Article:

Cook, A. orcid.org/0000-0003-0301-3125, Osler, T., Glance, L. et al. (8 more authors) (2018) Comparison of two prognostic models in trauma outcome. British Journal of Surgery, 105 (5). pp. 513-519. ISSN 0007-1323

https://doi.org/10.1002/bjs.10764

This is the peer reviewed version of the following article: Cook, A., Osler, T., Glance, L., Lecky, F., Bouamra, O., Weddle, J., Gross, B., Ward, J., Moore, F. O., Rogers, F. and Hosmer, D. (2018), Comparison of two prognostic models in trauma outcome. Br J Surg., which has been published in final form at https://doi.org/10.1002/bjs.10764. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

COMPARISON OF ISS-BASED INJURY SEVERITY COMPONENT OF THE TARN MODEL VERSUS TMPM

Alan Cook, MD, FACS[1]; Turner Osler, MD, MSc, FACS[2]; Laurent Glance, MD[3]; Fiona Lecky, MB, Ch B, FRCS, DA, MSc, PhD, FCEM[4]; Omar Bouamra, PhD[5]; Jo Weddle, MD[6]; Brian Gross, BS[7]; Jeanette Ward, MS-CR[1], Forrest O. Moore, III, MD[1]; Frederick Rogers, MD[7;] David Hosmer, PhD[8]

[1]Chandler Regional Medical Center, USA
[2]University of Vermont, USA
[3]University of Rochester, USA
[4]University of Sheffield, UK
[5]University of Manchester, UK
[6]Baylor University Medical Center, USA
[7]Lancaster General Hospital, USA
[8]University of Massachusetts, USA

Corresponding Author:
Alan Cook, MD, FACS
Chandler Regional Medical Center
485 S. Dobson Rd.
Suite 201
Chandler, AZ 85224
903-941-0229
Fax: 480-728-4747
adcookmd@gmail.com

Abstract

Background Outcome assessment is mandated by the American College of Surgeons and the Royal College of Surgeons of England. The Trauma Audit and Research Network (TARN) in the United Kingdom (UK) publicly reports hospital performance. The TARN risk adjustment model uses a fractional polynomial transformation of the Injury Severity Score (ISS) as the measure of anatomic injury severity. The Trauma Mortality Prediction Model (TMPM) has been shown superior to ISS. We compared the anatomic injury components of the TARN model to TMPM.

Methods Data are from the National Trauma Data Bank for 2011-2015. Probability of death was estimated for the TARN fractional polynomial transformation of ISS and compared to TMPM. The coefficients for the models were estimated using 80% of the data set, randomly selected. The remaining 20% of the data were used for model validation. TMPM and TARN were compared using calibration curves, measures of discrimination (area under ROC curves (AUROC)), proximity to the true model (Akaike Information Criterion (AIC)) and goodness of model fit (Hosmer-Lemeshow test (HL)). **Results**: N=438 058. There were no characteristic differences between derivation and validation groups. TMPM demonstrated preferable AUROC (0.882 for TMPM vs 0.845 for TARN) , AIC (18 204 vs 12 143), and better fit to the data (32.4 vs. 153.0) compared to TARN.

Conclusions: TMPM had greater discrimination, proximity to the true model and goodness-of-fit than the anatomic injury component of TARN, TMPM should be considered for the injury severity measure for the comparative assessment of trauma centers in the UK.

Keywords: Trauma, Injury Severity, Mortality Prediction, Anatomic Injury

Background

The need for valid and reliable performance measures is vital as payers, patients and accrediting bodies compare outcomes among hospitals or physicians. For trauma, outcome assessment is mandated by the American College of Surgeons (ACS) and the National Health Service of England (NHS). [1,2] The Trauma Audit and Research Network (TARN) in the United Kingdom (UK) and the Trauma Quality Improvement Program (TQIP) in the United States (US), each calculate a patient's baseline predicted probability of dying from their traumatic injuries.[3,4] Central to the predicted probability of death is the extent of the anatomic injury from the traumatic event.

