RESEARCH

A systematic review of studies comparing health outcomes in Canada and the United States

Gordon H. Guyatt, MSc, MD; P.J. Devereaux, MD; Joel Lexchin, MSc, MD; Samuel B. Stone, MD: Armine Yalnizyan, MA; David Himmelstein, MD; Steffie Woolhandler, MD, MPH; Qi Zhou, PhD; Laurie J. Goldsmith, PhD, MSc; Deborah J. Cook, MSc, MD; Ted Haines MSc, MD; Christina Lacchetti, MHSc; John N. Lavis, MD, PhD; Terrence Sullivan, PhD, MSc; Ed Mills, DPH MSc; Shelley Kraus; Neera Bhatnagar, BSc, MLIS

Dr. Guyatt is an internist and clinical epidemiologist at McMaster University, Hamilton, Ont.
Dr. Devereaux is a cardiologist and clinical epidemiologist at McMaster University, Hamilton, Ont.
Dr. Lexchin, is Professor of Health Policy, York University, Toronto, Ont.
Dr. Stone is a resident in the Department of Surgery, McMaster University, Hamilton, Ont.
Ms. Yalnizyan is an economist with the Canadian Centre for Policy Alternatives, Ottawa, Ont.
Dr. Himmelstein is a primary care physician and is Chief of Social and Community Medicine, Cambridge Hospital, Cambridge, Mass.
Dr. Woolhandler is a primary care physician at Cambridge Hospital and Associate Professor of Medicine, Harvard Medical School, Cambridge, Mass.
Dr. Zhou is a statistician at McMaster University, Hamilton, Ont.
Dr. Goldsmith is a PhD candidate in the Department of Health Policy and Administration, University of North Carolina at Chapel Hill, Chapel Hill, NC.
Dr. Cook is an internist and clinical epidemiologist at McMaster University, Hamilton, Ont.
Dr. Haines is an occupational health researcher at McMaster University, Hamilton, Ont.
Ms. Lacchetti is a research associate at McMaster University, Hamilton, Ont.
Dr. Lavis is a health policy researcher at McMaster University Hamilton, Ont.

Dr. Sullivan is Chief Executive Officer of Cancer Care Ontario, University of Toronto, Toronto, Ont.
Ed Mills is an epidemiologist at Simon Fraser University, Burnaby, BC.
Shelley Kraus is a student at McMaster University, Hamilton, Ont.
Neera Bhatnagar is a health sciences librarian at McMaster University, Hamilton, Ont.

Competing interests: None declared.

Correspondence: Dr. Gordon Guyatt, McMaster University, Faculty of Health Sciences, Clinical Epidemiology & Biostatistics, Room 2C12, 1200 Main Street West Hamilton, ON L8N 3Z5, (905) 525-9140, x22900; fax: (905) 524-3841; guyatt@mcmaster.ca

ABSTRACT

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

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

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 artery disease, chronic medical illnesses and surgical procedures. Of 10 studies that included extensive statistical adjustment and enrolled broad populations, 5 favoured Canada, 2 favoured the United States, and 3 showed equivalent or mixed results. Of 28 studies that failed one of these criteria, 9 favoured Canada, 3 favoured the United States, and 16 showed equivalent or mixed results. Overall, results for mortality favoured Canada (relative risk 0.95, 95% confidence interval 0.92–0.98, $p = 0.002$) but were very heterogeneous, and we failed to find convincing explanations for this heterogeneity. The only condition in which results consistently favoured 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 a national insurance program (medicare). Simultaneously, the United States implemented its Medicare program for elderly people.

Although 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] With respect to financing, Canada has virtually first-dollar, universal public coverage of hospital and physician services. With respect to 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 United States relies on a mixture of public and private insurance to finance health care, and leaves 16% of the population without coverage. Investor-owned for-profit providers play a substantial role.

The United States also spends far more on health care, i.e., approximately 15% of its gross domestic product versus about 10% in Canada. In 2003, Americans spent an estimated US$5,635 per capita on health care, while Canadians spent US$3,003.

How do these alternative approaches to health care financing and delivery affect health outcomes? Although 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 United States and Canada have made may influence access and quality of care, and hence morbidity and mortality. To inform debate on this issue we undertook a systematic review addressing the following question: Are there differences in health outcomes (mortality or morbidity) 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 papers and abstracts published 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 morbidity) in Canada and the United States for patients of any age with the same diagnosis. 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 either to a lower incidence of cancer or to 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 outcomes of pregnancy because we considered that prenatal 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 (N.B.) 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 (Fig. 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 period of observation, the number of participants, and the main outcomes.

We classified studies as being of high or low quality according to the following 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 as of 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 criterion. We considered studies that enrolled 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 rating of 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 United States 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 American states, or a random sample of patients from at least an entire province and two entire American 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 outcome of importance to patients (morbidity, mortality, or quality of life) but did not state statistical significance, we calculated associated *p* values using a threshold 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 1 and 6 months, we included data at 6 months, reasoning that if outcomes differ at 1 but not 6 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 either in Canada or 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 US patients had health insurance (in-hospital studies involving primarily those ≥65 years of age or any study undertaken in Veterans Administration facilities will have excluded most uninsured people)

- completeness of follow-up
- whether the US site included or was restricted to New England (hypothesized to have better outcome than in other areas of US)[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 favoured 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 (shaded).

Of the 5 studies that reported superior outcomes in the United States, 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 favouring the US presented results from administrative databases in the United States 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) 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 favouring the United States include populations of patients with rheumatoid arthritis[22] and patients after myocardial infarction (MI)[23]. In the latter study looking at only one Canadian and one US hospital, more aggressive treatment in the United States was associated with superior functional status, but not with any difference in recurrent MI or death. Another much larger observational study also found greater use of invasive treatments in the US 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 6 weeks, and at 1 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 3 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 US 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 Manitoban 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 US case-mix 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 end-stage renal disease treated in the province of Manitoba between 1983 and 1989.[29] Case-mix 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 case-mix and treatment variables (including likelihood of transplant) the relative mortality rate was 47% higher in the US population (95% confidence interval [CI] 16%–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 had

an impact 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 United States. Survival was similar in patients with higher-stage disease, but Canadian patients with lower-stage disease 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 partly 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 that used the same databases and focused on head and neck cancer showed mixed results[37] (Table 4). Other mixed findings from studies using these databases are described earlier in the Results. [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 that demonstrated 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 US 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 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 that relied on volunteer call-in found that Canadian women with nausea and vomiting of pregnancy had more depression and more adverse effects on marital relationships, but fewer lost hours of paid work, less hospitalization, and less weight loss than did American women suffering from the condition.[49] A study that relied on an administrative database from one US and one Canadian hospital found higher intensive care unit (ICU) admission rates and longer ICU stays, but shorter overall hospital stays, in US 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 favour Canada and values to the right of 0 favour the United States. The pooled relative risk of dying in Canada versus the United States was 0.95 (95% CI 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 favouring Canada without corresponding studies favouring 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 favour.

