

**A SYSTEMATIC REVIEW OF STUDIES COMPARING HEALTH OUTCOMES  
IN CANADA AND THE UNITED STATES**

**Gordon H. Guyatt<sup>§§\*</sup> MSc, M.D., P.J. Devereaux<sup>§§\*</sup> M.D, Joel Lexchin<sup>^</sup> MSc, MD,  
Samuel B. Stone MD<sup>§§</sup>, Armine Yalnizyan<sup>#</sup> MA, David Himmelstein<sup>@</sup> MD,  
Steffie Woolhandler<sup>@</sup> MD, MPH, Qi Zhou<sup>§§</sup> PhD, Laurie J. Goldsmith<sup>\*\*\*</sup> MSc,  
Deborah J. Cook<sup>§§\*</sup> MSc, M.D., Ted Haines<sup>§§‡</sup> MSc, M.D., Christina Lacchetti<sup>§§</sup>  
MHSc., John N. Lavis<sup>§§ £</sup> M.D, Ph.D., Terrence Sullivan<sup>‡</sup> Ph.D., MSc., Ed Mills<sup>§§</sup>  
DPH, MSc, Shelley Kraus<sup>§§</sup>, Neera Bhatnagar<sup>¶¶</sup> BSc., M.L.I.S.**

<sup>\*</sup>Department of Medicine, McMaster University, Hamilton, Ontario, Canada

<sup>§§</sup>Department of Clinical Epidemiology and Biostatistics, McMaster University,  
Hamilton, Ontario, Canada

<sup>@</sup> Department of Medicine, Cambridge Hospital and Harvard Medical School,  
Cambridge, Mass.

<sup>£</sup> Hamilton, Institute for Work & Health, Toronto, and Population Health  
Program, Canadian Institute for Advanced Research, Toronto, Ontario, Canada

<sup>‡</sup> Department of Health Policy Management and Evaluation, University of  
Toronto, Toronto, Ontario, Canada

Number range CHAPTER<sup>\*\*\*</sup> Department of Health Policy and Administration,  
University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

<sup>§§</sup> Departments of Medicine and Clinical Epidemiology and Biostatistics,  
McMaster University, Hamilton, Ontario, Canada

# Consulting Economist, Research Associate, Canadian Centre for Policy  
Alternatives, Ottawa, Ontario, Canada

<sup>¶¶</sup> Health Sciences Library, McMaster University, Hamilton, Ontario, Canada

<sup>^</sup> School of Health Policy and Management, York University, Toronto Ontario;  
Department of Emergency Medicine, University Health Network, Toronto  
Ontario.

Author for correspondence and reprint requests: Dr. Gordon Guyatt, McMaster  
University, Faculty of Health Sciences, Clinical Epidemiology & Biostatistics,  
Room 2C12, 1200 Main Street West Hamilton, ON, L8N 3Z5, Tel: (905) 525-9140,  
x22900, Fax: (905) 524-3841, email: [guyatt@mcmaster.ca](mailto:guyatt@mcmaster.ca)

## Abstract

**Context:** Differences in medical care in the United States compared to Canada, including greater reliance on private funding and for-profit delivery, and markedly higher expenditures, may result in different health outcomes.

**Objectives:** To systematically review studies comparing health outcomes in patients with similar underlying medical conditions treated in the United States and Canada.

**Methods:** We identified studies comparing health outcomes of patients in Canada and the United States by searching multiple bibliographic databases and resources. We masked study results before determining study eligibility. We abstracted study characteristics, including methodological quality and generalizability.

**Results:** We identified 38 studies comparing populations of patients in Canada and the United States. Studies addressed diverse problems including cancer, coronary disease, chronic medical illnesses and surgical procedures. Of 10 studies that included extensive statistical adjustment and enrolled broad populations, 5 favored Canada, 2 favored the United States, and 3 showed equivalent or mixed results. Of 28 studies that failed one of these criteria, 9 favored Canada, 3 favored the United States, and 16 showed equivalent or mixed results. Overall mortality favored Canada (relative risk 0.95 (95% confidence interval 0.92 to 0.98,  $p = 0.002$ ) but results were very heterogeneous, and we failed to find convincing explanations for

heterogeneity. The only condition in which results consistently favored one country was end-stage renal disease, in which Canadian patients fared better.

**Interpretation:** Available studies suggest that health outcomes may be superior in patients cared for in Canada versus the United States, but differences are not consistent.

## INTRODUCTION

Canada and the United States are similar in many ways, and until 40 years ago their health care systems were nearly identical. At that time Canada adopted Medicare, its national health insurance program. Simultaneously, the U.S. implemented its Medicare program for the elderly.

Though both nations continue to rely largely on private funding for drugs, they now differ substantially in both the financing and delivery of physician and hospital services[1]. In financing, Canada has virtually first-dollar, universal public coverage of hospital and physician services. In delivery, not-for-profit institutions provide almost all hospital services and large for-profit organizations are almost entirely excluded from the provision of physician services. In contrast, the US relies on a mixture of public and private insurance to finance care, and leaves 16% of the population without coverage. Investor-owned for-profit providers play a substantial role.

The US also spends far more on health care, approximately 15% of its gross domestic product vs. about 10% in Canada. In 2003, Americans spent an estimated \$5,635 per capita on health care, while Canadians spent \$3,003 (both in U.S. dollars).

How do these alternative approaches to health care financing and delivery affect health outcomes? While a number of factors beyond the health care system influence the health of populations, for conditions amenable to medical treatment the health care system is a major determinant of outcomes. [2, 3] The choices the US and Canada have made may influence access and

quality of care, and hence morbidity and mortality. To inform this issue we undertook a systematic review addressing the following question: are there differences in health outcomes (mortality or morbid events) in patients suffering from similar medical conditions treated in Canada versus those treated in the United States?

