

Derivation and validation of predictive indices for 30-day mortality after coronary and valvular surgery in Ontario, Canada

Louise Y. Sun MD SM, Anna Chu MHSc, Derrick Y. Tam MD PhD, Xuesong Wang MSc, Jiming Fang PhD, Peter C. Austin PhD, Christopher M. Feindel MD, Garth H. Oakes PhD, Vicki Alexopoulos MGA, Natasa Tusevjak MSc, Maral Ouzounian MD PhD, Douglas S. Lee MD PhD; for the CorHealth Ontario Cardiac Surgery Risk Adjustment Task Group

■ Cite as: *CMAJ* 2021 November 22;193:E1757-65. doi: 10.1503/cmaj.202901

Abstract

Background: Coronary artery bypass grafting (CABG) and surgical aortic valve replacement (AVR) are the 2 most common cardiac surgery procedures in North America. We derived and externally validated clinical models to estimate the likelihood of death within 30 days of CABG, AVR or combined CABG + AVR.

Methods: We obtained data from the CorHealth Ontario Cardiac Registry and several linked population health administrative databases from Ontario, Canada. We derived multiple logistic regression models from all adult patients who underwent CABG, AVR or combined CABG + AVR from April 2017 to March

2019, and validated them in 2 temporally distinct cohorts (April 2015 to March 2017 and April 2019 to March 2020).

Results: The derivation cohorts included 13 435 patients who underwent CABG (30-d mortality 1.73%), 1970 patients who underwent AVR (30-d mortality 1.68%) and 1510 patients who underwent combined CABG + AVR (30-d mortality 3.05%). The final models for predicting 30-day mortality included 15 variables for patients undergoing CABG, 5 variables for patients undergoing AVR and 5 variables for patients undergoing combined CABG + AVR. Model discrimination was excellent for the CABG (c-statistic

0.888, optimism-corrected 0.866) AVR (c-statistic 0.850, optimism-corrected 0.762) and CABG + AVR (c-statistic 0.844, optimism-corrected 0.776) models, with similar results in the validation cohorts.

Interpretation: Our models, leveraging readily available, multidimensional data sources, computed accurate risk-adjusted 30-day mortality rates for CABG, AVR and combined CABG + AVR, with discrimination comparable to more complex American and European models. The ability to accurately predict perioperative mortality rates for these procedures will be valuable for quality improvement initiatives across institutions.

Coronary artery bypass grafting (CABG) and surgical aortic valve replacement (AVR) are 2 of the most common cardiac surgical procedures in North America.¹ Accurate risk models of perioperative mortality for CABG and AVR are not only useful for operative decision-making,² but also valuable for quality improvement initiatives across surgeons and institutions.

In North America, the most widely used 30-day mortality risk score is the Society of Thoracic Surgeons (STS)–Predicted Risk of Mortality tool, derived from more than 1000 hospitals in the United States and encompassing more than 50 variables.³ An ideal risk model should be built and validated on the patient population in which it will be applied. Although the STS–

Predicted Risk of Mortality tool was derived from a large surgical population, regional differences in patient sociodemographics and health care delivery systems may preclude this model from performing optimally in the health system where cardiac surgery is publicly funded. Furthermore, collecting more than 50 variables is resource intensive and is not feasible for all institutions. Similar limitations apply to the EuroSCORE II, which was derived from a population-based cohort in Europe.⁴ Given these limitations, we developed a more parsimonious model using readily available, linked clinical and administrative data sets in Ontario, Canada, to efficiently and accurately calculate risk-adjusted 30-day mortality rates for the purpose of province-wide quality improvement after CABG, AVR and combined CABG + AVR.

Methods

Study design and population

We conducted a retrospective analysis of patients aged 18 years and older and eligible for Ontario's public health insurance plan, who were identified in the CorHealth Ontario Cardiac Registry as having had CABG, AVR or combined CABG + AVR surgery between Apr. 1, 2015, and Mar. 31, 2020.^{5,6} CorHealth is a provincial organization with a mandate to collect health data from all patients undergoing cardiac procedures and to provide strategic leadership to improve cardiac, stroke and vascular care in Ontario. This mandatory registry contains demographic, clinical and perioperative information on all patients who undergo major cardiovascular procedures and related cardiac interventions in Ontario.⁷⁻⁹

Our derivation cohort consisted of patients who underwent cardiac surgery between Apr. 1, 2017, and Mar. 31, 2019. Two temporally distinct validation cohorts included patients who underwent procedures between Apr. 1, 2015, and Mar. 31, 2017, and between Apr. 1, 2019, and Mar. 31, 2020. For each patient, we considered only the first surgical procedure in a given fiscal year. We confirmed surgical procedures performed using Canadian Classification of Health Interventions procedure codes, through linkage to the Canadian Institute for Health Information Discharge Abstract Database, which contains demographic, diagnostic and procedural information from the discharge abstracts of all acute care hospital admissions in Ontario; and the Ontario Health Insurance Plan Physician Claims Database, which contains information from nearly all physician encounters, diagnostic tests and outpatient laboratory services performed in Ontario. We excluded surgical procedures not recorded in the Discharge Abstract Database or Ontario Health Insurance Plan database, or for which CABG, AVR or combined CABG + AVR was performed concurrently with other cardiac procedures.