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 favour the US); 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 favour Canada; cardiology and other studies tended to favour the US). Neither the univariate models, nor the multivariate 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 although Canadian outcomes were more often superior to US 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 Canadian women. In data 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 who were aware of the 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 2–4).

Even studies that meet these criteria, and meet the more rigorous criterion of utilizing population-based registries, present challenges with respect to their 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 early identification and treatment, 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 a 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 survival 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 US cities suggest higher breast cancer survival in low-income Canadian women than in their American 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 US 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 US 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 multivariate model identified apparent sources of heterogeneity and provided adjusted estimates of relative risk (Table 5), the results were inconsistent between univariate and multivariate 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 US. 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 the limitations of the available studies, some robust conclusions are possible from our systematic review. These results are incompatible with the hypothesis that American patients receive consistently better care than Canadians. Americans are not, therefore, getting value for money; the 89% higher per-capita expenditures on health care in the United States 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 a much lower cost. Canada's single-payer system for physician and hospital care yields large administrative efficiencies in comparison with 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.

Acknowledgements

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, to Venessa Yu and Marco Venettilli for help with eligibility adjudication and data abstraction, and to Pam Leece for help with article blinding.

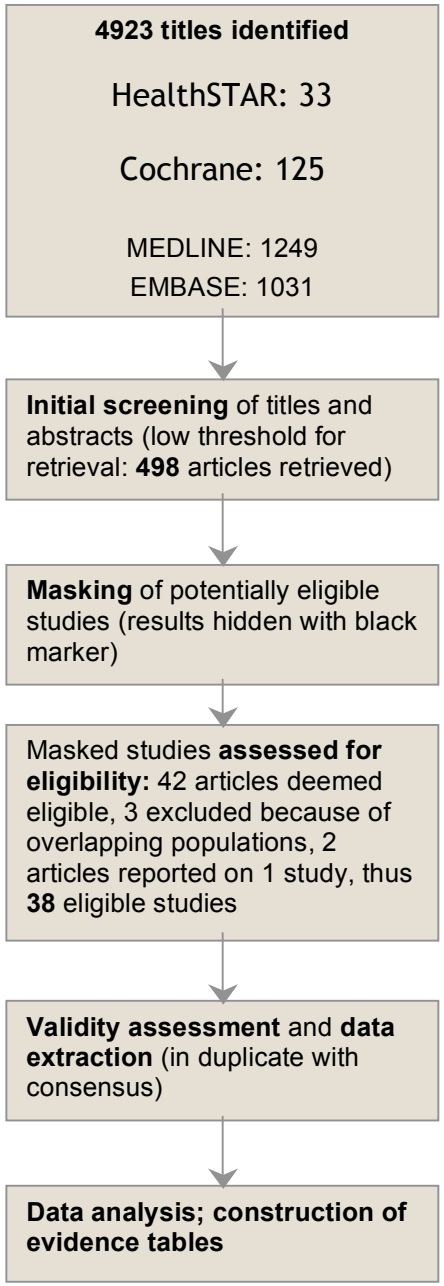
4923 titles identified
HealthSTAR: 33
Cochrane: 125
MEDLINE: 1249
EMBASE: 1031
Initial screening of titles and abstracts (low threshold for retrieval: 498 articles retrieved)
Masking of potentially eligible studies (results hidden with black marker)
Masked studies assessed for eligibility: 42 articles deemed eligible, 3 excluded because of overlapping populations, 2 articles reported on 1 study, thus 38 eligible studies
Validity assessment and data extraction (in duplicate with consensus)
Data analysis; construction of evidence tables

Fig. 1: Methodological steps in systematic review

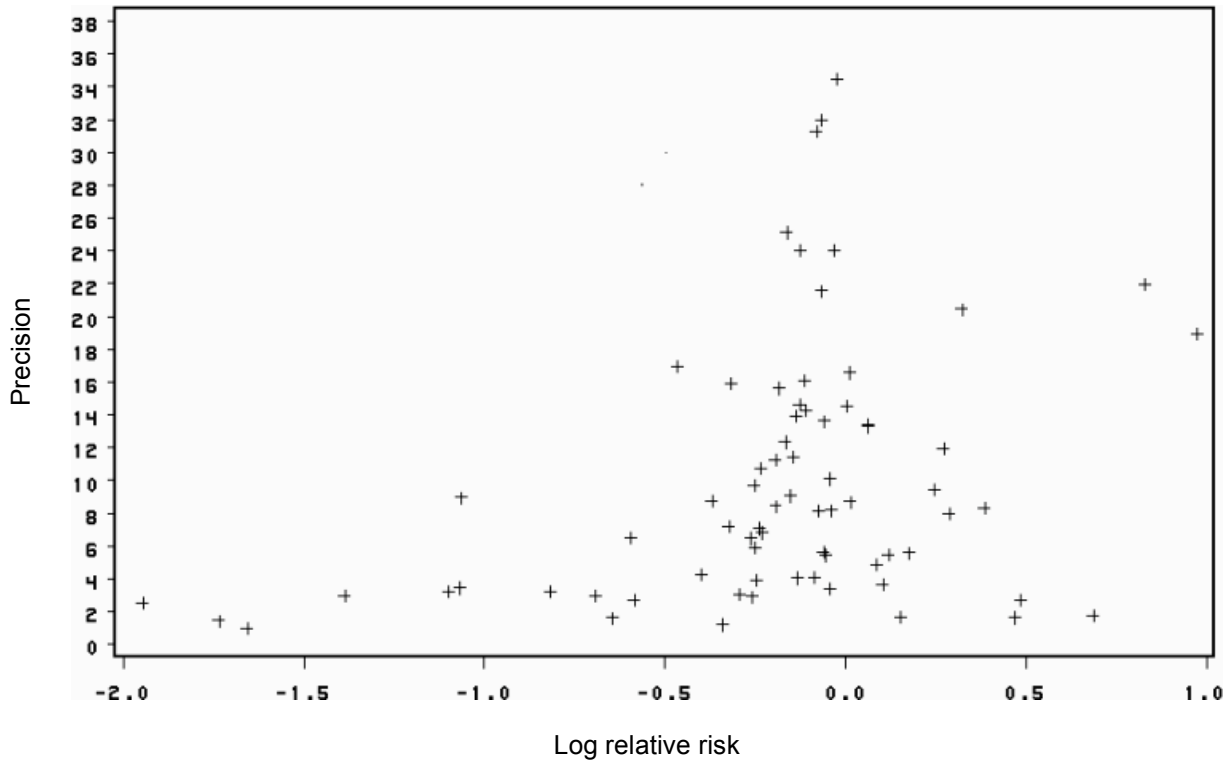

Fig. 2: Funnel plot for all-cause mortality, US versus Canadian studies.