## **METHODS**

Interested readers can obtain the detailed protocol for this review from the corresponding author. In brief, the formal search included published papers and abstracts up to the end of 2002. The process was standard for systematic reviews: definition of eligibility criteria; a broad search identifying possibly eligible titles and abstracts; selection of titles and abstracts that might possibly be eligible; selection of eligible reports from review of full documents; and abstraction of descriptive information, validity, and outcome data.

### **Eligibility criteria**

We included published and unpublished prospective or retrospective observational studies comparing health outcomes (mortality or morbid events) for patients of any age with the same diagnosis in Canada and the United States. We excluded randomized trials, studies that identified the patients on the basis of the occurrence of one of the adverse health outcomes of interest, and national disease-specific mortality studies that failed to define the

population at risk (that is, those with the disease of interest). For instance, we excluded studies of national rates of death from cancers because lower mortality may be due to either a lower incidence of cancer or better care for those with the disease.

The review process required many methodological decisions not fully anticipated in the initial protocol. These included issues regarding eligibility. For instance, we considered whether or not to consider low-birth-weight a disease. We decided not to do so because it has a wide variety of social and medical causes with associated differences in prognosis. On the other hand, we decided to include studies of the outcome of pregnancy because we considered that pre-natal and obstetrical care were potentially important types of care that we could legitimately assess. We discussed whether to include studies that evaluated critically ill patients with an array of diagnoses. We decided to do so on the basis that acute illness severity scores are very powerful predictors of outcome across a range of critically ill populations.

Only members of our team who were both blinded to the results of the studies in question and had expertise in the clinical issue at hand participated in these decisions.

### **Study identification**

A professional librarian (NB) conducted a search for the studies in bibliographic databases that included EMBASE (1980-Feb. 2003), MEDLINE (1966-Feb. 2003), HealthSTAR (1975-Feb. 2003), EBM Reviews - Cochrane

Central Register of Controlled Trials (2003, Issue 1) and Dissertation Abstracts Ondisc (1969-Feb. 2003). The search included an iterative process to refine the search strategy through testing of several search terms and incorporation of new search terms as new relevant citations were identified.

We further conducted a “cited reference search” in Web of Science on the relevant papers and used the “related articles feature” in PubMed. After reviewing 1,357 of the “related articles” and “cited reference” search results and finding only one potentially (but not ultimately) eligible article, we discontinued that part of the search.

### **Screening process**

Our initial search identified 4,923 potentially eligible studies (Figure 1). Teams of two reviewers independently evaluated titles and, when available, abstracts to determine whether or not the articles might meet eligibility criteria. If either reviewer concluded that there was any possibility that the article would fulfill eligibility criteria, we obtained the full text publication.

### **Assessment of study eligibility**

Research staff masked the results (blacked out the results in tables and text) of all studies identified for full evaluation in the screening process. Teams of two reviewers independently assessed all studies identified for full evaluation and resolved disagreements by discussion. Reviewers never assessed the same report at the title/abstract stage and at the full report stage.



For papers deemed eligible, two data abstractors, with access to the unmasked paper, reviewed the eligibility decision. If the data abstractors had questions about eligibility, the pair of reviewers who initially adjudicated the full blinded paper was informed of the reason for the concern and, still blind to results, reevaluated their initial decision. Their decision after this second review was deemed final.

### **Methodological quality assessment and data abstraction**

Teams of two reviewers independently assessed the methods and abstracted data from all eligible studies; they resolved disagreements through discussion. Information relevant to the methodological quality of the studies included the study design, the populations selected (criteria for diagnosis, similarity of patient groups in the two nations and the degree to which the studied population was representative of the wider universe of patients with the diagnosis), measurement of outcome (that is, the extent to which the outcome measures were defined similarly, and monitored similarly), loss to follow-up, and the extent of risk adjustment for confounders that might affect prognosis. Other data we abstracted included the geographic region in which the study was conducted, the time period of observation, the number of participants, and the main outcomes.

We classified studies as high or low quality according to two criteria. 1) Did the investigators adequately adjust for prognostic differences? Specifically, we considered adjustment inadequate if either disease severity or comorbidity

were not considered in the analysis. In the case of cancer, this decision resulted in only studies documenting cancer stage being rated high quality. 2) Did the investigators enroll a sufficiently diverse and representative population that it is plausible that the outcomes in patients studied are representative of the outcomes in the country at large? Studies might enroll similar populations, and adjust for prognostic differences, but only examine one delivery site in each country, or only sites in a single state. Such studies would fail the second criteria. We considered studies enrolling patients from a number of regions, or from a very large population within a region, as meeting this criterion.

For each study two reviewers, blinded to outcome, independently made the decision concerning high or low quality. If we identified apparently contradictory decisions across pairs of reviewers (for instance, if one set of reviewers rated a study using Canadian and US cancer databases as high quality, and another team rated a different study using the same databases as low quality), we informed reviewers of the inconsistency. The reviewers resolved the issue through discussion.

In response to editorial suggestions, we further evaluated the issue of representativeness with more rigorous and explicit criteria. We considered studies as fully representative only if samples in both countries were drawn from similar population-based registries that included at least one entire Canadian province, and at least two entire U.S. states, or a random sample of patients from at least an entire province and two U.S. states.

