Comparison of cardiac surgery mortality reports using administrative and clinical data sources: a prospective cohort study

Cedric Manlhiot PhD, Vivek Rao MD PhD, Barry Rubin MD PhD, Douglas S. Lee MD PhD

Abstract

Background: Outcomes for coronary artery bypass surgery are of broadening interest, but the impact of data type on quality reporting has not been fully examined. We compared the performance of administrative and clinical data–based risk adjustment models at a tertiary–quaternary care hospital.

Methods: We used a prospective study design to test two risk adjustment models, one from administrative (Canadian Institute for Health Information [CIHI] Cardiac Care Quality Indicator) and one from clinical data (Society of Thoracic Surgeons), on cardiac surgical procedures performed between 2013 and 2016 (n = 1635). Our primary outcome was in-hospital mortality within 30 days of surgery. Model performance was established by comparing predicted and observed mortality, model calibration and handling of critical covariates.

Results: Observed mortality was 1.96%, which was the same as that predicted by the Society of Thoracic Surgeons model (1.96%), but significantly higher than that predicted by the CIHI model (1.03%). Despite both models having similar C statistics (0.756 CIHI; 0.758 Society of Thoracic Surgeons), the CIHI model showed significant underestimation of mortality among patients at higher risk. There was significant miscalibration of risk associated with 7 covariates: New York Heart Association class IV, congestive heart failure, ejection fraction less than 20%, atrial fibrillation, acute coronary insufficiency, cardiac compromise (shock, myocardial infarction < 24 h, intra-aortic balloon pump, cardiac resuscitation or preprocedure circulatory support) and creatinine concentration of 100 mg/dL or more. Together, these factors accounted for 84% of the difference in predicted mortality between the administrative and clinical models.

Interpretation: Risk prediction using administrative data underestimated risk of death, potentially inflating observed-to-predicted mortality ratios at hospitals with patients who are more ill. Caution is warranted when hospital reports of cardiac surgery outcomes are based on administrative data alone.

Reporting the outcomes of cardiac surgical procedures to hospitals and providers is increasingly common. However, whether to release cardiac surgical outcomes to the public is a matter of debate.1 An important consideration in both private and public release of report cards is accuracy of risk adjustment to account for case-mix differences between hospitals.2 Specifically, tertiary or quaternary referral centres may provide surgery to patients at high-risk of death, whereas other centres’ case-mixes may be lower risk overall.1 Valid benchmarking of hospitals’ performance is contingent on risk adjustment models that can adequately account for patients’ characteristics in the full range of sickness. Reporting of surgical outcomes can be performed using administrative or clinical data. Clinical databases are costly and require greater effort for data collection,4 whereas administrative data are routinely collected and inexpensive.5,6 Although the reporting of surgical outcomes using administrative databases has expanded, these data were not originally designed for evaluating the quality of surgical care. However, differences in reports generated using administrative and clinical data sources and the reasons for these differences have not been elucidated.

We examined 2 methods for risk adjustment for cardiac surgical outcomes: one from the Canadian Institute for Health Information (CIHI) and one from the Society of Thoracic Surgeons.

Competing interests: Dr. Rao reports personal fees from Medtronic, personal fees from Abbott Labs, outside the submitted work. No other competing interests were declared.

This article has been peer reviewed.

Correspondence to: Douglas Lee, dlee@ices.on.ca

CMAJ Open 2018. DOI:10.9778/cmajo.20180072
Surgeons. The Cardiac Care Quality Indicators Report from CIHI uses administrative data, whereas the Society of Thoracic Surgeons’ risk adjustment model incorporates clinical variables. Our objective was to compare these models using data from the same cohort of patients from a tertiary–quaternary care hospital. We hypothesized that the two models would differ in their performance because they contain different types of data and variables.

Methods

Data sources and abstraction
Canadian Institute for Health Information’s discharge abstract database was the primary source of administrative data for this study and contains data from all hospital admissions. Submission of discharge data to CIHI is completed upon patient discharge (or death) and includes patient demographic characteristics, admission information, diagnoses, treatments and status at discharge. Clinical data were obtained from the Division of Cardiac Surgery clinical database at the University Health Network. These data were prospectively collected by cardiac surgeons and experienced data managers and include detailed preprocedural, intraoperative and inhospital outcomes.