References

1. Maioni A. *Parting at the crossroads: the emergence of health insurance in the United States and Canada.* Princeton (NJ): Princeton University Press; 1998.
2. Poikolainen K, Eskola J. 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):199-202.
3. Velkova A, Wolleswinkel-van den Bosch JH, Mackenbach JP. The East–West life expectancy gap: differences in mortality from conditions amenable to medical intervention. *Int J Epidemiol* 1997;26(1):75-84.
4. Juday T. Crossnational comparisons of FSDR treatment and outcomes. In: Mohr P, Rajan S, Rowell P. Newschaffer C, Bernardin M, Juday T. Technology and outcomes: international comparisons. Final Report. Bethesda (MD): Project HOPE Center for Health Affairs; 1993:i-8-8.
5. Rowe BH, Bota GW, Pollack E, Pollack CV, Emond SD, Comargo CA. Comparison of Canadian versus US emergency department visits for acute pediatric asthma [abstract]. *Acad Emerg Med* 1999;6(5):497.
6. Fleiss J. The statistical basis of meta-analysis. In *Statistical Methods in Medical Research*. London (UK): Hodder Arnold Journals; 1993. p 121-45.
7. Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J Health Serv Res Policy* 2002:7(1):51-61.
8. Egger M, Davey Smith G, Schneider M, Minder C. *Bias in meta-analysis detected by a simple, graphical test. BMJ* 1997;1315(7109):629-34.
9. Himmelstein DU, Woolhandler S, Hellander I, Wolfe SM. Quality of care in investor-owned vs not-for-profit HMOs. *JAMA* 1999;1282(2):159-63.
10. Gorey KM, Kliewer E, Holowaty EJ, Laukkanen E, Ng EY. An international comparison of cancer survival: Toronto, Ontario, and Detroit, Michigan, metropolitan areas. *Am J Public Health* 1997;187(7):1156-63.
11. Groome PA, MacKillop WJ, Rothwell DM, O'Sullivan B, Irish JC, Hall SF, et al. Management and outcome of glottic cancer: a population-based comparison between Ontario, Canada and the SEER areas of the United States. *J Otolaryngol* 2000;29(2):67-77.
12. Roos LL, Fisher ES, Sharp SM, Newhouse JP, Anderson G, Bubolz TA. Postsurgical mortality in Manitoba and New England. *JAMA* 1990;1263(18):2453-8.
13. Norregaard JC, Hindsberger C, Alonso J, Bellan L, Bernth-Petersen P, Black 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;1116(8):1095-100.
14. Norregaard JC, Bernth-Petersen P, Bellan L, Alonso J, Black C, Dunn E, et al. Intraoperative clinical practice and risk of early complications after cataract extraction in the United States, Canada, Denmark, and Spain. *Ophthalmology* 1999;1106(1):42-8.
15. Ho V, Hamilton BH, Roos LL. Multiple approaches to assessing the effects of delays for hip fracture patients in the United States and Canada. *Health Serv Res* 2000;134(7):1499-518.
16. Roos LL, Fisher ES, Brazauskas R, Sharp SM, Shapiro E. Health and surgical outcomes in Canada and the United States. *Health Aff (Millwood)* 1992;11(2):56-72.
17. Roos LL, Walld RK, Romano PS, Roberecki S. Short-term mortality after repair of hip fracture. Do Manitoba elderly do worse? *Med Care* 1996;134(4):310-26.

18. Keller DM, Peterson EA, Silberman G. Survival rates for four forms of cancer in the United States and Ontario. *Am J Public Health* 1997;187(7):1164-7.

19. Boyd C, Zhang-Salomons JY, Groome PA, Mackillop WJ. Associations between community income and cancer survival in Ontario, Canada, and the United States. *J Clin Oncol* 1999;17(7):2244-55.

20. Gorey KM, Holowaty EJ, Fehringer G, Laukkanen E, Richter NL, Meyer CM. An international comparison of cancer survival: relatively poor areas of Toronto, Ontario and three US metropolitan areas. *J Public Health Med* 2000;122(3):343-8.

21. Gorey KM, Holowaty EJ, Fehringer G, Laukkanen E, Richter NL, Meyer CM. An international comparison of cancer survival: metropolitan Toronto, Ontario, and Honolulu, Hawaii. *Am J Public Health* 2000;90(12):1866-72.

22. Wolfe F, Mitchell DM, Sibley JT, Fries JF, Block DA, Williams CA, et al. The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;137(4):481-94.

23. Pilote L, Racine N, Hlatky MA. Differences in the treatment of myocardial infarction in the United States and Canada. A comparison of two university hospitals. *Arch Intern Med* 1994;1154(10):1090-6.

24. Yusuf S, Flather M, Pogue J, Hunt D, Varigos J, Piegas L, et al. OASIS (Organisation to Assess Strategies for Ischaemic Syndromes) Registry Investigators.Variations between countries in invasive cardiac procedures and outcomes in patients with suspected unstable angina or myocardial infarction without initial ST elevation. *Lancet* 1998;1352(9127):507-14.

25. Anderson HV, Gibson RS, Stone PH, Cannon CP, Aguirre F, Thompson B, 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):1441-6.

26. Tu JV, Pashos CL, Naylor CD, Chen E, Normand SL, Newhouse JP, 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;1336(21):1500-5.

27. Gilpin EA, Koziol JA, Madsen EB, Henning H, Ross J Jr. Periods of differing mortality distribution during the first year after acute myocardial infarction. *Am J Cardiol* 1983;152(3):240-4.

28. Koyama H, Cecka JM, Terasaki PI. Kidney transplants in black recipients. HLA matching and other factors affecting long-term graft survival. *Transplantation* 1994;157(7):1064-8.

29. Hornberger JC, Garber AM, Jeffery JR. Mortality, hospital admissions, and medical costs of end-stage renal disease in the United States and Manitoba Canada. *Med Care* 1997;135(7):686-700.

30. Moran J. Changes in the dose of peritoneal dialysis: Have these independently affected outcomes? *Am J Kidney Dis* 1998;132(6 Suppl 4):S52-7.

31. Mohr PR, Rajan S, Rowell P, Newschaffer C, Bernardin M, Juday T. *Technology and outcomes: international comparisons*. Bethesda (MD); Project HOPE Center for Health Affairs; 1993.

32. Hussey PS, Anderson GF, Osborn R, Feek C, McLaughlin V, Millar J, et al. How does the quality of care compare in five countries? *Health Aff (Millwood)* 2004;123(3): 89-99.

33. Churchill DN, Thorpe KE, Vonsh EF, Keshaviah PR. 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;18(6):965-71.

34. Groome PA, O'Sullivan B, Irish JC, Rothwell DM, Schulze K, Warde PR, 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;121(3):496-505.

35. Groome PA, O'Sullivan B, Irish JC, Rothwell DM, Math KS, Bissett RJ, 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;154(3):301-15.

36. Gorey KM, Holowaty EJ, LaukkanenE, Fehringer G, Richter NL. An international comparison of cancer survival: advantage of Toronto's poor over the near poor of Detroit. *Can J Public Health* 1998;189(2):102-4.

