, residential care facilities and community care – should be included. In order to adjust for possible random fluctuations in spending over time, we take the average of three years of data (2007–2009), where possible. Also, we measure spending on a per capita basis by dividing the spending data by regional population size.

To account for the likely spatial dependence that may bias estimates of efficiency at the regional level (Felder and Tauchmann, 2013), we adjust the spending estimates for patient flow before entering them into the DEA. A patient flow ratio was calculated for hospital services as: the total number of separations (discharges or deaths) from acute care facilities within a given region, multiplied by the estimated cost of each separation (using a resource intensity weight for specialist spending and the average cost per weighted case for the hospital where care was delivered), divided by the number of acute care separations that were only by residents of that particular region (again, multiplied by the estimated costs). Thus, health regions with patient flow ratios greater than one had their per capita estimates adjusted downwards to account for the fact that more was spent on treating patients than was accounted for by the dollars spent to treat the geographically defined population (the region uses resources to provide access to residents of other regions). These ratios were used to adjust two inputs in the DEA: hospital spending, and specialist physician spending, given the majority of specialists work in hospitals.

As stated by Coelli *et al.* (2005), not accounting for the environmental differences between health regions may result in misleading efficiency measurements. However, there is no generally accepted methodology to account for environmental factors (to decide which ones to include in the DEA in the first stage; Coelli *et al.*, 2005). We include a selection of environmental variables as inputs in the DEA estimation consistent with previous studies (Afonso and St. Aubyn, 2006; Liu, 2008), and then incorporate the remaining 'environmental' (non input) variables in a second-stage analysis.

The environmental adjustors were included in the DEA on the basis that they can be considered to be outside of the responsibility of health systems but they significantly affect the outcome measure (treatable mortality). These included the proportion of the population aged 25 to 29 who have a secondary school graduation certificate or equivalent, the proportion of the population who immigrated within the past 10 years, and the proportion of the population who were not Aboriginal. These three measures were based on self-report to the 2006 Census (Statistics Canada, 2013b). Alternate DEA specifications including one, or two, of these environmental adjustors were also conducted, and the results were relatively robust to the choice of the number of adjustors included as inputs (see results). There are other possible environmental variables that could have been included as inputs, such as the proportion of the population that is over the age of 65, or population density, because they can be considered beyond the control of

health system leaders. However, these variables were not found to be significantly associated with the outcome measure, so they were included as factors in the second stage of the analysis.

3.2 Identifying factors related to health system efficiency

The factors that relate to efficiency could be divided into those that relate to what the decisions managers make (managerial factors), and those that reflect the environment in which managers operate (environmental factors). In the former category, we consider indicators of health care being provided that is ineffective, inappropriate, or harmful, public health factors, and indicators of care that is provided that is effective yet more costly than it ought to be (Bentley *et al*., 2008; McGlynn *et al*., 2003; Smith *et al*., 2012). In the latter category, consistent with the literature, we consider the following environmental factors: the distribution of age and gender within the population; physical and geographical characteristics (e.g. population density and the presence of a teaching hospital in the region); and socio-economic conditions (e.g. income inequality and average income). Equitable access to care was also included, and was measured by the income-related inequity in the likelihood of visiting a physician in the past year, based on the concentration index approach (van Doorslaer and Masseria, 2004). Table A1 summarises the factors that were considered in this study, as well as their data source and the time period the data are available.

4. Results

4.1 Summary of regions, inputs and outputs data

There is considerable variation across health regions in their size and characteristics of the populations served. On average, population density across health regions was 249/100 km^2, ranging from 0.13 to 5679. Roughly 14% of population was aged above 65 years, with a range from 8 to 22%. Also 20% of health regions (17 of 84) had teaching hospitals.

Descriptive statistics of the output, inputs and environmental factors used in the analysis are presented in Table 1. Treatable PYLL per 100,000 varies across the health regions even after adjusting for different age structures across the regional populations. Hospitals represent the largest component of spending in the data set with $1719 average per capita expenses, and here again, spending per capita varies across region.

4.2 First stage: DEA results

Seven different DEA models were run to calculate robust efficiency estimates ranging from 0.65 to 0.82 on average across the sample of health regions (Table 2). Model 3, which includes 1/PYLL at age 80, five spending inputs and three

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

Table 1. Description of the health system input and output variables used in DEA, 84 regions

			Range	
Variables	Mean	SE	Minimum	Maximum
Inputs – spending per capita ($)				
Hospitals	1,718.93	520.4	951.32	3,826.39
Prescription drugs	545.6	123.5	288.53	884.25
Physicians	471.15	122.42	177.01	816.72
Residential care facilities	336.42	164	74.2	901.83
Community nurses	54.49	18.51	19.59	98.68
Inputs – environmental adjustors				
Education (per cent with high school or more)	82.33	6.85	63.3	94
Recent immigrants (per cent)	3.16	4.21	0.1	16.7
Non-Aboriginal (per cent)	92.74	9.21	49.5	99.6
Output – age standardised, per 100,000 population				
PYLL from treatable causes (before age 80)	1,666.34	317.92	1,066.60	2,452.60
PYLL from treatable causes (before age 75)	1352.53	367.23	796.00	3302.80
PYLL from treatable causes (before age 85)	2139.17	381.31	1420.80	3066.40
Mortality from treatable causes (before age 80)	88.63	14.33	63.60	136.60
Survival rate from treatable causes (before age 80)	4971.03	932.25	3260.84	7499.47

Survival rate = (8,000,000 – PYLL)/PYLL.