Outcome

Our primary outcome was 30-day all-cause mortality, captured from Ontario's Registered Persons Database. This is a registry maintained by the Ontario Ministry of Health, containing demographic information about every individual who has ever been registered for the Ontario Health Insurance Plan, including their eligibility and dates of death. Registration is required to access publicly funded health care services in the province.

Candidate variable selection

We identified potential variables to be included in our mortality model from a review of predictors in previously published models or those deemed clinically important by our co-author group.^{4,10-13} In addition to key demographic variables (age, sex and ethnicity), we developed a list of 63 variables and forwarded it to members of the CorHealth Ontario Cardiac Surgery Risk Adjustment Task Group for further selection through a modified Delphi process.^{14,15} The task group comprises clinical-, administrative- and system-level leadership, with representatives from cardiac surgery centres across the province. It serves to advise CorHealth Ontario on risk-adjustment models for key quality

indicators and clinical variables to be used in the monitoring and reporting of quality of care and outcomes of cardiac surgery. We first asked respondents to rate each of the variables as important or not in the risk stratification process (Appendix 1, Supplemental Table 1, available at www.cmaj.ca/lookup/doi/10.1503/cmaj.202901/tab-related-content). If an organization had more than 1 representative in the task group, we asked that 1 electronic survey be returned on behalf of all its members. Respondents were also able to suggest variables not already on the list. We then reviewed a summary of results from responses received from 7 of 15 organizations (47% response rate) in a subsequent task group teleconference, where a final list of 57 candidate variables was created through consensus-based discussion. Further refinement to combine similar variables — for example, previous stroke with previous transient ischemic attack — resulted in 49 candidate variables for model development (48 for the CABG model, owing to exclusion of endocarditis) (Appendix 1, Supplemental Table 2).

Data sources

Data sources for candidate variables are provided in Appendix 1, Supplemental Table 2. We used the CorHealth Registry, the Discharge Abstract Database, the National Ambulatory Care Reporting System and the Ontario Health Insurance Plan database to obtain baseline demographics and comorbidities in addition to identifying our study population.^{16,17} Other data sources included the Ontario Laboratories Information System for laboratory information; the Canadian Institute for Health Information Same-day Surgery database for day procedure history; the Ontario Cancer Registry for cancer and radiation treatment history; and the Ontario Visible Minority Database for ethnicity.¹⁸ These data sets were linked using unique, encoded identifiers and analyzed at ICES (formerly Institute for Clinical Evaluative Sciences). Administrative codes and definitions used for variables, validation study results (where available) and variable formats are provided in Appendix 1, Supplemental Table 2.

Statistical analysis

In each model, we first performed unadjusted logistic regression to select potential predictors of 30-day mortality for each procedure of interest separately. We then entered candidate variables into a multivariable logistic regression model with backward selection and a significance threshold of < 0.05 .¹⁹ Where missing, we imputed values using the procedure and sex-specific cohort mean (Appendix 1, Supplemental Table 2). We reviewed resulting models for face and content validity and selected final covariates based on statistical and clinical importance, as determined by the task group. For continuous variables, we examined their association with 30-day mortality using cubic spline analyses with 5 knots at the 5th, 27.5th, 50th, 72.5th and 95th percentiles. We entered linear variables (age, body surface area, hematocrit, leukocytes) into the models as continuous values, but treated nonlinear variables (Hospital Frailty Risk Score, body mass index, platelets) categorically based on their distribution in tertiles and clinically meaningful ranges.^{20,21} We report odds ratios, 95% confidence intervals (CIs) and *p* values for final covariates in each model.

In both derivation and validation cohorts, we evaluated model discrimination using the c-statistic. For internal validation in the derivation sample, we computed optimism-corrected c-statistics using 250 bootstrap samples. We assessed calibration using the Hosmer–Lemeshow χ^2 statistic, Brier score, calibration slope and a calibration curve, comparing observed versus expected mortality rates across deciles of expected risk. We also assessed the performance of the STS model in our derivation and validation cohorts. Roughly half of the hospitals were collecting STS data at the time, and these data were available to us. For all other hospitals, we mapped as many of the STS variables as possible to existing data sources at ICES. This was then used to estimate risk based on STS, ultimately including the entire cohort in the STS calculation. We conducted all analyses using SAS version 9.4 (SAS Institute, Cary NC).

Ethics approval

The use of these data was authorized under section 45 of Ontario's *Personal Health Information Protection Act*, which does not require review by a Research Ethics Board.

Results

The derivation cohorts included 13 435 patients who underwent CABG, 1970 patients who underwent AVR, and 1510 patients who underwent combined CABG + AVR (Figure 1). The sample size, number of deaths and proportion of patients who died in the derivation and validation cohorts are shown in Table 1. The baseline characteristics were similar between the derivation and validation cohorts across all groups (Appendix 1, Supplemental Tables 3–5).