37. Skarsgard DP, Groome PA, Mackillop WJ, Zhou S, Rothwll D, Dixon PF, et al. Cancers of the upper aerodigestive tract in Ontario, Canada, and the United States. *Cancer* 2000;188(7):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;173(1):27-40.

39. Jones SE, Miller TP, Connors JM. 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;17(9):1186-91.

40. LoCoco S, Covens A, Carney M, Franssen E, Dogde R, Rosen B, et al. Does aggressive therapy improve survival in suboptimal stage IIIc/IV ovarian cancer? A Canadian–American comparative study. *Gynecol Oncol* 1995;159(2):194-9.

41. Lai HC, Corey M, FitzSimmons S, Kosorok MR, Farrell PM. Comparison of growth status of patients with cystic fibrosis between the United States and Canada. *Am J Clin Nutr* 1999;169(3):531-8.

42. Corey M, McLaughlin FJ, Williams M, Levison H. A comparison of survival, growth, and pulmonary function in patients with cystic fibrosis in Boston and Toronto. *J Clin Epidemiol* 1988;141(6):583-91.

43. Chan KC, Galli RA, Montaner JS, Harrigan PR. Survival rates after the initiation of antiretroviral therapy stratified by CD4 cell counts in two cohorts in Canada and the United States. *AIDS* 2002;116(12):1693-5.

44. Grumbach K, Anderson GM, Luft HS, Roos LL, Brook R. Regionalization of cardiac surgery in the United States and Canada. Geographic access, choice, and outcomes. *JAMA* 1995;1274(16):1282-8.

45. Katz SJ, Armstrong RW, LoGerfo JP. The adequacy of prenatal care and incidence of low birthweight among the poor in Washington State and British Columbia. *Am J Public Health* 1994;184(6):986-91.

46. Rapoport J, Teres D, Barnett R, Jacobs P, Shustack A, Lemeshow S, et al. A comparison of intensive care unit utilization in Alberta and western Massachusetts. *Crit Care Med* 1995;123(8):1336-46.

47. Duckworth GS, Kedward HB, Bailey WF. Prognosis of mental illness in old age: a four year follow-up study. *Can J Psychiatry* 1979;124(7):674-82.

48. Sherrer YS, Block DA, Mitchell DM, Roth SH, Wolfe F, Fries JF. Disability in rheumatoid arthritis: comparison of prognostic factors across three populations. *J Rheumatol* 1987;114(4):705-9.

49. Mazzotta P, Maltepe C, Navioz Y, Magee LA, Korne G. Attitudes, management and consequences of nausea and vomiting of pregnancy in the United States and Canada. *Int J Gynaecol Obstet* 2000;170(3):359-65.

50. Boulanger BR, McLellan BA, Sharkey PW, Rizoli S, Mitchell K, Rodriguez A. A comparison between a Canadian regional trauma unit and an American level I trauma center. *J Trauma* 1993;135(2):261-6.

51. Szick S, Angus 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?* Toronto: Institute for Clinical and Evaluative Sciences; 1999.

52. Katz SJ, Zemencuk JK, Hofer TP. Breast cancer screening in the United States and Canada, 1994: socioeconomic gradients persist. *Am J Public Health* 2000;190(5):799-803.

53. Finkelstein MM. Do factors other than need determine utilization of physicians' services in Ontario? *CMAJ* 2001;1165(5):565-70.

54. Katz SJ, Hofer TP. Socioeconomic disparities in preventive care persist despite universal coverage. Breast and cervical cancer screening in Ontario and the United States. *JAMA* 1994;1272(7):530-4.

55. Davis K. Inequality and access to health care. Milbank Q 1991;169(2):253-73.

56. Andersen RM, Davidson PL. Improving access in America: individual and contextual indicators. In *Changing the US health care system: key issues in health services policy and management.* Andersen RM, Rice TH, Kominski GF, editors. San Francisco: Jossey-Bass; 2001. p 3-30.

57. Ross NA, Wolfson MC, Dunn JR, Berthelot HM, Kaplan GA, Lynch JW. Relation between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics. *BMJ* 2000;1320(7239):898-902.

58. Sekkarie M, Cosma M, Mendelssohn D. Nonreferral and nonacceptance to dialysis by primary care physicians and nephrologists in Canada and the United States. *Am J Kidney Dis* 2001;138(1):36-41.

59. Devereaux PJ, Schunemann HJ, Ravindran N, Bhandari M, Garg AX, Choi PT, 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):2449-57.

60. Woolhandler S, Campbell T, Himmelstein DU. Costs of health care administration in the United States and Canada. *N Engl J Med* 2003;1349(8):768-75.

61. Devereaux PJ, Heels-Ansdell D, Lacchetti C, Haines T, Burns KE, Cook DJ, et al. Payments for care at private for-profit and private not-for-profit hospitals: a systematic review and meta-analysis. *CMAJ* 2004;1170(12):1817-24.

62. Devereaux PJ, Choi PT, Lacchetti C, Weaver B, Schunemann HJ, Haines T, 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;1166(11):1399-406.

Table 1: Summary of findings

	High-quality studies	Low-quality studies
Results favoured United States	2	3
Results favoured Canada	5	9
Mixed or equivocal results	3	16

Table 2: Studies with results favouring the United States

Author (year)	Data dates	Population, total number of patients	Setting, basis of selection	Primary data sources	Statistical adjustment	Monitoring start	Follow-up duration	Losses to follow-up	Outcomes
Wolfe (1994)	US 1965–1990 Canada 1955–1990	Rheumatoid arthritis Adult patients, most < 65 y Canada: symptoms start after age 16 US: 2596 Canada: 905	Wichita, KS Stanford, CA Santa Clara, CA Saskatoon, SK Availability of data	Patient charts single observer only, independent review of death certificates, administrative database	Age at entry, sex, survival in underlying populations, year of entry into study, duration of follow-up	Diagnosis: uncertain timing in course of illness	Up to 35 y	US: 14% Canada: 4.5%	**Mortality** US: lower mortality (US 20.39% vs Canada 53.13% , $p < 0.0001$) when calculated from study data
Norregaard (1998, 1999)	US 1991–1992 Canada 1992–1993	Cataracts 50 y or older No previous cataract surgery, live in area, access to telephone, speak local language US: 772 Canada: 159	Columbus, OH St. Louis, MO Houston, TX Manitoba Differences in organization of care and predominant surgical techniques	Patient charts, face-to-face and phone interviews	Age, sex, comorbidity, preoperative visual acuity, general health	Scheduling of surgery	4 months	≈ 5%	**Intra- and postoperative adverse events, visual acuity** US: intraoperative adverse events significantly lower (2.1 times less likely) Better mean 4-month visual acuity (US 0.74 vs Canada 0.66, $p < 0.001$) Probability of postoperative adverse events similar
Pilote (1994)	US 1989–1991 Canada 1989–1991	MI Excluded hospital transfers Adult patients, most > 65 y US: 233 Canada: 285	Stanford University, Palo Alto, CA McGill University, Montreal, QC Similar population, size of city and involvement with academic institutions, full range of cardiac services	Phone interview with patients, mail survey of patients, post-discharge questionnaires, patient charts	Age, sex, race, comorbidity, smoking, Q wave and anterior infarctions, peak creatine kinase levels	Hospital admission	Median follow-up of 20 months	US: < 1% in-patients, 23 post discharge Canada: < 1% as in-patients, 12 % post discharge	**Recurrent MI, mortality, functional status** Recurrent MI rates in hospital and after discharge similar (13% for Stanford, 8% for McGill, $p > 0.05$) Mortality rates similar (28% at Stanford, 27% at McGill, $p > 0.05$) US: functional status of patients higher at Stanford (mean DASI score US 28.8 vs Canada 22.9, $p =$

