Canadian benchmarks for acute injury care

Lynne Moore, PhD
David Evans, MD, MSc
Natalie L. Yanchar, MD, MSc
Jaimini Thakore, MBA
Henry Thomas Stelfox, MD, PhD
Sayed Morad Hameed, MD, MPh
Richard Simons, MD
John Kortbeek, MD
Julien Clément, MD
François Lauzier, MD, MSc
Alexis F. Turgeon, MD, MSc

Accepted Apr. 24, 2017; Early-released Oct. 1, 2017; subject to revision

Correspondence to:
L. Moore
CHU de Québec – Université Laval Research Centre
1401, 18e rue, local H-012A
Québec QC G1J 1Z4
lynne.moore@fmed.ulaval.ca

DOI: 10.1503/cjs.002817

Background: Acute care injury outcomes vary substantially across Canadian provinces and trauma centres. Our aim was to develop Canadian benchmarks to monitor mortality and hospital length of stay (LOS) for injury admissions.

Methods: Benchmarks were derived using data from the Canadian National Trauma Registry on patients with major trauma admitted to any level I or II trauma centre in Canada and from the following patient subgroups: isolated traumatic brain injury (TBI), isolated thoracoabdominal injury, multisystem blunt injury, age 65 years or older. We assessed predictive validity using measures of discrimination and calibration, and performed sensitivity analyses to assess the impact of replacing analytically complex methods (multiple imputation, shrinkage estimates and flexible modelling) with simple models that can be implemented locally.

Results: The mortality risk adjustment model had excellent discrimination and calibration (area under the receiver operating characteristic curve 0.886, Hosmer–Lemeshow 36). The LOS risk-adjustment model predicted 29% of the variation in LOS. Overall, observed:expected ratios of mortality and mean LOS generated by an analytically simple model correlated strongly with those generated by analytically complex models ($r > 0.95$, κ on outliers > 0.90).

Conclusion: We propose Canadian benchmarks that can be used to monitor quality of care in Canadian trauma centres using Excel (see the appendices, available at canjsurg.ca). The program can be implemented using local trauma registries, providing that at least 100 patients are available for analysis.

Contexte : L’issue des traitements dispensés dans les services de traumatologie d’urgence varie substantiellement d’une province canadienne et d’un centre de traumatologie à l’autre. Notre but était d’établir des valeurs de référence pour suivre la mortalité et la durée des séjours hospitaliers en traumatologie au Canada.

Méthodes : Les paramètres ont été sélectionnés à partir des données du Registre national des traumatismes concernant les grands polytraumatisés admis dans tout centre de traumatologie de niveau I ou II au Canada et selon les catégories de patients suivantes : traumatisme crânien isolé (TCI), traumatisme thoraco-abdominal isolé, traumatisme plurisystémique fermé, âge de 65 ans ou plus. Nous avons évalué la validité prédictive à l’aide de critères discriminants et de paramètres d’étalonnage et nous avons procédé à des analyses de sensibilité pour évaluer l’impact du remplacement de méthodes analytiques complexes (imputation multiple, estimations par contraction des coefficients et modélisation flexible) par des modèles simples applicables à l’échelle locale.

Résultats : Le modèle d’ajustement du risque de mortalité s’est révélé doté d’un pouvoir discriminant et d’un étalonnage excellents (aire sous la courbe de la fonction d’efficacité du récepteur [ROC] 0,886, test de Hosmer–Lemeshow 36). Le modèle d’ajustement du risque pour la durée du séjour hospitalier a permis de prédire 29 % de sa variation. De plus, les rapports observés:attendus pour la mortalité et la durée moyenne des séjours hospitaliers générés par un modèle analytique simple ont été en étroite corrélation avec les rapports générés par les modèles analytiques complexes ($r > 0,95$, κ pour valeurs aberrantes > 0,90).

Conclusion : Nous proposons des valeurs de référence canadiennes qui peuvent être utilisées pour faire le suivi de la qualité des soins dans les centres de traumatologie canadiens à l’aide d’un simple programme Excel (voir les annexes, accessible à l’adresse canjsurg.ca). Le programme peut être appliqué à l’aide des données des registres de traumatologie locaux à la condition qu’au moins 100 patients y soient accessibles pour analyse.