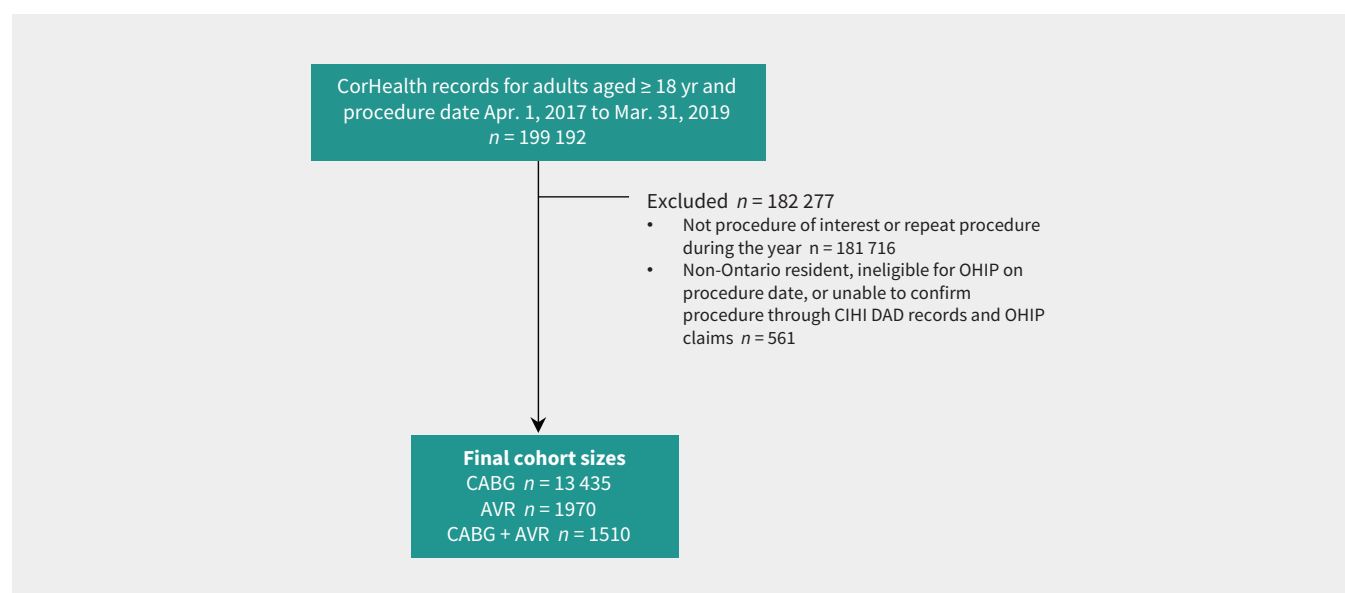


Figure 1: Flow diagram for the derivation cohort. Note: AVR = aortic valve replacement, CABG = coronary artery bypass graft, CIHI DAD = Canadian Institute for Health Information Discharge Abstract Database, OHIP = Ontario Health Insurance Plan. Note: Owing to small cell sizes, some exclusion criteria have been collapsed.

Table 1: Population size, number of deaths and proportion who died in the derivation and validation cohorts, by procedure category

Cohort	CABG		AVR		CABG + AVR	
	Population size	No. (%) of patients who died within 30 days	Population size	No. (%) of patients who died within 30 days	Population size	No. (%) of patients who died within 30 days
Derivation cohort						
FY2017	6734	105 (1.56)	966	22 (2.28)	759	23 (3.03)
FY2018	6701	128 (1.91)	1004	11 (1.10)	751	23 (3.06)
Overall	13 435	233 (1.73)	1970	33 (1.68)	1510	46 (3.05)
Validation cohorts						
FY2015–FY2016	13 447	208 (1.55)	1946	26 (1.34)	1601	60 (3.75)
FY2019	6430	89 (1.38)	777	10 (1.29)	555	20 (3.60)

Note: AVR = aortic valve replacement, CABG = coronary artery bypass graft, FY = fiscal year.

Predictors of 30-day mortality after isolated CABG

History of percutaneous coronary intervention and left ventricular ejection fraction were forced into the model on the basis of clinical significance. Of the candidate covariates evaluated, older age, female sex, Hospital Frailty Risk Score,²⁰ renal insufficiency, thrombocytopenia, atrial arrhythmia, chronic lung disease, peripheral arterial disease, cerebrovascular disease, previous CABG, percutaneous coronary intervention within 1 day before surgical revascularization, thoracic aortic disease, preoperative cardiogenic shock and moribund status²² were predictors of 30-day CABG mortality (Table 2).

The c-statistic was 0.888 in the derivation data set (optimism-corrected 0.866), indicating excellent discrimination,²³ and the Hosmer–Lemeshow χ^2 statistic *p* value was 0.2, indicating that there was no lack of model fit. These metrics of performance remained robust in both validation cohorts (Table 3). Supplemental Figure 1a in Appendix 1 shows the calibration plot of observed versus expected rates of 30-day CABG mortality according to each decile of risk, and Table 3 shows the predicted probability for the derivation and validation samples. The observed and predicted numbers of deaths were similar across all except the highest risk decile, in which the model tended to overestimate mortality.

In comparison, when fitted to the derivation cohort, the c-statistic of the STS model was 0.816, and the Hosmer–Lemeshow *p* value was 0.6. The c-statistic of the STS model was 0.841 and 0.863 in each of the validation cohorts.

Predictors of 30-day mortality after isolated AVR

Sex was forced into the model on the basis of clinical significance. The multivariable predictors of 30-day mortality were frailty, leukocytosis, liver disease and preoperative cardiogenic shock (Table 4).

The c-statistic was 0.850 in the derivation data set (optimism-corrected 0.762) and the Hosmer–Lemeshow χ^2 statistic *p* value was 0.08. These metrics of performance remained robust in both validation cohorts (Table 3). Supplemental Figure 1b (Appendix 1) shows the calibration plot of observed versus expected rates of 30-day AVR mortality according to each decile of risk. The observed and predicted numbers of deaths were similar across all risk deciles.

