The impact of interhospital transfer on mortality benchmarking at Level III and IV trauma centers: A step toward shared mortality attribution in a statewide system

Daniel N. Holena, MD, MSCE, Elinore J. Kaufman, MD, MSHP, Justin Hatchimonji, MD, Brian P. Smith, MD, Ruiying Xiong, MS, Thomas E. Wasser, PhD, MEd, M. Kit Delgado, MD, MS, Douglas J. Wiebe, PhD, Brendan G. Carr, MD, MSHP, *and* **Patrick M. Reilly, MD**, *Philadelphia, Pennsylvania*

AAST Continuing Medical Education Article

Accreditation Statement

This activity has been planned and implemented in accordance with the Essential Areas and Policies of the Accreditation Council for Continuing Medical Education through the joint providership of the American College of Surgeons and the American Association for the Surgery of Trauma. The American College Surgeons is accredited by the ACCME to provide continuing medical education for physicians.

AMA PRA Category 1 Credits™

The American College of Surgeons designates this journal-based CME activity for a maximum of 1 AMA PRA Category 1 Credit™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Of the AMA PRA Category 1 Credit™ listed above, a maximum of 1 credit meets the requirements for self-assessment.

Credits can only be claimed online

AMERICAN COLLEGE OF SURGEONS

Inspiring Quality: Highest Standards, Better Outcomes

100+years

Objectives

After reading the featured articles published in the *Journal of Trauma and Acute Care Surgery*, participants should be able to demonstrate increased understanding of the material specific to the article. Objectives for each article are featured at the beginning of each article and online. Test questions are at the end of the article, with a critique and specific location in the article referencing the question topic.

Claiming Credit

To claim credit, please visit the AAST website at http://www.aast.org/ and click on the "e-Learning/MOC" tab. You must read the article, successfully complete the post-test and evaluation. Your CME certificate will be available immediately upon receiving a passing score of 75% or higher on the post-test. Post-tests receiving a score of below 75% will require a retake of the test to receive credit.

Disclosure Information

In accordance with the ACCME Accreditation Criteria, the American College of Surgeons, as the accredited provider of this journal activity, must ensure that anyone in a position to control the content of J Trauma Acute Care Surg articles selected for CME credit has disclosed all relevant financial relationships with any commercial interest. Disclosure forms are completed by the editorial staff, associate editors, reviewers, and all authors. The ACCME defines a `commercial interest' as "any entity producing, marketing, re-selling, or distributing health care goods or services consumed by, or used on, patients." "Relevant" financial relationships are those (in any amount) that may create a conflict of interest and occur within the 12'months preceding and during the time that the individual is engaged in writing the article. All reported conflicts are thoroughly managed in order to ensure any potential bias within the content is eliminated. However, if you'perceive a bias within the article, please report the circumstances on the evaluation form.

Please note we have advised the authors that it is their responsibility to disclose within the article if they are describing the use of a device, product, or drug that is not FDA approved or the off-label use of an approved device, product, or drug or unapproved usage.

Disclosures of Significant Relationships with Relevant Commercial Companies/Organizations by the Editorial Staff

Ernest E. Moore, Editor: PI, research support and shared U.S. patents Haemonetics; PI, research support, Instrumentation Laboratory, Inc.; Co-founder, Thrombo Therapeutics. Associate Editors David Hoyt, Ronald V. Maier and Steven Shackford have nothing to disclose. Editorial staff and Angela Sauaia have nothing to disclose.

Author Disclosures

Daniel Holena: Shareholder: Enteropass; M. Kit Delgado: Travel costs: United Health Group.

Reviewer Disclosures

The reviewers have nothing to disclose.

Cost

For AAST members and Journal of Trauma and Acute Care Surgery subscribers there is no charge to participate in this activity. For those who are not a member or subscriber, the cost for each credit is $25.

System Requirements

The system requirements are as follows: Adobe® Reader 7.0 or above installed; Internet Explorer® 7 and above; Firefox® 3.0 and above, Chrome® 8.0 and above, or Safari™ 4.0 and above.

Questions

If you have any questions, please contact AAST at 800-789-4006. Paper test and evaluations will not be accepted.

BACKGROUND: Many injured patients presenting to Level III/IV trauma centers will be transferred to Level I/II centers, but how these transfers influence benchmarking at Level III/IV centers has not been described. We hypothesized that the apparent observed to expected (O:E) mortality ratios at Level III/IV centers are influenced by the location at which mortality is measured in transferred patients.

METHODS: We conducted a retrospective study of adult patients presenting to Level III/IV trauma centers in Pennsylvania from 2008 to 2017. We used probabilistic matching to match patients transferred between centers. We used a risk-adjusted mortality model to estimate predicted mortality, which we compared with observed mortality at discharge from the Level III/IV center (O^1) or observed mortality at discharge from the Level III/IV center for nontransferred patients and at discharge from the Level I/II center for transferred patients (O^2).

RESULTS: In total, 9,477 patients presented to 11 Level III/IV trauma centers over the study period (90% white; 49% female; 97% blunt mechanism; median Injury Severity Score, 8; interquartile range, 4–10). Of these, 4,238 (44%) were transferred to Level I/II centers, of which 3,586 (85%) were able to be matched. Expected mortality in the overall cohort was 332 (3.8%). A total of 332 (3.8%) patients died, of which 177 (53%) died at the initial Level III/IV centers (O^1). Including posttransfer mortality for transferred patients in addition to observed mortality in nontransferred patients (O^2) resulted in worse apparent O:E ratios for all centers and significant differences in O:E ratios for the overall cohort (O^1:E, 0.53; 95% confidence interval, 0.45–0.61 vs. O^2:E, 1.00, 95% confidence interval, 0.92–1.11; $p < 0.001$).