For all eligible studies, we sent the original authors our summary of the information abstracted from their article and asked them to correct and complement as they saw fit (11 authors, representing 16 studies, responded). When authors provided additional specific information or corrections, we incorporated these in our descriptive tables. For two eligible abstracts[4, 5], we requested and received a complete description of the study from the authors.

### **Data analysis**

When studies reported any patient-important outcome (morbidity, mortality, or quality of life) but did not state statistical significance, we calculated associated p-values using a threshold p of 0.05 for significance.

Because it was the most reliably and consistently measured outcome, we restricted the meta-analyses to the outcome of total mortality. When studies presented outcome data at one and six months, we included data at six months reasoning that if outcomes differ at one but not six months this is likely to be of limited importance to patients.

The statistical analysis included each non-overlapping study that provided the proportion of patients who died in both Canada and the United States, along with the associated variance (or data that allowed its calculation). We pooled the results using a random-effects model. We assess heterogeneity in results using the Cochrane's Q test[6], and calculated the  $I^2$ [7]. Relative risk was used as the summary statistic. When articles reported

separate procedures (for instance, mortality for different operations; mortality for different cancers), we treated each patient population as if it came from a separate study. Similarly, if an article reported major sub-populations within a patient group (such as low and high income), we treated these groups as coming from separate studies. We created funnel plots to provide graphical evaluation of publication bias and used a statistical technique suggested by Egger to provide a quantitative evaluation of the likelihood of publication bias[8].

To try to explain heterogeneity in effect estimates from individual studies, we conducted meta-regression analyses in which an additive between-study variance component of residual heterogeneity was used in accordance with the random-effects. The dependent variable was the log of the relative risk. The independent variables were based on the following a priori hypotheses explaining heterogeneity:

- overall study quality based on adequacy of adjustment for potential confounders and representativeness of the sample;
- source of the data (primary data collection versus administrative database)
- whether care was primarily out-patient or in-patient;
- the extent to which U.S. patients had health insurance (in-hospital studies involving primarily those  $\geq 65$  years of age or any study undertaken in Veterans Administration facilities will have excluded most uninsured people)
- completeness of follow-up;

- whether the U.S. site included or was restricted to New England (hypothesized to have better outcome than in other areas of U.S.)[9];
- the underlying health problem (renal failure, cardiology, cancer, surgery, and other);
- data collection before or after the median date of 1986 (we initially considered the key date for Canada before or after all provinces entered into Medicare [1970], and for the United States before or after the introduction of Medicare and Medicaid [July 1, 1966]; this choice, however, would have led to insufficient variability - almost all the data came from after 1970).

## RESULTS

As presented in Figure 1, of the 4,923 titles and abstracts identified, 498 appeared potentially eligible on initial review, and 42 of these proved eligible on review of the full article. We excluded three of these publications because the data overlapped substantially with those in another report that was eligible and included[10-12]. One study was reported in two complementary articles[13, 14].

Table 1 summarizes the results in terms of high and low quality studies, and whether results favored the United States, Canada, or showed mixed findings or no difference. Tables 2 to 4 present key methods and results beginning with the highest quality studies from population registries with adequate adjustment (unshaded); then the intermediate quality studies that were reasonably representative and had adequate adjustment (lightly shaded);

and finally the low quality studies in which the populations were unrepresentative or adjustment was inadequate.

Of the 5 studies that reported superior outcomes in the US, we classified 2 as high quality (one of which utilized population registries) and 3 as low quality (Table 2). Of the 2 high-quality studies, one presents results from a population-based registry that showed higher 30-day post-operative mortality after hip fracture in Manitoba and Quebec in comparison to several American states[15]. Canadians had longer wait times for surgery, longer post-operative lengths of stay, and higher inpatient mortality. Differences in mortality were not, however, attributable to differences in wait times for surgery. Furthermore, the increase in mortality did not persist over time, and Canadian outcomes proved superior for several other surgical procedures [16, 17] (Table 4).

The second high quality study was prospectively designed to examine outcomes of cataract surgery in a number of countries, including Canada and the United States[13, 14]. The two reports of this study fail to describe the mix of insured and uninsured patients in the US sample.

The first of the low-quality studies favoring the U.S. presented results from administrative databases in the US and Ontario and showed similar survival in patients with colon and lung cancer and Hodgkin's lymphoma, but superior survival in American breast cancer patients[18]. Another study using the same databases over a somewhat different (but overlapping) time period showed similar results for breast cancer and Hodgkin's disease, but found an

overall survival advantage for American patients in colon cancer and Canadian patients in lung cancer[19] (Table 4). Two studies that used the same database but restricted their analysis to Toronto versus American cities that the authors considered comparable showed a significant advantage[20], or a trend[21], toward superior survival in breast cancer patients in Canada versus the United States (Table 3).

Other low-quality studies favoring the US include populations of patients with rheumatoid arthritis [22] and patients following myocardial infarction (MI) [23]. In the latter study looking at only one Canadian and one U.S. hospital, more aggressive treatment in the US was associated with superior functional status, but no difference in recurrent MI or death. Another much larger observational study also found greater use of invasive treatments in the U.S. with superior functional status, but similar death and reinfarction (though higher stroke) rates [24] (Table 4). These results are not completely consistent across studies. Indeed, one study that included 14 American and 4 Canadian sites and over 2,000 patients demonstrated similar rates of invasive procedures in patients who experienced non-Q wave MI and unstable angina, with a lower rate of recurrent ischemia in hospital, at six weeks, and at one year in Canadian patients[25] (Table 3). The finding of similar rates of cardiovascular deaths in MI patients, with the exception of slightly lower death rates in American elderly patients in the first three months after MI[26], does appear consistent[27] (Table 4).