									0.006), adjusted DASI 6.5 points greater for US, p = 0.0007 up to 2.5 y follow-up
Ho (2000)	US 1990–1992 Canada 1990–1994	Hip fracture Patients 45 y or older, hospitalized for 365 d or less US: 25,919 Canada: 13,597	California Massachusetts Quebec Manitoba Availability of data	Administrative databases US: Nationwide Inpatient Sample Canada: Quebec Ministry of Health, Manitoba Health File	Age, sex, comorbidity, day of week, fracture type, volume of surgery in hospital	Hospital admission	End of hospital stay	< 1%	**Post-surgery length of stay, mortality** US: shorter post-surgery length of stay (Manitoba 31.1 d, Quebec 28.2 d vs Massachusetts 13.7 d, California 8.6), lower inpatient mortality (> 5% in Canada and < 5% in US)
Keller (1997)	US 1978–1990 Canada 1978–1990	Breast, colon, lung cancers, Hodgkin's disease Pediatric and adult patients US and Canada: uncertain number of patients	US: SEER program areas Canada: Ontario	Administrative databases US: SEER program areas Canada: Ontario Cancer Registry	Age, sex, region, survival differences in under-lying populations, year of observation	Diagnosis: uncertain timing in course of illness	4–13 y	< 1%	**Relative survival** Similar cumulative relative survival rates for colon cancer, lung cancer and Hodgkin's disease US: cumulative survival rate for breast cancer consistently 4 to 5 percentage points higher starting at the fourth year

Unshaded: studies deemed high quality from population registries with adequate adjustment.

Lightly shaded: studies deemed of intermediate quality that were reasonably representative and had adequate adjustment.

Shaded: studies deemed of low-quality in which the populations were unrepresentative or adjustments were inadequate.

Much of the data of Gorey (Table 3), Boyd (Table 4), and Keller (above) overlap; however, Keller does not stratify subjects by socioeconomic status and does analyze data spanning a greater number of years. All articles were deemed eligible and were included.

MI = myocardial infarction
SEER = National Cancer Institute's Surveillance, Epidemiology, and End Results Program
DASI = Duke Activity Status Index

Table 3: Studies with results favouring Canada

Author (year)	Data dates	Population, total number of patients	Setting, basis of selection	Primary data sources	Statistical adjustment	Monitoring start	Follow-up duration	Losses to follow-up	Outcomes
Juday (1994) Mohr (1993)	US 1978–1989 Canada 1978–1989	Renal failure Mixed pediatric and adult US: uncertain Canada: uncertain	Entire US, all of Canada	Administrative databases	Age, sex,	Treatment for end-stage renal disease	Not specified	Uncertain	**Mortality** US: higher mortality
Moran (1998)	US 1987–1993 Canada 1981–1996	Renal failure, peritoneal dialysis Adult patients US: uncertain Canada: uncertain	Entire US, all of Canada	Administrative databases US: US Renal Data System (USRDS) Canada: Canadian Organ Replacement Register (CORR)	Age, sex, race, geographic region, consistent adjustors	Treatment for end-stage renal disease	3 y	Uncertain	**Mortality** Canada: lower mortality and “more marked” improvement in survival rates
Churchill (1997)	US 1990–1993 Canada 1990–1993	Renal failure, peritoneal dialysis Adult patients, most < 65 y US: 81 Canada: 491	Milwaukee, WI Columbia, MO Philadelphia, PA Augusta, GA Halifax, NS Hamilton, ON Mississauga, ON Ottawa, ON (two hospitals) Winnipeg, MB Calgary, AB Edmonton, AB	Active data collection as part of prospective cohort study	Age, sex, comorbidity, underlying renal disease, diabetes, albumin, cardiovascular disease at baseline, heart failure, nutritional status	Start of chronic ambulatory peritoneal dialysis	2 y	Uncertain	**Mortality, cardiovascular morbidity, peritonitis** Canada: higher 2 y survival : US 63.2% vs Canada 79.7% , $p < 0$ .05, RR 1.87 (95% CI 1.09–3.19) Higher 2 y survival without a nonfatal cardiovascular event (65.9% vs 53.5%, $p = 0.02$; RR 1.80, 95% CI 1.21–2.67) No difference in time to first episode of peritonitis (2 y peritonitis free US 46.5% vs Canada 39.2%, $p = 0.69$)
Koyama (1994)	US 1987– ? Canada 1987– ?	Renal transplant recipients Mixed pediatric and adult, most < 65 y US: 5282 black,	Entire US (30 centres), all of Canada Availability of data, locations with greater than 50 transplants of each	United Network for Organ Sharing Scientific Renal Transplant Registry UCLA	None	Time of surgery	3 y	Uncertain	**Comparative graft survival between black and white patients** US: 1 y and 3 y graft survival rates significantly lower among black recipients as compared to