CONCLUSION: Apparent O:E mortality ratios at Level III/IV centers are influenced by the timing of measurement. To provide fair and accurate benchmarking and identify opportunities across the continuum of the trauma system, a system of shared attribution for outcomes of transferred patients should be devised. (*J Trauma Acute Care Surg.* 2020;88: 42–50. Copyright © 2019 American Association for the Surgery of Trauma.)

KEY WORDS: Trauma outcomes; epidemiology; interhospital transfer.

The acute care of the injured patient spans a timeline from the point of injury to hospital discharge. Efforts to improve care for injured patients such as the American College of Surgeons Trauma Quality Improvement Program (ACS-TQIP) have to date primarily focused on the role of Level I and Level II trauma centers, but there has been less investigation into the role Levels III and IV centers play in patient outcomes. In largely rural states such as Pennsylvania, Levels III and IV centers may play an integral role in the early care of injured patients, with approximately 33% of patients arriving at Levels I and II hospitals after transfer from outside hospitals.

The fact that many patients treated at Levels III and IV centers will be ultimately transferred to Levels I and II centers has implications for typical risk-adjusted mortality modeling strategies, in which the outcome of interest is in-hospital mortality. Because expected survival until reaching the receiving center is a precondition of transfer in most cases, use of in-hospital mortality as a basis for risk-adjusted modeling at Levels III and IV centers may lead to observed to expected mortality ratios which may not reflect ultimate patient outcomes. For instance, a critically ill patient presenting to a Level III trauma center may have a high predicted probability of mortality based on age, physiology, and injury severity. If transferred to a Level I trauma center, from the perspective of the Level III center, this patient would represent an unexpected survivor since the outcome of interest is survival to hospital discharge despite the fact that such a patient might go on to die at the Level I trauma center. This phenomenon has implications for apparent observed-to-expected mortality rates at both centers.

In the recently published report on population-based trauma outcomes, the National Quality Forum recommended advancing models of shared attribution that promote improved planning and coordination within regional trauma networks in order to promote shared accountability across relevant stakeholders and accelerate quality improvement in trauma care.[1] With the long-term goal of developing a system of shared outcomes attribution between centers, in this study, we sought to understand the influence of interhospital transfer from Levels III and IV to Levels I and II centers on risk-adjusted mortality at Levels III and IV centers. Because the sickest patients tend to be transferred from Levels III and IV centers to Levels I and II centers, we hypothesized that risk-adjusted mortality for Levels III and IV centers (as measured by observed to expected mortality ratios) would appear worse when mortality for transferred patients was attributed to referring centers versus receiving centers.

MATERIALS AND METHODS

After approval from the University of Pennsylvania Institutional Review Board, we performed a retrospective analysis of the Pennsylvania Trauma Outcomes Study (PTOS) database from 2008 to 2017. The Pennsylvania Trauma Systems Foundation (PTSF) is the governing body that accredits all trauma centers in the Commonwealth of Pennsylvania. To maintain accreditation, Pennsylvania trauma centers must abstract, collect, and submit data on all trauma patients meeting inclusion criteria for the PTOS registry.

Submitted: June 8, 2019, Revised: July 17, 2019, Accepted: July 19, 2019, Published online: September 13, 2019.

From the Division of Traumatology, Surgical Critical Care and Emergency Surgery (D.N.H., E.J.K., J.H., B.P.S., P.M.R.), Center for Clinical Epidemiology and Biostatistics (D.N.H., J.H., R.X., M.K.D., D.J.W.), Perelman School of Medicine at the University of Pennsylvania, Philadelphia; Pennsylvania Trauma Systems Foundation (T.E.W.), Harrisburg; Department of Emergency Medicine (M.K.D.), Perelman School of Medicine at the University of Pennsylvania, Philadelphia; and Department of Emergency Medicine (B.G.C.), Kimmel School of Medicine at the Thomas Jefferson University, Philadelphia, Pennsylvania.

Meetings at which this work will be presented: 78th Annual Meeting of AAST and Clinical Congress of Acute Care Surgery in Dallas, Texas.

Address for reprints: Daniel Holena, MD, MSCE, Department of Surgery, University of Pennsylvania School of Medicine, 51 N 39th St, 1st floor, MOB Bldg, Philadelphia, PA 19104; email: holenad@upenn.edu.

DOI: 10.1097/TA.0000000000002491

The population of interest in this study was all patients presenting to Levels III and IV trauma centers in Pennsylvania over the study period. The PTSF establishes standards for accrediting trauma centers based on a hospital's resources and capabilities based on the American College of Surgeons Committee on Trauma Resources for Optimal Care of the Injured Patient 2014[2] and publishes standards for Level I, Level II, Level III, and Level IV adult trauma centers, which are available online.[3,4] Patients younger than 18 years and patients with burns as a primary mechanism of injury were excluded from the study. Data were inspected for missingness, and multiple imputation was used to impute missing data found to be missing not completely at random. The PTOS registry, like the American College of Surgeon's National Trauma Data Bank, is an instance-based registry that does not contain a primary key variable that allows for identification of unique patients. This means that if an injured patient is transferred between two trauma centers, there will be two rows of data representing that patient in the data set but there is no existing variable to track this patient longitudinally. For this reason, we used probabilistic matching to identify patients transferred from Levels III and IV centers to Levels I and II centers so that their records could be linked. We matched on patient age, sex, birth dates, zip codes, and dates of admissions using the "dtalink" package[5] in Stata 15.0 (College Station, TX) using the "calcweight" option to generate recommended weights for matching variables. Because no gold standard data set exists against which to check the accuracy of our matching, we compared the characteristics of linked and unlinked transfers using standardized differences.[6]

The primary exposure of interest in this study was interhospital transfer between a Level III or Level IV center to a Level I or Level II center. The primary outcome of interest was center level observed to expected risk-adjusted in-hospital mortality. Because in transferred patients there are two opportunities to observe this outcome, we measured center level outcomes twice for transferred patients: O^1, mortality status at discharge from the Level III/IV center; and O^2, mortality status at discharge from the Level I/II center (see Fig. 1—conceptual diagram).