Important improvements in mortality and morbidity associated with injury have been achieved with the introduction of trauma systems.[1–3] However, variations in patient outcomes have been observed even among level I trauma centres in the United States[4] and Canada,[5] and suboptimal quality of care has been documented in up to 50% of trauma system admissions.[6,7] Research has suggested that benchmarking activities stimulate quality improvement, which in turn improves patient outcomes.[8]

Institutional evaluation of injury care is often limited to preventable morbidity and mortality case conferences or observed:expected (O:E) mortality ratios derived using the outdated Trauma Injury Severity (TRISS) model.[9] Hospitals worldwide now have the opportunity to participate in the Trauma Quality Improvement Program (TQIP) of the American College of Surgeons, which produces quality reports, including indicators of mortality and major complications.[10] However, TQIP costs preclude participation among low-volume or budget-constrained centres. Furthermore, while international comparisons are certainly informative, given the differences in patient populations and health care systems, there is also a need to compare injury outcomes nationally. In order to drive national, provincial, and institutional efforts to improve the quality of injury care in Canada, there is a need to develop Canadian benchmarks that will enable trauma centres nationwide to evaluate their injury outcomes using routinely collected data.

The objective of this study was to develop Canadian benchmarks to evaluate acute injury care in terms of mortality and hospital length of stay (LOS).

Methods

Study setting, population and data

Benchmarks were based on injury admissions to level I and II trauma centres[11] across Canada between Apr. 1, 2006, and Mar. 31, 2012. Patient-level data were extracted from the National Trauma Registry comprehensive data set, which contains information on sociodemographic characteristics, circumstances of injury, anatomic injury descriptions, physiologic parameters, interventions, and outcomes for patients admitted with major trauma to level I and II trauma centres across Canada. Data are abstracted from patient files by trained data coders in each hospital according to a standardized data dictionary. Anatomic injury is coded using the Abbreviated Injury Scale (AIS) according to guidelines published by the Association for the Advancement of Automotive Medicine (AAAM).[12] The registry was centralized at the Canadian Institute of Health Information (CIHI), where data quality was managed.

Patients

All adults (age ≥ 16 yr) with major trauma, defined as an injury severity score (ISS) greater than 12, were included in the study. Patients coded as dead on arrival and patients who arrived with no vital signs and died within 30 minutes were excluded. For benchmarking LOS, only live discharges were included. Previous analyses of the same data showed no significant changes in benchmarking results when deaths were included in a competitive risks framework.[5] The research ethics board of the CHU de Québec approved our study.

Outcomes

Mortality was defined as any deaths occurring between arrival at the emergency department and hospital discharge. Length of stay in the index hospital was calculated as the number of days between admission and discharge, truncated at 90 days; any LOS longer than 90 days was attributed 90 days to minimize the influence of extreme values. Length of stay has a right-skewed distribution and was therefore modelled using a natural log transformation. The natural log transformation generates geometric means that are approximately equivalent to the median.[13] Previous research has suggested that additional acute care days owing to interfacility transfer do not influence benchmarking results.[14]

Derivation of risk-adjustment models

Mortality was modelled using logistic regression and LOS using a log-linear model. Candidate risk-adjustment variables (i.e., baseline patient demographic and clinical characteristics) were determined a priori using previous work by our research team.[15,16] Variable selection was based on bootstrap resampling.[17,18] Variables that were statistically significant in more than 70% of 500 bootstrap samples of size *n* were included in the final risk-adjustment model.[17]

Models were derived for the whole patient population (all injury) and for 4 injury subgroups:

- isolated traumatic brain injury (TBI), defined as patients with any of the AIS codes 115299.9, 115999.9, 115099.9, 113000.6, 116002.3, 116004.5, 100099.9, 100999.9, 113000.6, 116000.3, 116002.3, 116004.5, 120099.3–22899.3, 130202.2–132699.2, 140202.5–140799.3, 150000.2–150408.4, 161000.1–161013.5 and no injury in another body region with an AIS of 3 or higher;
- isolated thoracoabdominal injury, defined as an injury to the thorax or abdomen with an AIS of 3 or higher and no injury in another body region with an AIS of 3 or higher;
- blunt multisystem injury, defined as injuries in at least 2 body regions with an AIS of 3 or higher; and
- patients aged 65 years or older.

In the database used to produce risk-adjustment models (Canadian National Trauma Registry, 2006–2012), injury coding was based on AIS90 (updated to 98).[12] Most trauma registries across Canada converted to AIS05 shortly after 2012. We therefore produced risk-adjustment models for AIS90 and for AIS05. For the latter we converted AIS90 codes to AIS05 codes using validated conversion algorithms recommended by the AAAM.[19]

Validation of models

The predictive performance of mortality risk–adjustment models derived with 2006–2011 data (derivation sample) was evaluated with 2012 data (validation sample). We assessed discrimination with the area under the receiver operating characteristic curve (AUC)[20] and calibration with the Hosmer–Lemeshow statistic and Cox's calibration intercept and slope.[21] We evaluated overall model fit with Nagelkerke's R^2.[22] Risk-adjustment models for mortality were considered to have excellent discrimination if $0.8 \le AUC < 0.9$ and outstanding discrimination if $AUC \ge 0.9$.[23] We evaluated risk-adjustment models for LOS with r^2 statistics.