		14,917 white Canada: uncertain	black and white recipients	International Transplant Registry					white recipients (76.3% vs 80.6% and 56.8% vs 70.2%, $p < 0.001$) Canada: graft survival rates similar between blacks and whites
Hornberger (1997)	US 1986–1992 Canada 1983–1989	Renal failure Mixed pediatric and adult patients, most < 65 y, excluded those dying within 90 d of diagnosis US: 6770 Canada: 549	US: entire Canada: Manitoba Availability of data, Manitoba has outcomes similar to rest of Canada	Administrative databases US: Renal Data System Casemix Severity Study Canada: patient charts, Manitoba Health Services Commission	Age, sex, disease stage/ severity, race, comorbidity	Treatment of end-stage renal disease	Until April 1992	Uncertain	**Mortality** Canada: mortality rate significantly lower: adjusted relative mortality in US 1.47 (95% CI 1.16–1.87)
Anderson (1997)	US 1990–1993 Canada 1990–1993	Non Q-wave MI, unstable angina, most < 65 y US: 1733 Canada: 642	14 American and 4 Canadian tertiary care centers participating in the Thrombolysis in Myocardial Ischemia (TIMI) registry	In-person and phone interviews of patients, doctors and relatives, and chart review	Age, sex	Hospital admission	6 weeks and 1 y	Uncertain	**Death, acute MI, recurrent ischemia** Similar mortality and infarction rates during hospital stay, 6 weeks and 1 y Canada: incidence of death, acute MI or recurrent ischemia lower Hospital admission US 12.9% vs Canada 10.0%, $p < 0.001$ 6 weeks: US 18.4% vs Canada 13.9%, $p = 0.004$ 1 y: US 30.3% vs Canada 27.3%, $p = 0.046$
Groome (2003)	US 1988–1995 Canada 1988–1996	Supraglottic cancer Adults of all ages US: 1643 Canada: 265	Georgia, California, Iowa, Louisiana, New Jersey, New Mexico, Washington State, Conneticut, Utah, Kentucky (SEER program areas) Canada: Ontario Availability of data	Administrative database supplemented by chart review in Canada) US: SEER registries, Linked Medicare–Tumor Registry Database Canada: Ontario Cancer Registry	Age, sex, disease stage/ severity, differences in survival in underlying populations	Diagnosis: first mention in the charts or biopsy-proven squamous cell carcinoma of the larynx (whichever came first)	5 y	< 1%	**Survival, laryngectomy** 5 y survival: Canada: significantly higher for stage I/II (74.3% vs 55.8%, $p = 0.01$) No significant difference in stage III and IV: Laryngectomy: 3 y actuarial rates lower in Canada: stage I/II Canada 3% vs US 35%, $p < 0.001$, stage III Canada 30% vs US 54%, $p = 0.03$, stage

									IV Canada 33% vs US 64%, $p = 0.002$
Gorey (2000)	US 1986–1995 Canada 1986–1995	Breast and prostate cancer 25 y or older US: 1783 breast cancer, 1112 prostate cancer Canada: 5807 breast cancer, 3383 prostate cancer	Honolulu, HI (Hawaii minimizes risk of being uninsured) Canada: Toronto Similar demographics, availability of data and services	Administrative databases US: SEER program Canada: Ontario Cancer Registry	Age, income	Diagnosis in first 3 months of illness, death certificate	5 y	< 1%	**Survival** Canada: significantly higher 5 y survival in low-income groups: SRR = 1.06 (1.00, 1.12) for breast cancer, SRR = 1.10 (1.00, 1.22) for prostate cancer, lowest income groups SRR = 1.20 (1.06, 1.36) for breast cancer, SRR = 1.24 (1.01, 1.53) for prostate cancer Survival in middle- and high-income groups not significantly different
Gorey (2000)	US 1986–1993 Canada 1986–1993	Breast, colon, rectal, lung, bladder, oral, cervical, stomach, pancreas, kidney, ovarian, uterus, central nervous system, prostate cancers, lymphoma, Adult patients, most > 65 y US: 37,329 Canada: 23,437	Seattle, WA San Francisco, CA Hartford, CT Toronto, ON Availability of data, US cities "most advantaged," greater than 1 million people	Administrative databases Canada: Ontario Cancer Registry US: SEER Census reports	Age, income	Diagnosis in first 3 months of illness, death certificate	5 y	< 1%	**Survival** Canada: 5 y survival advantage for lowest-income third of population: SRR = 1.35 (95% CI 1.30–1.40), even greater for those diagnosed before age 65: SRR = 1.46 (95% CI 1.40–1.52) No significant difference in cancer survival in middle- and high-income groups
Gorey (1998)	US 1986–1993 Canada 1986–1993	Breast, colon, rectal, bladder, oral, cervical, prostate cancers Adult, low socio-economic status US: 3709 Canada: 2702	Detroit, MI Toronto, ON Similar mix of urban and rural, availability of data, prior study done using these locales	Administrative databases US: SEER Canada: Ontario Cancer Registry	Age, income	Diagnosis in first 3 months of illness, death certificate	5 y	< 1%	**Survival** Canada: lowest income quintile better in Canada vs second-lowest income quintile in Detroit; SRR for 5 y is 1.20 (95% CI 1.16–1.24)
Groome (2001)	US 1988–1994 Canada 1982–1995	Glottic cancer Adults US: 3921 Canada: 3295	US: SEER Program areas Canada: Ontario Availability of data	Administrative database supplemented by chart review in Canada US: SEER registries, Linked Medicare–Tumor	Differences in survival in underlying populations; analyses by disease stage for some patients	Diagnosis: first mention in the charts or biopsy-proven squamous cell carcinoma	5 y	< 1%	**Survival, laryngectomy** No overall survival difference Canada: better relative survival for localized disease at 2, 3 and 4 y

				Registry Database Canada: Ontario Cancer Registry		of the larynx — whichever came first			Lower overall rate of laryngectomy: US 25.3% vs Canada 21.3%, p = 0.01
Chan (2002)	US 1994– ? Canada 1994– ?	AIDS Mixed pediatric and adult patients, 13 y and older US: 5110 Canada: 1219	US: 11 cities, 9 specified (Atlanta, Dallas, Denver, Los Angeles, San Antonio, Houston, New Orleans, Seattle, Detroit) Canada: British Columbia Availability of data	Administrative databases Canada: British Columbia Drug Treatment Program US: Centers for Disease Control and Prevention Adult and Adolescent Spectrum of Disease Project	Baseline CD4 count	Initiation of double or triple therapy	2 y	Uncertain	**Survival** Canada: higher overall 2 y survival rate 93.4% (95% CI 92.3–94.5) vs 80.9% (data unavailable to calculate CI)
Corey (1988)	US 1972–1982 Canada 1972–1982	Cystic fibrosis Mixed pediatric and adult patients, < 65 y US: 499 Canada: 534	Boston, MA Toronto, ON Similar weather, demographics in population, population size of city, involvement with academic institutions, size of institutions, volume of patients	Patient charts, administrative database (Cystic Fibrosis Patient Registry)	Age, sex	Diagnosis: uncertain timing in course of illness	10 y	Uncertain	**Survival, growth, pulmonary function** Canada: age of survival significantly higher (30 y vs 21 y), mean percentile height greater (male Canada average at the 42nd percentile vs US 33rd percentile, p < 0.001, female Canada 44th vs US 33rd, p < 0.01) Similar age-specific pulmonary function
Lai (1999)	US 1992–1994 Canada 1992–1994	Cystic fibrosis Mixed pediatric and adult patients, < 65 y US: 20,610 Canada: 3145	Entire US, all of Canada	Administrative databases US: Cystic Fibrosis Patient Registry Canada: Canadian Patient Data Registry	Sex only in multivariate analysis but effects consistent across genotypes, comorbidities	1992	2 y	Uncertain	**Height, weight** Canada: for children, mean height 4–5 percentiles higher (p < 0.01) For adults significantly higher weight (26th vs 21st percentiles) and percentage of ideal weight (93% vs 90%)

Unshaded: studies deemed high quality from population registries with adequate adjustment.