Recently, the Centers for Medicare and Medicaid Services (CMS), the largest healthcare payer in the US, adopted the Merit-based Incentive Payment System. This program determines payments, based in part, on healthcare quality measures, including data submission to qualified clinical data registries.[5] Similarly, beginning in 2012, the NHS requires all major trauma centers submit data to TARN within 25 days following a patient's discharge from the hospital as part of the National Tariff Payment System.[6,7] In both the UK and US, the largest healthcare payers are the federal governments. The NHS incentivizes providers and hospitals to, "… adequately reimburse care that is high quality and cost effective."[8] In the US, the Hospital Value-Based Purchasing (VBP) program was incorporated as part of the Affordable Care Act.[9] The VBP provides incentive payments for meeting performance standards.[9,10] Injury severity plays an important role in healthcare finance because injury severity measurement adjusts for differences in patient case mix across hospitals caring for trauma patients. For example, the NHS best practices for trauma, anatomic injury severity differentiates two levels of payment by anatomic injury severity, ISS more than 8 and ISS 16 or greater.[8] While CMS has outlined several measures specific to trauma care to be reported in a risk-stratified manner.[11]

Over the past 40 years, several methods of injury severity measurement have been proposed.[12-16] However, the first such method, the Injury Severity Score (ISS)[17], remains the most widely applied injury severity measure. The ISS is familiar to clinicians and researchers, which is probably the greatest contributor to its longevity. The ISS has functioned as a stand-alone severity measure, and has also been incorporated into the leading models of trauma mortality, such as [18] TARN. Although it has remained the leading measure of injury severity since it was introduced, ISS has four major limitations. First, the ISS is based on the Abbreviated Injury Scale (AIS) severity values which are based on expert consensus, as opposed to being empirically derived. Second, ISS only accommodates the three worst injuries from three separate body regions. As such, it cannot account for two or more serious injuries in the same body region. Third, many hospitals do not assign AIS codes to their patients' injuries de novo. Instead, they convert International Classification of Disease, 9th Revision, Clinical Modification (ICD-9) codes to AIS codes, thus reducing the accuracy of injury descriptions.[19] Finally, Kilgo et al., described the ISS as "choppy" due to sharp increases or decreases with respect to mortality over incremental increases in ISS values.[20] Of note, TARN uses a mathematical transformation of ISS to achieve better performance of their model.

The Trauma Mortality Prediction Models (TMPM) was developed as an empirically based alternative to ISS without the limitations of ISS. TMPM incorporates the patient's five worst injuries as predictors of a patient's probability of death using a logistic regression model. Prior work comparing TMPM to ISS and ISS-based injury severity models showed that TMPM was better able to predict mortality.[21-23] The current analysis, expands on previous work by comparing TMPM to the anatomic injury elements of the prediction model used in TARN. We hypothesized the TMPM would better predict survivors from fatalities given the limitations of the ISS.

Methods

Data Source

After obtaining approval from the institutional review board of Chandler Regional Medical Center, data from the National Trauma Data Bank (NTDB) for the years 2011 to 2015 were used in a retrospective cohort study. The NTDB contains data for nearly 7 million patient visits in more than 900 hospitals. The NTDB data are anonymized (i.e., all identifying information has been removed) to ensure confidentiality for patients, physicians, and hospitals participating in the NTDB.[24] The outcome of interest was in-hospital death.

Patients were excluded if they were younger than 18 years old, were burn victims, their discharge status (alive or dead) was unknown, they were dead on arrival, or were missing ISS values, or AIS codes. We also excluded patients for whom the TMPM probability of death could not be calculated. The reliability and quality of injury documentation in the AIS lexicon is essential to this study. To this end, AIS codes were taken from the RDS_AISPCODE files as these were submitted by each hospital and least likely to contain codes mapped from ICDMAP-90. Two additional exclusion criteria were applied using methods previously described.[21] Patients were excluded if they were from hospitals that admitted less than 300 patients per year or that used fewer than 20% of the available AIS codes.

Comparisons of surviving patients with those who expired were performed using the χ^2 and Kruskal-Wallis test statistics, as appropriate. Summary measures are presented as means with 95% confidence intervals (95% CI), or as medians and interquartile ranges (IQR).