Benchmarking

Benchmarking is based on Agency for Healthcare Research and Quality (AHRQ) methodology.[24] For mortality, we calculated the O:E ratio by applying coefficients from the risk-adjustment models described previously (derived using 2006–2011 data) to hospital-specific patient characteristics (2012 data). Benchmarking on LOS was based on the ratio of the sum of observed hospital days to expected hospital days.

Sensitivity analyses

Our goal was to derive an analytically simple benchmarking method that would enable trauma centres to calculate their own benchmarking statistics locally. We therefore performed extensive sensitivity analyses to evaluate the impact of analytically complex methods, considered to be the gold standard in benchmarking methodology. We performed multiple imputations for missing data, shrinkage estimates to handle unstable estimates for low-volume centres and used flexible modelling techniques for continuous covariates.

Handling missing data

Glasgow Coma Scale (GCS) scores were missing for 20% of data observations. In most benchmarking analyses, observations with missing GCS are excluded. This may introduce bias, as the GCS is not missing completely at random; rather, it is frequently missing for patients with minor extracranial injury and for patients who are sedated and intubated on arrival.[25] An alternative method is to attribute a GCS value based on the value of variables highly correlated to the GCS, such as maximum head AIS (single imputation). However, this method does not account for the complex covariance structure of the analysis or the uncertainty surrounding the missing value and, therefore, tends to underestimate variance (too many outliers). The current gold standard for handling missing data in benchmarking analysis is multiple imputation. However, this method requires complex analytical skills. We compared O:E ratios generated using single imputation to those generated using multiple imputation. The former was based on the mean GCS observed for maximum head AIS categories. The latter was based on the Markov Chain Monte Carlo method with a noninformative prior and a single chain used to generate 5 imputations for each missing data value.[25,26] Each multiple imputation model (1 per outcome) included all independent and dependent variables used in the respective analysis model. In total 3% of observations had missing data on systolic blood pressure. These observations were attributed a "normal" blood pressure (i.e., ≥ 90 mm Hg).

Handling unstable estimates for low-volume centres

For low-volume centres, O:E estimates may be subject to regression to the mean bias, whereby punctual estimates based on few observations are too far from the mean and too often declared outliers. To address this bias, multilevel modelling can be used to generate shrinkage estimates, whereby O:E ratios are multiplied by a shrinkage factor based on the inverse of their variance.[27] We compared O:E ratio estimates to which a shrinkage factor was applied (random intercept in a multilevel model) to those with no shrinkage factor (fixed effect in a single-level model).

Flexible modelling continuous covariates

Continuous covariates, such as age, systolic blood pressure (SBP), GCS score and the new injury severity score (NISS), may not have linear associations with the outcome of interest. For example, the risk of death in injury populations is generally stable until 55 years of age and increases exponentially thereafter. A simple technique for handling nonlinear associations is to model categories of continuous covariates using dummy variables (e.g., model age 16–54, 55–64, 65–74, 75–84 and ≥ 85 yr). However, this approach leads to the difficulty of choosing appropriate cut-points and to unintuitive step increases in risk that assume constant risk within categories (e.g., all patients in the 75–84 yr age group are assumed to have the same risk of death).

Flexible modelling strategies, such as splines and fractional polynomials,[28] can be used to address the limitations of categories. However, they are computationally complex. We compared O:E ratios generated by a model based on categories of continuous covariates to those generated by a model based on fractional polynomials.

Additional sensitivity analyses

Information on comorbidities is often unavailable in trauma registries and is subject to underreporting and variable coding across centres. In addition, previous research has shown that with robust adjustment for age, information on comorbidities does not significantly improve injury mortality prediction models. We evaluated the impact of adjustment for comorbidities in observations that could be matched to hospital discharge data. Finally, mortality should theoretically be evaluated over a fixed period of time, and the proportion of injury-related deaths occurring after 30 days varies substantially across Canadian trauma centres (0%–15% in our sample). However, time to death among patients admitted for injuries may be influenced by factors that are unrelated to injury severity and quality of care, such as do-not-resuscitate directives. We evaluated the impact of using 30-day in-hospital mortality rather than mortality evaluated over the whole hospital stay.