Table 2. Average robust efficiency scores from seven alternate DEA models

Model description	Mean	Minimum	P25	P50	P75	Maximum
Model 1	0.6836	0.4462	0.6190	0.7019	0.7317	0.9115
Model 2	0.7119	0.4721	0.6403	0.7402	0.7461	0.9195
Model 3 (baseline)	0.7248	0.4732	0.6542	0.7563	0.7616	0.9225
Model 4	0.7590	0.5112	0.6892	0.7926	0.7982	0.9367
Model 5	0.6483	0.2134	0.6035	0.6489	0.6892	0.8836
Model 6	0.7248	0.4732	0.6521	0.7562	0.7626	0.9223
Model 7	0.8236	0.5930	0.7618	0.8538	0.8660	0.9644

DEA = data envelopment analysis.
Model 1: spending inputs and recent immigrants; 1/PYLL at age 80.
Model 2: spending inputs, recent immigrants and non-Aboriginal; 1/PYLL at age 80.
Model 3: spending inputs, higher education, recent immigrants and non-Aboriginal; 1/PYLL at age 80.
Model 4: baseline modified to use 1/PYLL at age 85.
Model 5: baseline modified to use 1/PYLL at age 75.
Model 6: baseline modified to use survival rate based on PYLL at age 80.
Model 7: baseline modified to use age-standardised mortality rate based on PYLL at age 80.

environmental adjustors, was used as the baseline model and it yielded a robust estimate that averaged 0.73.

It is important to note that the value of an approach such as DEA is less in the extraction of precise estimates of efficiency, which one should report cautiously,

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

than as an exploratory analytic tool for further analysis (Jacobs *et al.*, 2006). That said, the robust estimates of efficiency did not appear to be sensitive to changes in model specification. Correlations between the efficiency estimates produced by the seven models that varied the input and output variables were high (0.70 to 0.99). Correlation coefficients for the estimates and rankings produced by the seven different models are presented in Table A2. It is also worth noting that this average efficiency score of 0.73 in health regions in Canada is not far from the 0.80 estimated efficiency found for Canada by the OECD, based on a comparison across countries (Joumard *et al.*, 2010).

In order to test the extent to which our methods of adjusting for patient flow were effective in reducing the possible bias arising from spatial dependence (Felder and Tauchmann, 2013), we compared the average and range of efficiency scores by regions with and without a teaching hospital, and between regions that are relatively more urbanised compared with those that are more rural. Table 3 shows that there is no systematic pattern, either nationally, or by province, of less efficient urban regions surrounded by more highly efficient regions.

4.3 Second stage: log-linear regression results

Table 3 reports the results of the backward step-wise regression of managerial and environmental factors on efficiency scores produced from the seven models. Table A1 lists all the variables that were considered in this second stage analysis, as well as the correlation coefficients between each variable and the robust efficiency estimate.

As shown in Table 4, among the environmental factors, a higher average income of residents in a region was negatively related with efficiency in five of the models. This implies that in richer regions there may be more spending on health services that does not necessarily translate into reductions in treatable causes of death, after controlling for variations in education, immigrant and Aboriginal concentration (as these were included as environmental adjustors in the calculation of DEA estimates). Also there was a negative association between inequitable access to physician services and efficiency in four of the models: regions with greater inequity in favour of higher income groups were less likely to be efficient than regions with a more equitable distribution of physician services. In other words, health system managers do not need to trade-off equity for efficiency; these results suggest that they can improve performance by improving access to care for individuals with lower income.

Several public health factors were associated with efficiency. The results from the regressions suggested that an increase in the prevalence of smoking by 10 percentage points would decrease efficiency by 10%; an increase in the proportion physically inactive by 10 percentage points would decrease efficiency by a range of 5–9%; an increase in the proportion with multiple chronic conditions by 10 percentage points would decrease efficiency by between 10 and 18%; an

Table 3. Efficiency scores by province, presence of a teaching hospital and urban/rural regions

	Comparison by presence of teaching hospitals						Comparison by rural and urban					
	With teaching hospital			No teaching hospital			Rural			Large urban population centre		
Province	Mean	Minimum	Maximum	Mean	Minimum	Maximum	Mean	Minimum	Maximum	Mean	Minimum	Maximum
AB	0.691	0.645	0.737	0.669	0.653	0.682	0.669	0.653	0.682	0.691	0.645	0.737
BC	0.801	0.801	0.801	0.822	0.654	0.923	0.81	0.654	0.923	0.837	0.801	0.87
MB	0.542	0.542	0.542	0.711	0.597	0.762	0.711	0.597	0.762	0.542	0.542	0.542
NB	0.640	0.640	0.640	0.756	0.700	0.860	0.736	0.640	0.860			
NL	0.501	0.501	0.501	0.722	0.651	0.76	0.722	0.651	0.76	0.501	0.501	0.501
NS	0.727	0.727	0.727	0.692	0.473	0.835	0.692	0.473	0.835	0.727	0.727	0.727
ON	0.637	0.587	0.731	0.67	0.53	0.758	0.599	0.53	0.686	0.702	0.599	0.758
QC	0.795	0.675	0.866	0.739	0.576	0.827	0.739	0.576	0.866	0.772	0.675	0.844
SK	0.680	0.659	0.700	0.745	0.614	0.811	0.745	0.614	0.811	0.680	0.659	0.700
Canada	0.678	0.501	0.866	0.737	0.473	0.923	0.722	0.473	0.923	0.731	0.501	0.870

Table 4. Regression results with environmental and managerial factors as predictors