We based our risk-adjusted mortality model on the PTSF risk-adjusted mortality model[7] which is in turn based on the published description of the American College of Surgeons Trauma Quality Improvement risk-adjusted mortality model.[8] This model contains 17 variables including sex, age, mechanism of injury, maximum head Abbreviated Injury Scale (AIS), maximum lower extremity AIS, patient physiology (presenting Glasgow Coma Scale [GCS] motor scores, systolic blood pressure, and pulse rate), and patient comorbidities. Because of the long timescale of the study, we included time in years as a candidate variable in our model to account for changes in quality of care over time. We used our final multivariable model to generate an expected probability of mortality for each patient presenting to a Level III/IV center which were then summed by center to generate a center-level expected mortality (E). We then calculated the observed mortality for each Level III/IV center using two different methods. First, we calculated observed mortality at each Level III/IV center by summing observed in-hospital mortality at discharge from the Level III/IV center for both nontransferred and transferred patients (O^1). Next, we calculated observed mortality for each Level III/IV center by summing the observed mortality at discharge from the Level III/IV center for nontransferred patients and the observed mortality at discharge from the Level I/II center for transferred patients (O^2).

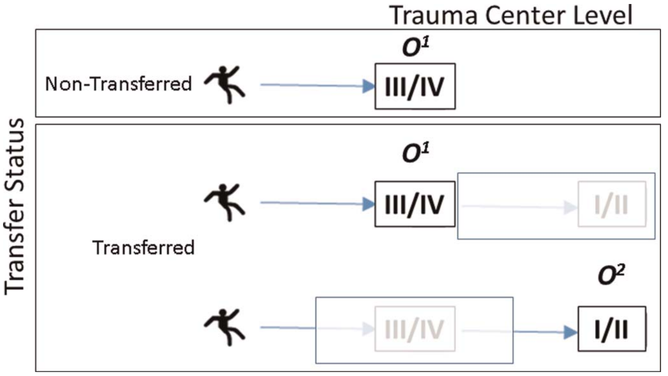

Figure 1. Conceptual diagram. Patients presenting to Level III/IV centers may receive definitive care at these centers or may be transferred to Level I/II centers. Apparent risk-adjusted outcomes at Level III/IV centers may be influenced by the transfer rate and the outcome measured: in-hospital mortality at the level III/IV center (O^1) versus in-hospital mortality at the Level I/II center (O^2).

To examine differences in demographic, injury, and physiologic variables between transferred and nontransferred patients presenting to Level III/IV centers, we used Mann-Whitney test to compare nonnormally distributed continuous variables and two-sample t-test to compare normally distributed continuous variables. Chi-squared test was used to compare categorical variables. The methodology underlying the development of our risk-adjusted mortality model is described elsewhere.[7] We used Stata v15.0 (College Station, TX) and R v3.5.2 for all statistical analyses.

RESULTS

In total, 9,477 patients presented to 11 Levels III and IV trauma centers over the study period, of whom 8,825 were included in the final analysis. A flow diagram of the study can be seen in Figure 2. Overall, the cohort was 90% white 49% female, 97% blunt mechanism, and had a median Injury Severity Score [ISS] of 8 (interquartile range [IQR], 4–10) (Table 1). Of these, 4,238 (45%) were transferred to Levels I and II centers. Using probabilistic matching, we were able to successfully match 3,586 (85%) of these patients. Examination of standardized differences (SD) between age, sex, injury mechanism, physiology, and injury severity of matched and unmatched transferred patients was generally reassuring, with only ISS demonstrating a moderate risk of bias (SD, −0.56), indicating that matched patients were, on average, more severely injured than unmatched patients.

Characteristics of patients presenting to Levels III and IV trauma centers stratified by transfer status can be seen in Table 1. Compared with patients who were not transferred, transferred patients were younger (age, 60 years; IQR, 39–77 years vs. 69, IQR, 50–83 years, $p < 0.001$), more likely to be male (59% vs. 46%, $p < 0.001$), and were more likely to be severely injured (ISS > 15) (18.1% vs. 13.0%, $p < 0.001$). Transferred patients were more likely to be head injured (52% vs. 27%, $p < 0.001$) and less likely to have extremity injuries (44% vs. 55%, $p < 0.001$).

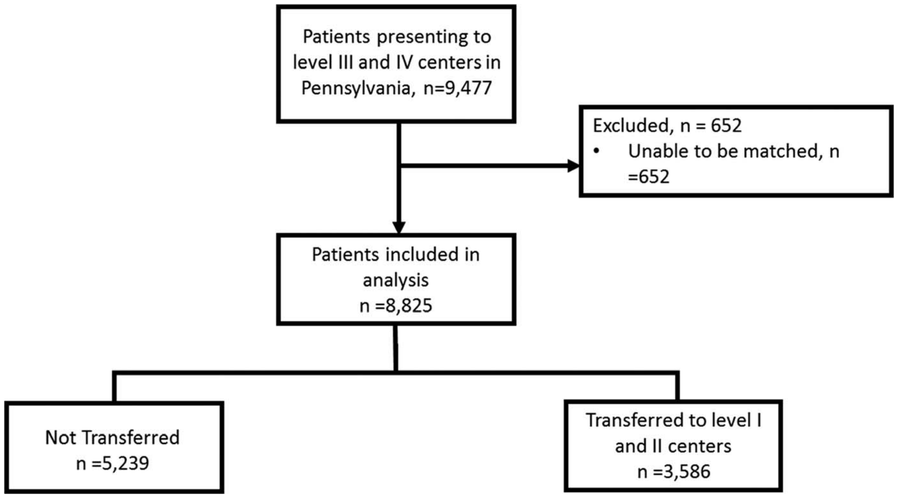

Figure 2. Flow diagram of patients included in the study.