To evaluate the agreement on benchmarking results under each sensitivity analysis, we calculated Pearson correlation coefficients on O:E ratios and weighted κ coefficients on outliers. We considered correlation coefficients greater than 0.95 and κ coefficients greater than 0.7 to convey acceptable agreement.[29]

All analyses were performed using SAS software version 9.4 for Windows (SAS Institute Inc.). We created a series of Excel spreadsheets to calculate O:E ratios for mortality (Appendix 1, canjsurg.ca/002817-a1 for AIS-1998 and Appendix 2, canjsurg.ca/002817-a2 for AIS-2005) and LOS (Appendix 3, canjsurg.ca/002817-a3 for AIS-1998 and Appendix 4, canjsurg.ca/002817-a4 for AIS-2005) along with detailed instructions (Appendix 5, canjsurg.ca/002817-a5). The algorithm is based on coefficients derived from 2006–2012 data and has been tested on hospital registry data by Trauma Services BC (J.T.).

Results

Study population

Between 2006 and 2012, 80 353 major adult trauma admissions to level I and II trauma centres across Canada were included in the NTR. After exclusion of 968 patients who were dead on arrival (1.2%) and 578 patients admitted for observation with undetermined injuries (0.7%), 78 807 (98.1% of those eligible) patients were included in the final study sample and used for analyses on in-hospital mortality. Of these 70 425 (87.6%) patients were discharged alive and were included in analyses on LOS.

In the whole study population (all years, all types of injury), the mean age was 51 years, and 71% of patients were men. One-third of the study population had isolated TBI, one-third injuries to multiple body regions and one-sixth isolated thoracoabdominal injuries (Table 1).

Globally, 8382 (10.6%) patients died during their hospital stay. Crude mortality increased with age and injury severity and was highest among patients with severe isolated TBI (Table 1). Crude mortality varied little over the study period but did vary across provinces.

Mean (geometric) LOS was 16.7 ± 18.1 (median 10) days. Mean LOS increased with age and injury severity and was longer for women than men and for patients with TBI than those with other injuries. Crude LOS decreased over the study period and varied across provinces.

Risk-adjustment models

Risk-adjustment models for mortality had an AUC greater than 0.8 in the whole validation sample and in every injury subgroup, indicating excellent discrimination (Table 2).[23] Calibration intercepts close to 0 and slopes close to 1 also indicate excellent agreement between observed and predicted mortality probabilities, and high R^2 values indicated good overall model fit. The risk-adjustment model for LOS predicted 29% of the variation in hospital stay in the whole validation sample, 26% in multiple blunt injury, 22% in isolated TBI, 20% in patients aged 65 years or older, and 18% in isolated thoracoabdominal injury.

Benchmarking results

The O:E mortality ratios varied between 0.19 and 1.55 across Canadian trauma centres for the whole sample (Fig. 1). Three centres had observed mortality significantly higher than expected and 4 lower than expected according to the average mortality experience across Canada. The O:E ratios for hospital stay varied between 0.91 and 1.10 across trauma centres (Fig. 2). Two hospitals had a mean LOS significantly higher than expected according to their case mix and 2 had significantly lower mean LOS.

Sensitivity analyses

In the whole sample and in all injury subgroups, O:E ratios of mortality and LOS generated under multiple imputation, fractional polynomial modelling and with shrinkage estimates correlated strongly with analytically simple estimates (Table 2). The only exception was a weak agreement on outliers with the model derived under multiple imputation for isolated TBI (κ = 0.32).

The O:E mortality ratios calculated with adjustment for comorbidities had very strong agreement with those calculated using no adjustment for comorbidities in the whole sample and in all injury subgroups. Again, the only exception was a weak agreement on outliers for isolated TBI (κ = 0.49). However, O:E ratios on LOS with adjustment for comorbidities had low agreement with those generated by a model without comorbidities in the whole sample and in all subgroups.

