A retrospective assessment of prognostication in 456,685 patients undergoing elective major non-cardiac surgery

Une évaluation rétrospective du pronostic de 456 685 patients subissant une chirurgie non cardiaque et non urgente majeure

Daniel McIsaac, MD, MPH · Luke T. Lavallée, MD, MSc · Carl van Walraven, MD, MSc

Received: 5 December 2016 / Revised: 22 February 2017 / Accepted: 8 June 2017 / Published online: 5 July 2017
© Canadian Anesthesiologists' Society 2017

Abstract

Background The accuracy of patient prognostication varies extensively in studies of select populations. Prognostication is an important component of patient selection for elective surgery. This retrospective study assessed patient prognostication prior to elective surgery by comparing the observed with the expected number of deaths in such patients.

Study design We used population-based administrative data to identify all adults undergoing one of 13 most common elective major non-cardiac gender-neutral surgeries in Ontario, Canada from 2002–2014. Survival status within one year of surgery was determined by linking to vital statistics. Expected death risk was determined with health administrative data and a previously derived and externally validated index.

Results We identified 456,685 patients of which 17,266 (3.8%) died within one year of surgery. Patients whose expected one-year death risk was > 25% accounted for 5.3% of the entire cohort ($n = 24,178$) but 51.7% of all deaths ($n = 8,927$). The overall observed death risk was significantly lower than expected (standardized mortality ratio [SMR], 0.72; 95% confidence interval, 0.71 to 0.73; $P < 0.0001$). The SMRs were significantly < 1 (values ranged from 0.54 [partial liver resection] to 0.93 [total knee replacement]) in 11/13 (85%) surgery types. Improved outcomes were especially notable in patients with a higher expected death risk. Only 35/5,539 (0.6%) surgeons had one-year patient death risks that exceeded the population average.

Conclusions The observed number of deaths within one year of elective surgery is significantly lower than expected, with minimal inter-surgeon variation. These results suggest that patient selection for major elective non-cardiac surgery identified individuals with better than expected survival and whose survival was not adversely influenced by their surgery.

This article is accompanied by an editorial. Please see Can J Anesth 2017; 64: this issue.

D. McIsaac, MD, MPH

Departments of Anesthesiology, University of Ottawa, Ottawa Hospital Research Institute, Institute for Clinical Evaluative Sciences, Ottawa, ON, Canada

L. T. Lavallée, MD, MSc

Departments of Surgery, University of Ottawa, Ottawa Hospital Research Institute, Institute for Clinical Evaluative Sciences, Ottawa, ON, Canada

C. van Walraven, MD, MSc (✉)

Departments of Medicine, University of Ottawa, Ottawa Hospital Research Institute, Institute for Clinical Evaluative Sciences, ASB1-003 1053 Carling Ave, Ottawa, ON K1Y 4E9, Canada
e-mail: carlv@ohri.ca

Résumé

Contexte L'exactitude de la pose de pronostic chez les patients varie considérablement selon les études auprès de populations choisies. Le pronostic est une composante importante de la sélection des patients pour une chirurgie non urgente. Cette étude rétrospective a évalué la pose de pronostic avant une chirurgie non urgente en comparant le nombre de décès observé au nombre de décès prévu de tels patients.

Plan de l'étude Nous avons utilisé les données administratives fondées sur la population afin d'identifier tous les adultes subissant l'une des 13 chirurgies majeures

les plus fréquentes n'incluant pas la chirurgie cardiaque, les urgences et non liées au genre en Ontario, Canada, entre 2002 et 2014. Le statut de survie un an après la chirurgie a été déterminé en liant ces données aux statistiques de l'état. Le risque attendu de décès a été déterminé à l'aide des données administratives de santé et d'un indice dérivé et validé au préalable à l'externe.

Résultats Nous avons identifié 456 685 patients, parmi lesquels 17 266 (3,8 %) sont décédés dans l'année suivant leur chirurgie. Les patients dont le risque anticipé de décès à un an était > 25 % représentaient 5,3 % de la cohorte totale ($n = 24\,178$), mais 51,7 % de tous les décès ($n = 8927$). Le risque global de décès observé était significativement plus bas qu'anticipé (taux de mortalité standardisé [TMS], 0,72; intervalle de confiance 95 %, 0,71 à 0,73; $P < 0,0001$). Les TMS étaient significativement < 1 (les valeurs se situant entre 0,54 [hépatectomie partielle] et 0,93 [arthroplastie totale du genou]) dans 11/13 types de chirurgie. De meilleurs pronostics ont été particulièrement notables chez les patients dont le risque de décès anticipé était plus élevé. Le risque de décès à un an des patients de 35/5539 (0,6 %) chirurgiens seulement était supérieur à la moyenne de la population.

Conclusion Le nombre observé de décès à un an après une chirurgie non urgente est significativement plus bas qu'anticipé, avec une variation minimale entre chirurgiens. Ces résultats suggèrent que la sélection des patients pour les chirurgies non cardiaques et non urgentes majeures a permis d'identifier les personnes ayant des chances de survie plus élevées qu'anticipées et dont la survie n'a pas été négativement influencée par la chirurgie.