Of the 14 studies that demonstrated superior outcomes in Canada, we classified 5 as high quality (3 from population-based registries including all patients from at least one Canadian province and two U.S. states) and 9 as low quality (Table 3). Five studies, two high quality (one from a population-based registry) and three low quality, showed consistently lower mortality in Canadian than American patients with renal failure (Table 3). These studies included administrative database studies of black patients receiving renal transplants[28], of Manitoba and American patients receiving either hemodialysis or peritoneal dialysis[29], and of the entire Canadian and American populations receiving peritoneal dialysis[30] or any dialysis[4, 31]. Another study that almost certainly used similar data sources but did not report their methods as thoroughly also suggested lower mortality in Canadian than American patients receiving dialysis or renal transplants[32]. The strongest study from a data collection and adjustment point of view (though with a small number of American patients and not drawn from a population-based registry), a prospective cohort study in which the investigators were responsible for data collection, showed lower mortality in Canadian patients undergoing peritoneal dialysis[33].

The most rigorous of the dialysis studies, taking into account both sampling and adjustment, used data from 5,192 patients in the U.S. casemix severity study (a random sample of all Americans who began dialysis in 1986 or 1987). The investigators complemented these data with clinical and administrative records from the Henry Ford Hospital in Detroit, Michigan, and



review of charts of all patients (549) with endstage renal disease treated in the province of Manitoba between 1983 and 1989[29]. Casemix adjustment included age, sex, and a wide range of comorbidity (including diabetes, coronary artery disease, heart failure, respiratory disease, and cancer). After adjustment for both casemix and treatment variables (including likelihood of transplant) the relative mortality rate was 47% higher in the U.S. population (95% confidence interval 16% to 87%). One could argue that treatment variables should not have been included in the adjustment. If so, the increased risk of death in the American population would have been even higher. By far the biggest treatment-related variable that impacted on mortality was dialysis (relative mortality 0.53 in those transplanted). Transplantation rates were 35% in Manitoba and 17% in the American sample.

A series of reports used the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) and the Ontario Cancer Registry (OCR) to compare cancer patients' outcomes. Two of these population-based studies also conducted chart reviews in a sample of Canadian patients to obtain staging information not available in the OCR database. These investigations showed lower mortality rates in lower stage supraglottic and glottic cancer in Canadian patients, along with lower rates of laryngectomy[34, 35]. The stronger of these studies, focusing on patients with glottic cancer, supplemented electronic data from population-based cancer registries with chart review, hospital discharge data, and clinical databases and was able to adjust for stage, age, and sex. Laryngectomy rates across all stages were 5% in Canada

and 13.9% in the U.S. Survival was similar in higher stage patients, but lower stage Canadian patients showed a statistically significant survival advantage in years 2, 3, and 4.

The other studies utilizing these databases are weaker because they do not adjust for cancer stage or severity. One set of reports compared Toronto to a number of American cities, and suggested that poorer Canadian patients fared better than their American peers[20, 21, 36]. These results were only in part consistent with a report from the entire SEER database and the entire province of Ontario that supported the finding of better outcomes in poorer Canadians than Americans, but also suggested that wealthier Americans with cancer may fare better than wealthier Canadians[19] (Table 4). Another study using the same databases focusing on head and neck cancer showed mixed results[37] (Table 4). A prior section in the Results has described other mixed findings from studies using these databases[18, 19]. Three smaller studies of cancer patients that relied on chart review showed no differences in outcomes between Canada and the United States (Table 4)[38-40].

A high quality population-based study that looked at the entire cystic fibrosis population in both countries showed apparent benefits in height and weight from Canadian care[41] (Table 3). A second study restricted to one Canadian and one US institution suggested higher survival in Canadian cystic fibrosis patients[42]. A study comparing AIDS patients in British Columbia to those in a number of American cities suggested lower death rates in Canadian patients; the only adjustment was for baseline CD-4 count[43].

Of the 19 studies demonstrating comparable or mixed outcomes, we classified 3 as high quality (two using population-based registries) and 16 as low quality (Table 4). We have described some of these studies in the context of studies included in Tables 2 and 3. High quality studies relying on administrative databases of broad populations have shown equivalent mortality in Canada and the U.S. in coronary artery bypass grafting[44], lower mortality in Canada in a variety of low and moderate risk surgeries, and higher short but not long-term mortality in high risk surgeries including hip fracture repair[16, 17]. Lower quality studies have suggested a similar incidence of low birth weight infants[45], no difference in outcomes in asthmatic patients presenting to the emergency departments[5], no difference in outcomes in critically ill patients[46] or demented patients admitted to hospital[47], and no differences in functional status in patients with rheumatoid arthritis[48]. A study relying on volunteer call-in found that Canadian women with nausea and vomiting of pregnancy had more depression, more adverse effects on marital relationships, but less lost hours of paid work, less hospitalization, and less weight loss than did American women suffering from the condition[49]. A study relying on an administrative database from one U.S. and one Canadian hospital found higher ICU admission rates and longer ICU stays, but shorter overall hospital stays, in U.S. patients hospitalized for trauma[50].

## Statistical Analysis

The statistical analysis was based on results of 83 populations in 23 studies that reported all-cause mortality with sufficient completeness for inclusion[15-18, 20-23, 25-27, 29, 33, 37-40, 42, 43, 46, 47, 50]. In Figure 2, which depicts the distribution of the log of the relative risk against the precision of the estimates (the inverse of the standard deviation of the log RR), values to the left of 0 favor Canada and values to the right of 0 favor the United States. The pooled relative risk of dying in Canada versus the United States was 0.95 (95% confidence interval 0.92 to 0.98,  $p = 0.002$ , heterogeneity  $p < 0.0001$ ,  $I^2 = 0.94$ ). The plot suggests some asymmetry, with a number of low precision studies favoring Canada without corresponding studies favoring the United States. This is consistent with the statistical analysis, which suggested rejecting the null hypothesis of no asymmetry ( $p = 0.02$ ). One possible explanation for this result is publication bias in Canada's favor.