Discussion

We have developed Canadian benchmarks for mortality and hospital LOS for acute injury care. These benchmarks

Table 1. Characteristics of the study population

Characteristic	No. (%)	Mortality, %	Mean LOS, d	Characteristic	No. (%)	Mortality, %	Mean LOS, d
All patients	78 807 (100)	10.6	16.7	GCS score*			
Age, yr				3–8	6383 (8.0)	37.0	31.5
< 55	44 773 (56.8)	6.5	15.6	9–12	3908 (5.0)	16.3	22.1
55–64	10 190 (12.9)	9.3	17.0	13–15	52 528 (66.7)	4.0	14.0
65–74	8778 (11.1)	13.1	18.5	Missing	15 988 (20.3)	20.4	—
75–84	9803 (12.4)	19.8	19.1	NISS			
≥ 85	5263 (6.7)	27.3	20.2	12–24	24 529 (31.1)	3.2	11.8
Sex				25–40	32 197 (40.9)	6.8	15.3
Male	56 267 (71.4)	10.2	16.5	41–49	9979 (12.7)	11.5	21.4
Female	22 540 (28.6)	11.8	17.5	50–66	10 171 (12.9)	30.0	30.3
Mechanism of injury				67–75	1931 (2.4)	63.0	47.1
MVC	32 476 (41.2)	8.2	17.4	Year			
Fall from height	12 709 (16.1)	15.4	16.2	2006	10 779 (13.7)	11.0	17.3
Fall from own height	17 378 (22.1)	13.3	16.9	2007	11 345 (14.4)	10.4	17.4
Penetrating	3320 (4.2)	11.3	14.5	2008	11 192 (14.2)	10.9	16.9
Other	12 924 (16.4)	8.5	16.0	2009	11 138 (14.1)	10.5	16.4
Type of injury				2010	11 353 (14.4)	10.6	16.3
Multiple blunt	25 126 (31.9)	11.9	22.4	2011	12 156 (15.4)	10.3	16.4
Isolated TBI	29 395 (37.3)	14.7	14.5	2012	10 844 (13.8)	10.8	16.6
Isolated thoracoabdominal	13 178 (16.7)	3.6	10.5	Province			
Other	10 612 (13.5)	5.3	18.2	British Columbia	11 663 (14.8)	10.5	18.1
Transfer				Alberta	15 689 (19.9)	8.3	15.3
No	40 774 (51.7)	12.2	17.5	Manitoba	3024 (3.8)	10.3	16.2
Yes	38 033 (48.3)	9.0	16.0	Ontario	26 888 (34.1)	11.6	16.2
SBP*				Quebec	18 229 (23.1)	11.1	17.7
≥ 90	72 946 (92.6)	9.7	16.3	Nova Scotia	2665 (3.4)	12.1	17.3
< 90	3045 (3.9)	33.4	30.4	Newfoundland & Labrador	649 (0.8)	9.7	24.3
Missing	2816 (3.6)	10.4	—				

GCS = Glasgow Coma Scale; LOS = length of stay in hospital; MVC = motor vehicle collision; NISS = new injury severity score; SBP = systolic blood pressure; TBI = traumatic brain injury
*Measured on arrival in the emergency department.

Table 2. Accuracy of risk adjustment models for predicting mortality in the whole study population and injury subgroups

	Group				
Variable; model	Whole sample	Isolated TBI	Isolated thoracoabdominal injury	Multiple blunt injuries	Age ≥ 65 yr
Mortality					
No. of patients	10 671	4257	1849	3205	3612
AUC (95% CI)	0.886 (0.876–0.895)	0.842 (0.827–0.857)	0.854 (0.794–0.915)	0.878 (0.861–0.896)	0.817 (0.799–0.835)
Hosmer–Lemeshow statistic	36.3	18.9	7.6	17.3	12.1
Calibration intercept; slope	0.00253; 0.97645	0.00635; 0.95829	–0.00242; 1.08462	0.00253; 0.97921	0.00194; 0.98949
Nagelkerke's R^2	0.4008	0.3410	0.2708	0.3893	0.3197
Length of stay					
No. of patients	9524	3613	1794	2829	2940
R^2	0.2898	0.2204	0.1780	0.2611	0.2088

AUC = area under the receiver operating characteristic curve; CI = confidence interval.