Our risk-adjusted mortality model (Table 2) had good characteristics, with excellent discrimination (AUC, 0.93; 95% confidence interval [CI], 0.92–0.95) and calibration. The Hosmer-Lemeshow test was not significant ($p = 0.12$), indicating no significant difference between the observed and expected mortality and visual inspection of Hosmer-Lemeshow plots indicated good model calibration. As in prior work, the strongest predictors of mortality included low GCS motor scores, deranged systolic blood pressure and heart rate, gunshot wound as a mechanism of injury, and higher abbreviated injury scores. Based on our model, expected mortality in the overall cohort was 332 (3.8%).

A total of 332 (3.8%) patients died during the study period, of which 177 (53%) patients died at the initial Levels III and IV centers (Table 3). An additional 155 (47%) transferred patients died at the Levels I and II centers to which they were transferred. The attribution of these mortalities to the receiving vs. the referring center was resulted significant difference in observed to expected (O:E) ratios for the overall cohort (O^1:E, 0.53; 95% CI, 0.45–0.61 vs. O^2:E, 1.00; 95% CI, 0.92–1.11; $p < 0.001$). For each Level III and IV center, including posttransfer mortality for transferred patients in addition to observed mortality in nontransferred patients (O^2) resulted in higher O:E ratios, in some cases, shifting the apparent performance of a center from "as expected" mortality to a significant outlier for higher than expected mortality (Fig. 3).

DISCUSSION

In this retrospective cohort study, we describe the impact of the frame of reference from which outcomes are observed on the apparent performance of Levels III and IV centers. Because many of the sickest patients will be transferred to Level I and Level II centers for definitive management, and all patients are transferred alive, methodology that examines observed to expected mortality at the center level will be influenced by these 'unexpected survivors' in proportion to the number that are transferred out. The net effect of this is that the performance of Levels III and IV centers may appear better than expected after considering patient factors associated with mortality. From the perspective of the receiving Level I or Level II center, we have previously shown that these transfers may negatively influence apparent center performance, especially if patient physiology at the time of arrival to the Level I or Level II center is used for risk adjustment rather than physiology at the time of arrival at the Level III or Level IV center, and that this effect is proportional to the number of patients that such a center receives.[9]

Of the deaths that occur in patients presenting to Levels III and IV centers, we found that nearly half occurred after the patient had been transferred to a Level I or Level II center. This highlights the fact that in order for quality improvement efforts to be successful, they must span the continuum of care including interhospital transfer. It is incumbent upon referring centers to seek feedback on patients transferred, just as it is incumbent upon receiving centers to provide it. Such efforts have been shown to lead to improvements in the perception of care provided to injured patients, as well as the development of performance improvement efforts at referring centers.[10] As such, they are reflected in the American College of Surgeons Resources for the Optimal Care of the Injured Patient publication which states that receiving hospitals should have input from, provide feedback to, and maintain adequate communication with the personnel responsible for the transport process and the referring hospital to ensure that problems occurring during and associated with transport are addressed in a timely manner.[2]

This is the first study that we are aware of that examines benchmarking risk-adjusted mortality at Levels III and IV centers. To date, reports regarding benchmarking specific to non-Level I and II trauma centers have generally been descriptive and have not included risk-adjusted mortality models. In a single-state study in Ohio, one such study reported on the development of a benchmarking program for Level III centers. This effort included structural measures (emergency department volume, rurality), process measures (average surgeon response time, % of orthopedic fractures managed operatively within 24 hours), and outcome measures (raw mortality).[11] Another single-state study from Colorado described the development of performance improvement filters relevant to Levels III and IV trauma centers, including such process measures as adherence to trauma team activation criteria, response time for trauma team leaders, and time

TABLE 1. Demographics, Mechanism of Injury, Injury Severity, and Presenting Vital Signs Upon Trauma Center Presentation for Nontransferred and Transferred Patients