Patient population and outcomes
All patients who underwent cardiac surgery during fiscal years 2013/14 to 2015/16 (Apr. 1, 2013 to Mar. 1, 2016) were considered to emulate the CIHI Cardiac Care Quality Indicators Report. From this population, 3 surgical subgroups were identified: isolated coronary bypass graft surgery, isolated aortic valve surgery and combined coronary bypass graft and aortic valve surgery. All other types of surgery not listed above and reoperations within 30 days were excluded because our aim was to examine the administrative risk adjustment model while replicating the methods of the CIHI report, wherein 30-day reoperations were excluded. Dates of review, classification of patients in specific subgroups of procedures, inclusion/exclusion criteria and outcomes were selected to coincide with those used in the Cardiac Care Quality Indicators Report (see https://www.cihi.ca/en/indicatorlibrary for details). The primary outcome for this study was all-cause in-hospital death within 30 days of the procedure, as assessed through the CIHI database.

Model for risk adjustments
We compared the performance of the Society of Thoracic Surgeons model based on clinical data against the CIHI model, which is based on administrative data. Both models predict 30-day in-hospital mortality using multiple logistic regression, where the inverse logit function is used to derive the probability of death. Details of the CIHI model are provided in the Cardiac Care Quality Indicators Report Methodology and Technical Notes.7,8 Details of the Society of Thoracic Surgeons models have been published elsewhere.9–11 We used the aforementioned methodologies to generate predictions for 30-day in-hospital mortality after cardiac surgery. The CIHI directly provided regression coefficients for each of the 3 risk models that they generated. Data to generate mortality probabilities with the Society of Thoracic Surgeons method were obtained from the prospectively maintained cardiovascular surgery database. Data to generate mortality probabilities with the CIHI methodology were obtained from the administrative data that are submitted to CIHI as part of mandatory reporting. The CIHI directly provided the data for 2 of the risk factors in their model: previous cardiac surgery and previous acute myocardial infarction within 365 days, because these data were not available in our internal administrative database.

Statistical analysis
Crude mortalities were reported as percentages and 95% confidence intervals (CIs). Differences between predicted mortality derived from CIHI and Society of Thoracic Surgeons risk models and factors associated with those differences were assessed using linear regression. We compared model fit for mortality within our study population using C statistics from logistic regression models, and standard calibration curves were constructed for both models. For both risk adjustment strategies, observed-to-predicted ratios were calculated overall and for various patient subgroups to assess any imbalance in risk adjustment. Finally, we simulated the change in the predicted number of deaths when recalibrating the CIHI model for various patient characteristics that were found to affect risk adjustment. To do this, we used the Society of Thoracic Surgeons–estimated effect on mortality to estimate the recalibration coefficient and calculate the effect on predicted mortality in our patient population. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC).

Ethics approval
Ethical approval for this study was obtained from the University Health Network Research Ethics Board.

Results

Patient population: inclusions and exclusions
During fiscal years 2013/14 to 2015/16 (Apr. 1, 2013 to Mar. 1, 2016), there were 4212 cardiac surgeries performed, including 1344 isolated coronary artery bypass grafts, 665 valve procedures, 334 combined coronary bypass and valve procedures and 1669 other procedures. As per the CIHI–prescribed method, patients who underwent valve procedures (other than aortic) were excluded. Furthermore, procedures were excluded owing to mismatches in patient identifiers or procedure date. After these exclusions, we identified 1635 cardiac procedures: 1341 isolated coronary bypass grafts, 143 isolated aortic valve replacements and 151 combined procedures (Figure 1).