We compared the anatomic injury component of the TARN model to TMPM. TARN 2014 model (Ps14) is based upon the ISS transformed by applying fractional polynomials. However, given the TARN Ps14 model was developed for data defined by the TARN inclusion criteria[25], coefficients for the fractional polynomial transformation if ISS was computed de novo using NTDB data for this study. This model is as follows:

$$\begin{aligned}&\left[\mathrm{ISS}_1=\left(\log_e(\mathrm{ISS}/10)-0.19499\right),\ \mathrm{ISS}_2=\left(\log_e\left(\mathrm{ISS}/10\right)^2-0.03802\right)\right]\\&\text{probit death }\mathrm{ISS}_1\ \mathrm{ISS}_2=\beta_1\mathrm{ISS}_1+\beta_2\mathrm{ISS}_2+\text{constant}\\&=\mathrm{ISS}_1\times0.762669+\mathrm{ISS}_2\times0.2587654-2.052421\end{aligned}$$

TMPM uses the worst five anatomic injuries coded as categorical variables along with two interaction terms.[22] Coefficients for these seven TMPM terms for were computed for the study data. For the present study, the AIS lexicon was applied as the descriptor of anatomic injury. Of note however, the TMPM only uses the six digit "predot" code to define injuries and does not incorporate the AIS severity value in its computation of the probability of death. In contrast, the ISS metric is based exclusively on the expert consensus-based severity value.[17]

Using 80% of the data, the fractional polynomial transformation of the ISS was computed for the TARN model and coefficients for the TARN and TMPM models were estimated. The models were validated using the remaining 20% of the data set.

Measures of model performance included area under the receiver operating characteristic curve (AUROC) (C statistic) [26] and the Akaike information criterion (AIC).[27] The AUROC is a measure of sensitivity over 1-specificity. This measures the ability of a model to discriminate subjects having the outcome of interest, here mortality, from those who do not. As AUROC values approach 1.0 the model’s discrimination improves.[28] The AIC is a means to rank competing models and to estimate which model most closely approximates the hypothetical “true” model of the phenomenon at hand. Generally, the model with the lowest associated AIC value is preferred. The Hosmer-Lemeshow (HL) goodness-of-fit test is a means to assess how well the model describes the data under analysis and inform the plausibility of the inferences drawn from the model.[26] The HL test statistic was calculated for each score with eight degrees of freedom specified in the development sample and eight in the validation sample. Calibration curves were constructed for each model to assess monotonicity of the severity measures. Monotonicity describes incremental increases in the observed outcome, here mortality, as a consistent function of incremental increases in the injury severity values. Observed mortality was plotted versus expected mortality for TARN and TMPM. Of note, ISS has 44 unique values for predicted probability of mortality, so the anatomic injury of the TARN model also has 44 distinct values. The TMPM produced 50 596 unique predicted probability values in the derivation group and 174 974 such values in the validation group. Given this level of granularity in the TMPM, the probabilities were grouped into deciles and plotted against the observed mortality for each decile. The 95% CI for the AUROC, AIC and HL for each severity score was computed using 1 000 bootstrap samples of the dataset. There were no missing values for any model in the study. All statistical analysis was performed using Stata/MP version 14.2 (Stata Corporation, College Station, TX, USA).

Results

After applying the exclusion criteria, 438 058 patients were included in the study. (Figure 1.) A total of 15 930 died (3.6%, 95% CI 3.6%-3.7%) during their hospitalization. The mean age was 49.8 years old and the non-survivors were older (58.3, 58.0-58.7). White males represented the largest race-gender demographic group in the cohort (38.6%). Blunt mechanisms of injury were predominant (85.6%). Of these, falls were the majority (59.2%, 59.0%-59.4) followed by motor vehicle crashes, (25.7%, 25.5%-25.8%). Firearm-related injuries were associated with the highest mortality (8.4%, 7.8-8.9). The mean (SD) ISS and TMPM pDeath were 12 (6.9) and 0.036 (0.092), respectively. (Table 1.) The range of predicted probabilities of death were 1.34×10^{-6} - 1.0 and 0.002 – 0.666 for TMPM and TARN, respectively.