Patient Characteristics	Overall (N = 8,825)	Nontransferred (n = 5,239)	Transferred (n = 3,586)	*p*
Age (y)	65 (IQR, 46–81)	69 (IQR, 50–83)	60 (39–77)	<0.001
Male (%)	4,488 (50.9)	2,391 (45.6)	2,097 (58.5)	<0.001
Race	256 (2.9)	139 (2.7)	117 (3.3)	0.004
White	7,915 (89.7)	4,719 (90.1)	3,196 (89.1)	
Black	437 (5.0)	275 (5.2)	162 (4.5)	
Asian	68 (0.8)	32 (0.6)	36 (1.0)	
Other	147 (1.7)	72 (1.4)	75 (2.1)	
Blunt mechanism	8,565 (97.1)	5,090 (97.2)	3,475 (96.9)	0.534
ISS				<0.001
<9	5,829 (47.0%)	4,366 (49.6%)	1,463 (40.7%)	
9–15	4,620 (37.3%)	3,218 (36.5%)	1,402 (39.0%)	
16–25	1,168 (9.4%)	739 (8.4%)	429 (11.9%)	
>25	627 (5.1%)	403 (4.6%)	224 (6.2%)	
Missing	167 (1.4%)	86 (0.9%)	81 (2.3%)	
Maximum AIS score, head				<0.001
0	5,512 (62.5)	3,796 (72.5)	1,716 (47.9)	
1	1,003 (11.4)	640 (12.2)	363 (10.1)	
2	886 (10.0)	556 (10.6)	330 (9.2)	
3	655 (7.4)	45 (0.9)	610 (17.0)	
4	265 (3.0)	19 (0.4)	246 (6.9)	
5	185 (2.1)	25 (0.5)	160 (4.5)	
6	1 (0.0)	1 (0.0)	0 (0.0)	
Missing	318 (3.6)	157 (3.0)	161 (4.5)	
Maximum AIS score, lower extremity				<0.001
0	4,745 (53.8)	2,374 (45.3)	2,371 (66.1)	
1	1,395 (15.8)	825 (15.7)	570 (15.9)	
2	1,555 (17.6)	1,204 (23.0)	351 (9.8)	
3	1,096 (12.4)	813 (15.5)	283 (7.9)	
4	5 (0.1)	3 (0.1)	2 (0.1)	
5	1 (0.0)	1 (0.0)	0 (0.0)	
Missing	28 (0.3)	19 (0.4)	9 (0.3)	
Systolic blood pressure (mm Hg)				0.221
0 to 40	3,394 (38.5)	2,008 (38.3)	1,386 (38.7)	
41 to 90	4,478 (50.7)	2,688 (51.3)	1,790 (49.9)	
91 to 120	521 (5.9)	287 (5.5)	234 (6.5)	
121 to 160	91 (1.0)	50 (1.0)	41 (1.1)	
> 160	341 (3.9)	206 (3.9)	135 (3.8)	
Pulse (beats per minute)				0.001
0 to 50	1,812 (20.5)	1,005 (19.2)	807 (22.5)	
51 to 100	941 (10.7)	562 (10.7)	379 (10.6)	
101 to 150	6,072 (68.8)	3,672 (70.1)	2,400 (66.9)	
GCS motor score				<0.001
6	7,789 (88.3)	4,663 (89.0)	3,126 (87.2)	
1	188 (2.1)	122 (2.3)	66 (1.8)	
2	14 (0.2)	4 (0.1)	10 (0.3)	
3	13 (0.1)	3 (0.1)	10 (0.3)	
4	75 (0.8)	28 (0.5)	47 (1.3)	
5	209 (2.4)	101 (1.9)	108 (3.0)	
Missing	537 (6.1)	318 (6.1)	219 (6.1)	

Data for nonparametric continuous variables expressed as median (IQR); parametric continuous variables expressed as mean (standard deviation); categorical values expressed as n (%). *p* values are for Mann-Whitney test for nonparametric continuous variables, *t* test for parametric continuous variables, and χ^2 test for categorical variables.

BP, blood pressure.

TABLE 2. Results of the Final PTOS-RAM Model to Estimate the Risk of Death Based on Patient Demographic and Clinical Characteristics

Patient Characteristics	OR	95% CI	
Age (y)	1.04	1.03	1.05
GCS motor			
6	reference		
1	113.67	70.61	186.94
2	8.85	1.96	36.94
3	19.27	4.51	81
4	4.82	2.17	10.1
5	3.52	2	5.97
Systolic blood pressure (mm Hg)			
91 to 121	reference		
0 to 40	2.07	1.08	4.28
41 to 90	1.26	0.67	2.61
121 to 160	3.59	1.11	10.8
>160	6.68	3.14	15.08
Pulse (beats per minute)			
51 to 100	reference		
0 to 50	3.79	2.24	6.64
101 to 150	1.33	0.81	2.28
Injury mechanism			
Fall	reference		
Pedestrian struck	4.47	1.61	10.87
MVC	2.26	1.44	3.52
MCC	3.16	1.28	7.24
GSW	6.07	2.3	15.9
Assault	0.87	0.35	1.88
Stab	1.01	0.22	4.03
Other	1.49	1	2.19
Maximum AIS score, head			
0	reference		
1	0.81	0.47	1.35
2	0.79	0.42	1.41
3	0.82	0.48	1.36
4	0.84	0.43	1.63
5	0.89	0.44	1.8
6	13,131.57	0	
Maximum AIS score, lower extremity			
0	reference		
1	0.74	0.49	1.11
2	0.93	0.52	1.62
3	0.76	0.46	1.25
4	3.11	0.11	35.61
5	161,841.00	0	
Maximum AIS score			
1	reference		
2	0.63	0.35	1.17
3	1.61	0.93	2.86
4	3.41	1.72	6.77
5	13.12	6.31	27.49
6	24.09	2.59	239.52
Heart disease	1.86	1.32	2.6
Cancer	2.09	0.78	4.75
Liver disease	2.35	0.96	5.2
Hypertension	0.87	0.62	1.21
Dialysis	3.41	1.05	9.02
Impaired sensorium	1.15	0.68	1.88
Functional dependence	0.92	0.44	1.74
Bleeding disorder	2.92	1.17	6.46
Peripheral vascular disease	1.57	0.46	4.16
Year of study	0.95	0.90	1.01

Continued next page

TABLE 2. *(Continued)*

MVC, motor vehicle collision; MCC, motorcycle collision; GSW, Gunshot wound.

to intubation for patients with GCS scores < 9.[12] Such efforts are invaluable to facilitate sharing of best practices and ongoing education of participants but are limited in their ability to provide meaningful feedback on risk-adjusted mortality between centers. Knowledge surrounding risk-adjusted outcomes at Level III centers will hopefully be improved by the recent stipulation from the American College of Surgeons Committee on Trauma requiring participation in a risk-adjusted mortality benchmarking program for Level III trauma center verification.