Expected patient survival influences many decisions in clinical medicine and can affect the utility of many diagnostic, therapeutic, and preventive interventions.¹ For people being considered for elective surgery, predicting the probability of their survival can be difficult because it can be influenced by multiple patient and procedural factors. It is important for surgeons to predict patient survival accurately because this may significantly influence decisions regarding surgery. This topic also has important implications in jurisdictions with limited healthcare resources. For example, if patients with a poor prognosis undergo an intervention in preference to patients with a better prognosis, the health of that population may suffer.

Given the practical importance of prognosticating patients, it is surprising that the accuracy of physician prognostication has not truly been measured in a large sample. The accuracy of physician prognostication has been studied primarily in patients with end-stage tumours and

those with critical illness. Systematic reviews of end-stage cancer patients suggest that the correlation between predicted and observed survival times, though significant, was low² and that physicians tend to over estimate survival (by a median value of 74%³ and 90%).⁴ In contrast, predictions of hospital survival for critically ill patients by experienced intensive care unit physicians were much more accurate. They were highly discriminative (C-statistic of 0.85),⁵ the strongest independent predictors of survival,⁶ and tended to outperform predictions based on multivariate scoring systems.⁷ The results of studies within these systematic reviews could have questionable generalizability because they were usually single-centred, were often prognostications elicited from a limited number of physicians (who may not have participated in the patient's care), and frequently used case summaries rather than actual patients.

In our view, a more realistic way to measure the accuracy of patient prognostication is to compare observed and expected death risks for a particular clinical scenario. Elective surgery is an ideal clinical scenario within which to measure physician prognostic accuracy. For instance, there is a clear start time (the date of surgery); patient prognosis is a key component for the selection of appropriate surgical patients,⁸ and there is time to include this information in the decision-making process (since the surgery is not emergent). In this population-based study, we explored how well patients undergoing elective major non-cardiac surgeries are being prognosticated by comparing their observed and expected one-year death risk within individual procedures and surgeons. Our overall objective for this study was to determine whether there were any signals to indicate that patients with a low prospect of survival were systematically selected for elective surgery. Such indicators could suggest that accurate tools for patient prognosis are needed to help physicians select those patients who are most likely to benefit from elective surgery.

Methods

Study design, setting, and datasets

For this cohort study, we used linked health administrative data in Ontario, Canada where hospital and physician services are provided to all residents through a publicly funded healthcare system, and records of care are collected in health administrative datasets using standardized methods. All data were linked deterministically using encrypted patient-specific identifiers. The Sunnybrook Health Sciences Research Ethics Board provided ethics approval (January 2016) and waived the need for written or oral consent due to the anonymized nature of the utilized data. All data were linked and analyzed at the Institute for Clinical Evaluative Sciences.

The datasets used for the primary analysis included the Discharge Abstract Database (DAD), which captures all surgeries, and the Registered Persons Database (RPDB), which captures all deaths for residents of Ontario. To calculate the expected death risk, we also required linkage to the National Ambulatory Care Reporting System (NACRS), which captures details of emergency and outpatient care, the Assistive Devices Program (ADP) database, which records funding for medical devices, and the Continuing Care Reporting System (CCRS), which records details of long-term and respite care.

Study cohort

We identified all episodes of surgical care for patients who were ≥ 18 years of age and had one of the following elective, intermediate- to high-risk, non-cardiac, gender-neutral surgeries: total hip replacement, total knee replacement, carotid endarterectomy, endovascular abdominal aortic aneurysm repair, open abdominal aortic aneurysm repair, peripheral arterial bypass, nephrectomy, radical cystectomy, large bowel or rectal resection, partial liver resection, pancreaticoduodenectomy, gastrectomy or esophagectomy, and pneumonectomy or lobectomy. These procedures were identified using validated codes (Appendix A)^{9,10} and have previously been used to study outcomes for surgical patients in Ontario.^{11–16} To ensure that surgeries were truly elective, we included only those patients whose admission was classified as “elective” and whose surgical date was the same as their admission date. Episodes of care were identified from April 1, 2002 to March 31, 2014. April 1, 2002 was chosen to coincide with the introduction of the International Classification of Diseases, Tenth Revision (to identify diagnoses) and the Canadian Classification of Interventions (to identify procedures). March 31, 2014 was the latest time at which all datasets were complete when we conducted the study. We analyzed only the first episode of surgical care for each patient within each surgical type to ensure a patient-level analysis.

Outcome

The primary outcome was all-cause death within one year of hospital admission. This was determined by linking to RPDB. All deaths were included as outcomes regardless of cause or location of the event.