There were 350 325 (80%) and 87 733 (20%) patients in the derivation and validation group, respectively. There were no significant differences between the derivation and validation groups with respect of age, sex, race/ethnicity, mechanisms of injury, median ISS, lengths of stay in the hospital or ICU, and rates of ICU admission and death.

Areas under the ROC curves (AUROC) were used to compare discrimination between survivors and non-survivors for the anatomic injury scores. TMPM exhibited better discrimination compared to the TARN score in the derivation group (0.880, 0.880-0.880) and (0.841, 0.840-0.841) for TMPM and TARN, respectively. Similarly the AUROC was higher for TMPM in the validation group, (0.882, 0.882-0.882) versus (0.841, 0.840-0.841) for the fractional polynomial transformation of ISS used by TARN.

The Akaike information criterion was used to estimate each models' proximity to a theoretical and unknown "ideal" probability of mortality model. Given a group of prediction models, the best model is the one with the smallest AIC value. TMPM performed better in comparison to the fractional polynomial transformation of ISS used by TARN. Similarly, the HL test indicated TMPM demonstrated a lower value and thus, is a better fit to the data in the current study. (Table2.)

Calibration curves were plotted for TMPM and TARN. When the observed percent mortality is plotted over the proportion of mortality predicted by each of the 44 the discrete TARN ISS values, a non-monotonic pattern is observed in both the derivation and validation groups. It is noteworthy that the maximum probability of mortality predicted by the fractional polynomial transformation of ISS is approximately 0.6. This is likely due to the inherent limitations of ISS as the mortality among patients with ISS values of 75 in this study was 60%. (Figures 2 and 3)

Discussion

TMPM exhibited greater discrimination over a broader range of injury severity compared to the ISS-based anatomic injury component of TARN. Additionally, TMPM had the lower value for AIC indicating it more closely approximates a hypothetical "ideal" model. Moreover, the goodness-of-fit for TMPM surpassed that of TARN. In sum, the inherent limitations of the ISS which have been well documented in the literature[20,21,29], are not overcome by mathematical transformation. Illness severity in trauma patients is measured using the probability of injury-related mortality. Unlike TMPM in which injury severity is estimated empirically using the AIS so-called "predot" codes, the basis for the injury severity component of the TARN model is the AIS injury severity based on expert consensus. Robust measures of injury severity are essential for performance benchmarking in order to improve outcomes in injured patients, and for incentivizing higher quality care in the NHS National Tariff Payment System and the Merit-Based Incentive Payment System and VBP from CMS. Since anatomic injury is the fundamental element of trauma, the credible measurement of injury severity is essential to the evaluation of trauma care quality. In our study cohort, the AUROC was four percentage points greater for TMPM. Thus, TMPM accurately predicted mortality in 589 more patients than did the injury severity component of TARN. The cumulative weight of our findings suggest TARN should consider using TMPM to quantify injury severity instead of relying on an ISS-based severity measure.

Developed over 40 years ago by Baker, et al., ISS was adopted as the gold standard for measuring anatomic injury severity.[17] It is the sum of the squared AIS severity values for the worst injury in each of three separate body regions. AIS severity values are based on expert consensus, as opposed to being empirically derived. This algorithm results in 56 combinations of squared AIS severity scores that yield 44 possible unique ISS values. The ISS is relatively simple to calculate. This simplicity is very likely the key to its longevity. However, the ISS is also known for its non-monotonic nature due to steep peaks and valleys in mortality as the ISS value gets larger.[20] More importantly, trauma care has advanced considerably in the 43 years since its publication as evidenced by the limitation of ISS in predicting mortality in the most severely injured patients. The TARN mortality model applies a sophisticated polynomial transformation of the ISS as the measure of anatomic severity. This improved the performance of the TARN model was demonstrated by Boumara, et al. in 2015.[25] More recently, the TMPM was developed using an empiric regression-based approach to estimate injury severity, instead of relying on the expert-based estimates of injury severity used in ISS-based severity measures.

TMPM provides empirically derived probabilities of mortality using the AIS lexicon. Previous work comparing TMPM to ISS, maximum AIS score, New Injury Severity Score and the ICD-9-Based Injury Severity Score (ICISS) demonstrated that TMPM more accurately predicts trauma mortality.[21] Although somewhat technical in our approach, it is important to assess the adequacy of the models prior to interpreting their results.[30] Our current analysis is the first study to compare the anatomic component of TARN to TMPM and to find that TMPM performed best .