Table 5 presents the results of the univariable and multivariable regressions. The results show no variables as significant in the univariable model, whereas several are significant in the multivariable model: study quality (higher quality studies tend to favor the U.S.); whether New England was included (inclusion of New England tends to an estimate of lower mortality in Canada); and disease category (renal failure, cancer, and surgery tended to favor Canada, cardiology and other studies tended to favor the U.S.). Neither the univariable models, nor the multivariable model (despite apparently explaining 49% of the variance) were stable. For instance, omission of two

relatively large studies that represented outliers resulted in very different results.

## DISCUSSION

In this systematic review, we demonstrated that while Canadian outcomes were more often superior to U.S. outcomes than the reverse, neither the United States nor Canada can claim hegemony in terms of quality of medical care and the resultant patient-important outcomes. In virtually all areas, study results have demonstrated some apparent advantages for Canada and others for the United States. In cancer, where a number of strong studies have used population-based registries, Canadian outcomes appear superior in head and neck cancer, and possibly for low-income patients with a variety of cancers; American women with breast cancer appear to have better survival rates than Canadians. In data that comes from population-based registries, Canadians enjoy better risk-adjusted survival after a variety of surgeries, but American outcomes appear superior after hip fracture repair and cataract surgery. Studies that do not utilize population-based registries suggest that Americans have, possibly as a result of more aggressive interventions, less angina after MI, but the benefit may come at the price of increased strokes and bleeding. There is one area in which Canadian outcomes appear consistently superior - end-stage renal failure. Even here, however, as we shall discuss, one cannot be certain that superior medical care is responsible for the differences.

The strengths and limitations of this systematic review bear on its interpretation. We established a team that included expertise in medicine, clinical epidemiology, health economics, health policy, and health services research in both Canada and the United States, developed explicit eligibility criteria, and conducted a comprehensive search that uncovered a number of eligible articles not included in a previous systematic review[51]. We excluded studies, such as randomized trials of medical interventions in which Canadian investigators recruited some patients and American investigators others, in which care would be idiosyncratic or atypical of care in usual clinical practice. Our thorough examination of each study addressed issues of validity (selection of populations, adjustment for confounders, loss to follow-up) and generalizability (breadth of samples, including specifying studies that came from population-based registries).

Reviewers who determined eligibility and judged validity and generalizability were blind to the results of the study. In decision making regarding methodologic issues that arose as the review progressed, we recused investigators aware of study results. We made explicit a priori hypotheses regarding possible sources of heterogeneity, and tested these hypotheses in a thorough statistical analysis. Our results are consistent with those of a prior systematic review that completed its search (less comprehensive than ours) in 1997, conducted a limited assessment of study validity, and failed to conduct a formal meta-analysis[51].

The main limitation of our review is in the uneven quality of the original studies, and the threats to validity that remain even in those studies of high quality. There were two key ways a study could fail to adequately address our question: either the population might be small or narrow, or the investigators might not carry out statistical adjustment for potential differences in underlying prognosis. Most of the studies we identified failed one of these two criteria (Tables 2 to 4).

Even studies that meet these criteria, and meet the more rigorous criterion of utilizing population-based registries, leave challenges to interpretation. In general, a health care system can improve outcomes in two ways. One is to facilitate early entry to care, including preventive care, and thus avoid unnecessary morbidity and mortality. For instance, if access to primary care is easy and without financial obstacles, one might expect superior outcomes in hypertension, e.g. fewer strokes. Alternatively, a system might generate better outcomes by better treatment of serious morbidity once it arises. For instance, stroke patients may be more likely to receive early thrombolysis, thromboprophylaxis, and multidisciplinary rehabilitation.

If a health system does better in the early identification and treatment of the ill, diseased patients in that system will appear less ill. Statistical adjustment for severity of illness is in general appropriate - one wouldn't want to attribute to better care what is in fact due to patients' better prognosis. The risk, however, is that the adjustment will obscure the benefits of early identification and treatment.

Such issues become relevant in comparisons of outcomes between Canada and the United States. For instance, the United States does a better job of screening women for breast cancer[52]. To the extent that early diagnosis reduces breast cancer deaths, one would expect a mortality advantage for American women. At the same time, any apparent increase in longevity may be largely, or even completely, due to the length and lead-time biases inherent in observational studies of screening.

A number of studies using the American National Cancer Institute's Surveillance, Epidemiology, and End Results Program (SEER) and the Ontario Cancer Registry (OCR) have addressed breast cancer outcomes. Although studies using these databases and examining Toronto versus a number of U.S. cities suggest higher breast cancer survival in low-income Canadian women than in their U.S. counterparts[20, 21, 36], several studies using the entire database have suggested superior overall breast cancer survival in American women[18, 19, 32]. We rated these studies as low quality because of failure to adjust for disease stage. If higher screening rates or better self-detection in the U.S. result in the identification of earlier stage histologic cancers that would have remained asymptomatic and dormant, studies would demonstrate superior survival despite equivalent medical care. On the other hand, perhaps there is a true American advantage that results from higher rates of screening[52] or from superior care after diagnosis. The data do not allow assessment of the relative likelihood of these possible explanations.