In comparison, when fitted to the derivation cohort, the c-statistic of the STS model was 0.861, and the Hosmer–Lemeshow *p* value was 0.3. The c-statistic of the STS model was 0.846 in the first validation cohort. There was an insufficient number of events to validate the STS model in the second validation cohort.

Predictors of 30-day mortality after combined CABG + AVR

Sex and a history of previous CABG were forced into this model on the basis of clinical significance. Other multivariable predictors of 30-day mortality were frailty, anemia, a history of previous CABG and preoperative cardiogenic shock (Table 5).

The c-statistic was 0.84 in the derivation data set (optimism-corrected 0.764) and the Hosmer–Lemeshow χ^2 statistic *p* value was 0.7. These metrics of performance remained robust in both

validation cohorts (Table 3). Supplemental Figure 1c in Appendix 1 shows the calibration plot of observed versus expected rates of 30-day combined CABG + AVR mortality according to each decile of risk. The observed and predicted numbers of deaths were similar across all except the middle risk decile, in which the model tended to overestimate mortality.

In comparison, when fitted to the derivation cohort, the c-statistic of the STS model was 0.828, and the Hosmer–Lemeshow *p* value was 0.3. The c-statistic of the STS model was 0.881 in the first validation cohort. There was an insufficient number of events to validate the STS model in the second validation cohort.

Interpretation

We found that multidimensional data sources consisting of readily available clinical registry and administrative health databases can be used to develop 30-day mortality risk models for CABG, AVR and combined CABG + AVR, with excellent performance. We found that the Ontario CABG model was the best-performing model with a c-statistic of 0.888, while those for AVR and combined CABG + AVR also predicted well, with c-statistics of 0.850 and 0.844, respectively, and performed consistently in the validation cohorts. By comparison, the STS CABG model did not perform as well in Ontario, with a c-statistic of 0.816, while its performance for AVR (c-statistic 0.861) and CABG + AVR (c-statistic 0.828) was comparable to the Ontario models in the derivation data set.¹²

Several aspects of our models are novel, compared with existing perioperative mortality models. First, the incorporation of frailty in our models represents a major advance in the field. Indeed, cardiac surgery literature^{7,24,25} cites the exclusion of frailty as a major limitation of commonly used cardiac surgery risk scores. Second, the Ontario models achieved parsimony without sacrificing performance, which allows for efficient assessment of the quality of surgical care. Our CABG model included only 15 predictors as compared with more than 50 in the STS model, while our CABG + AVR and AVR models each included only 5 predictor variables. The large number of variables needed to risk-adjust using the STS model is a limitation, and only half of Ontario surgical hospitals participate in the STS data collection. Third, we were able to derive these models using routinely collected data that are readily available across all cardiac care institutions, without loss to follow-up. In contrast, collecting the data elements necessary for the STS model would require additional infrastructure, resources and personnel time to be put in place across all cardiac centres. Lastly, our models were developed by an interdisciplinary team with complementary expertise in cardiac surgery, cardiac anesthesiology, cardiology and clinical administration, for the purpose of quality assessment across centres. This differs from the other risk scores, which were derived primarily for preoperative risk assessment and operative decision-making.

Our modelling methodology has additional unique strength. In contrast with other commonly used universal mortality prediction models, such as the EuroSCORE II and the American College of Surgeons National Surgical Quality Improvement

Table 2: Multivariable model after backward selection for predicting 30-day mortality in patients who underwent coronary artery bypass graft surgery

Covariate	Parameter	Standard error	OR (95% CI)
Intercept	-8.8215	1.0805	0 (0)
Age on procedure date	0.04832	0.0087	1.05 (1.03–1.07)
Sex, female	0.4509	0.1622	1.57 (1.14–2.16)
Hospital Frailty Risk Score			
0 to < 1.0	Ref.		Ref.
1.0 to 3.0	0.6181	0.2840	1.86 (1.06–3.24)
> 3.0	1.3357	0.2571	3.80 (2.30–6.29)
Left ventricular ejection fraction			
≥ 35%	Ref.		Ref.
< 35%	0.3189	0.1979	1.38 (0.93–2.03)
Missing	0.5230	0.2762	1.69 (0.98–2.90)
Serum creatinine			
Missing or < 120	Ref.		Ref.
120–179	0.1629	0.2083	1.18 (0.78–1.77)
180+	0.7859	0.2228	2.19 (1.42–3.40)
Platelet count			
< 130	Ref.		Ref.
130–400	-0.7033	0.2798	0.49 (0.29–0.86)
> 400	-1.6244	0.6846	0.20 (0.05–0.75)
Atrial fibrillation or flutter			
None	Ref.		Ref.
Recent onset (≤ 30 d)	0.3801	0.2240	1.46 (0.94–2.27)
Remote (> 30 d)	0.9369	0.2792	2.55 (1.48–4.41)
Chronic lung disease	0.6986	0.2020	2.01 (1.35–2.99)
Peripheral artery or vascular disease	0.5658	0.2274	1.76 (1.13–2.75)
Previous stroke or transient ischemic attack	0.5786	0.2323	1.78 (1.13–2.81)
History and timing of PCI			
No previous PCI	Ref.		Ref.
PCI within 1 day	3.5112	1.2196	33.49 (3.07–365.63)
PCI > 1 day prior	-0.4232	0.2177	0.65 (0.43–1.00)
Previous CABG	1.7743	0.3965	5.90 (2.71–12.83)
Number of diseased coronary vessels			
Unknown	Ref.		Ref.
1	0.7170	0.8820	2.05 (0.36–11.54)
2	-0.0829	0.8366	0.92 (0.18–4.74)
3	0.3807	0.8230	1.46 (0.29–7.34)
Thoracic aorta disease	1.6117	0.6746	5.01 (1.34–18.80)
Moribund*	0.7936	0.2182	2.21 (1.44–3.39)
Shock; preoperative ECMO, IABP or CBA; or resuscitation within 1 day or on admission	2.7808	0.1832	16.13 (11.27–23.10)
<p>Note: CABG = coronary artery bypass graft, CBA = catheter-based assist device, CI = confidence interval, ECMO = extracorporeal membrane oxygenation, IABP = intra-aortic balloon pump, OR = odds ratio, PCI = percutaneous coronary intervention, Ref. = reference category.</p> <p>*Moribund is defined by American Society of Anesthesiologists (ASA) Physical Classification status 5, as a patient who is not expected to survive for > 24 hours with or without surgery.²²</p>			