This study has several limitations. First, the NTDB is not population-based but is based on a self-selected group of trauma centers. Thus, our findings may not be generalizable to all trauma and non-trauma centers. However, a priori, there is no reason to believe our finding that empiric measures of injury severity out-perform measures of injury severity based on expert consensus would be limited to this sample of trauma patients. Second, the complete TARN model includes variables for age, Glasgow Coma Scale score and terms for the Charlson Comorbidity Index.[25,31,32]. Thus, the results presented in this paper represent the anatomic injury components only and no inference should be made regarding the predictive capabilities of the complete TARN model. However, we believe that the basis for accurate risk adjustment in trauma mandates accurate specification of the anatomic component of injury severity, before including measures of physiologic derangement and comorbidities.

Conclusion

In the era of ubiquitous benchmarking of providers and institutions, risk adjustment in trauma based on the most accurate measure of injury severity is essential. TMPM more accurately captures the severity of traumatic injury than the ISS-based injury severity component of the TARN risk-adjustment model. The comparison of trauma systems, including those of separate nations, is only possible using a common injury lexicon and should be based on the most accurate method to quantify injury severity. The replacement of ISS with TMPM for national benchmarking efforts should be considered in the United Kingdom.

Table 1. Characteristics of 438 058 patients*

	All Patients		Survivors		Non-Survivors	
Total, per category, n (%)	438 058		422 128	(96.4)	15 930	(3.6)
Age, mean (95% CI)	49.8	(49.8-49.9)	49.5	(49.4-49.6)	58.3	(58.0-58.7)
Sex, Male, n (%)	284 764	(65.0)	273 655	(96.1)	11 109	(3.9)
Race/Ethnicity						
White, non-Hispanic, n (%)	277 794	(63.4)	267 350	(96.2)	10 444	(3.8)
Black, non-Hispanic, n (%)	52 946	(12.1)	51 336	(97.0)	1 610	(2.9)
Hispanic/Latino, n (%)	66 976	(15.3)	65 004	(97.1)	1 972	(3.0)
Race/other, n (%)	40 342	(9.2)	38 438	(95.3)	1 904	(4.7)
Alcohol Present, n (%)	56 203	(12.8)	54 437	(96.9)	1 766	(3.1)
Mechanism of Injury						
Blunt, n (%)	268 605	(85.6)	258 574	(96.3)	10 031	(3.7)
Fall, n (%)	159 053	(59.2)	152 407	(95.8)	6 646	(4.2)
Motor vehicle crash, n (%)	68 956	(25.7)	66 217	(96.0)	2 739	(4.0)
Penetrating, n (%)	32 867	(10.5)	31 761	(96.6)	1 106	(3.4)
Firearm-related, n (%)	9 948	(30.3)	9 117	(91.7)	831	(8.4)
Other Mechanisms, n (%)	12 365	(3.9)	131 819	(96.5)	4 795	(3.5)
Payer Status at Discharge						
Medicare/Medicaid, n (%)	142 424	(32.5)	135 832	(95.4)	6 592	(4.6)
Private insurance, n (%)	118 565	(27.1)	115 632	(97.5)	2 942	(2.5)
Other insurance, n (%)	33 850	(7.7)	32 422	(95.8)	1 428	(4.2)
Uninsured, n (%)	77 266	(17.6)	74 608	(96.6)	2 658	(3.4)
Insurance status unknown, n (%)	65 953	(15.1)	63 643	(96.5)	2 310	(3.5)
Length of Stay, mean (SD)	6.5	(9.3)	6.5	(9.2)	6.9	(11.9)
ICU Admission, n (%)	138 837	(31.7)	125 697	(90.5)	13 140	(9.5)
ICU length of stay, days mean (SD)	5.6	(7.8)	5.6	(7.7)	5.8	(8.9)
Mechanical ventilation, days mean (SD)	6.4	(9.2)	6.7	(9.4)	5.2	(8.7)
Injury Severity Score, mean (SD)	12.1	(9.6)	11.6	(8.9)	27.4	(14.5)
Mortality Prediction models, mean (SD)						
TMPM, probability of death	0.036	(0.092)	0.027	(0.062)	0.271	(0.269)
TARN, probability of death	0.011	(0.036)	0.011	(0.032)	0.106	(0.146)