Lightly shaded: studies deemed of intermediate quality that were reasonably representative and had adequate adjustment.

Shaded: studies deemed of low-quality in which the populations were unrepresentative or adjustments were inadequate.

The population analyzed in Churchill's article (794) is a subpopulation of that of Moran; the information analyzed, however, does differ. All articles were deemed eligible and were included.

The Toronto data in the 3 articles by Gorey overlap – the cities' data to which the Toronto data is compared, however, does differ. All articles were deemed eligible and were included.

The Ontario data employed by Boyd (above) overlaps those of Gorey's articles (Table 3); Boyd, however, compares Ontario data to the entire US and has other methodologic differences. All articles were deemed eligible and were included.

Much of the data of Gorey (Table 3), Boyd (above), and Keller (Table 2) overlap; however, Keller does not stratify subjects by socioeconomic status and does analyze data spanning a greater number of years. All articles were deemed eligible and were included.

Due to the partial overlap of data among articles, the glottic cancer data in Groome (2001) and supraglottic cancer data in Groome (2003) is considered in this analysis; we excluded Groome (2000) because of overlap with the other 2 studies.

MI = myocardial infarction
SEER = National Cancer Institute's Survellance, Epidemiolocy, and End Results program
SRR = survival rate ratio
DASI Duke Activity Status Index

Table 4: Studies with equivalent, equivocal or mixed results

Author (year)	Data dates	Population, total number of patients	Setting, basis of selection	Primary data sources	Statistical adjustment	Monitoring start	Follow-up duration	Losses to follow-up	Outcomes
Campbell (1966)	US 1953–1962 Canada 1953–1962	Cervical cancer Adults of all ages US: uncertain number of patients Canada: uncertain number of patients	Entire US, all of Canada	Patient charts independent review of death certificates	None	Diagnosis in first 3 months of illness	5 y	US: 1.4% Canada: 0.9%	**Survival** No significant difference in 5 y survival: US 52% vs Canada 53% when calculated from study data
Jones (1989)	US 1971–1988 Canada 1980–1988	Lymphoma Adults US: 64 Canada: 78	Tucson, AZ Vancouver, BC Availability of data	Patient charts	None	Diagnosis in first 3 months of illness	2 y to < 5 y	< 1%	**Relapse-free survival** No significant difference in relapse-free surviv No treatment-related mortality or instances of leukemia or late cardiac toxicity
Boyd (1999)	US 1987–1994 Canada 1987–1994	Breast, colon, rectal, lung, bladder, cervical, stomach, pancreas, ovarian, central nervous system, esophageal, head and neck, uterine, testicular cancers, lymphoma US: uncertain Canada: uncertain	US: SEER areas Canada: Ontario	Administrative databases US: SEER, 1990 census Canada: Ontario Cancer Registry, 1991 census	Age, sex, year of diagnosis	Diagnosis in first 3 months of illness, death certificate	5 y	< 1%	**Survival** Canadian overall survival advantage ($p < 0.05$ in head and neck, lung, pancreas, ovary. American survival advantage in colon and rectum, breast, prostate, and bladder. Canadians in lower income communities tended to do better than Americans in lower income communities, but Americans in higher income communities tended to do better than Canadians in higher income communities.
LoCoco (1995)	US 1987–1993 Canada 1987–1993	Suboptimally debulked Stage IIIc or IV ovarian cancer Adults, most < 65 y US: 68 Canada: 61	Durham, NC Toronto, ON Tertiary care cancer centres	Patient charts single observer only	Age, sex, comorbidity, histology, disease stage, surgeon specialty	Diagnosis and classification of disease stage	October 1993	Uncertain	**Complications, survival** Proportion of significant complications similar (US 10% vs Canada 6%, p = NS) Similar 5 y survival (US 11% vs Canada 10%)
Skarsgard (2000)	US 1982–1994 Canada 1982–1994	Oral cavity, oropharyngeal, nasopharyngeal, hypopharyngeal, laryngeal, sinus, nasal cavity, middle ear, auditory tube, and	Connecticut, Iowa, New Mexico, Utah, Hawaii, Detroit MI, San Franciso–Oakland CA, Seattle–Puget Sound WI, Atlanta	Administrative databases US: SEER Canada: Ontario Cancer Registry	Age, sex, difference in survival in underlying populations, year of diagnosis	Diagnosis in first 3 months of illness	10 y	< 1%	**Relative survival** **Canada:** higher 5 y relative survival for cancer of the oral cavity: Canada 55.8 % (95% CI 54.0–57.7) vs US 51.8% (50.6-53.0), p = 0.0001

		mastoid cancers Adults of all ages US: 42,990 Canada: 16,577	GA Canada: Ontario Availability of data		US: adjustment for race				Cancer of the nasopharynx: Canada 54.5% (50.5-58.5) vs US 49.2% (46.3-52.1), p = 0.05888 **US:** higher relative survival for cancer of the supraglottis: US 48.6% (46.8-50.4) vs Canada 45.5% (42.5-48.5), p = 0.0435 No significant difference for other cancers
Hussey (2004)	US ? Canada: ?	Breast, colon, cervical cancers, lymphoma, renal failure, liver transplant recipients US: uncertain Canada: uncertain	Entire US, all of Canada	Administrative databases	Uncertain	Diagnosis in first 3 months of illness	variable	Uncertain	**Survival** **Canada:** 5 y relative survival rates higher for colorectal cancer (5%), childhood leukemia (8%), kidney transplant (13 %), liver transplan (21%) **US:** 5 y relative survival rates higher for brea cancer (10%), cervical cancer (2 %), non-Hodgkin's lymphoma (2%)
Rapoport (1995)	US 1990–1991 Canada 1990–1991	Diagnosis related groups for which most ICU care provided Adult patients US: uncertain Canada: uncertain	Massachusetts Alberta	Administrative databases, other study (European-North American study of severity) Canada: Alberta Ministry of Health US: Massachusetts health Data Consortium	Age	Hospital admission	End of hospital stay	< 1%	**Mortality** Hospital mortality rates similar
Yusuf (1998)	US 1995–1996 Canada 1995–1996	Admitted within 48 h of acute ischemic cardiac chest pain US: 918 Canada: 1626	US: 11 hospitals, uncertain criteria Canada: 11 hospitals, uncertain criteria Broad range of practice patterns and varied availability of catheterization facilities	Patient charts single observer only	Age, comorbidity, heart rate, systolic blood pressure, abnormal ECG	Hospital admission	6 months	Uncertain	**MI, cardiovascular death, refractory angina stroke, major bleeding** Rates of MI or cardiovascular death similar: **Canada:** less stroke and major bleeding **US:** lower rates of refractory angina at 7 d and readmission for unstable angina at 6 months
Grumbach (1995)	US 1987–1989 Canada 1987–1989	Coronary artery disease with coronary artery bypass grafting US: uncertain Canada: uncertain	New York California Manitoba, Ontario, British Columbia All non-federal hospitals, availability of data, comparison with	Administrative databases Computerized hospital discharge abstracts	Age, sex	Time of surgery	End of hospital stay or 14 d after coronary artery bypass grafting	< 1%	**Mortality** No significant difference in mortality when calculated from study data