Table 3: Predicted probability of 30-day mortality in the derivation and validation cohorts, by procedure category

Performance metric	CABG	AVR	CABG + AVR
Derivation cohort: FY2017 and FY2018			
Predicted probability of 30-day mortality, %	1.80	1.95	4.15
C-statistic	0.8882	0.8500	0.8400
Optimism-corrected c-statistic	0.8659	0.7618	0.7644
Hosmer–Lemeshow χ^2 statistic <i>p</i> value	0.2	0.08	0.7
Brier score*	0.0150	0.0136	0.0269
Calibration slope	0.8988	0.6932	0.6634
Validation cohort: FY2015–FY2016			
Predicted probability of 30-day mortality, %	1.35	0.95	3.80
C-statistic	0.8801	0.7882	0.7591
Optimism-corrected c-statistic	Not applicable	Not applicable	Not applicable
Hosmer–Lemeshow χ^2 statistic <i>p</i> value	0.2	0.06	0.001
Brier score*	0.0142	0.0145	0.0320
Calibration slope	0.9672	0.8426	0.9218
Validation cohort: FY2019			
Predicted probability of 30-day mortality, %	2.65	1.10	1.55
C-statistic	0.8587	0.9008	0.8175
Optimism-corrected c-statistic	Not applicable	Not applicable	Not applicable
Hosmer–Lemeshow χ^2 statistic <i>p</i> value	0.7	0.8	0.01
Brier score*	0.0136	0.0128	0.0309
Calibration slope	0.9491	0.9596	1.0061

Note: AVR = aortic valve replacement, CABG = coronary artery bypass graft, FY = fiscal year.
*The Brier score is the average squared prediction error, where lower values indicate better model performance.

Table 4: Multivariable model after backward selection for predicting 30-day mortality in patients who underwent aortic valve replacement

Covariate	Parameter	Standard error	OR (95% CI)
Intercept	-6.4791	0.6230	0 (0.00–0.01)
Sex, female	0.3500	0.4091	1.42 (0.64–3.16)
Hospital Frailty Risk Score			
0 to < 1.0	Ref.		
1.0 to 3.0	0.0775	0.7554	1.08 (0.25–4.75)
> 3.0	1.4117	0.5332	4.10 (1.44–11.67)
Leukocytes, per 10 ³	0.1191	0.0459	1.13 (1.03–1.23)
Liver disease	1.8277	0.7377	6.22 (1.46–26.41)
Shock; preoperative ECMO, IABP or CBA; resuscitation within 1 day or on admission; or ventricular tachycardia or fibrillation within 30 days	3.3448	0.4437	28.36 (11.88–67.66)

Note: CBA = catheter-based assist device, CI = confidence interval, ECMO = extracorporeal membrane oxygenation, IABP = intra-aortic balloon pump, OR = odds ratio, Ref. = reference category.

Table 5: Multivariable model after backward selection for predicting 30-day mortality in patients who underwent combined coronary artery bypass graft surgery and aortic valve replacement

Covariate	Parameter	Standard error	OR (95% CI)
Intercept	-1.5823	1.3567	0.21 (0.01–2.94)
Sex	0.1718	0.3568	1.19 (0.59–2.39)
Hospital Frailty Risk Score			
0 to < 1.0	Ref.		
1.0 to 3.0	0.8814	0.7403	2.41 (0.57–10.30)
> 3.0	2.0618	0.6146	7.86 (2.36–26.22)
Hematocrit, per 10%	-0.9301	0.3026	0.39 (0.22–0.71)
Previous CABG	0.9574	0.6593	2.60 (0.72–9.48)
Shock; preoperative ECMO, IABP, or CBA; resuscitation within 1 day or on admission; or ventricular tachycardia or fibrillation within 30 days	2.2722	0.4094	9.70 (4.35–21.64)

Note: CABG = coronary artery bypass graft, CBA = catheter-based assist device, CI = confidence interval, ECMO = extracorporeal membrane oxygenation, IABP = intra-aortic balloon pump, OR = odds ratio, Ref. = reference category.