	Model 1	Model 2	Model 3 (baseline)	Model 4	Model 5	Model 6	Model 7
Contextual factors							
Average income (logarithm)	−0.239 (−0.113)**	−0.198 (−0.11)*	−0.304 (−0.098)***	−0.187 (−0.09)**		−0.301 (−0.098)**	
Inequity in physician visit			−1.737 (−0.862)**	−2.237 (−0.761)**		−1.712 (−0.862)*	−1.703 (−0.685)**
Aboriginal (per cent)	−0.005 (−0.002)**	−0.001 (−0.002)					
Male (per cent)					7.263 (−4.117)*		5.273 (−1.636)**
Seniors (per cent 65 and over)					1.863 (−0.912)**		0.908 (−0.416)**
No teaching hospital in the region							0.097 (−0.028)**
Managerial and public health factors							
Smoking (per cent)			−0.010 (−0.004)**			−0.01 (−0.004)**	
Physical inactivity (per cent)	−0.009(−0.003)***	−0.008 (−0.003)***	−0.007 (−0.002)***	−0.005 (−0.002)**	−0.007 (−0.004)**	−0.007 (−0.002)**	
3+ chronic conditions (per cent)			−0.013 (−0.004)***	−0.01 (−0.004)**	−0.02 (−0.007)**	−0.013 (−0.004)**	
Obese (per cent)	−0.008 (−0.004)**	−0.012 (−0.004)***		−0.008 (−0.004)**			−0.01 (−0.002)***
Overweight (per cent)	−0.012 (−0.005)**				−0.015 (−0.008)*		−0.008 (−0.004)**
30-day Readmission (per 100)	−0.016 (−0.011)	−0.018 (−0.01)*	−0.021 (−0.009)**			−0.021 (−0.009)**	
GPs (per cent of all physicians)	0.006 (−0.001)***	0.005 (−0.001)***	0.005 (−0.001)***	0.004 (−0.001)***	0.005 (−0.002)**	0.005 (−0.001)***	
Average ALC length of stay in acute hospital (days)	−0.002 (−0.001)**	−0.002 (−0.001)***	−0.002 (−0.001)***	−0.001 (−0.001)**	−0.004 (−0.001)***	−0.002 (−0.001)**	
ALC cases (per cent total) inpatient cases)					0.009(0.004)**		
Nursing hours per weighted inpatient case		0.002 (−0.002)		0.002 (−0.002)			
Administrative services expense (per cent total hospital expenses)					−0.027 (0.016)		−0.021 (−0.009)**
Constant	2.925 (−1.304)**	2.116 (−1.272)*	3.619 (−1.126)***	2.054 (−1.052)*	−3.147 (2.08)	3.584 (−1.126)***	−2.516 (−0.812)**
R^2	44%	39%	46.10%	44.40%	51.90%	46.20%	47.20%

***Indicates statistical significance at $p < 0.001$ level, **indicates $p < 0.05$ level.

increase in the proportion of obesity in the population by 10 percentage points would decrease efficiency by 8–11% and an increase in the proportion of overweight by 10 percentage points would decease efficiency by 8–11%.

The results also suggest that several managerial factors were significantly associated with efficiency. The overall rate of unplanned readmissions to hospital within 30 days was inversely related to efficiency: an increase in the rate of 30-day readmission per 100 patients of 10 percentage points was associated with a reduction in efficiency of between 16 and 19%. Other factors include the relative density of GPs in a region compared with specialist physicians (an increase by 10 percentage points in that relative density would improve efficiency by 4–6%), the average length of stay among patients designated as 'alternate level of care' (ALC; an increase by 10 days of the ALC length of stay decreased efficiency by 1–4%) and the proportion of ALC cases of total hospital patient cases (an increase by 10 percentage points in the proportion of ALC cases in total inpatient cases decreased efficiency by 9%). Finally, an increase by 10 percentage points in the acute hospital administrative services expenses was associated with a decrease in efficiency of 19%.

5. Discussion

In this paper, we adopted a 'top-down' approach to measuring health system efficiency with the objective of facilitating uptake of this evidence by decision makers and system managers. We followed best practice in the field in order to move towards the standardisation of efficiency measurement (Jacobs, 2006). In this work, we made all decisions systematic and documented, rather than being content to analyse data as they existed and adjusting concepts to existing observations. Decisions to be made relate to the definition of outputs (what is the objective of the health care system that efficiency should be measured against?), inputs (what resources are available and in what units should they be measured?), decision-making units and the empirical strategy to estimate the frontier.

As far as the first three decisions (outputs, inputs, DMU) are concerned, two options are available: one is to force definitions from without, based on what the analyst thinks ought to be the right definition; the other is to rely on a *sui generis* definition and to ask stakeholders themselves how they would define the outputs and resources. Our study at the Canadian level shows that stakeholders have different views but a clearly dominant theme emerged from the interviews and a consensus was relatively easily reached during the stakeholder dialogue: it was clear that 'access to timely care when sick' was the objective they should be assessed against.

In order to generate results that are meaningful to policy makers, our main objective in this study was not so much to estimate the level of inefficiency *per se* as it was to understand how it varies across DMUs and what seems to explain these variations. Moreover, in order to inform decision makers we categorised these factors into those that are within the control of managers, which can be used to help local decision makers improve performance, and those that are characteristics

of the operating conditions, which would be relevant to decision makers at a higher level interested in allocating resources to compensate regions operating in more challenging environments.