These studies raise another important limitation of the current data. Canada has largely[53] (though not completely[52, 54]) eliminated gradients in access to care by socioeconomic status that remain in the United States[55, 56], and this may contribute to Canada's smaller socioeconomic gradients in health outcome[57]. If this were so, one would expect that studies focused on poorer individuals would reveal superior outcomes in Canada, whereas differences might be obscured in studies of entire populations. Indeed, the cancer studies by Gorey and colleagues[20, 21, 36] and by Boyd[19] suggest this may be the case. At the same time, it is possible that being able to pay for better care might lead to better outcomes in those with high incomes in the U.S. versus Canada. Indeed one of the studies in cancer patients suggested this possibility[19]. Unfortunately, these are the only studies that explore gradients in outcome across socioeconomic status.

Although the overall effect in the meta-analysis may be of some interest (a 5% reduction in relative risk of all-cause mortality in Canada versus the United States) the large variability in study results (heterogeneity  $p < 0.0001$ ,  $I^2$  94%, Figure 2) makes the pooled estimate difficult to interpret. Our primary reason for conducting the statistical analysis was, through meta-regression, to explore possible explanations of variability in results, and provide adjusted estimates of relative risk. This exploration proved difficult to interpret. Although the multivariable model identified apparent sources of heterogeneity and provided adjusted estimates of relative risk (Table 5), the results were inconsistent between univariable and multivariable approaches, and both the

univariable and multivariable models were very unstable. Thus, we do not feel confident that the statistical modeling has provided either a satisfactory explanation for the study-to-study variability in results or credible estimates of adjusted relative risk.

One group of patients fared consistently better in Canada than in the U.S., those with end-stage renal disease[4, 28-33]. Whether in hemodialysis programs, peritoneal dialysis, or after receipt of renal transplants, Canadians survive longer. The larger proportion of Americans than Canadians who begin dialysis treatment confounds interpretation of this finding. Perhaps Americans fare worse because a larger number of sicker patients enter dialysis. On the other hand, it may be that the larger proportion of Americans on dialysis reflects a lower threshold to start dialysis, and thus a less sick dialysis population. The limited available evidence suggests that thresholds for dialysis are in fact similar in the two countries[58]. Furthermore, two high quality studies that included extensive adjustment for comorbidity[29, 33] still show substantially lower mortality in Canadian patients, suggesting that imbalance in risk cannot explain superior Canadian outcomes.

Nevertheless, the weight of the evidence strongly suggests that Canadian end stage renal patients truly have higher survival than those in the U.S. The explanation for this difference may lie in differences in the ownership of dialysis facilities. Virtually all Canadian dialysis care is not-for-profit, while for-profit providers deliver approximately 75% of American care for end-stage

renal failure. A systematic review has shown a higher mortality in patients undergoing dialysis in for-profit centres[59].

Despite limitations of the available studies, some robust conclusions are possible from our systematic review. These results are incompatible with the hypothesis that patients in the U.S. receive consistently better care than those in Canada. Americans are not, therefore, getting value for money; the U.S.' 89% higher per capita expenditures on health care does not buy superior outcomes for the sick.

Canadian health care has many well-publicized limitations. Nevertheless, it produces health benefits similar, or perhaps superior, to those of the US health system, but at much lower cost. Canada's single payer system for physician and hospital care yields large administrative efficiencies in comparison to the American multi-payer model[60]. Not-for-profit hospital funding results in appreciably lower payments to third party payers in comparison to for-profit hospitals[61] while achieving lower mortality rates[62]. Policy debates and decisions regarding the direction of health care in both Canada and the United States should consider the results of our systematic review: Canada's single payer system, which relies on not-for-profit delivery, achieves health outcomes that are at least equal to those in the United States at two-thirds the cost.

## Acknowledgments

We are grateful to Laurel Raftery and particularly to Denise Healey for their crucial role in organizing and administering all aspects of the study. We offer thanks to David Churchill, Elizabeth Gilpin, Kevin Gorey, Patti Groome, Timothy Juday, Penny Mohr, Les Roos, David Skarsgard, Brian Rowe, and Jack Tu, who provided clarification and additional information regarding their eligible studies, and to Venessa Yu and Marco Venettilli for help with eligibility adjudication and data abstraction and Pam Leece with article blinding.