Determining expected one-year death risk

We measured each person’s expected one-year all-cause death risk by calculating their Hospitalized-patient One-

year Mortality Risk (HOMR) score (Appendix B).¹⁷ This index consists of ten variables that were calculated using population-based health administrative data available *at patient admission*. These covariates include patient demographics (age, sex, and living status); health burden (measured with the Charlson score,¹⁸ home oxygen status, and number of emergency department visits and admissions to hospital by ambulance in the previous year); and acuity of illness (admission urgency and hospital service, direct admission to an intensive care unit, or whether this was an urgent readmission to hospital). The latter was also quantified using the Diagnostic Risk Score,¹⁷ which quantified diagnosis-specific death risk at admission beyond that explained by the other covariates. Discrete values for each of these covariates were given a specific number of points which were then summed to create the HOMR score. For one-year mortality, the HOMR score has been shown to be highly discriminative [C-statistic, 0.923 (0.922–0.924)] and well calibrated. The mean *relative* difference between observed and expected death risk was 2.0%, ranging from 0.0–7.0%.¹⁷ The HOMR score was externally validated in more than 3 million adults from Ontario, Alberta, and Boston in whom the C-statistic ranged from 0.89–0.92 and calibration remained excellent.¹⁹ The HOMR score has also been validated when calculated using chart-abstracted data²⁰ as well as when using data available within a hospital data warehouse at patient admission.²¹

Statistical analysis

We determined the expected death risk associated with each HOMR score using death rates from the HOMR external validation study.¹⁹ Within each surgical type, these risk estimates were summed to determine the expected number of deaths within a year. We compared the observed number of deaths with this value using the standardized mortality ratio (SMR)—i.e., the ratio of observed to expected death totals—the statistical significance of which was determined by calculating 95% confidence intervals (CIs) using methods described by Rothman and Greenland.²²

We then used generalized linear mixed models (GLMMs) to determine if surgeon-specific adjusted one-year death risks differed significantly from the population average within each surgical type (Appendix C). These models were stratified by surgery and clustered patients with their respective surgeon. We adjusted for patient-specific death risk by including each patient’s HOMR score (centred on its average value) as a fixed covariate. For each surgeon, we calculated the absolute difference between the

one-year patient death risk for that surgeon and the population average one-year death risk for that procedure. This was calculated as:

$$\frac{e^{int-population+int-surgeon}}{1 + e^{int-population+int-surgeon}} - \frac{e^{int-population}}{1 + e^{int-population}}$$

where *int-population* (represented as β_0 in Appendix C) is the mean one-year death risk for all patients within a particular surgical type after accounting for surgeon effects and patient HOMR scores, and *int-surgeon* (represented as s_j in Appendix C) is the mean one-year death risk for a specific surgeon after accounting for patient HOMR scores. The GLMM for carotid endarterectomy did not converge, thereby prohibiting surgeon-specific results for this procedure.

Results

From April 1, 2002 to March 31, 2014, we identified 567,256 major non-cardiac gender-neutral surgeries conducted on Ontarians ≥ 18 years of age with a valid health card. We excluded 110,571 (19.4%) surgeries because patients had more than one operation within a particular procedure ($n = 86,936$, 78.6%); surgeries were not conducted on the date that the patient was admitted, which makes the designation of elective status questionable ($n = 22,527$, 20.4%); or records were missing a surgeon identifier ($n = 1,108$, 1.0%).

This left 456,685 major non-cardiac gender-neutral surgeries (Table 1). Although these patients were elderly, no diagnostic code was entered in the Charlson Comorbidity Index for almost two-thirds of them. Total hip and knee replacement surgeries were the most common procedure (comprising 62% of all surgeries). Almost 90% of patients had an American Society of Anesthesiologists' (ASA) physical status score of III or less. Three-quarters of patients had an expected one-year death risk of $< 5\%$, and only 1.1% had an expected death risk $> 50\%$.

There were 17,266 (3.8%) patient deaths within one year of their surgery. Compared with survivors, the patients who died were older, more likely to be male, had a higher Charlson Comorbidity Score, and were more likely to have been admitted urgently in the previous year (Table 1). Surgeries with the highest crude one-year death risk included radical cystectomy (1,098/4,969, 22.1%) and pancreaticoduodenectomy (527/ 2,696, 19.5%). Death risk was also prominent in patients with an ASA physical score greater than III. Observed death risk increased notably as expected death risk increased. Patients with an expected one-year death risk $> 25\%$ accounted for 5.3% of the cohort but 51.7% of all deaths.

The expected one-year death risk varied extensively between surgical types (Table 2). The discrimination of the HOMR score remained strong in this population (C-statistic 0.822; 95% CI, 0.819 to 0.825). Surgeries with the lowest expected death risk included total hip and knee replacement and carotid endarterectomy, each with a median expected risk of $\leq 1\%$. In contrast, patients' prognoses for surgeries that were frequently related to neoplastic processes (including radical cystectomy, pancreaticoduodenectomy, gastrectomy, and pneumonectomy) had much higher median expected death risks ranging from 13.6–23.2%. Almost one-fifth of patients undergoing pancreaticoduodenectomy had an expected one-year risk of death $> 50\%$.

The observed death risk in patients undergoing elective major surgery was significantly lower than expected (Table 2). For all surgeries combined, the observed number of deaths within one-year of surgery was 28% lower than expected (SMR, 0.72; 95% CI, 0.71 to 0.73; $P < 0.0001$). When stratified by procedure, the SMR was significantly lower than expected in 11/13 surgeries, with values ranging from 0.54 (partial liver resection) to 0.93 (total knee replacement). The SMR significantly exceeded unity for radical cystectomy only (SMR, 1.29; 95% CI, 1.22 to 1.37; $P < 0.0001$).