			previous study						
Tu (1997)	US 1991– Canada 1991	MI Patients > 65 y US: 224,258 Canada: 9,444	US: entire Canada: Ontario Availability of data	Administrative databases US: Health Care Financing Administration's Health Insurance database Canada: Canadian Institute for Health Information (CIHI database	Age, sex	Hospital admission	1 y	< 1%	**Mortality** **US:** significantly lower 30 d mortality rates (US 21.4% vs Canada 22.3%, $p = 0.03$) Similar 1 y mortality rates (US 34.3% vs Canada 34.4%, $p = 0.94$)
Gilpin (1983)	US 1968–1979 Canada 1977-1980	MI, admission within 24 h of symptom onset US: 346 Canada: 704	San Diego, CA Vancouver, BC	Research databases	None	Hospital admission	6 months to 2 y	7%	**Mortality** Similar 24 h and 1 yr mortalities (US 6% vs Canada 5% for 24 h, both 21% for 1 y)
Rowe (2001)	US 1996–1998 Canada 1996–1998	Asthma presenting to the emergency department Pediatric and adult < 65 y US: 2,876 Canada: 155	US: 69 emergency departments across the country Canada: 8 emergency departments in Quebec, Alberta, Ontario, British Columbia Data availability	Interview of patient at visit, phone interview with patients, patient charts single observer only	Age, sex, disease stage/ severity, income	Arrival in emergency department	2 weeks	US: 24% Canada: 12%	**Hospital admission and length of stay, relapse events, severity** Similar long stay in emergency department: U 30% vs Canada 27%) Similar relapse rates: odds ratio for Canada v US 1.6 (95% CI 0.7 to 3.3), $p = 0.25$
Mazzotta (2000)	US 1996–1998 Canada 1996–1998	Women < 65 with nausea and vomiting of pregnancy US: 611 Canada: 833	Entire US, all of Canada	Phone interview with patients at approximately 8 and 20 weeks' gestation	None	Diagnosis in first 3 months of illness	12 weeks	Uncertain	**Severity of symptoms, impact of nausea and vomiting of pregnancy on emotional and social functioning** **Canada:** more depression, greater adverse effects on marital relationships, fewer lost hou of paid work, less hospitalization, and less weight loss
Duckworth (1979)	US 1972? Canada 1972–1977	Dementia > 65 y US: 50 Canada: 100 patients	Two public hospitals in New York, NY Three hospitals in Toronto, ON	Patient charts	None	Hospital admission	US: 90 d Canada: 4 y	Uncertain	**Mortality, disposition to community** No significant difference in mortality or rates c discharge
Sherrer (1987)	US 1976–1983	Rheumatoid arthritis Mixed paediatric and	Phoenix, AZ Wichita, KS Saskatoon, SK	Patient charts	Age, sex	Diagnosis: uncertain timing in	US: average 1.7 to 3.4 y	< 10%	**Functional disability**

	Canada 1966–1983	adult patients US: 1654 Canada: 892 patients	Availability of data			course of illness	Canada: 12 y		When calculated from study data, no significa difference in functional disability
Roos (1996)	US 1984–1985 Canada 1979–1992	Hip fracture > 65 y US: 16,206 Canada: 10,007	US: New England Canada: Manitoba Population-based registry, similar rates of repair and re-operation	Administrative databases US: Medicare Canada: discharge abstracts	Age, sex, comorbidity, pre-operative delay	Diagnosis in first 3 months of illness	3 y	US: <1% Canada: < 2%	**Mortality** **Canada:** Higher 30 d postsurgical mortality: adjusted odds ratio 1.35 (95% CI 1.21–1.49), but survival curves cross at 2 y (Canadian survival superior after 2 y, significance not reported)
Roos (1992)	US 1984–1988 Canada 1980–1989	Hip replacement, cholecystectomy, prostatectomy, carotid endarterectomy, coronary artery bypass grafting, heart valve replacement, hip fracture repair US: 59,720 Canada: 17,358	US: New England Canada: Manitoba Similar numbers of physicians and beds per capita	Administrative databases US: Medicare Canada: Manitoba Health Services Commission database	Age, sex, comorbidity, operation type (emergency, scheduled, delayed)	Hospital admission	3 y	US: < 1% Canada: < 2%	**Mortality** **Canada:** low mortality procedures: lower 30 d and 1 y mortality (OR 0.73, NS, and 0.76, $p <$ 0.05) Moderate mortality procedures: lower 30 d an 1 y mortality (OR 0.87, NS, and 0.76, $p <$ 0.05 High mortality procedures: higher 30 d mortality (OR 1.40, P < 0.05), and lower 3 y mortality (NS)
Katz (1994)	US 1988–1989 Canada 1987–1989	Exclusively pediatric Singlet births in non-federal hospitals US: 95,401 Canada: 79,638	US: Washington State Canada: British Columbia Similar demographics, organization of obstetric care	Administrative databases	Maternal age, number of previous pregnancies, marital status, race	Hospital admission	End of hospital stay	< 1%	**Low birthweight** Similar risk for low birth weight (< 2500 g)
Boulanger (1993)	US 1986–1990 Canada 1986–1990	Trauma, motor vehicle crash victims Exclude former trauma elective readmissions < 14 y, most < 65 y US: 4632 Canada: 1263	Baltimore, MD (1 hospital) Toronto, ON (1 hospital) Similar types / severity of illness, "renowned centres," "representative of optimal care"	Administrative database	None	Diagnosis in first 3 months of illness, hospital admission	End of hospital stay	< 1%	**Mortality, ICU admission, acute hospital cost** Similar mortality rates (15.6% in Canada and 17.7% in US, p = not significant) US: higher ICU admission (74% vs 35%, $p <$ 0.01) and longer stays in ICU (15.4 d vs 8.4 d $p <$ 0.01) Canada: hospital stays longer (26.2 d vs 18.5 d, $p <$ 0.01)

Unshaded: studies deemed high quality from population registries with adequate adjustment.

Lightly shaded: studies deemed of intermediate quality that were reasonably representative and had adequate adjustment.

Shaded: studies deemed of low-quality in which the populations were unrepresentative or adjustments were inadequate.

The Ontario data employed by Boyd overlaps those of Gorey's articles (in Table 3); Boyd, however, compares Ontario data to the entire US and has other methodologic differences. All articles were deemed eligible and were included.

Much of the data of Gorey (Table 3), Boyd, and Keller (Table 2) overlap; however, Keller does not stratify subjects by socioeconomic status and does analyze data spanning a greater number of years. All articles were deemed eligible and were included.

ICU = intensive care unit
OR = odds ratio
NS = not significant
ICU = intensive care unit