## REFERENCES:

1. Maioni, A., *Parting at the crossroads: the emergence of health insurance in the United States and Canada*. 1998, Princeton, NJ: Princeton University Press.
2. Poikolainen, K. and J. Eskola, *The effect of health services on mortality: decline in death rates from amenable and non-amenable causes in Finland, 1969-81*. *Lancet*, 1986. **1**(8474): p. 199-202.
3. Velkova, A., J.H. Wolleswinkel-van den Bosch, and J.P. Mackenbach, *The East-West life expectancy gap: differences in mortality from conditions amenable to medical intervention*. *Int J Epidemiol*, 1997. **26**(1): p. 75-84.
4. Juday, T. *Crossnational comparisons of FSDR treatment and outcomes*. in *Annual Meeting of International Society of Technology Assessment in Health Care*. 1994.
5. Rowe, B.B., GW. Pollack, E. Pollack, CV. Emond, SD. Comargo, CA., *Comparison of Canadian versus US Emergency Department Visits for Acute Asthma*. 2005.
6. Fleiss, J., *The statistical basis of meta-analysis.*, in *Statistical Methods in Medical Research*. 1993, Hodder Arnold Journals: London. p. 121-145.
7. Higgins, J., et al., *Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice*. *J Health Serv Res Policy*, 2002. **7**(1): p. 51-61.
8. Egger, M., et al., *Bias in meta-analysis detected by a simple, graphical test*. *Bmj*, 1997. **315**(7109): p. 629-34.
9. Himmelstein, D.U., et al., *Quality of care in investor-owned vs not-for-profit HMOs*. *Jama*, 1999. **282**(2): p. 159-63.
10. Gorey, K.M., et al., *An international comparison of cancer survival: Toronto, Ontario, and Detroit, Michigan, metropolitan areas*. *Am J Public Health*, 1997. **87**(7): p. 1156-63.
11. Groome, P.A., et al., *Management and outcome of glottic cancer: a population-based comparison between Ontario, Canada and the SEER areas of the United States*. *Surveillance, Epidemiology and End Results*. *J Otolaryngol*, 2000. **29**(2): p. 67-77.
12. Roos, L.L., et al., *Postsurgical mortality in Manitoba and New England*. *Jama*, 1990. **263**(18): p. 2453-8.
13. Norregaard, J.C., et al., *Visual outcomes of cataract surgery in the United States, Canada, Denmark, and Spain. Report From the International Cataract Surgery Outcomes Study*. *Arch Ophthalmol*, 1998. **116**(8): p. 1095-100.
14. Norregaard, J.C., et al., *Intraoperative clinical practice and risk of early complications after cataract extraction in the United States, Canada, Denmark, and Spain*. *Ophthalmology*, 1999. **106**(1): p. 42-8.
15. Ho, V., B.H. Hamilton, and L.L. Roos, *Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada*. *Health Serv Res*, 2000. **34**(7): p. 1499-518.
16. Roos, L.L., et al., *Health and surgical outcomes in Canada and the United States*. *Health Aff (Millwood)*, 1992. **11**(2): p. 56-72.
17. Roos, L.L., et al., *Short-term mortality after repair of hip fracture. Do Manitoba elderly do worse?* *Med Care*, 1996. **34**(4): p. 310-26.
18. Keller, D.M., E.A. Peterson, and G. Silberman, *Survival rates for four forms of cancer in the United States and Ontario*. *Am J Public Health*, 1997. **87**(7): p. 1164-7.
19. Boyd, C., et al., *Associations between community income and cancer survival in Ontario, Canada, and the United States*. *J Clin Oncol*, 1999. **17**(7): p. 2244-55.
20. Gorey, K.M., et al., *An international comparison of cancer survival: relatively poor areas of Toronto, Ontario and three US metropolitan areas*. *J Public Health Med*, 2000. **22**(3): p. 343-8.
21. Gorey, K.M., et al., *An international comparison of cancer survival: metropolitan Toronto, Ontario, and Honolulu, Hawaii*. *Am J Public Health*, 2000. **90**(12): p. 1866-72.
22. Wolfe, F., et al., *The mortality of rheumatoid arthritis*. *Arthritis Rheum*, 1994. **37**(4): p. 481-94.
23. Pilote, L., N. Racine, and M.A. Hlatky, *Differences in the treatment of myocardial infarction in the United States and Canada. A comparison of two university hospitals*. *Arch Intern Med*, 1994. **154**(10): p. 1090-6.
24. Yusuf, S., et al., *Variations between countries in invasive cardiac procedures and outcomes in patients with suspected unstable angina or myocardial infarction without initial ST elevation*.