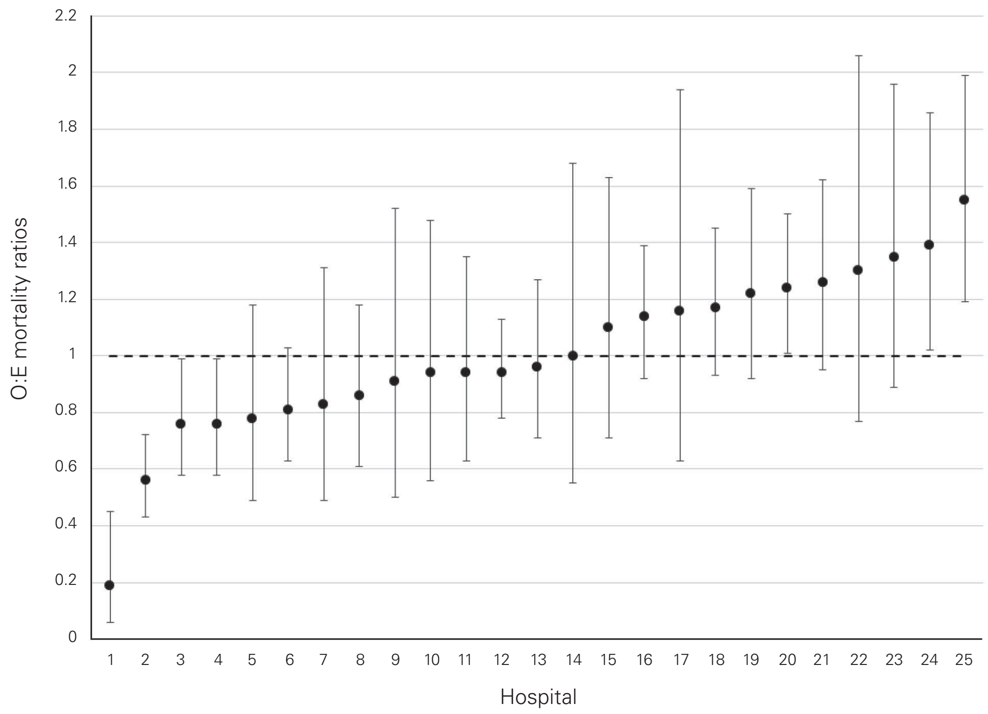

Fig. 1. Observed:expected (O:E) ratios of mortality with 95% confidence intervals for Canadian trauma centres.

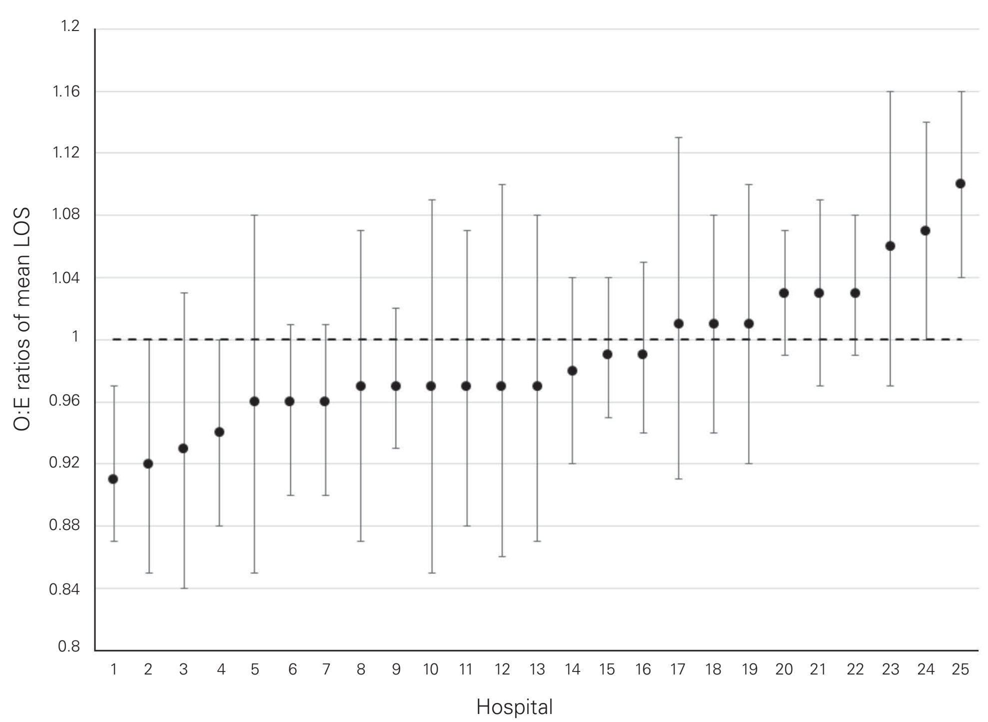

Fig. 2. Observed:expected (O:E) ratios of mean (geometric) length of stay (LOS) with 95% confidence intervals for Canadian trauma centres.

will enable Canadian trauma centres to evaluate quality using a simple Excel algorithm that can be easily applied using trauma registry data (see the appendices for algorithms and detailed instructions). This information can be used by local trauma programs to inform quality-improvement activities. The benchmarks could also be used in trauma systems outside Canada for international comparisons. They may be particularly useful in countries with similar health care systems (e.g., United Kingdom, Australia).