Overall, this two-stage approach to measuring regional-level efficiency in Canada allowed us to identify several important managerial and environmental factors related to system efficiency. Consistent with other studies of efficiency across countries (Ryltseva, 2010; Hadad *et al.*, 2013) and within US and Canadian states and provinces (Liu, 2008), public health factors were significantly and negatively associated with efficiency scores: our main conclusion is therefore that the health care system must spend more to prevent the same number of premature deaths when the population is less healthy, either because treatments of treatable causes of death are more expensive when patients are less healthy, or because the system has to spend resources on other things than preventing premature deaths because more individuals are not healthy.

A number of managerial factors were identified as being significantly associated with efficiency in this study. These include factors related to appropriate and effective health care (as measured by fewer readmissions to hospital) as well as the use of overly costly inputs (as measured by the ratio of GPs to total physicians in a region), and the use of hospitals among patients who have been identified as being better suited to another, less expensive care setting such as long-term care institution or at home. These particular indicators have received increasing attention in recent years. For instance, studies show that while readmissions relate in part to characteristics and management practices of hospitals (Joynt and Jha, 2011; Stukel *et al.*, 2012), they could potentially be reduced through improved coordination efforts and partnerships between hospital and community care (Boutwell *et al.*, 2013), as well as timely and effective follow-up care with a physician (Jencks *et al.*, 2009). Therefore, continued efforts to make improvements here could have the effect of improving efficiency at the system level.

This study also identified several characteristics of the environment in which health systems are operating that were significantly associated with efficiency in this study. Most of the models identified higher average income of the population to be negatively associated with efficiency, in contrast to studies across the 191 countries in the WHO region (Greene, 2004). In other words, after adjusting for the level of education in the region as well as other characteristics associated with socio-economic status included in the health production function such as the proportion of the population who recently immigrated to Canada and who do not identify as an aboriginal, higher average income does not increase efficiency. This implies that in richer regions there may be more spending on health service that does not necessarily translate into saving more lives. The results of the study also suggest there may be efficiency gains associated with more equitable access to physician care; therefore, it can also be argued that more effective health systems provide more equitable access and, as a result, treat the most vulnerable who are also those who can benefit the most from interventions.

Our study is a first step toward top-down analyses of efficiency of health systems that are able to help decision makers make improvements. One key aspect of making our conclusions helpful is the systematic way we made decisions at each step. One limitation of the present study is that such a systematic approach could not be maintained all along and some decisions had to be made based on inconclusive information. For instance, even though we received a clear mandate from stakeholders on the definition of inputs to be entered in the first step, there are no clear rules in the literature regarding which environmental factors should be included as well, and which ones should be used as factors in the second stage. The rationale for including some environmental characteristics in the first step (that calculates efficiency scores) is to ensure DMUs are comparable; however, there is some arbitrariness involved in the definition of comparable and no clear systematic rule would allow the analyst to know what variables to control. A rule of thumb is that everything not included in the first step is implicitly considered to be the responsibility of the health system: for instance, including smoking rates in the first step would amount to consider that a health care system has to prevent as many treatable premature deaths it can, knowing it has a given proportion of smokers in its population. Decreasing the smoking rate would not improve its efficiency score. On the other hand, entering smoking rates in the second stage relies implicitly on the idea that a DMU could improve its efficiency score by decreasing the prevalence of smoking. We decided to include education, the proportion of recent immigrants and the proportion of non-Aboriginals in the population of the region as three characteristics we wanted to control in the first step. This may seem reasonable as no one would argue that DMUs should be held responsible for them. However, we left income and the proportion of seniors out of the first step and entered them as factors in the second step, even though DMUs cannot reasonably be held responsible for their values. The decision was essentially pragmatic, as a limited number of environmental variables only could be entered in the DEA and we focussed on those that had a significant association with the outcome measure.

Another pragmatic decision we had to make regarded the choice of estimation strategy for the frontier (DEA vs SFA). Our decision was partly based on systematic reasons making DEA a better alternative in the sense that it imposes almost no non-testable specification assumptions (whereas SFA relies on strong non-testable specification assumptions) and its main drawbacks (lack of randomness) can now be overcome by the Simar Wilson method that we applied in this work. However, the decision was also partly based on pragmatic considerations related to the data.

Future research in this area could examine whether the factors related to efficiency in Canada are similar to those found in other jurisdictions. Also given that the findings from the regression are largely exploratory, qualitative research could be used to conduct more in-depth analysis of best performers in order to share best practices across regions. Finally, future research could improve the

outcome variable by measuring preventable years of life lost both to death and to poor quality of life that better reflects the concept of timely access to care when sick, since not getting access to care can either kill or reduce the quality of life of those who survive.

Acknowledgments

This study was funded by the Canadian Institute for Health Information (CIHI). The authors would like to acknowledge the tremendous support and guidance provided by CIHI, and especially by Jeremy Veillard, the Vice President of Research and Analysis, CIHI. The authors are also grateful to the two anonymous reviewers for their helpful comments and suggestions.

References

Abelson, J. and D. Pasic (2011), *Health System Efficiency Project: A Qualitative Study of Provincial and Territorial Health Ministry Perspectives*, Ottawa, ON: CIHI.

Afonso, A. and M. St. Aubyn (2006), Relative Efficiency of Health Provision: A DEA Approach With Non-Discretionary Inputs. http://www.iseg.utl.pt/departamentos/economia/wp/wp332006deuece.pdf [19 September 2013].

Badunenko, O., D. J. Henderson and S. B. Kumbhakar (2012), 'When, where and how to perform efficiency estimation', *Journal of the Royal Statistical Society*, **175**(4): 863–892.