The extent to which the observed death risk was less than expected varied with the latter procedure (Fig. 1). Patients with expected death risks $< 10\%$ had observed death risks that appeared to vary around the expected, although observed death risks for patients undergoing radical cystectomy exceeded the expected. Nevertheless, as the expected death risk increased, observed death risks in all surgical types decreased notably compared with the expected. When expected death risks exceeded 40%, the observed one-year death risks for all surgical types were less than expected (Fig. 1).

Figure 2 summarizes surgery-specific one-year death risks for particular surgeons. The number of individual surgeons in the analysis varied by surgical type (from 91 for pancreaticoduodenectomy to 1,227 for bowel or rectal resection), as did the maximum number of cases per individual surgeon (from 290 for radical cystectomy to 2,851 for knee replacement). In patients with a HOMR score of 23 (the population average), the procedure-specific one-year death risk ranged from 1.0% (open abdominal aortic aneurysm [AAA] repair) to 9.0% (radical cystectomy). For five surgical types (hip replacement, knee replacement, endovascular AAA repair, open AAA repair, and arterial bypass), none of the surgeon-specific adjusted one-year death risks deviated (in absolute terms) more than 1% from the adjusted provincial average. Pancreaticoduodenectomy had the largest range in

Table 1 Study cohort by death status

Variable	Value	Death at one year		Overall <i>n</i> = 456,685
		No <i>n</i> = 439,419 (96.2%)	Yes <i>n</i> = 17,266 (3.8%)	
Mean age (SD)		66.4 (11.7)	71.6 (11.1)	66.6 (11.7)
Female		231,979 (52.8%)	6,721 (38.9%)	238,700 (52.3%)
Rural		71,795 (16.3%)	2,704 (15.7%)	74,499 (16.3%)
Charlson Score ¹⁸	0	295,721 (67.3%)	4,412 (25.6%)	300,133 (65.7%)
	1	69,688 (15.9%)	2,397 (13.9%)	72,085 (15.8%)
	2-3	45,397 (10.3%)	3,678 (21.3%)	49,075 (10.7%)
	4+	28,613 (6.5%)	6,779 (39.3%)	35,392 (7.7%)
Urgent hospitalizations*	0	391,665 (89.1%)	12,112 (70.1%)	403,777 (88.4%)
	1	38,790 (8.8%)	3,702 (21.4%)	42,492 (9.3%)
	2+	8,964 (2.0%)	1,452 (8.4%)	10,416 (2.3%)
Procedure type	Total hip replacement	105,862 (24.1%)	1,593 (9.2%)	107,455 (23.5%)
	Total knee replacement	174,107 (39.6%)	1,827 (10.6%)	175,934 (38.5%)
	Carotid endarterectomy	10,259 (2.3%)	342 (2.0%)	10,601 (2.3%)
	Endovascular AAA repair	4,099 (0.9%)	260 (1.5%)	4,359 (1.0%)
	Open AAA repair	8,957 (2.0%)	616 (3.6%)	9,573 (2.1%)
	Peripheral arterial bypass	13,020 (3.0%)	1,186 (6.9%)	14,206 (3.1%)
	Nephrectomy	18,371 (4.2%)	1,099 (6.4%)	19,470 (4.3%)
	Radical cystectomy	3,871 (0.9%)	1,098 (6.4%)	4,969 (1.1%)
	Large bowel and rectal	73,911 (16.8%)	4,613 (26.7%)	78,524 (17.2%)
	Partial liver resection	4,964 (1.1%)	664 (3.8%)	5,628 (1.2%)
	Pancreaticoduodenectomy	2,099 (0.5%)	527 (3.1%)	2,626 (0.6%)
	Gastrectomy/esophagectomy	7,582 (1.7%)	1,555 (9.0%)	9,137 (2.0%)
	Pneumonectomy/lobectomy	12,317 (2.8%)	1,886 (10.9%)	14,203 (3.1%)
ASA physical status score	II	174,510 (39.7%)	2,985 (17.3%)	177,495 (38.9%)
	III	221,630 (50.4%)	9,042 (52.4%)	230,672 (50.5%)
	IV	42,978 (9.8%)	5,124 (29.7%)	48,102 (10.5%)
	V	301 (0.1%)	115 (0.7%)	416 (0.1%)
Expected one-year death risk	<5%	338,801 (77.1%)	4,338 (25.1%)	343,139 (75.1%)
	5-<10%	39,736 (9.0%)	2,667 (15.4%)	42,403 (9.3%)
	10-<25%	41,547 (9.5%)	5,418 (31.4%)	46,965 (10.3%)
	25-<50%	15,553 (3.5%)	3,511 (20.3%)	19,064 (4.2%)
	>50%	3,782 (0.9%)	1,332 (7.7%)	5,114 (1.1%)

* In year prior to admission. AAA = abdominal aortic aneurysm; ASA = American Society of Anesthesiologists; SD = standard deviation

absolute surgeon-specific deviations, extending from 3.1% below to 4.6% above the adjusted average. For all surgical types combined, one-year patient death risks for 281/5,539 (5.1%) surgeons deviated significantly from the adjusted average. This varied from 0% of 642 surgeons for hip replacement to 7.5% of 199 surgeons for pneumonectomy. There were 246 (4.4%) surgeons with adjusted one-year procedure-specific death risks that were significantly *lower* than the population average. In contrast, only 35 surgeons (0.63% of all surgeons) had adjusted one-year procedure-

specific death risks that significantly *exceeded* the population average.