*$p<0.05$ for all comparisons

Table 2. Area under the ROC curve and AIC values for TARN and TMPM

	AUROC (95% CI)		**AIC (95% CI)**		**HL (95% CI)**	
			DERIVATION			
TARN	0.841	(0.840-0.841)	86 980	(86 942-87 019)	684.6	(681.3-687.8)
TMPM	0.880	(0.880-0.880)	74 509	(74 773-74545)	90.6	(89.4-91.7)
			VALIDATION			
TARN	0.845	(0.845-0.846)	21 163	(21 143-21 182)	153.0	(151.6-154.3)
TMPM	0.882	(0.882-0.882)	18 204	(18 186-18 223)	32.4	(31.7-33.1)

References

1. Committee on Trauma. Resources for optimal care of the injured patient 2014. Chicago, Ill: American College of Surgeons; 2014.
2. Intercollegiate Group on Trauma Standards. Regional Trauma Systems: Interim Guidance for Commissioners. London, England: The Royal College of Surgeons of England; 2009.
3. Hemmila MR, Nathens AB, Shafi S, et al. The Trauma Quality Improvement Program: Pilot Study and Initial Demonstration of Feasibility. J Trauma. 2010;68(2):253-262.
4. Lecky F, Woodford M, Yates D. Trends in trauma care in England and Wales 1989–97. The Lancet. 2000;355(9217):1771-1775.
5. Centers for Medicare Medicaid Services. CMS Quality Measure Development Plan: Supporting the Transition to the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). 2016.
6. NHS England. D15/S/a NHS Standard Contract For major Trauma (All ages) Schedule 2 - the services - A Service specifications [Internet]. 2013; http://www.england.nhs.uk/wp-content/uploads/2013/06/d15-major-trauma.pdf
7. NHS England. 2016/17 National Tariff Payment System. London: Monitor; 2016.
8. NHS England. Payment by results guidance for 2012-13. Leeds, UK 2012.
9. Medicare program; hospital inpatient value-based purchasing program. Final rule. Fed Regist. 2011;76(88):26490-26547.
10. Centers for Medicare and Medicaid Services. Hospital value-based purchasing. 2014; https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/Value-Based-Programs.html. Accessed April 21, 2017.
11. Centers for Medicare and Medicaid Services. Physician Quality Reporting System Qualified Clinical Data Registries. 2016; https://www.cms.gov/Medicare/Quality-Initiatives-Patient.../2016QCDRPosting.pdf. Accessed April 21, 2017.
12. Champion H, Copes W, Sacco W, et al. A new characterization of injury severity. J Trauma. 1990;30(5):539-546.
13. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. Trauma score. Crit Care Med. 1981;9(9):672-676.
14. Champion HR, Sacco WJ, Lepper RL, Atzinger EM, Copes WS, Prall RH. An anatomic index of injury severity. J Trauma. 1980;20(3):197-202.
15. Levy PS, Mullner R, Goldberg J, Gelfand H. The estimated survival probability index of trauma severity. Health Serv Res. 1978;13(1):28-35.

16. Osler T, Rutledge R, Deis J, Bedrick E. ICISS: an international classification of disease-9 based injury severity score. J Trauma. 1996;41(3):380-386; discussion 386-388.

17. Baker S, O'Niell B, Haddon W, Long W. The injury severity score:A method for describing patients with multiple injuries and evaluating emergency care. J Trauma. 1974;14(3):187-196.

18. Boyd C, Tolson, M. . Evaluating trauma care:The TRISS method. J Trauma. 1987;27(4):370-378.

19. MacKenzie E, Steinwachs D, Shankar B. An ICD-9-CM to AIS conversion table: Development and application. Paper presented at: Annual Conference of the American Association for Automotive Medicine; October 6–8, 1986, 1986; Montreal, Quebec.