Bentley, T., R. Effros, K. Palar and E. B. Keeler (2008), 'Waste in the U.S. health care system: a conceptual framework', *The Milbank Quarterly*, **86**(4): 629–659.

Boutwell, A. E., M. Bihrle Johnson, P. Rutherford, S.R. Watson, N. Vecchioni, B. S. Auerbach, P. Griswold, P. Noga and C. Wagner (2013), 'An early look at a four-state initiative to reduce avoidable hospital readmissions', *Health Affairs*, **30**(7): 1272–1280.

Canadian Institute for Health Information (CIHI) (2012a), *Developing a Model for Measuring the Efficiency of the Health System in Canada*, Ottawa, ON: CIHI.

Canadian Institute for Health Information (CIHI) (2012b), *Health Indicators 2012*, Ottawa, ON: CIHI.

Charnes, A., W. Cooper and E. Rhodes (1978), 'Measuring the efficiency of decision-making units', *European Journal of Operational Research*, **2**: 429–444.

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell and G. E. Battese (2005), *An Introduction to efficiency and productivity analysis*, 2nd edn, New York: Springer.

Felder, S. and H. Tauchmann (2013), 'Federal state differentials in the efficiency of health production in Germany: an artifact of spatial dependence?' *European Journal of Health Economics*, **14**(1): 21–39.

Fried, H. O., K. C. A. Lovell and S. S. Schmidt (2008), *The Measurement of Productive Efficiency and Productivity Growth*, New York: Oxford University Press Inc.

Greene, W. (2004), 'Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems', *Health Economics*, **13**: 959–980.

Hadad, S., Y. Hadad and T. Simon-Tuval, Ben-Gurion U. Negev (2013), 'Determinants of healthcare system's efficiency in OECD countries', *European Journal of Health Economics*, **14**(2): 253–265.

Häkkinen, U., T. Iversen, M. Peltola, T. T. Seppälä, A. Malmivaara, É. Belicza, G. Fattore, D. Numerato, R. Heijink, E. Medin and C. Rehnberg (2013), 'Health care performance comparison using a disease-based approach: the EuroHOPE project', *Health Policy*, **112**(1–2): 100–109.

Hollingsworth, B. (2008), 'The measurement of efficiency and productivity of health care delivery', *Health Economics*, **17**(10): 1101–1128.

Hollingsworth, B. (2012), 'Revolution, evolution, or status quo? Guidelines for efficiency measurement in health care', *Journal of Productivity Analysis*, **37**: 1–5.

Hollingsworth, B. and P. Smith (2003), 'The use of ratios in data envelopment analysis', *Applied Economics Letters*, **10**(11): 733–735.

Hollingsworth, B. and S. J. Peacock (2008), *Efficiency Measurement in Health and Health Care*, New York: Routledge.

Jacobs, R., P. C. Smith and A. Street (2006), *Measuring Efficiency in Health Care: Analytic Techniques and Health Policy*, Cambridge, UK: Cambridge University Press.

Jencks, S. F., M. V. Williams and E. A. Coleman (2009), 'Rehospitalisations among patients in the Medicare fee-for-service program', *New England Journal of Medicine*, **360**(14): 1418–1428.

Joumard, I., P. Hoeller, C. André and C. Nicq (2010), *Health Care Systems: Efficiency and Policy Settings*, Paris, France: OECD Publishing.

Joynt, K. E. and A. K. Jha (2011), 'Who has higher readmission rates for heart failure, and why?: implications for efforts to improve care using financial incentives', *Circulation: Cardiovascular Quality and Outcomes*, **4**: 53–59.

Lavis, J. N. (2011), *Dialogue Summary: Measuring Health System Efficiency in Canada*, Hamilton, ON: McMaster Health Forum.

Liu, C. (2008), *Three Essays on Health Economics*, Guelph, ON: University of Guelph.

McCallion, G. M., J. C. Glass, R. Jackson, C. A. Kerr and D. G. McKillop (2000), 'Investigating productivity change and hospital size: a nonparametric frontier approach', *Applied Economics*, **32**: 161–174.

McGlynn, E. A., S. M. Asch, J. Adams, J. Kessey, J. Hicks, A. DeCristofaro and E. A. Kerr (2003), 'The quality of health care delivered to adults in the United States', *New England Journal of Medicine*, **348**(26): 2635–2645.

Mohammad, R. Alirezaee (1998), sampling size and efficiency bias in data envelopment analysis', *Journal of Applied Mathematics& Decision Sciences*, **2**(1): 51–64.

Newhouse, J. P. (1994), 'Frontier estimation: how useful a tool for health economics', *Journal of Health Economics*, **13**(3): 317–322.

Nolte, E., R. Scholz, V. Shkolnikov and M. McKee (2002), 'The contribution of medical care to changing life expectancy in Germany and Poland', *Social Science and Medicine*, **55**: 1907–1923.

Nolte, E. and M. McKee (2008), 'Measuring the health of nations: updating an earlier analysis', *Health Affairs*, **27**: 58–71.

Organisation for Economic Cooperation and Development (OECD) (2011), Health at a Glance 2011: OECD Indicators. http://www.oecd.org/els/health-systems/49105858.pdf [24 September 2013].

Ryltseva, A. (2010), Determinants of health status and the difference in efficiency of the utilisation of health care inputs in countries of the former Soviet Union, CEU eTD Collection, http://www.etd.ceu.hu/2010/ryltseva_anna.pdf [21 November 2013].

Simar, L. and P. W. Wilson (1998), 'Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models', *Management Science*, **44**(1): 49–61.