Program (ACS NSQIP) score, which were based on pooled sets of diverse surgical procedures (whereby each procedure type is treated as a model covariate), our prediction models were procedure specific.²⁶ Universal models may be particularly useful when event rates are low. The ACS NSQIP model includes more than 100 procedures, encompassing a wide variety of surgical specialties. By combining procedures into sets, more covariates could be added to potentially improve model performance.²⁷ However, this approach also makes the limiting assumption that the effect of each predictor is the same across procedures, which may not reflect reality. In addition, the EuroSCORE II allows for calculation of the risk for almost any combination of cardiac surgical procedures, as each procedure is treated as a model covariate.²⁸ Although this method of modelling may produce a c-statistic that is acceptable overall, its predictive performance for individual procedures is poor, especially for procedures that differ technically.²⁹ The fact that guidelines recommend using the STS–Predicted Risk of Mortality score (Class I) over the EuroSCORE II (Class IIb) for the prediction of 30-day mortality after CABG reflects the potential importance of procedure-specific models.³⁰

Risk prediction models serve several purposes. In addition to informing treatment decision-making, they can be used for risk adjustment to allow evaluation of reporting on quality-of-care outcomes.^{31,32} Risk-adjusted rates for surgical procedures can be expressed as a ratio of the observed versus the expected number of deaths after patient characteristics have been adjusted for. These risk-adjusted measures enhance comparability of outcomes within and among institutions, and can be used to assess quality of surgical care. Reports of risk-adjusted mortality rates are now part of the standard repertoire to help facilitate high-quality surgical care and quality improvement initiatives.

Our research was motivated by a province-wide initiative to improve cardiac surgery quality that includes the provision of outcomes reports on key quality indicators for all cardiac centres in Ontario. Although these reports are not released to the public, each cardiac surgery program sees the outcomes of all other surgical centres in an identifiable manner. It should be noted that the practice of public reporting is controversial, as the observed outcomes are influenced by practice variations in patient selection, as well as the fact that even a small excess of adverse events could have a large impact on rates of rare outcomes.³³ Interestingly, a population-based cluster randomized trial by Tu and colleagues showed that the public release of hospital-specific quality indicators did not improve outcomes after acute myocardial infarction and congestive heart failure.³⁴ Conversely, in the setting of the Michigan Society of Thoracic and Cardiovascular Surgeons Quality Collaborative, where 33 hospitals participated in quarterly presentation of unblinded data for the purpose of quality improvement through enhanced feedback, a substantial reduction in the rate of postoperative pneumonia was shown after intervention.³⁵ Further studies are needed to determine whether an enhanced feedback system could reduce operative mortality after cardiac surgery.

Limitations

Our study has several limitations. First, important socio-demographic risk factors such as low socioeconomic status are difficult to capture using administrative data sets. Although we included neighbourhood income as a surrogate measure for socioeconomic status, we did not adjust for other determinants of socioeconomic status in this analysis. Second, certain physiologic details, such as specific lesion locations and exact percentage stenoses of coronary lesions, were unavailable in the

data sets used. There is evidence that the inclusion of coronary anatomic complexity may improve mortality risk prediction.³⁶ Third, we relied on administrative data and physician billing codes to derive covariates of interest, but the data sources used in this study and associated codes have been previously validated or published.^{18,37,38} Fourth, our models apply to the 3 most commonly performed cardiac surgery procedures, and the incremental risk of concomitant procedures — such as aortic root enlargement ascending aorta replacement — was not captured. Fifth, the low event rates for AVR and combined CABG + AVR precluded us from entering a large number of covariates during the modelling process. Despite this, our models performed well in 2 separate validation cohorts. Sixth, our study is limited by a lack of validation outside Ontario. Future opportunities to evaluate the ability of these models to benchmark national cardiac surgery performance are warranted, using data sources such as the Canadian Institute for Health Information. Lastly, continuous model updates are also warranted, to accommodate evolving patient demographics and indications for CABG and AVR.³⁹

Conclusion

Accurate computation of 30-day mortality risk for CABG, AVR and combined CABG + AVR can be achieved parsimoniously using routinely collected multidimensional administrative and clinical registry data sets, with comparable performance to more complex models derived from large, clinical data-derived US and European registries. The parsimonious Ontario cardiac surgery risk scores are a product of province-wide interdisciplinary collaboration among cardiac surgeons, cardiac anesthesiologists, cardiologists and clinical administrators. Hybridization (using a hybrid of clinical registry and administrative data sources) of routinely collected multidimensional data sources represents an efficient approach to data collection that has utility in system-wide quality of care evaluation and reporting.