20. Kilgo PD, Meredith JW, Hensberry R, Osler TM. A note on the disjointed nature of the injury severity score. J Trauma. 2004;57(3):479-485; discussion 486-477.

21. Cook A, Weddle J, Baker S, et al. A comparison of the Injury Severity Score and the Trauma Mortality Prediction Model. J Trauma Acute Care Surg. 2014;76(1):47-52; discussion 52-43.

22. Osler T, Glance L, Buzas JS, Mukamel D, Wagner J, Dick A. A trauma mortality prediction model based on the anatomic injury scale. Ann Surg. 2008;247(6):1041-1048.

23. Glance LG, Osler TM, Mukamel DB, Meredith W, Wagner J, Dick AW. TMPM-ICD9: a trauma mortality prediction model based on ICD-9-CM codes. Ann Surg. 2009;249(6):1032-1039.

24. American College of Surgeons. National Trauma Data Bank Annual Report 2014. 2014; https://www.facs.org/quality%20programs/trauma/ntdb/docpub. Accessed November 10, 2015.

25. Bouamra O, Jacques R, Edwards A, et al. Prediction modelling for trauma using comorbidity and ‘true’30-day outcome. Emergency Medicine Journal. 2015;32(12):933-938.

26. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Assessing the fit of the model. Applied logistic regression. Hoboken, New Jersey: John Wiley & Sons; 2013:173-182.

27. Akaike H. A new look at the statistical model identification. Automatic Control, IEEE Transactions on. 1974;19(6):716-723.

28. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143:29-36.

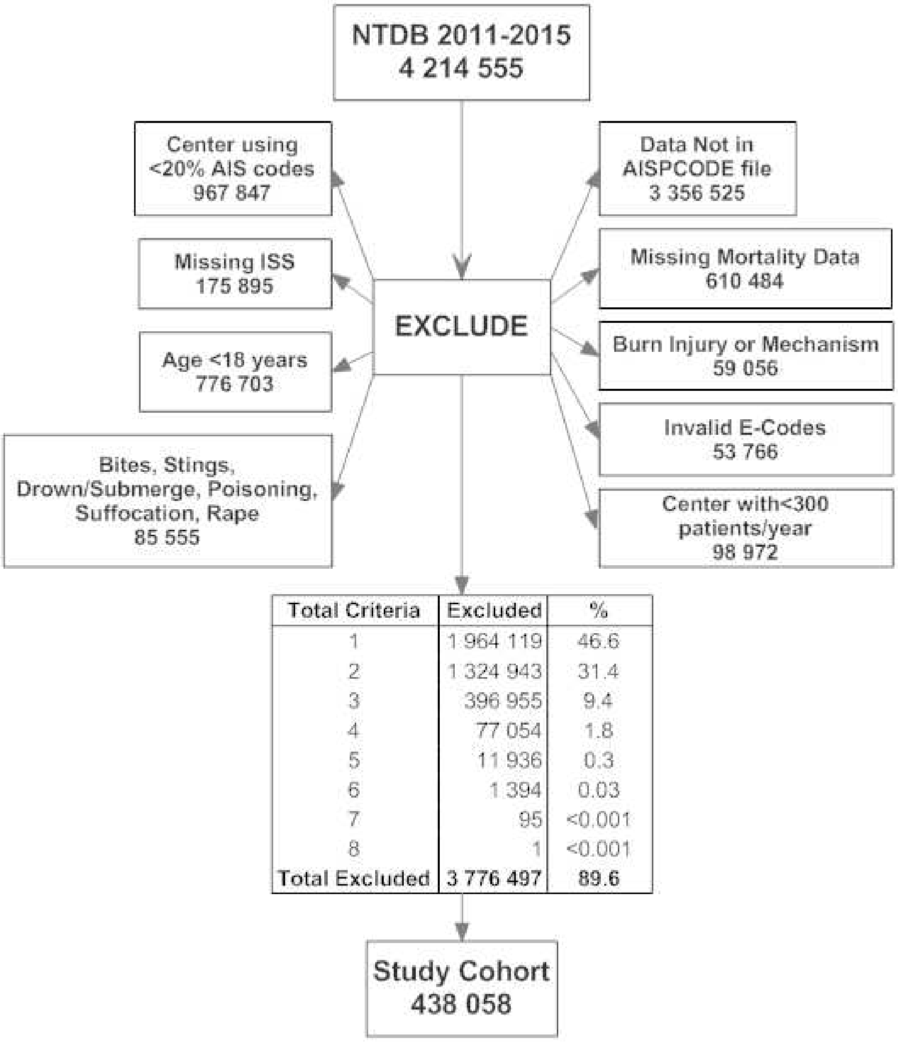
29. Haider AH, Villegas CV, Saleem T, et al. Should the IDC-9 Trauma Mortality Prediction Model become the new paradigm for benchmarking trauma outcomes? J Trauma Acute Care Surg. 2012;72(6):1695-1701.

30. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2013.

31. Charlson M, Pompei P, Ales K, MacKenzie C. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Disease. 1987;40(5):373-383.

32. Dr Foster Intelligence. Understanding HSMRs: A Toolkit on Hospital Standardised Mortality Ratios. London2014.

Figure 1. Flow diagram of patients excluded by criteria



Total Criteria	Excluded	%
1	1 964 119	46.6
2	1 324 943	31.4
3	396 955	9.4
4	77 054	1.8
5	11 936	0.3
6	1 394	0.03
7	95	<0.001
8	1	<0.001
Total Excluded	3 776 497	89.6

Figure 2. Calibration curve for the anatomic components of TARN and TMPM models: Derivation

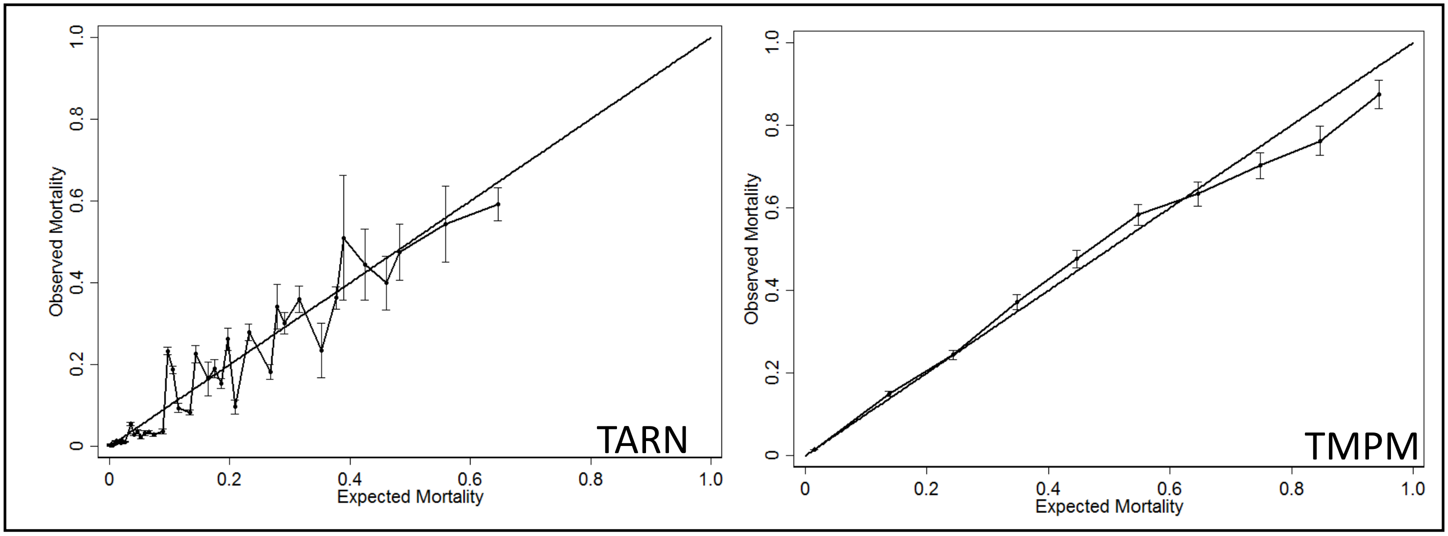


Figure 3. Calibration curve for the anatomic components of TARN and TMPM models: Validation