Simar, L. and P. W. Wilson (2007), 'Estimation and inference in two-stage, semi-parametric models of production processes', *Journal of Econometrics*, **136**: 31–64.
Smith, P. (1997), 'Model misspecification in data envelopment analysis', *Annals of Operations Research*, **73**: 233–252.
Smith, P. and A. Street (2005), 'Measuring the efficiency of public services: the limits of analysis', *Journal of the Royal Statistical Society*, **168**(2): 401–417.
Smith, M., R. Saunders, L. Stuckhardt and J. M. E. McGinnis (2012), *Best Care at Lowest Cost: The Path to Coninuously Learning Health Care in America*, Washington, DC: National Academies Press.
Statistics Canada (2013a), Table 102-4311, Premature and potentially avoidable mortality, three-year average, Canada, provinces, territories, health regions and peer groups, occasional, CANSIM (database). 12 April. Statistics Canada, Ottawa, ON. 16 September.
Statistics Canada (2013b), Table 109-0300, Census indicator profile, Canada, provinces, territories, health regions (2001 boundaries) and peer groups. http://www5.statcan.gc.ca/cansim/a26?lang=eng&id=1090300 [21 November 2013].
Steyerberg, E. W., M. J. C. Eijkemans, F. E. Harrell and J. D. F. Habbema (2000), 'Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methodsin small data sets', *Statistics in Medicine*, **19**: 1059–1079.
Stukel, T., E. S. Fisher, D. A. Alter, A. Guttmann, D. T. Ko, K. Fung, W. P. Wodchis, N. N. Baxter, C. C. Earle and D. S. Lee (2012), 'Association of hospital spending intensity with mortality and readmission rates in Ontario hospitals', *Journal of the American Medical Association*, **307**(10): 1037–1045.
Valdmanis, V. G., M. D. Rosko and R. L. Mutter (2008), 'Hospital quality, efficiency, and input slack differentials', *Health Services Research*, **43**(5 Pt 2): 1830–1848.
Van Doorslaer, E. and C. Masseria (2004), *Income-Related Inequality in the Use of Medical Care in 21 OECD Countries*, Paris, France: OECD.
Wilson, P. W. (1993), 'Detecting outliers in deterministic nonparametric frontier models with multiple outputs', *Journal of Business and Economic Statistics*, **11**(3): 319–323.
Wranik, D. (2012), 'Healthcare policy tools as determinants of health-system efficiency: evidence from the OECD', *Health Economics, Policy and Law*, **7**: 197–226.
Zere, E., D. McIntyre and T. Addison (2001), 'Technical efficiency and productivity of public sector hospitals in three South African provinces', *South African Journal of Economics*, **69**(2): 336–358.

Appendix
Detailed methods

In this study, we employ an output-orientation approach to DEA, based on the assumption that ministries of health are interested in getting closer to achieving their objectives given a fixed budget. Moreover, we assume variable returns to scale (VRS) which is considered to be appropriate for analysis in the health sector, because of evident diminishing marginal returns (adding more resources increases output by less) when one examines the relationship between spending on aggregate and health outcomes (OECD, 2011), and for technical reasons relating to the use of ratio data (such as spending per capita; Hollingsworth and Smith, 2003).

The output-oriented DEA model under VRS specification for the DMU_0 is specified as the following mathematical formula:

$$\min_{\varnothing,z} \varnothing_0$$

$$\text{Subject to :} \qquad \frac{-y_{r0}}{\varnothing_0} + \sum_{j=1}^{n} z_j y_{rj} \geq 0 \qquad r = 1, 2, ..., m$$

$$x_{i0} - \sum_{j=1}^{n} z_j x_{ij} \geq 0 \qquad i = 1, 2, ..., k$$

$$z_j \geq 0, \qquad j = 1, 2, ..., n$$

$$\sum_{j=1}^{n} z_j = 1;$$

where y_{rj} is the vector of outputs for DMU_j, x_{ij} the vector of inputs for DMU_j and the values of outputs and inputs should be nonnegative. z_j represents a vector of weights attached to each DMU_j in the comparison group from the perspective of DMU_j (Charnes *et al.*, 1978) and is determined by the above linear programming problem. $\varnothing$ is the efficiency score that measures technical efficiency (TE) of a DMU and satisfies $\varnothing \leqslant 1$. A TE score 1 indicates that the DMU lies on the production frontier, i.e., it is technically efficient. With the TE <1, the DMU is inside the frontier, i.e., it is technically inefficient, and the more inefficient, the lower the TE score.

We then applied a statistical outlier detection method developed by Wilson (1993). This approach adopts the influence function based on the geometric volume spanned by the sample observations and the sensitivity of this volume with respect to deleting suspicious observations from the sample (Wilson, 1993). The results can be graphically analysed where the influence function (log-ratios) is expressed as a function of the number (i) of deleted observations (l). The larger the distance from the smallest ratio, the more likely there is an outlier among the remaining observations. The detailed mathematical description of the method can be found in Wilson (1993) or Fried *et al.* (2008).

The process of bootstrap DEA (output orientation) proceeds in the following steps:

(1) Employ DEA to the original data to calculate efficiency scores (point estimate): $\widehat{\varnothing}_j$

(2) Generate a random pseudo sample with replacement of size n from the empirical distribution of efficiency scores (F distribution) $\widehat{\varnothing}_j$, providing $\varnothing^*_{1b,}...\varnothing^*_{nb}$.

(3) Obtain a bootstrap set of pseudo-inputs ($x^*_{jb} = \frac{\hat{\varnothing}_j x_j}{\varnothing^*_{jb}}, j = 1, ...n$)using the ratio of the original efficient input level, i.e., the product of original efficiency score and the original input ($\widehat{\varnothing}_j x_j$) and the pseudo-efficiency scores ($\varnothing^*_{jb}$).