Discussion

This novel study represents an extensive examination of the accuracy of physician prognostication for individual patients. In adult Ontarians undergoing major non-cardiac gender-neutral surgeries, observed death risks were 28% lower than

Table 2 Expected one-year death risk by surgical type

Surgical Group	n	Median expected death risk [interquartile range]	% of surgical group in specific strata of expected one-year death risk					Deaths		
			<1%	1-10%	10-25%	25-50%	≥50%	Observed	Expected	Ratio (95% CI)
All surgeries	456,685	1.0 [0.2 - 4.9]%	48.0%	36.4%	10.3%	4.2%	1.1%	17 266	23 879.3	0.72 (0.71 to 0.73)
Total hip replacement	107,455	1.0 [0.2-1.7]%	47.6%	47.8%	3.0%	1.3%	0.3%	1593	2511.2	0.63 (0.60 to 0.67)
Total knee replacement	175,934	0.2 [0.2-0.2]%	83.6%	13.9%	1.8%	0.7%	0.1%	1827	1957.2	0.93 (0.89 to 0.98)
Carotid endarterectomy	10,601	0.2 [0.2-0.2]%	78.9%	9.8%	7.2%	3.5%	0.6%	342	354.4	0.97 (0.87 to 1.07)
Endovascular AAA repair	4359	9.4 [6.3-13.6]%	0.6%	50.5%	39.1%	8.4%	1.4%	260	559.4	0.46 (0.41 to 0.52)
Open AAA repair	9573	9.4 [6.3-13.6]%	0.5%	53.7%	40%	5.4%	0.4%	616	1111.9	0.55 (0.51 to 0.60)
Peripheral arterial bypass	14,206	6.3 [3.3-11.4]%	2.4%	67.9%	22.3%	6.4%	1.1%	1186	1365.5	0.87 (0.82 to 0.92)
Nephrectomy	19,470	3.3 [1.7-7.7]%	10.5%	71.8%	12.7%	4.1%	0.8%	1099	1301.6	0.84 (0.80 to 0.90)
Radical cystectomy	4969	13.6 [6.3-23.2]%	2.0%	38.6%	38.3%	16%	5.1%	1098	848.3	1.29 (1.22 to 1.37)
Large bowel and rectal resection	78,524	6.3 [2.2-11.4]%	11.7%	61.3%	20.1%	5.8%	1.1%	4613	6863.2	0.67 (0.65 to 0.69)
Partial liver resection	5628	19.9 [7.7-32.4]%	5.7%	25.2%	32.7%	27.2%	9.2%	664	1237.6	0.54 (0.50 to 0.58)
Pancreaticoduodenectomy	2626	23.2 [9.4-42.8]%	2.3%	24.3%	25.3%	29.6%	18.5%	527	732.8	0.72 (0.66 to 0.78)
Gastrectomy/esophagectomy	9137	19.9 [4.9-37.2]%	4.7%	30.1%	23.7%	29.3%	12.2%	1555	2124.4	0.73 (0.70 to 0.77)
Pneumonectomy/lobectomy	14,203	16.1 [9.4-27.1]%	2.9%	23.9%	44.4%	22.5%	6.3%	1886	2911.8	0.65 (0.62 to 0.68)

AAA = abdominal aortic aneurysm; CI = confidence interval

expected, with the distinction between observed and expected death risks being most notable for higher-risk patients. The difference between observed and expected death risk varied little between surgeons, with only 0.6% of individual surgeons having one-year patient deaths that significantly *exceeded* the adjusted provincial average. These results suggest that, without using a formal patient prognostication tool, surgeons are accurately identifying those patients who will most likely survive following elective surgery.

In our view, the most likely explanation for these results is excellent prognostication of patients having major elective surgery in Ontario. This approach achieved better results than would have been achieved with the sole use of an externally validated multivariate prognostic score. Undoubtedly, surgeons informally use many of the factors that contribute to the HOMR score, such as patient age and comorbidities, to prognosticate individual patients when determining the clinical utility of surgery. Our data show that, more often than not, surgeons outperformed the HOMR score model, with the observed number of deaths being significantly lower than expected. This points to their ability to select those patients whose actual death risk was less than that projected by the HOMR. It is possible that this was achieved with the use of other important patient prognostic factors (such as functionality), which is

significantly associated with mortality risk independent of the HOMR score.²⁰

Given these results, we question whether adding an accurate one-year patient-specific death-risk prognostication tool to the clinician's clinical impression could further improve patient selection in those being considered for elective surgery. It is noteworthy that patients whose expected one-year death risk exceeded 25% comprised about 5% of the population but contributed > 50% of all deaths in this cohort. These patterns were also identified in a large cohort of British general surgery patients.²³ Whether or not the provision of a patient-specific prognostication tool from a multivariate model improves patient selection (especially medically complex patients or those whose physiologic reserve is difficult to assess clinically, such as the frail elderly),¹⁶ we suggest that this approach should be explored further. By itself, the HOMR score *underestimates* survival in patients who have cleared the hurdles for elective surgery.