- OASIS (Organisation to Assess Strategies for Ischaemic Syndromes) Registry Investigators. Lancet, 1998. 352(9127): p. 507-14.*
25. Anderson, H.V., et al., *Management of unstable angina pectoris and non-Q-wave acute myocardial infarction in the United States and Canada (the TIMI III Registry).* Am J Cardiol, 1997. **79**(11): p. 1441-6.
  26. Tu, J.V., et al., *Use of cardiac procedures and outcomes in elderly patients with myocardial infarction in the United States and Canada.* N Engl J Med, 1997. **336**(21): p. 1500-5.
  27. Gilpin, E.A., et al., *Periods of differing mortality distribution during the first year after acute myocardial infarction.* Am J Cardiol, 1983. **52**(3): p. 240-4.
  28. Koyama, H., J.M. Cecka, and P.I. Terasaki, *Kidney transplants in black recipients. HLA matching and other factors affecting long-term graft survival.* Transplantation, 1994. **57**(7): p. 1064-8.
  29. Hornberger, J.C., A.M. Garber, and J.R. Jeffery, *Mortality, hospital admissions, and medical costs of end-stage renal disease in the United States and Manitoba, Canada.* Med Care, 1997. **35**(7): p. 686-700.
  30. Moran, J., *Changes in the dose of peritoneal dialysis: have these independently affected outcomes?* Am J Kidney Dis, 1998. **32**(6 Suppl 4): p. S52-7.
  31. Mohr, P.R., S. Rowell, P. Newschaffer, C. Bernardin, M. Juday, T., *Tecnology and Outcomes: International comparisons.* 1993, Project HOPE Center for Health Affairs: Bethesda.
  32. Hussey, P.S., et al., *How does the quality of care compare in five countries?* Health Aff (Millwood), 2004. **23**(3): p. 89-99.
  33. Churchill, D.N., et al., *Lower probability of patient survival with continuous peritoneal dialysis in the United States compared with Canada. Canada-USA (CANUSA) Peritoneal Dialysis Study Group.* J Am Soc Nephrol, 1997. **8**(6): p. 965-71.
  34. Groome, P.A., et al., *Management and outcome differences in supraglottic cancer between Ontario, Canada, and the Surveillance, Epidemiology, and End Results areas of the United States.* J Clin Oncol, 2003. **21**(3): p. 496-505.
  35. Groome, P.A., et al., *Glottic cancer in Ontario, Canada and the SEER areas of the United States. Do different management philosophies produce different outcome profiles?* J Clin Epidemiol, 2001. **54**(3): p. 301-15.
  36. Gorey, K.M., et al., *An international comparison of cancer survival: advantage of Toronto's poor over the near poor of Detroit.* Can J Public Health, 1998. **89**(2): p. 102-4.
  37. Skarsgard, D.P., et al., *Cancers of the upper aerodigestive tract in Ontario, Canada, and the United States.* Cancer, 2000. **88**(7): p. 1728-38.
  38. Campbell, H., *Cancer of the cervix--survival comparative study of 5-year survival rates from cancer of the cervix in 14 countries from 1953-1957 followed to 1962.* J Obstet Gynaecol Br Commonw, 1966. **73**(1): p. 27-40.
  39. Jones, S.E., T.P. Miller, and J.M. Connors, *Long-term follow-up and analysis for prognostic factors for patients with limited-stage diffuse large-cell lymphoma treated with initial chemotherapy with or without adjuvant radiotherapy.* J Clin Oncol, 1989. **7**(9): p. 1186-91.
  40. LoCoco, S., et al., *Does aggressive therapy improve survival in suboptimal stage IIIc/IV ovarian cancer? A Canadian-American comparative study.* Gynecol Oncol, 1995. **59**(2): p. 194-9.
  41. Lai, H.C., et al., *Comparison of growth status of patients with cystic fibrosis between the United States and Canada.* Am J Clin Nutr, 1999. **69**(3): p. 531-8.
  42. Corey, M., et al., *A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto.* J Clin Epidemiol, 1988. **41**(6): p. 583-91.
  43. Chan, K.C., et al., *Survival rates after the initiation of antiretroviral therapy stratified by CD4 cell counts in two cohorts in Canada and the United States.* Aids, 2002. **16**(12): p. 1693-5.
  44. Grumbach, K., et al., *Regionalization of cardiac surgery in the United States and Canada. Geographic access, choice, and outcomes.* Jama, 1995. **274**(16): p. 1282-8.
  45. Katz, S.J., R.W. Armstrong, and J.P. LoGerfo, *The adequacy of prenatal care and incidence of low birthweight among the poor in Washington State and British Columbia.* Am J Public Health, 1994. **84**(6): p. 986-91.
  46. Rapoport, J., et al., *A comparison of intensive care unit utilization in Alberta and western Massachusetts.* Crit Care Med, 1995. **23**(8): p. 1336-46.
  47. Duckworth, G.S., H.B. Kedward, and W.F. Bailey, *Prognosis of mental illness in old age: a four year follow-up study.* Can J Psychiatry, 1979. **24**(7): p. 674-82.

48. Sherrer, Y.S., et al., *Disability in rheumatoid arthritis: comparison of prognostic factors across three populations*. J Rheumatol, 1987. **14**(4): p. 705-9.
49. Mazzotta, P., et al., *Attitudes, management and consequences of nausea and vomiting of pregnancy in the United States and Canada*. Int J Gynaecol Obstet, 2000. **70**(3): p. 359-65.
50. Boulanger, B.R., et al., *A comparison between a Canadian regional trauma unit and an American level I trauma center*. J Trauma, 1993. **35**(2): p. 261-6.
51. Szick, S.A., DE. Nichol, G. Harrison, MB. Page, J. Moher, D., *Health care delivery in Canada and the United States: Are there relevant differences in health outcomes?* 1999, Institute for Clinical and Evaluative Sciences: Toronto.
52. Katz, S.J., J.K. Zemencuk, and T.P. Hofer, *Breast cancer screening in the United States and Canada, 1994: socioeconomic gradients persist*. Am J Public Health, 2000. **90**(5): p. 799-803.
53. Finkelstein, M.M., *Do factors other than need determine utilization of physicians' services in Ontario?* Cmaj, 2001. **165**(5): p. 565-70.
54. Katz, S.J. and T.P. Hofer, *Socioeconomic disparities in preventive care persist despite universal coverage. Breast and cervical cancer screening in Ontario and the United States*. Jama, 1994. **272**(7): p. 530-4.
55. Davis, K., *Inequality and access to health care*. Milbank Q, 1991. **69**(2): p. 253-73.
56. Andersen, R.D.P., *Improving access in America: Individual and contextual indicators.*, in *Changing the US health care system*, R.R. Andersen, TH. Kominski, GF., Editor. 2001, Jossey-Bass: San Francisco. p. 3 - 30.
57. Ross, N.A., et al., *Relation between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics*. Bmj, 2000. **320**(7239): p. 898-902.
58. Sekkarie, M., M. Cosma, and D. Mendelssohn, *Nonreferral and nonacceptance to dialysis by primary care physicians and nephrologists in Canada and the United States*. Am J Kidney Dis, 2001. **38**(1): p. 36-41.
59. Devereaux, P.J., et al., *Comparison of mortality between private for-profit and private not-for-profit hemodialysis centers: a systematic review and meta-analysis*. Jama, 2002. **288**(19): p. 2449-57.
60. Woolhandler, S., T. Campbell, and D.U. Himmelstein, *Costs of health care administration in the United States and Canada*. N Engl J Med, 2003. **349**(8): p. 768-75.
61. Devereaux, P.J., et al., *Payments for care at private for-profit and private not-for-profit hospitals: a systematic review and meta-analysis*. Cmaj, 2004. **170**(12): p. 1817-24.
62. Devereaux, P.J., et al., *A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals*. Cmaj, 2002. **166**(11): p. 1399-406.