Comparing models (Appendix 1, available at www.cmajopen.ca/content/6/3/E316/suppl/DC1), the Society of Thoracic Surgeons risk adjustment model includes more data on cardiac function, cardiac pathologies and patient status immediately before surgery. The CIHI risk adjustment model includes more details on noncardiac comorbidities.
In-hospital mortality

Of 1635 patients, 32 died within 30 days of surgery (crude mortality 1.96%, 95% CI 1.40% to 2.75%). Comparisons of observed versus predicted mortality are shown in Figure 2. For all surgery types, crude mortality was either at or below the rates predicted by the Society of Thoracic Surgeons model. Conversely, crude mortalities for isolated coronary bypass graft and isolated aortic valve replacement were significantly higher than those predicted by the CIHI model. For combined procedures, crude mortality was lower than either of the model-predicted rates, though not significantly, given the small number of events.

Risk model performance and calibration

Both risk prediction models had virtually identical C statistics (0.756 for CIHI v. 0.758 for Society of Thoracic Surgeons). Despite this similarity, CIHI-predicted mortality was significantly lower than Society of Thoracic Surgeons-predicted mortality, with a mean absolute difference of −1.03% (95% CI −1.19% to −0.86%, p < 0.001). Using the CIHI model, 77% of patients had lower predicted mortality. In relative terms, the CIHI model underestimated risk of death by a mean of 23.7% (95% CI 19.6% to 27.7%, p < 0.001). The underestimation was more pronounced among patients who

Figure 1: Identification of patient population. Note: AVR = aortic valve replacement, CABG = coronary artery bypass graft.

Figure 2: Crude mortality (and 95% confidence interval [CI]) compared with rates predicted by Canadian Institute for Health Information (teal bars) and Society of Thoracic Surgeons (blue bars) models. Note: AVR = aortic valve replacement, CABG = coronary artery bypass graft.
died within 30 days of surgery (mean –2.47% [95% CI –5.24% to 2.92%]) than among those who did not (–1.00% [95% CI –1.16% to –0.84%]), which was significant (p = 0.01). Underestimation of CIHI-predicted mortality was more pronounced in patients at higher predicted risk, as evidenced by the calibration plots presented in Figure 3.

The CIHI model predicted that there would be 17 deaths within 30 days of surgery among 1635 cardiac surgeries. With 32 observed deaths (26 isolated coronary bypass, 4 isolated aortic valve replacement and 2 combined procedures), the CIHI observed-to-predicted ratio was statistically higher than expected, at 1.88. This was in sharp contrast with the Society of Thoracic Surgeons model, which predicted 34 deaths during the same period (observed-to-predicted ratio 0.94, no statistical differences). To further explore this discrepancy, we assessed all variables in the Society of Thoracic Surgeons model to determine whether any of them were statistically associated with the observed-to-predicted ratio in either model; such an association would suggest an imbalance in risk adjustment. This analysis, presented in Appendix 2 (available at www.cmajopen.ca/content/6/3/E316/suppl/DC1), shows that New York Heart Association class IV, congestive heart failure, an ejection fraction of less than 20%, atrial fibrillation, acute coronary insufficiency, cardiac compromise (shock, myocardial infarction within 24 h, intra-aortic balloon pump, cardiac resuscitation or preprocedure circulatory support) and a creatinine level of 100 mg/dL or more were most strongly associated with an imbalance in the CIHI risk adjustment model. Other risk factors (e.g., age, sex, body surface area, hypertension, number of diseased coronary arteries and previous surgeries), although associated with mortality, were not associated with an imbalance in risk adjustment.

We then merged the 7 risk factors identified in Appendix 2 that were associated with an imbalance in risk adjustment into 3 risk groups: cardiac compromise and end-stage heart disease (ejection fraction < 20% or New York Heart Association class IV); refined cardiac diagnosis including congestive heart failure, atrial fibrillation and acute coronary syndrome or myocardial infarction between 24 hours and 1 week before cardiac surgery; and elevated creatinine concentration (> 100 mg/dL). Starting with the number of CIHI-predicted deaths, we investigated the additional number of deaths that would be predicted if the risk model was recalibrated using the mortality predicted by the Society of Thoracic Surgeons model for each of these 3 groups sequentially. These results are presented in Figure 4. Starting with the 17 deaths predicted by the CIHI risk model, an additional 6 deaths were predicted after adjustment for cardiac compromise and end-stage heart disease, 5 were predicted after refinement of the cardiac diagnoses and 3 were predicted after accounting for elevated creatinine concentration. This left 3 deaths predicted in the STS model that could be accounted for by other factors not listed above, and not included in the CIHI risk model. Overall, 84% of the difference in model-predicted mortality was accounted for by 7 risk factors: New York Heart Association class IV, congestive heart failure, an ejection fraction of less than 20%, atrial fibrillation, acute coronary insufficiency, cardiac compromise and a creatinine level of 100 mg/dL or more.