Several issues should be considered when interpreting this study. First, our results indicate that the HOMR score was poorly calibrated for patients undergoing common major elective non-cardiac gender-neutral surgery, because the observed death risk in these patients was significantly lower than expected. It is possible that these differences

Fig. 1 Observed vs expected one-year death risk by surgical type. The observed death risk (vertical axis) is plotted against the expected death risk (horizontal axis) for elective non-cardiac surgeries (legend). Point estimates below the reference line (red) indicate observed death probabilities that were expected death probabilities

were due to an intrinsic error in the HOMR score, such that these patients randomly had better than expected survival. We consider this unlikely because several of the factors that overtly define this patient population (e.g., elective hospital admissions and surgical service performing the surgery) were included in the HOMR score model (Appendix B). In addition, elective surgeries are a major component of a hospital's total admissions. As such, they strongly influence the final HOMR score, making poor calibration for these patients less likely. We suggest that our results more likely indicate the ability of physicians involved in the surgical selection process to identify patients whose survival would be better than expected—and doing so without the use of the twelve covariates included in the HOMR score (Appendix B). Second, several factors other than preoperative prognosis contribute to a patient's likelihood of succumbing within one year of surgery. A surgeon's technical skill, a patient's perioperative care, and post-discharge management could all influence patient outcomes. Therefore, issues other than

patient prognostication and case selection will influence the likelihood that patients will die within one year of their surgery. Nevertheless, while these factors will influence postoperative survival, preoperative patient prognosis will undoubtedly influence this outcome. This notion is supported by our finding that only 1.1% of elective surgical patients had a > 50% expected one-year mortality risk. We therefore suggest that elective surgery remains an excellent clinical context within which to study the quality of individual patient prognostication. Third, patient survival is not the only issue that will influence whether surgery is conducted. Cases will exist where surgery is attempted (despite knowledge of the patient's poor prognosis) because no other treatment options exist for the patient's disease. This likely explains why almost one-fifth of patients undergoing pancreaticoduodenectomy had an expected one-year death risk > 50%. Surgeries such as pancreaticoduodenectomy and cystectomy are indeed associated with a high mortality risk, but they also offer the sole chance of long-term survival for these patients. The

Fig. 2 Absolute deviation in surgeon-specific one-year death risk from adjusted provincial surgical average. Within each figure of this array, the absolute difference in surgeon-specific one-year patient death risk from the population average for that surgery (horizontal axis) is plotted for each surgeon against the total number of cases conducted by that surgeon (vertical axis). These risks were determined using generalized linear mixed models that clustered patients within surgeons and controlled for patient expected one-year death risk using the Hospitalized-patient One-year Mortality Risk

(HOMR) score. The procedure-specific average is that for patients with a HOMR score of 23 (the provincial average) and is presented in the upper left hand portion of each plot. Surgeons whose adjusted one-year death risk is statistically distinct from the population average are indicated with black markers (those who are not statistically distinct from the population average are in grey). The proportion of surgeons whose death risks are statistically distinct from the population average is presented in the title of each plot

high SMR seen in cystectomy may also reflect how surgery for such patients (who often have recurrent bleeding and pain) may be performed for palliative reasons. Despite these issues, the observed number of deaths in patients undergoing elective surgery remained significantly less than expected. Fourth, the surgeon is not solely responsible for deciding whether surgery should be attempted for a particular patient. This decision also reflects decisions by referring physicians, preoperative consultants, and the patients themselves. Despite their knowledge of a patient's poor prognosis, surgeons may proceed with a procedure because the group reached a consensus that surgery was the best treatment option for that patient.

Therefore, our positive results most correctly reflect the appropriate decision-making of a cadre of individuals, including the patient, referring physician, perioperative consultants, and the surgeon. Fifth, survival is not the only postoperative outcome. Other factors, such as patient function and quality of life, will also determine the appropriateness of surgical case selection. Sixth, we used hierarchical modelling to cluster patients within surgeons when estimating surgeon-specific one-year death risks. These methods tend to "shrink" surgeon-specific risks to the population mean, especially when surgeons have a small number of procedures.²⁴ While this approach avoids incorrectly classifying surgeons as deviant from the

population mean, it will partly explain the paucity of surgeons with exceptionally high or low one-year mortality rates. Finally, our study was large but limited to a single province. Nevertheless, we have no reason to think that the quality of case selection would be inferior in other medical jurisdictions, but this issue should be studied.

In summary, these data show that the observed risk of one-year death in a large group of Ontario patients undergoing elective non-cardiac gender-neutral surgery was significantly less than that expected from an accurate externally validated multivariable death risk prediction index. In our view, these data indicate excellent prognostication in these patients.

Acknowledgment 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 paper 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. Parts of this material are based on data and information compiled and provided by the Canadian Institute for Health Information (CIHI). However, the analyses, conclusions, opinions, and statements expressed herein are those of the author and not necessarily those of CIHI.

Conflicts of interest None declared.

Editorial Responsibility This submission was handled by Dr. Hilary P. Grocott, Editor-in-Chief, *Canadian Journal of Anesthesia*.

Author contributions Daniel McIsaac and Luke T. Lavallée contributed substantially to all aspects of the paper. Carl van Walraven contributed substantially to all aspects of the paper, except the analysis.