The way in which outcomes of transferred patients are optimally attributed in trauma populations has yet to be discovered, but analogous work exists in other emergency care sensitive conditions.[13] For instance, most hospital-based performance metrics surrounding acute myocardial infarction attribute the outcome of transferred patients to the referring hospital,[14] and correspondingly, the outcomes for transferred patients are excluded when calculating performance metrics for receiving centers. This methodology has been critiqued, however, because it may paint an incomplete picture of the quality of care for transferred patients and because apparent performance of the receiving center will change based on whether or not outcomes of transferred patients are attributed to receiving centers or excluded.[15] Regardless of the center to which the outcome is attributed, systems in which outcomes are attributed solely to either the referring or to the receiving center do not adequately capture the relative contributions of these two centers to the patient's care. Methods of shared attribution, such as using marginal structural models to treat hospital quality as a time-varying covariate[16] have been proposed but have not been widely adopted.

While defining a method of shared attribution between two hospitals is challenging, defining a system of shared attribution that spans the continuum of care is more challenging still. Since the prehospital system, acute care hospitals, and postacute care settings all contribute to patient outcomes, an ideal methodology would allow stakeholders to understand how they perform relative to peers while accounting for the variability present in other phases of care. While such methodology could prove invaluable in efforts to improve performance, another approach that has been suggested is the population health-based concept of "coopetition."[17] Under this rubric, patterns of hospital use for emergency care-sensitive conditions (including but not limited to trauma) are used to define emergency care service regions which are then benchmarked referent to other emergency care service regions. Because under this system hospitals would be incentivized and penalized at the level of the emergency care

TABLE 3. Observed Mortality, Expected Mortality and Apparent O:E Ratios When Using Patient Status at Discharge From the Level III/IV Center (O^1) for Both Transferred and Nontransferred Patients or Patient Status at Discharge From the Level III/IV Center for nontransferred Patients and Patient Status at Discharge From the Level I/II Center (O^2) for Transferred Patients

Center	n	O^1	O^2	E	O^1:E	95% CI	O^2:E	95% CI	ΔO:E	95% CI
1	5256	111	204	205.4	0.54	0.44–0.64)	0.99	0.89–1.13	0.45	0.36–0.54
2	393	5	12	11.5	0.43	0.05–0.82)	1.04	0.66–1.63	0.61	0.16–1.06
3	294	8	13	12.2	0.65	0.20–1.11)	1.06	0.61–1.64	0.41	0.05–0.77
4	205	3	9	8.4	0.36	−0.05 to 0.76	1.07	0.67–1.77	0.72	0.14–1.29
5	322	5	12	6.6	0.76	0.09–1.42)	1.82	1.16–2.85	1.06	0.28–1.85
6	874	18	30	37.7	0.48	0.26–0.70)	0.80	0.58–1.08	0.32	0.14–0.50
7	598	12	23	23.2	0.52	0.22–0.81	0.99	0.70–1.40	0.47	0.19–0.76
8	50	3	4	3.0	0.99	−0.13 to 2.10	1.31	0.20–2.60	0.33	−0.32 to 0.97
9	33	0	1	1.0	0.00	0.00–0.00	1.00	1.00–2.96	1.00	−0.96 to 2.96
10	654	9	19	18.1	0.50	0.17–0.82	1.05	0.72–1.52	0.55	0.21–0.89
11	146	3	5	4.9	0.62	−0.08 to 1.31	1.03	0.33–1.92	0.41	−0.16 to 0.98
Total	8,825	177	332	332.0	0.53	0.45–0.61	1.00	0.92–1.11	0.47	0.39–0.54

service region rather than the individual center, this approach would in theory foster collaboration between stakeholders that may currently be inhibited by local market forces.

Our study has limitations and suffers from the inherent weaknesses of a retrospective investigation. We chose to use a statewide data set, and so our findings may not be generalizable to other regions with dissimilar rurality. However, national instance-based registries, such as the NTDB, do not contain unique identifiers for individual patients which makes direct tracking of patients transferred between centers impossible. Probabilistic matching could in theory be employed to identify patients transferred between centers, but only if both the referring and receiving centers contribute NTDB data. Given the absence of geographic identifiers and the lack of uniform reporting across the United States, this approach is unlikely to be effective at this time but may become more feasible as more centers contribute data. Alternatively, many statewide data sets, such as the statewide inpatient data set and the statewide emergency data set, contain unique patient identifiers that allow for direct tracking of individual patients between centers, but as administrative data sets do not contain physiologic data which is necessary for valid risk-adjustment in trauma populations. The PTOS overcomes these challenges because all trauma centers in Pennsylvania are required to contribute data, including physiology on arrival. We cannot discern from the data set the motivation for transferring patients from Level III/IV centers to Level I/II but note that this flow of patients is consistent with how a tiered trauma system should ideally function, with patients who will likely need the highest levels of care flowing to those centers with the greatest resources. We also were only able to match approximately 85% of transferred patients. Use of probabilistic matching must balance the competing priorities of complete matching vs. accurate matching. As constraints for matching are relaxed, the chance of incorrect matching rises which may limit the validity of findings, whereas tighter constraints result in more accurate matching at the expense of number of cases matched. Because