References

1. ElBardissi AW, Aranki SF, Sheng S, et al. Trends in isolated coronary artery bypass grafting: an analysis of the Society of Thoracic Surgeons adult cardiac surgery database. *J Thorac Cardiovasc Surg* 2012;143:273-81.
2. Tam DY, Bakaeen F, Feldman DN, et al. Modality selection for the revascularization of left main disease. *Can J Cardiol* 2019;35:983-92.
3. Shahian DM, Jacobs JP, Badhwar V, et al. The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 1 – Background, design considerations, and model development. *Ann Thorac Surg* 2018;105:1411-8.
4. Nashef SA, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg* 2012;41:734-44, discussion 744-5.
5. Johnston A, Mesana TG, Lee DS, et al. Sex differences in long-term survival after major cardiac surgery: a population-based cohort study. *J Am Heart Assoc* 2019;8:e013260.
6. Tam DY, Dharma C, Rocha R, et al. Long-term survival after surgical or percutaneous revascularization in patients with diabetes and, ultrivessel coronary disease. *J Am Coll Cardiol* 2020;76:1153-64.
7. Sun LY, Spence SD, Benton S, et al. Age, not sex, modifies the effect of frailty on long-term outcomes after cardiac surgery. *Ann Surg* 2020 June 11 [Epub ahead of print]. doi: 10.1097/SLA.0000000000004060.
8. Sun LY, Bader Eddeen A, Ruel M, et al. Derivation and validation of a clinical model to predict intensive care unit length of stay after cardiac surgery. *J Am Heart Assoc* 2020;9:e017847.
9. Sun LY, Gaudino M, Chen RJ, et al. Long-term outcomes in patients with severely reduced left ventricular ejection fraction undergoing percutaneous coronary intervention vs coronary artery bypass grafting. *JAMA Cardiol* 2020;5:631-41.
10. Hannan EL, Szypulski Farrell L, Wechsler A, et al. The New York risk score for in-hospital and 30-day mortality for coronary artery bypass graft surgery. *Ann Thorac Surg* 2013;95:46-52.
11. Hannan EL, Racz M, Culliford AT, et al. Risk score for predicting in-hospital/30-day mortality for patients undergoing valve and valve/coronary artery bypass graft surgery. *Ann Thorac Surg* 2013;95:1282-90.
12. O'Brien SM, Feng L, He X, et al. The Society of Thoracic Surgeons 2018 Adult Cardiac Surgery Risk Models: Part 2 – Statistical methods and results. *Ann Thorac Surg* 2018;105:1419-28.
13. Ristic V, de Rook S, Mesana TG, et al. The impact of preoperative risk on the association between hypotension and mortality after cardiac surgery: an observational study. *J Clin Med* 2020;9:2057.
14. Sun LY, Rodger J, Duffett L, et al. Derivation of patient-defined adverse cardiovascular and noncardiovascular events through a modified Delphi process. *JAMA Netw Open* 2021;4:e2032095.
15. Lee DS, Tran C, Flintoft V, et al.; Canadian Cardiovascular Outcomes Research Team/Canadian Cardiovascular Society Heart Failure Quality Indicator Panel. CCORT/CCS quality indicators for congestive heart failure care. *Can J Cardiol* 2003;19:357-64.
16. Hayatsu Y, Ruel M, Bader Eddeen A, et al. Single versus multiple arterial revascularization in patients with reduced renal function: long-term outcome comparisons in 23,406 CABG patients from Ontario, Canada. *Ann Surg* 2020 June 24 [Epub ahead of print]. doi: 10.1097/SLA.0000000000003908.
17. Sun LY, Tu J, Bader Eddeen A, et al. Prevalence and long-term survival after coronary artery bypass grafting in men and women with heart failure and preserved vs reduced ejection fraction. *J Am Heart Assoc* 2018;7:e008902.
18. Shah BR, Chiu M, Amin S, et al. Surname lists to identify South Asian and Chinese ethnicity from secondary data in Ontario, Canada: a validation study. *BMC Med Res Methodol* 2010;10:42. doi: 10.1186/1471-2288-10-42
19. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
20. Gilbert T, Neuburger J, Kraindler J, et al. Development and validation of a Hospital Frailty Risk Score focusing on older people in acute care settings using electronic hospital records: an observational study. *Lancet* 2018;391:1775-82.
21. Clinical laboratory tests: adult normal values. Ottawa: Medical Council of Canada; 2020. Available: <https://www.mcc.ca/objectives/normal-values/> (accessed 2021 Oct. 18).
22. Hurwitz EE, Simon M, Vinta SR, et al. Adding examples to the ASA-physical status classification improves correct assignment to patients. *Anesthesiology* 2017;126:614-22.
23. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010;5:1315-6.
24. Otto CM, Nishimura RA, Bonow RO, et al. 2020 ACC/AHA guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association joint committee on clinical practice guidelines. *Circulation* 2021;143:e72-227.
25. Tran DTT, Tu JV, Dupuis J-Y, et al. Association of frailty and long-term survival in patients undergoing coronary artery bypass grafting. *J Am Heart Assoc* 2018;7:e009882.
26. Bilimoria KY, Liu Y, Paruch JL, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg* 2013;217:833-42.e1-3.
27. Tam DY, Frenes SE. Commentary: One size doesn't always fit all. *J Thorac Cardiovasc Surg* 2020;160:180-1.
28. Ad N, Holmes SD, Patel J, et al. Comparison of EuroSCORE II, original EuroSCORE, and the Society of Thoracic Surgeons Risk Score in cardiac surgery patients. *Ann Thorac Surg* 2016;102:573-9.
29. Mori M, Shahian DM, Huang C, et al. Surgeons: buyer beware—does “universal” risk prediction model apply to patients universally? *J Thorac Cardiovasc Surg* 2020;160:176-9.e2.
30. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS guidelines on myocardial revascularization. *EuroIntervention* 2019;14:1435-534.