Appendix A: Procedural codes used to identify candidate surgeries

Procedure	CCI codes
ORTHOPEDIC	
Total hip replacement	1.VA.53.LA-PN; 1.VA.53.PN-PN
Total knee replacement	1.VG.53
VASCULAR	
Carotid endarterectomy	1.JE.57 <u>without</u> 1.IJ.76 on the same date (exclude concurrent CABG surgery)
Endovascular Abdominal aortic aneurysm repair	1.KA.80.GQ-NRN or 1.KA.50.GS-BD or 1.KA.50.LA-BD or 1.KE.50.GQ-OA or 1.KE.51.GQ-GE [must have ICD-10 code I71.4 on the same admission]
Open Abdominal aortic aneurysm repair	1.KA.80.LA-XXN or 1.KA.76.NB-XXN or 1.KA.76.MZ-XXN [must have ICD-10 code I71.4 on the same admission]
Peripheral arterial bypass surgery	1.KG.76; 1.JM.76.MI; 1.KT.76; 1.KG.80 or CCI code 1.KA.76 <u>without</u> ICD-10 codes I71.3 or I71.4 or I71.9 on the same admission
UROLOGIC	
Partial or total nephrectomy	1.PC.87; 1.PC.89; 1.PC.91
Radical cystectomy	1.PM.89; 1.PM.90; 1.PM.91; 1.PM.92
INTRAPERITONEAL	
Large bowel and rectal surgery	1.NM.87 (all <u>except</u> 1.NM.87.BA) or 1.NM.89; 1.NM.91; 1.NQ.89; 1.NQ.90 or 1.NQ.87 (all <u>except</u> 1.NQ.87.BA)
Partial liver resection	1.OA.87
Pancreaticoduodenectomy	1.OK.87; 1.OK.89; 1.OK.91
Gastrectomy or esophagectomy	1.NF.87; 1.NF.89; 1.NF.90; 1.NF.91; 1.NF.92; 1.NA.88; 1.NA.89; 1.NA.90; 1.NA.91; 1.NA.92; 1.NA.87
THORACIC	
Pneumonectomy and lobectomy	1.GR.89; 1.GR.91; 1.GT.89; 1.GT.91; 1.GT.87

CCI = Canadian Classification of Health Interventions; ICD-10 = 10th revision of the International Statistical Classification of Diseases and Related Health Problems

Appendix B: The hospitalized-patient one-year mortality risk (HOMR) score

Variable	Value	Points	Variable	Value	Points			
Sex	Female	0	ED visits**	0	0			
	Male	1		1+	1			
Home O ₂	No	0	Admissions by ambulance**	0	0			
	Yes	4		1	3			
Diagnostic Risk Score*	Same			2	4			
				3+	5			
Service	General medicine	10	Service	General surgery	8			
	Cardiology	8		Cardiovascular surgery	9			
	GI/Nephro/Neuro	9		Neurosurgery	10			
	Palliative Care	28		Orthopedic, Plastic Surgery	7			
	Hematology/Oncology	14		Thoracic / Transplant	7			
	Ante-,intra-,postpartum	0		Trauma	8			
	Gynecology	7		Urology	6			
Variable	Level	Points						
		Charlson Score						
		0	1	2	3	4	5	6
Age	20-24.9	0	3	5	7	8	9	10
	25-29.9	2	5	7	9	10	11	11
	30-34.9	4	7	9	11	12	12	13
	35-39.9	7	9	11	12	13	14	15
	40-44.9	8	11	13	14	15	15	16
	45-49.9	10	13	14	15	16	17	17
	50-54.9	12	14	16	17	17	18	18
	55-59.9	14	16	17	18	19	19	20
	60-64.9	15	17	18	19	20	20	21
	65-69.9	17	19	20	21	21	22	22
	70-74.9	18	20	21	22	22	23	23
	75-79.9	20	21	22	23	23	24	24
	80-84.9	21	23	23	24	24	25	25
	85-89.9	23	24	25	25	25	26	26
	90-94.9	24	25	26	26	26	27	27
	95+	25	26	27	27	27	28	28
Variable	Level	Points						
		Admissions by ambulance						
		0	1	2	3+			
Living Status	Independent	0	0	0	0			
	Rehabilitation	3	3	2	2			
	Home care	4	3	3	3			
	Nursing Home	4	4	4	3			
	Chronic Hospital	8	6	5	5			
Admission Urgency	Elective	0	0	0	0			
	ED, no ambulance	3	1	0	0			
	ED, ambulance	5	2	1	0			

**In previous year. ED = emergency department

Appendix C: Details regarding generalized linear mixed models (GLMMs)

A model was created for each procedure type. The Hospitalized-patient One-year Mortality Risk (HOMR) score was centred (i.e., the patient's actual HOMR score was subtracted from the population average so that the mean score was zero). In each model, patients (i) were clustered within surgeons (j). The level 1 model (i.e., the model at the patient level) was:

$$Y_{ij} = \alpha_0 + \alpha_1 * \text{HOMR} - \text{score} + \text{residual}_{ij}$$

in which Y_{ij} was the one-year risk of death; α_0 was the patient-specific intercept (or the expected one-year death risk for a patient with an average HOMR score); α_1 was the parameter estimate for the influence of HOMR score on one-year death risk; and residual_{ij} was the difference between a particular patient's expected death risk (calculated as $\alpha_0 + \alpha_1 * \text{HOMR} - \text{score}$) and their observed death risk.