Interpretation

The O:E mortality ratio is the ratio of the number of deaths observed in a trauma centre to the number that would be expected if that centre had the mortality experience of the average Canadian trauma centre. Centres with a lower O:E 95% confidence limit above the national average have higher mortality than expected according to their patient case mix. Local trauma committees can look into reasons behind the observed difference in terms of structures and processes. Although the risk-adjustment models derived here cannot be used to pass judgment on the appropriateness of care for individual patients, they can be used to identify patients with unexpected outcomes for discussion (e.g., deaths with an expected mortality probability < 20%). Similarly, patients with a large difference between observed and expected LOS may be targeted for review by the trauma committee. Centres with a lower O:E 95% confidence limit below the national average have lower mortality than expected according to their patient case mix and can be used for emulation. The possibility that high or low outliers are due to data quality problems, such as underestimation of injury severity, should also be considered and may stimulate improvements in trauma registry data quality. Hospitals can use O:E ratios plotted over time to monitor trends in outcomes within their centres.

Limitations

We used analytically simple O:E ratios because their calculation does not require specialized expertise or data analysis tools. However, they must be used in consideration of their limits.[30] First, O:E ratios should be used only to compare each hospital to the benchmark. They are not, in theory, suitable for comparisons across hospitals with different patient case mixes, as they are standardized to the case mix of the hospital under evaluation rather than to a common case mix pattern.[30] Second, like in many trauma registries, a high percentage of GCS data was missing. Although our sensitivity analyses suggest that simple imputation generates results similar to those obtained under multiple imputation, neither are a substitute for real data. Furthermore, agreement on outliers for isolated TBI was low between models with simple and multiple imputation strategies. Trauma centres should therefore make efforts to optimize the collection of these data in their trauma registries. Third, because shrinkage estimates were not used, we do not recommend calculating O:E ratios when fewer than 100 observations are available. Low-volume centres should accumulate 2 or more years of data for benchmarking analysis. Fourth, sensitivity analyses indicate that comorbidities may play a role in risk adjustment when benchmarking LOS. However, underreporting of comorbidities is a widespread problem, and the low agreement in O:E ratios with and without adjustment for comorbidities may be associated with differences in reporting standards across trauma centres. Nevertheless, we should aim for standardized collection of comorbidity data in all trauma registries to improve risk adjustment of LOS models. Fifth, the validity of benchmarking results depends on the quality of data used to derive them. Each trauma centre is responsible for the quality of their data, and CIHI promotes uniform data collection across Canadian centres using a standardized data dictionary and data quality checks. Data quality is central to the benchmarking process. Finally, to be timely, coefficients should be updated by recalibrating models in new data at least every 5 years. This study shows the potential of a Canadian trauma centre benchmarking system, but a sustainable source of national injury data is essential to its viability.

CONCLUSION

The benchmarks proposed in this study provide trauma centres with the opportunity to evaluate quality of care in terms of mortality and LOS locally. This information should be used to inform local quality-improvement initiatives, not for accountability. Users should be aware that the validity and precision of O:E estimates rely on the availability of large sample sizes (at least 100 observations) and on the quality of data in the trauma registry.

Affiliations: From the Department of Social and Preventative Medicine, Université Laval, Québec, Que. (Moore); the Axe Santé des Populations et Pratiques Optimales en Santé, Traumatologie–Urgence–Soins intensifs, Centre de Recherche du CHU de Québec, Hôpital de l'Enfant-Jésus, Université Laval, Québec, Que. (Moore, Lauzier, Turgeon); the Department of Surgery, University of Calgary, Calgary, Alta. (Yanchar); the Department of Surgery, University of British Columbia, Vancouver, BC (Evans, Thakore, Hameed); the Department of Critical Care Medicine, Medicine and Community Health Sciences (Stelfox), O'Brien Institute for Public Health, University of Calgary, Calgary, Alta. (Stelfox); the Department of Surgery, Division of General Surgery and Division of Critical Care, University of Calgary, Calgary, Alta. (Kortbeek); the Institut national d'excellence en santé et en services sociaux, Québec, Que. (Clément); the Department of Surgery, Université Laval, Québec, Que. (Clément); and the Department of Anesthesiology and Critical Care Medicine, Division of Critical Care Medicine, Université Laval, Québec, Que. (Lauzier, Turgeon).

Funding: L. Moore, H. Stelfox and A. Turgeon are supported by a Canadian Institutes of Health Research New Investigator Award; F. Lauzier is supported by a Fonds de Recherche du Québec — Santé award; and L. Moore is supported by a Canadian Institutes of Health Research operating grant (#110996).