(4) Adapt DEA to this new set of observations, composed of the pseudo-inputs (x^*_{jb}) from step 3 and the same set of outputs then calculate the bootstrapped efficiency scores ($\widehat{\varnothing}^*_{j,b}$).

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

(5) Repeat steps 1–4 B times to generate a distribution of size B of bootstrapped scores for statistical inference.

The estimated bootstrap bias is calculated by B times Monte-Carlo simulations using:

$$\widehat{bias_j} = \frac{1}{B}\sum_{b=1}^{B} \widehat{\emptyset}^*_{j,b} - \widehat{\emptyset}_j = \overline{\emptyset}^*_j - \widehat{\emptyset}_j$$

A bias corrected estimator of efficiency $\emptyset_j$ is $\widetilde{\emptyset}_j = \widehat{\emptyset}_j - \widehat{bias_j} = 2\widehat{\emptyset}_j - \overline{\emptyset}^*_j$

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

Table A1. Summary of health regions number and functions, by province

Province	Health regions	Brief description of regional responsibilities
Newfoundland and Labrador	4 Regional Integrated Health Authorities	Authorities are responsible for delivering direct care to individuals in hospitals, long-term care facilities, community-based offices and clinics, and through public health and community support services. They manage and allocate resources, including funds provided by government for health services and community services, in accordance with legislation and provincial spending frameworks
Prince Edward Island (PEI)	None	Health PEI is responsible for the operation and delivery of publicly funded health services in Prince Edward Island, as of 2010
Nova Scotia	9 District Health Authorities (grouped into 6 due to data availability)	The roles of the District Health Authorities are to govern, plan, manage, monitor, evaluate and deliver health services in a health district; and to maintain and improve the health of the residents of the health district
New Brunswick	7 Health Regions (recently merged into 2 Health Networks)	Health Networks are responsible for managing and delivering a variety of services including: Community Health Centre Services, Extra Mural Services and most Public Health Services. Services are offered in a variety of settings; at hospitals on both an inpatient and outpatient basis, at home, in schools, in clinics and in other community
Quebec	18 Régions sociosanitaires (15 regions due to data availability)	Regional authorities are responsible for coordinating and implementing health services and social services in their region, especially with regard to funding, human resources deployment and access to specialised services. They also facilitate the development and management of local health and social services resources. Finally, they have responsibilities in regional public health, including monitoring of health and well-being, promotion, prevention and protection
Ontario	14 Local Health Integration Networks (LHINs)	LHINs work with local health providers and community members to determine the health service priorities of regions. They plan, integrate and fund local health services, monitor quality and access in their geographically defined areas

Table A1. (*Continued*)

Province	Health regions	Brief description of regional responsibilities
Manitoba	10 Health Regions (5 Health Authorities since 2012)	A regional health authority is responsible for providing for the delivery of and administering health services to meet the health needs in its health region
Saskatchewan	13 Regional Health Authorities (grouped into 11 due to data availability)	The regional health authorities were established under The Regional Health Services Act. The Act came into force on 1 August 2002 and sets out the powers and responsibilities for both the Minister of Health and the regional health authorities. Health services are primarily delivered through these regional health authorities and their affiliates
Alberta	1 Health Authority, with 5 Zones	Alberta Health Services (AHS) is the provincial health authority responsible for overseeing the planning and delivery of health supports and services. Health system performance is measured and monitored at the level of the 5 Zones
British Columbia	5 Health Authorities, with 16 Health Service Delivery Areas	Five regional health authorities deliver a full continuum of health services to meet the needs of the population within their respective geographic areas. Within these five regional health authorities are 16 health service delivery areas, with boundaries that reflect the province's geography, as well as patient and physician referral patterns. An additional health authority operates at the provincial level (Provincial Health Services Authority) and is responsible for managing the quality, coordination and accessibility of selected province-wide health programs and services
Canada	97 regions (89 regions in the study due to data availability)	

For most data sources in Canada, health region level data are not available for some health regions, so data are grouped with neighbouring regions.
The provincial ministries of health are generally responsible for setting policy, legislation and standards for the provincial health system, paying physicians, organising provincial prescription drug plans, allocating funding to regional entities.

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

Table A2. Summary of the spending input data sources and limitations

Input category	Data source	Year	Exceptions/limitations
Hospitals	Canadian MIS Database (CMDB)	2007/2008–2009/2010	Quebec regional expenditures are estimated using a Quebec-specific methodology as hospital and non-hospital expenditures are not disaggregated in the expenditure data that CIHI receives
Physicians	National Physicians Database (NPDB) and Scott's Medical Database (SMDB)	2007/2008–2009/2010	Physician distributions from SMDB are used to allocate alternative payments to regions for all provinces, and to allocate fee-for-service (FFS) payments to regions in 1 province
Residential care facilities (RCF)	Statistics Canada Residential Care Facilities Survey	2008	Data from 2007 and 2009 had a relatively high proportion of missing geographic identifiers. To ensure comparability across jurisdictions, we include facilities providing all levels of care, including those that provide no medical or nursing supervision
Community care	Statistics Canada, Census	2006	Earnings of nurses are used to approximate community care spending
Prescription drugs	IMS Brogan Canada	2010	Data prior to 2010 is not available

https://doi.org/10.1017/S1744133115000274 Published online by Cambridge University Press