Then we let the intercepts (α_0) vary between surgeons. Therefore, the level 2 model (i.e., the model at the surgeon level) was:

$$\alpha_0 = \beta_0 + s_j$$

in which β_0 was the grand mean for surgeon effect and s_j was the surgeon-specific intercept. Therefore, the final GLMM within each procedure was:

$$Y_{ij} = \beta_0 + \alpha_1 * \text{HOMR} - \text{SCORE} + \text{residual}_{ij} + s_j$$

Estimation was accomplished using pseudo-likelihood techniques based on residual likelihood in which the expansion locus of the Taylor series expansion was the vector of random effects solutions. Since the outcome was one-year death, the response distribution was binary and the link function used was logit.

References

1. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. *JAMA* 2012; 307: 182-92.
2. Viganò A, Dorgan M, Buckingham J, Bruera E, Suarez-Almazor ME. Survival prediction in terminal cancer patients: a systematic review of the medical literature. *PalliatMed* 2000; 14: 363-74.
3. Cheon S, Agarwal A, Popovic M, et al. The accuracy of clinicians' predictions of survival in advanced cancer: a review. *Ann Palliat Med* 2016; 5: 22-9.
4. Chow E, Harth T, Hruby G, Finkelstein J, Wu J, Danjoux C. How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review. *Clin Oncol (R Coll Radiol)* 2001; 13: 209-18.
5. Knaus WA, Wagner DP, Lynn J. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. *Science* 1991; 254: 389-94.
6. Rocker G, Cook D, Sjøkvist P, et al. Clinician predictions of intensive care unit mortality. *Crit Care Med* 2004; 32: 1149-54.
7. Sinuff T, Adhikari NK, Cook DJ, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. *Crit Care Med* 2006; 34: 878-85.
8. 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.
9. Canadian Institute for Health Information. Technical supplement: Health Care in Canada 2005. Available from URL: <https://secure.cihi.ca/estore/productSeries.htm?pc=PCC64> (accessed February 2017).
10. Bourne R, DeBoer DP, Hawker G, Kreder H, Mahomed M, Paterson J. Total hip and knee replacement. In: Tu JV, Pinfold SP, McColgan P, Laupacis A, editors. *Access to Health Services in Ontario: An ICES Atlas*. Toronto: Institute for Clinical Evaluative Sciences; 2005. p. 156-71.
11. Wijeyesundera DN, Austin PC, Beattie WS, Hux JE, Laupacis A. Variation in the practice of preoperative medical consultation for major elective noncardiac surgery: a population-based study. *Anesthesiology* 2012; 116: 25-34.
12. Wijeyesundera DN, Wijeyesundera HC, Yun L, et al. Risk of elective major noncardiac surgery after coronary stent insertion: a population-based study. *Circulation* 2012; 126: 1355-62.
13. Wijeyesundera DN, Beattie WS, Karkouti K, Neuman MD, Austin PC, Laupacis A. Association of echocardiography before major elective non-cardiac surgery with postoperative survival and length of hospital stay: population based cohort study. *Br Med J* 2011; 342: d3695.
14. Wijeyesundera DN, Austin PC, Beattie WS, Hux JE, Laupacis A. A population-based study of anesthesia consultation before major noncardiac surgery. *Arch Intern Med* 2009; 169: 595-602.
15. McIsaac I, Bryson GL, van Walraven C. Elective, major noncardiac surgery on the weekend: a population-based cohort study of 30-day mortality. *Med Care* 2014; 52: 557-64.
16. McIsaac DI, Bryson GL, van Walraven C. Association of frailty and 1-year postoperative mortality following major elective noncardiac surgery: a population-based cohort study. *JAMA Surg* 2016; 151: 538-45.
17. van Walraven C. The hospital-patient one-year mortality risk score accurately predicted long-term death risk in hospitalized patients. *J Clin Epidemiol* 2014; 67: 1025-34.
18. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987; 40: 373-83.
19. van Walraven C, McAlister FA, Bakal JA, Hawken S, Donze J. External validation of the Hospital-patient One-year Mortality Risk (HOMR) model for predicting death within 1 year after hospital admission. *CMAJ* 2015; 187: 725-33.
20. Casey G, van Walraven C. Using the the Hospitalized patient One-year Mortality Risk (HOMR) score for prognostication using information from the medical record. *J Am Geriatr Soc* 2017 (in press).
21. van Walraven C, Forster AJ. HOMR-now! A modification of the HOMR score that predicts 1-year death risk for hospitalized patients using data immediately available at patient admission. *Am J Med* 2017 (in press).
22. Greenland S, Rothman KJ. Introduction to categorical statistics. In: Rothman KJ, Greenland S, editors. *Modern Epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1998. p. 253-80.
23. Pearse RM, Harrison DA, James P, et al. Identification and characterisation of the high-risk surgical population in the United Kingdom. *Crit Care* 2006; 10: R81.
24. Glance LG, Li Y, Dick AW. Quality of quality measurement: impact of risk adjustment, hospital volume, and hospital performance. *Anesthesiology* 2016; 125: 1092-102.