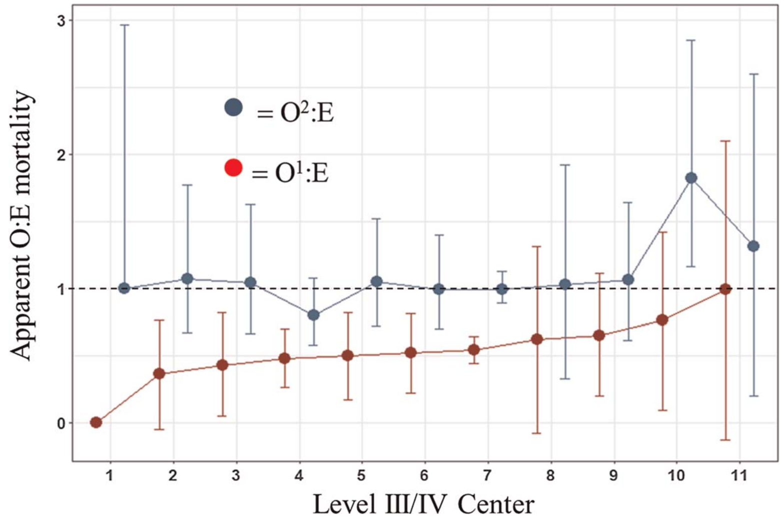

Figure 3. Caterpillar plots of observed to expected mortality at Level II and IV trauma centers when mortality of transferred patients is attributed to the receiving Level I/II center (O^1) or to transferring Level III/IV center (O^2).

of the relatively low number of unmatched patients and the generally low standardized differences between patient characteristics in the matched versus unmatched sets of transferred patients, we believe the risk of bias introduced by our unmatched transferred patients is not likely to substantially influence the results of our study.

CONCLUSION

This is the first study that we are aware of in trauma populations to consider observed and actual mortality stemming from presentation to Level III and Level IV trauma patients. The apparent O:E mortality ratio at Levels III and IV centers is influenced by the time at which mortality is measured in transferred patients. To provide fair and accurate benchmarking for Levels III and IV centers and identify opportunities across the continuum of the acute hospital setting, a system of shared attribution for outcomes of transferred patients should be devised.

DISCLOSURE

The authors declare no conflicts of interest. Portions of this work were supported by a training grant through the National Heart, Lung, and Blood Institute. (DNH; K08HL131995).

REFERENCES

1. Population-based trauma outcomes. National Quality Forum, 2019 5/22/2019. Report No.
2. Resources for the optimal care of the injured patient 2014 [updated 5/1/2019]. 6th:[Available from: https://www.facs.org/~/media/files/quality%20programs/trauma/vrc%20resources/resources%20for%20optimal%20care.ashx.
3. Standards of accreditation, adult level IV: Pennsylvania Trauma Systems Foundation; 2019 [updated 04/01/2019]. Available from: http://digital.graphcompubs.com/publication/?i=267434&pp=1#{%22issue_id%22:267434,%22numpages%22:%221%22,%22page%22:1}.
4. Standards of accreditation, adult levels I-III: Pennsylvania Trauma Systems Foundation; 2019 [updated 04/01/2019]. Available from: http://digital.graphcompubs.com/publication/?i=267432&pp=1#{%22issue_id%22:267432,%22numpages%22:%221%22,%22page%22:1}.
5. Kranker K. DTALINK: Stata module to implement probabilistic record linkage: Boston College Department of economics; 2019. Available from: https://EconPapers.repec.org/RePEc:boc:bocode:s458504.
6. Harron KL, Doidge JC, Knight HE, Gilbert RE, Goldstein H, Cromwell DA, van der Meulen JH. A guide to evaluating linkage quality for the analysis of linked data. *Int J Epidemiol*. 2017;46(5):1699–1710.
7. Wiebe DJ, Holena DN, Delgado MK, McWilliams N, Altenburg J, Carr BG. The Pennsylvania Trauma Outcomes Study risk-adjusted mortality model: results of a statewide benchmarking program. *Am Surg*. 2017;83(5):445–452.
8. Newgard CD, Fildes JJ, Wu L, et al. Methodology and analytic rationale for the American College of Surgeons Trauma quality improvement program. *J Am Coll Surg*. 2013;216(1):147–157.
9. Holena DN, Wiebe DJ, Carr BG, Hsu JY, Sperry JL, Peitzman AB, Reilly PM. Lead-time bias and interhospital transfer after injury: trauma center admission vital signs underpredict mortality in transferred trauma patients. *J Am Coll Surg*. 2017;224(3):255–263.
10. Byrnes MC, Irwin E, Becker L, Thorson M, Beilman G, Horst P, Croston K. A trauma outreach program provided by a level I trauma center is an effective way to initiate peer review at referring hospitals and foster process improvements. *J Trauma*. 2010;68(4):778–782.
11. Lang CL, Simon D, Kilgore J. A statewide collaboration: Ohio Level III Trauma Centers' approach to the development of a benchmarking system. *J Trauma Nurs*. 2016;23(6):376–379.
12. Coniglio R, McGraw C, Archuleta M, et al. A multicenter performance improvement program uses rural trauma filters for benchmarking: an evaluation of the findings. *J Trauma Nurs*. 2018;25(2):139–145.
13. Carr BG, Branas CC, Metlay JP, Sullivan AF, Camargo CA. Access to emergency care in the United States. *Ann Emerg Med*. 2009;54(2):261–269.
14. Barbash IJ, Zhang H, Angus DC, Reis SE, Chang CCH, Pike FR, Kahn JM. Differences in hospital risk-standardized mortality rates for acute myocardial infarction when assessed using transferred and nontransferred patients. *Med Care*. 2017;55(5):476–482.
15. Kosseim MN, Mayo NE, Scott S, Hanley JA, Brophy J, Gagnon B, Pilote L. Ranking hospitals according to acute myocardial infarction mortality. 2006; 44(7):664–670.
16. Sato T, Matsuyama Y. Marginal structural models as a tool for standardization. *Epidemiology*. 2003;14(6):680–686.
17. Carr BG, Kilaru AS, Karp DN, Delgado MK, Wiebe DJ. Quality through coopetition: an empiric approach to measure population outcomes for emergency care-sensitive conditions. *Ann Emerg Med*. 2018;72(3):237–245.