Table A3. Description of variables that were included in the second stage step-wise regression, and correlations with efficiency scores (n = 84)

Contextual factors	Year(s)	Mean	SD	Minimum	Maximum	Correlation
Men	2006	0.5	0.01	0.48	0.52	0.11
Population aged 65 years and older	2006	0.14	0.03	0.08	0.22	0.03
Population density	2006	249.08	896.03	0.13	5,679.00	− 0.05
Long-term unemployment	2006	4.1	2.85	1.3	16	− 0.19
Average income ($)	2006	32,164	5,287	23,611	50,111	− 0.02
Income-related inequality in likelihood of a physician visit	2007–2008	0.02	0.01	− 0.02	0.05	− 0.29*
Income inequality (Gini index)	2007–2008	0.26	0.14	0.04	0.65	0.16
No teaching hospitals in the region	FY2008/2009	0.8	0.4	0	1	0.26*
Clinical indicators						
Daily smoking (per cent of population aged 12 and over)	2007–2008	18.78	3.59	10.3	26.6	− 0.27*
Obese (per cent of population aged 18 and over)	2007–2008	19.91	5.06	6.3	30.8	− 0.31*
Overweight (per cent of population aged 18 and over)	2007–2008	35.65	3.31	23.7	43.7	− 0.19
3 or more chronic conditions (per cent of population aged 12 and over)	2007–2008	24.34	3.7	16.5	32.2	− 0.24*
Physically inactive (per cent of population aged 12 and over)	2007–2008	49.16	5.85	29.4	61.3	− 0.24*
ACSC admissions per 100,000 population	FY 2007/2008–2009/2010	415.66	151.74	185.67	880.33	− 0.2
Repeat hospitalisations for mental illness (per cent of patients with at least one hospitalisation for mental illness)	FY 2007/2008–2009/2010	10.93	2.68	4.1	18.1	0.1
C-Sections (per cent of total births)	FY 2007/2008–2009/2010	23.58	7.04	0	37.36	− 0.03
VBAC rate (per 100 births)	FY 2007/2008–2009/2010	15.64	7.39	2.97	34.71	− 0.01
Overall 30-day readmission rate (per cent of all hospital discharges)	FY 2009/2010	9.33	1.84	6.24	15.96	− 0.02
30-day readmissions (surgical)	FY 2009/2010	6.72	1.57	1.2	14.12	− 0.01
30-day readmissions (pediatric)	FY 2009/2010	5.54	1.46	1.56	9.44	− 0.22*
30-day readmissions (medical)	FY 2009/2010	13.64	1.64	10.38	18.53	0.06

Table A3. (*Continued*)

Contextual factors	Year(s)	Mean	SD	Minimum	Maximum	Correlation
Operational indicators						
GPs (per cent of total physicians)	2007–2009	63.75	14.37	34.53	98.93	0.19
Nursing inpatient services total worked hours per inpatient case	FY 2007/2008–2009/2010	50.99	7.43	39.63	72.74	– 0.11
Average typical length of stay in acute hospital (days)	FY 2007/2008–2009/2010	4.77	1.61	2.55	11.68	0.07
Average ALC length of stay in acute hospital (days)	FY 2007/2008–2009/2010	9.28	8.73	3.84	68.67	0.01
ALC cases (per cent total inpatient cases)	FY 2007/2008–2009/2010	4.86	5.52	0.45	33.4	0.002
Average occupancy rate in acute hospitals	FY 2007/2008–2009/2010	81.1	12.89	22.4	96.71	– 0.04
Average spending on administration as a per cent of total hospital spending	FY 2007/2008–2009/2010	5.32	1.23	3.2	8.9	– 0.23*
Average cost per weighted case ($), acute hospitals	FY 2009/2010	5123.27	711.34	3555.22	7197.14	– 0.12

*Is statistically significant at $p < 0.05$.

Table A4. Correlations between efficiency scores from seven model specifications

	Model 1	Model 2	Model 3 (baseline)	Model 4	Model 5	Model 6	Model 7
Correlation coefficients of efficiency scores							
Model 1: spending inputs + recent immigrants	1						
Model 2: spending inputs + recent immigrants + non-Aboriginal	0.94*	1					
Model 3: spending inputs + recent immigrants + non-Aboriginal + education	0.92*	0.97*	1				
Model 4: baseline modified to use 1/PYLL at age 85	0.91*	0.96*	0.99*	1			
Model 5: baseline modified to use 1/PYLL at age 75	0.78*	0.81*	0.83*	0.82*	1		
Model 6: baseline modified to use survival rate based on PYLL at age 80	0.92*	0.97*	1.0*	0.99*	0.83*	1	
Model 7: baseline modified to use ASMR at age 80	0.76*	0.81*	0.84*	0.90*	0.71*	0.84*	1
Spearman rank correlation coefficients of region rankings							
Model 1: spending inputs + recent immigrants	1						
Model 2: spending inputs + recent immigrants + non-Aboriginal	0.94*	1					
Model 3: spending inputs + recent immigrants + non-Aboriginal + education	0.89*	0.93*	1				
Model 4: baseline modified to use 1/PYLL at age 85	0.89*	0.93*	0.98*	1			
Model 5: baseline modified to use 1/PYLL at age 75	0.82*	0.86*	0.88*	0.87*	1		
Model 6: baseline modified to use survival rate based on PYLL at age 80	0.89*	0.94*	0.97*	0.96*	0.86*	1	
Model 7: baseline modified to use ASMR at age 80	0.75*	0.79*	0.82*	0.86*	0.77*	0.81*	1

*Is statistically significant at $p < 0.05$.