Competing interests: None declared.

Contributors: All authors designed the study. L. Moore acquired the data, which all authors analyzed. L. Moore wrote the article, which all authors reviewed and approved for publication.

References

1. Mackenzie EJ, Rivara FP, Jurkovich GJ, et al. The national study on costs and outcomes of trauma. *J Trauma* 2007;63:S54-67.
2. Moore L, Turgeon A, Lauzier F, et al. Evolution of patient outcomes over 14 years in a mature, inclusive Canadian trauma system. *World J Surg*. 2015;39:1397-405.
3. Gabbe BJ, Biostat GD, Lecky FE, et al. The effect of an organized trauma system on mortality in major trauma involving serious head injury: a comparison of the United kingdom and Victoria, Australia. *Ann Surg* 2011;253:138-43.
4. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. A national evaluation of the effect of trauma-center care on mortality. *N Engl J Med* 2006;354:366-78.
5. Moore L, Evans D, Hameed SM, et al. Mortality in Canadian trauma systems: a multicenter cohort study. *Ann Surg* 2017;265:212-17.
6. Copes WSSC, Konvolinka CW, Sacco WJ. American College of Surgeons audit filters: associations with patient outcome and resource utilization. *J Trauma* 1995;38:432-8.
7. Moore L, Lavoie A, Sirois MJ, et al. Evaluating trauma center process performance in an integrated trauma system with registry data. *J Emerg Trauma Shock* 2013;6:95-105.
8. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;6:CD000259.
9. Champion HR, Copes WS, Sacco WJ, et al. The Major Trauma Outcome Study: establishing national norms for trauma care. *J Trauma* 1990;30:1356-65.
10. Nathens AB, Cryer HG, Fildes J. The American College of Surgeons Trauma Quality Improvement Program. [x-xi.]. *Surg Clin North Am* 2012;92:441-54.
11. Trauma Association of Canada. Trauma system accredidation guidelines. Available: www.traumacanada.org/Resources/Documents/accreditation/Accreditation_Guidelines_2011.pdf (accessed 2016 Dec. 14).
12. American Association for the Advancement of Automotive Medicine. Abbreviated Injury Scale. Chicago (IL): AAAM Publications; 1990.
13. Daly L, Bourke GJ. Interpretation and uses of medical statistics. Oxford (UK): Blackwell Science; 2008.
14. Moore L, Stelfox HT, Turgeon AF, et al. Hospital length of stay after admission for traumatic injury in Canada: a multicenter cohort study. *Ann Surg* 2014;260:179-87.
15. Moore L, Lavoie A, Turgeon AF, et al. The trauma risk adjustment model: a new model for evaluating trauma care. *Ann Surg* 2009;249:1040-6.
16. Moore L, Stelfox HT, Turgeon AF, et al. Derivation and validation of a quality indicator of acute care length of stay to evaluate trauma care. *Ann Surg* 2014;260:1121-7.
17. Austin PC, Tu JV. Bootstrap methods for developing predictive models. *Am Stat* 2004;58:131-7.
18. 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.
19. American Association for the Advancement of Automotive Medicine. Convert between different AIS versions. Available: www.aaam.org/abbreviated-injury-scale-ais/ais-conversion-tool/ (accessed 2016 Dec. 14).
20. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
21. Cox DR. 2 further applications of a model for binary regression. *Biometrika* 1958;45:562-5.
22. Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691-2.
23. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2 ed. Hoboken (NJ): John Wiley & Sons, Inc; 2000.
24. Agency for Healthcare Research and Quality. Measuring and benchmarking clinical performance; 2013. Available: www.ahrq.gov/professionals/prevention-chronic-care/improve/system/pfhandbook/mod7.html (accessed 2016 June 10).
25. Moore L, Lavoie A, LeSage N, et al. Multiple imputation of the Glasgow Coma Score. *J Trauma* 2005;59:698-704.
26. Little RJ, Rubin DB. *Statistical analysis with missing data*. 2nd ed. New York (NY): Wiley; 2002.
27. Moore L, Hanley JA, Turgeon AF, et al. A new method for evaluating trauma centre outcome performance: TRAM-adjusted mortality estimates. *Ann Surg* 2010;251:952-8.
28. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964-74.
29. Altman DG. *Practical statistics for medical research*. Altman DG, editor. London (UK): Chapman & Hall; 1991.
30. Moore L, Hanley JA, Turgeon AF, et al. Comparing regression-adjusted mortality to standardized mortality ratios for trauma center profiling. *J Emerg Trauma Shock* 2012;5:333-7.