DISCUSSION

PETER FISCHER, M.D., M.Sc. (Memphis, Tennessee): I would like to congratulate Dr. Holena and colleagues on a well-written paper and excellent presentation.

The authors' goal was to determine if observed-to-expected ratios of mortality of Level III or IV centers would change if in this calculation, the ultimate mortality of the transferred patients was factored in.

This is a well-designed study that gets to an essential theme of all trauma patients, something that I think we all kind of struggle with, is alive at discharge really the model that we should be judging ourselves? I do have several questions.

A key part of your methods is a probabilistic matching of these patients after transfer from the Level III/IV center to the Level I/II center. Do you have any idea how good of a match this is, and could this same matching technique be used on a larger, national scale?

As expected, the performance of the lower level centers dropped when mortality of the transferred patients was included, but is this truly a marker of how well-functioning the Level III center is?

Maybe the Level III held onto the patient way too long and transferred the patients at death's doorstep. That would be a low-performing center.

But maybe they had a horrible injury and they did an amazing job and transferred a patient who probably was never supposed to survive anyway, and actually they're a high-performing center, and would this penalize them?

From a timing perspective, is there a cutoff point, where likely the Level III/IV center performance is not truly a factor and should not be calculated into that observed-to-expected ratio?

In other words, if a patient gets to the Level II center or the Level I center and stays for 120 days and ultimately dies, is the four hours that they spent at the Level III or IV center really a factor?

I really like the concept behind this study, because it's getting to what I think we all really want, which is trauma system performance, and what really are your true next steps with this to really identify and benchmark a trauma system?

I'd like to thank the AAST for the privilege of the podium

MARK L. GESTRING, M.D. (Rochester, New York): At least in our region, where we have a large coverage area, transfer patients tend to self-select.

In most cases, the ones who arrive from distant centers tend to be the ones that don't die. They might have severe injuries, but the ones that were truly going to die of something that we could fix in minutes, would not have survived the transfer process.

So my question is, of the ones who make it to the Level I trauma center that died, or Level I or Level II center that died, do you know what those injuries were?

Are these predominantly brain injury patients, or do you have an idea of what the actual mechanism of injury and possible cause of death was. Thank you.

DANIEL N. HOLENA, M.D. (Philadelphia, Pennsylvania): Thank you, Dr. Fischer and the rest of you, for your excellent questions.

The first question by Dr. Fischer was how good is the match. And with most of these probabilistic matching programs you can sort of set a threshold at which point you would say, you know, "This match is unlikely to have occurred by chance alone," and we can say with X-percentage probability that these two patients are in fact the same patient, and with this product, I believe the threshold that we set was 85 percent.

The second question was, could this be used for national models. And the answer to that is, conceptually, yes, but pragmatically, probably not, and that has to do with two reasons.

The first is that many national datasets don't actually contain variables that would be important for matching, like dates of admission and transfer. And the second is that the reason we were able to do this with our dataset is because our dataset contains the known universe of all trauma centers in the state; whereas national datasets like the NTDB and the TQIP dataset, they're voluntary datasets, and so, if you didn't match a patient, you could not be sure that it was because that patient simply was transferred to a center that's not in the dataset, or it could be an issue with your probabilistic matching, so that makes it quite a bit more challenging.

The third question had to do with timing perspective cutoff. If a patient was at a Level I or a Level II center for 120 days versus say six hours at a Level III or IV is it reasonable to attribute some of the mortality to the Level III or IV centers or whether you should just chalk it all up to the Level I or II. I think that's an interesting point.

I do think that, because trauma is a time-sensitive condition, it's hard to imagine a situation in which the Level III or IV center, the first center that gets the patient doesn't have some responsibility for the ultimate outcome; but where you draw that line, I think, is a very thorny issue.

The fourth question from Dr. Fischer is are we going to plan to continue to do future work on this, and the answer is a definitive yes.

I think that there's a lot of interesting work that you can do looking at trauma systems and regions in this; but if you think about it, most of us probably work in groups of providers.

And these same issues sort of are brought to bear on those groups as well where there may be an outcome and there are many providers who touch the patient and how do you attribute that outcome in a model where it's group care?

With respect to Dr. Gestring's question regarding, you know, what are these patients dying of? We really can't say for sure, although we do note that there is a higher incidence of head injuries in the group that is transferred to the Level I and II centers.

So if I had to guess, I would say that it's likely that many of these patients are patients with bad traumatic brain injuries that don't die immediately but will die in the days and weeks to follow the transfer.

Dr. Winchell, with respect to your comment about whether or not this is looking at comparing the transferring and the receiving center, I agree. There is some element of what could be perceived as gamesmanship in here.

But the true point of this, at least this work, is to sort of benchmark Level III and Level IV centers referent to each other, so it's not really about comparing the Level Is and the IIs versus the IIIs and the IVs. This is an effort to sort of provide accurate benchmarking for the Level III and Level IV centers.

Dr. Mains, you had said, a question about whether or not these patients were transported by ground or by air, and we don't have that information in the paper, but it's certainly something that we could pull out.

And finally, Dr. Harrington, you had mentioned about the appropriateness of transfer, and that's really, I think, one of the key limitations in this.

We're using a large dataset, and it's very difficult to know what was going through the minds of the providers when they said it's time to transfer this patient.

And so, we can't say whether these are significantly over-triage or under-triage – we just don't have that information. And I think if you really want to know the answers to those questions, you'd probably have to do a bit more in the way of prospective data collection.

Thank you very much.