**Interpretation**

We compared 2 different methods of risk adjustment for cardiovascular surgery. The results emanating from these approaches varied considerably. Risk-adjusted mortality using the CIHI model yielded lower predicted mortality than that yielded by the Society of Thoracic Surgeons’ clinical model. This difference would result in a higher observed-to-predicted mortality ratio using the CIHI model than using the Society of Thoracic Surgeons clinical model with the same patients. The discrepancy between the two methods may be due to the lack of adjustment for important indicators of disease severity that were not available in the CIHI analysis.
In our study, we found that there was similar overall model discrimination between the CIHI and the Society of Thoracic Surgeons models, which is consistent with previous observations. However, systematic underestimation of predicted risks among patients with clinically recognized high-risk features occurred in the CIHI model, because it did not include factors such as New York Heart Association class IV, congestive heart failure, ejection fraction of less than 20%, atrial fibrillation, acute coronary insufficiency, cardiac compromise and creatinine level of 100 mg/dL or more, which accounted for much of the difference between the two models. The consequence of the omission of these variables is that hospitals whose patients have greater prevalences of these features would be more likely to show underestimation of CIHI-predicted risk. Consequently, observed-to-predicted ratios would be artifactually increased.

Our study is relevant for the developers, recipients and consumers of cardiac surgery report cards. Cardiac surgical reports should account for predictors of risk that are clinically important, and accounting for these factors is important even though the overall model C statistic is deemed acceptable. Indeed, previous studies suggested that relying on the C statistic alone may not identify the best model for risk adjustment, which supports our assertion that the covariates in the model are important. If these high-risk indicators are not accounted for, there would be little impact on hospitals that select only the lowest risk cases. However, hospitals with a greater proportion of cases that are higher risk, such as tertiary referral centres or hospitals that provide mechanical circulatory support, would appear to underperform. This could lead to undesired effects; for example, patients who have the most potential to benefit from surgical revascularization might not be offered surgery in favour of potentially healthier patients with lower risk. The value of administrative databases for risk adjustment may differ based on the disease and outcome in question. For example, administrative databases were entirely adequate for analyses of risk-adjusted heart failure readmission. However, several studies have reported on the substantial improvement in discrimination of cardiac surgical outcomes with clinical data sources relative to administrative databases. Our data provide a potential explanation for these observations about administrative database report cards for cardiac surgery, reaffirming the limitations of administrative-only data sources for interinstitutional comparisons of the quality of cardiac surgical care. Before administrative databases can be used for such purposes, careful examination of the performance of risk adjustment models and the accuracy of coding should be compared with clinical databases that were designed for the evaluation of care quality to ensure that bias or the failure to collect data that could affect risk adjustment does not impede the validity of the results.

**Limitations**

Our study was conducted at a single-centre, tertiary–quaternary referral centre for mechanical therapies and cardiac transplantation during a 3-year period. Thus, our results should be verified in other centres and time periods. The determination of risk

---

**Figure 4:** Observed and predicted mortality determined by adjusting risk models for various categories of risk factors not available in administrative data. Note: CIHI = Canadian Institute for Health Information, ESHD = end-stage heart disease, STS = Society of Thoracic Surgeons, UHN = University Hospital Network.