31. Tu JV, Sykora K, Naylor CD. Assessing the outcomes of coronary artery bypass graft surgery: how many risk factors are enough? Steering Committee of the Cardiac Care Network of Ontario. *J Am Coll Cardiol* 1997;30:1317-23.
32. Guru V, Anderson GM, Fremes SE, et al.; Canadian CABG Surgery Quality Indicator Consensus Panel. The identification and development of Canadian coronary artery bypass graft surgery quality indicators. *J Thorac Cardiovasc Surg* 2005;130:1257.
33. Jha AK. Public reporting of surgical outcomes: Surgeons, hospitals, or both? *JAMA* 2017;318:1429-30.
34. Tu JV, Donovan LR, Lee DS, et al. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study – a randomized trial. *JAMA* 2009;302:2330-7.
35. Likosky DS, Harrington SD, Cabrera L, et al. Collaborative quality improvement reduces postoperative pneumonia after isolated coronary artery bypass grafting surgery. *Circ Cardiovasc Qual Outcomes* 2018;11:e004756.
36. Farooq V, van Klaveren D, Steyerberg EW, et al. Anatomical and clinical characteristics to guide decision making between coronary artery bypass surgery and percutaneous coronary intervention for individual patients: development and validation of SYNTAX score II. *Lancet* 2013;381:639-50.
37. Tam DY, Fang J, Tran A, et al. A clinical risk scoring tool to predict readmission after cardiac surgery: an Ontario administrative and clinical population database study. *Can J Cardiol* 2018;34:1655-64.
38. Braga JR, Austin PC, Ross HJ, et al. Importance of nonobstructive coronary artery disease in the prognosis of patients with heart failure. *JACC Heart Fail* 2019;7:493-501.
39. Englum BR, Ganapathi AM, Schechter MA, et al. Changes in risk profile and outcomes of patients undergoing surgical aortic valve replacement from the pre- to post-transcatheter aortic valve replacement eras. *Ann Thorac Surg* 2016;101:110-7.

Competing interests: Louise Sun received support from the Canadian Institutes of Health Research (CIHR) for article processing charges. Dr. Sun was named National New Investigator by the Heart and Stroke Foundation of Canada, and is supported by a Clinical Research Chair in Big Data and Cardiovascular Outcomes at the University of Ottawa. Douglas Lee is the Ted Rogers Chair in Heart Function Outcomes, University Health Network, University of Toronto. Dr. Lee also received a research grant from CorHealth Ontario and a foundation grant from the Canadian Institutes of Health Research (CIHR). Peter Austin is supported by a Mid-Career Investigator Award from the Heart and Stroke Foundation. Dr. Austin also reports receiving a CIHR Project Grant, paid to Sunnybrook Research Institute. No other competing interests were declared.

This article has been peer reviewed.

Affiliations: Division of Cardiac Anesthesiology (Sun), University of Ottawa Heart Institute, Ottawa, Ont.; ICES Central (Sun, Chu, Tam, Wang, Fang, Austin, Tusevlijak, Lee), Toronto, Ont.; School of Epidemiology and Public Health (Sun), University of Ottawa, Ottawa, Ont.; Sunnybrook Health Sciences Centre (Tam); University Health Network and Peter Munk Cardiac Centre (Feindel, Ouzounian, Lee); Departments of Cardiac Surgery (Tam, Feindel, Ouzounian), Physical Therapy (Chu), Surgery (Feindel), Cardiology (Lee), University of Toronto, Toronto, Ont.; Institute for Health Policy, Management and Evaluation (Austin), University of Toronto, Toronto, Ont.; CorHealth Ontario (Oakes, Alexopoulos), Toronto, Ont.

Contributors: Louise Sun, Anna Chu, Derrick Tam, Jiming Fang, Peter Austin, Garth Oakes and Douglas Lee contributed to the conception and design of the work. Louise Sun, Anna Chu, Derrick Tam, Xuesong Wang, Jiming Fang, Peter Austin, Natasa Tusevlijak, and Douglas Lee contributed to the acquisition and analysis of the data. All of the authors interpreted the data. Louise Sun, Anna Chu and Derrick Tam drafted the manuscript. All of the authors revised the manuscript critically for important intellectual content, gave final approval of the version to be published and agreed to be accountable for all aspects of the work.

Content licence: This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is non-commercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Funding: This study was funded by CorHealth Ontario as a part of a province-wide quality initiative, and by a Foundation grant from the Canadian Institutes of Health Research grant no. FDN 148446. It is also supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH). The authors acknowledge that the clinical registry data used in this analysis is from participating hospitals through CorHealth Ontario, which serves as an advisory body to the MOH, is funded by the MOH, and is dedicated to improving the quality, efficiency, access and equity in the delivery of the continuum of adult cardiac, vascular and stroke care in Ontario, Canada.

Data sharing: The data set from this study is held securely in coded form at ICES. While legal data sharing agreements between ICES and data providers (e.g., health care organizations and government) prohibit ICES from making the data set publicly available, the full data set creation plan and underlying analytic code are available from the authors upon request, understanding that the computer programs may rely upon coding templates or macros that are unique to ICES and are therefore either inaccessible or may require modification.

Disclaimer: Parts of this material are based on data and/or information compiled and provided by the Canadian Institute for Health Information (CIHI) and Cancer Care Ontario (CCO). The analyses, conclusions, opinions and statements expressed in the manuscript are those of the authors, and do not necessarily reflect those of the above agencies. No endorsement by CIHI or CCO is intended or should be inferred.

Accepted: Sept. 23, 2021

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