Previous studies have suggested that quality reports based on administrative and clinical data sources differ. Some studies have reported that the models employed contribute importantly to the discrepancies in quality reports. Others have suggested that the type of data, specifically administrative claims or clinical data sources, may be the primary reason for these differences. In our study, we compared an administrative model using claims data and a clinical model that required medical records, and found an opposite directionality of the observed-to-predicted ratios using these 2 approaches.

Other studies have suggested less marked differences between administrative and clinical surgical report cards. Comparing the EuroSCORE with an administrative database-based report card for mortality associated with coronary bypass surgery, Ugolini and Nobilio found that the latter could perform adequately provided that variables accounting for clinical complexity were included, such as emergency procedures, and provided that hospital data could be linked to multiple episodes of care up to 1 year before surgery. Another study reporting correlations between administrative and clinical report cards examined noncardiac surgical procedures, but found that the positive predictive value for identifying outlier hospitals was only 51%, whereas the negative predictive value was high. Thus, the adequacy of risk adjustment may depend on the type of surgery, but administrative risk adjustment models should be rigorously compared with clinical models.
factors and comorbidities in the administrative data set was limited to conditions documented within the index hospital admission. For this reason, our administrative data set might underestimate comorbidity burden. As mentioned previously, reoperations within 30 days were not considered because we followed the CIHI method. However, early reoperations have an increased procedural risk and could be considered a negative outcome of the initial procedure. In addition, although the Society of Thoracic Surgeons risk adjustment model has been validated in the United States, it has not been widely validated in the Canadian population. Finally, the study population represents a highly selected subset of patients undergoing cardiac surgeries; thus, the conclusions of this study might not apply to all types of cardiac surgery.

Conclusion
Cardiac surgery report cards showed different results when an administrative and a clinical model were used in their determination. The administrative model underestimated predicted risks in surgical patients at high risk, largely because many high-risk indicators were not included in the model or such variables were not available in claims data. Risk adjustment using a clinical model that adequately accounted for higher patient acuity yielded different results, and the predicted risks were well-matched to observed rates. We caution the use of administrative models for quality reporting, and such models should be considered exploratory if there are discrepancies in the results for high-risk patients when compared with a clinical database that has been collected and designed for the evaluation of care quality.

References

Affiliations:
Peter Munk Cardiac Centre, University Health Network (Manhliot, Rao, Rubin, Lee), Divisions of Cardiac Surgery (Manhliot, Rao), Vascular Surgery (Rabin) and Cardiology (Lee), Institute of Health Policy, Management and Evaluation (Lee), and Institute for Clinical Evaluative Sciences (Lee), University of Toronto (Manhliot, Rao, Rubin, Lee), Toronto, Ont.

Contributors:
Cedric Manhliot and Douglas Lee were responsible for the data analysis and interpretation. All of the authors contributed to the conception and design of this study, and to the drafting and editing of the manuscript. All of the authors approved the final version and agreed to act as guarantors of the results.

Funding:
This study was supported by a Foundation Grant from the Canadian Institutes of Health Research (grant no. FDN 148446). Barry Rubin is supported by the Peter Munk Medical Director Chair, University Health Network. Vivek Rao is supported by the Munk Chair in Advanced Cardiac Therapeutics, University Health Network. Douglas Lee is supported by a midcareer investigator award from the Heart and Stroke Foundation and the Ted Rogers Chair in Heart Function Outcomes, a joint Hospital–University Chair of the University Health Network and the University of Toronto. The funding agencies were not involved in the design, performance or interpretation of this study, nor were they involved in manuscript preparation.

Acknowledgements:
The authors thank Kira Leeb and Chantal Couris from the Canadian Institute for Health Information (CIHI) for providing the CIHI surgical risk adjustment model coefficients for isolated coronary artery bypass graft, isolated aortic valve replacement, and combined coronary artery bypass graft and aortic valve replacement surgeries, and any additional information relevant to this manuscript.

Disclaimer:
This study was supported by the Institute for Clinical Evaluative Sciences (ICES), which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results and conclusions reported in this article are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred.

Supplemental information:
For reviewer comments and the original submission of this manuscript, please see www.cmajopen.ca/content/6/3/E316/suppl/DC1.