

Trauma Early Mortality Prediction Tool (TEMPT) for assessing 28-day mortality

Ryan C Kunitake,¹ Lucy Z Kornblith,¹ Mitchell Jay Cohen,² Rachael A Callcut¹

¹Department of Surgery, University of California, San Francisco, California, USA

²Department of Surgery, Denver Health Medical Center, Denver, Colorado, USA

Correspondence to

Dr Rachael A Callcut, Department of Surgery, University of California, San Francisco, San Francisco CA 94110, USA; rachael.callcut@ucsf.edu

Received 2 October 2017

Revised 5 November 2017

Accepted 14 November 2017

ABSTRACT

Background Prior mortality prediction models have incorporated severity of anatomic injury quantified by Abbreviated Injury Severity Score (AIS). Using a prospective cohort, a new score independent of AIS was developed using clinical and laboratory markers present on emergency department presentation to predict 28-day mortality.

Methods All patients (n=1427) enrolled in an ongoing prospective cohort study were included. Demographic, laboratory, and clinical data were recorded on admission. True random number generator technique divided the cohort into derivation (n=707) and validation groups (n=720). Using Youden indices, threshold values were selected for each potential predictor in the derivation cohort. Logistic regression was used to identify independent predictors. Significant variables were equally weighted to create a new mortality prediction score, the Trauma Early Mortality Prediction Tool (TEMPT) score. Area under the curve (AUC) was tested in the validation group. Pairwise comparison of Trauma Injury Severity Score (TRISS), Revised Trauma Score, Glasgow Coma Scale, and Injury Severity Score were tested against the TEMPT score.

Results There was no difference between baseline characteristics between derivation and validation groups. In multiple logistic regression, a model with presence of traumatic brain injury, increased age, elevated systolic blood pressure, decreased base excess, prolonged partial thromboplastin time, increased international normalized ratio (INR), and decreased temperature accurately predicted mortality at 28 days (AUC 0.93, 95% CI 0.90 to 0.96, P<0.001). In the validation cohort, this score, termed TEMPT, predicted 28-day mortality with an AUC 0.94 (95% CI 0.92 to 0.97). The TEMPT score performed similarly to the revised TRISS score for severely injured patients and was highly predictive in those having mild to moderate injury.

Discussion TEMPT is a simple AIS-independent mortality prediction tool applicable very early following injury. TEMPT provides an AIS-independent score that could be used for early identification of those at risk of doing poorly following even minor injury.

Level of evidence Level II.

INTRODUCTION

Despite advances in care, between 2000 and 2010 deaths from trauma increased by 22.8%, a number far exceeding the parallel rise in US population (9.7%) and in stark contrast to a decrease in deaths from cancer and heart disease during this same period.¹ The exact cause of this increase remains unknown, but it may in part be explained by the

aging baby boomer population; older age is associated with increased morbidity and mortality after trauma, and patients aged 45–64 years exhibited the greatest rise in death rates across age groups (28.5%).^{1,2} Implementing a precision-based approach to trauma care requires an updated system of mortality prediction that reflects the evolving trauma population and is easy to use for practical application early in care.

Predictive modeling can facilitate effective resource allocation, estimation of patient outcomes, and trauma quality assessment. Algorithms based on Abbreviated Injury Scales (AIS) such as Trauma Injury Severity Score (TRISS), Revised Trauma Score (RTS), and Injury Severity Score (ISS) are plentiful. Yet, these scoring systems primarily focus on characterizing large patient populations post hoc rather than guiding real-time clinical decision-making.^{3–6} In fact, the most well-established scores are nearly impossible to adopt clinically, requiring complex mathematical calculation impractical for the chaotic trauma bay as well as necessary data points unattainable during the early stages of a patient's hospital course. Further, these scores largely predate balanced resuscitation and have been shown to demonstrate poorer prediction.^{7–9} Today, TRISS is the most widely used trauma mortality scoring system and is constrained by all the aforementioned limitations.

Advancing trauma outcome prediction is critical to meet the rapidly changing trauma population and resuscitation methodologies. While clinical acumen can indeed facilitate management for many trauma patients, we argue that the development of a standardised score calculable early in the care pathway could additionally serve to identify high-risk patients who might otherwise go undetected, more effectively personalise resuscitation, and focus high-value care on those who benefit the most from targeted interventions. With these principles in mind, we generated an AIS-independent prediction algorithm using simple clinical and laboratory values measured on admission and compared this to the predictive ability of historical scoring systems.

METHODS

All level I trauma patients presenting to San Francisco General Hospital between 2005 and 2015 who were simultaneously enrolled in an ongoing prospective observational cohort study entitled Activation of Coagulation and Inflammation in Trauma (ACIT) were selected for inclusion. ACIT enrolls adult patients who are highest level trauma activations in whom initial intravenous access is obtained in the emergency department (ED) for

To cite: Kunitake RC, Kornblith LZ, Cohen MJ, et al. *Trauma Surg Acute Care Open* 2018;**3**:1–6.

routine clinical care. The original ACIT study aimed to elucidate mechanisms of coagulopathy and inflammation among trauma patients, and this secondary analysis examined comprehensive demographic, injury, laboratory, and in-hospital outcome data collected prospectively to 28 days. Not all patients had all laboratory samples obtained in routine clinical care and no specific laboratory tests were mandated by enrollment in the observational study. To ensure accuracy of presence of traumatic brain injury (TBI), ED notes and imaging reports for studies obtained in the ED were reviewed by physician adjudicators. Patients were classified as having TBI if they had evidence of skull fractures, epidural hematomas, subdural hematomas, subarachnoid hemorrhages, midline shifts, brain contusions, intraventricular hemorrhages, diffuse axonal injury, or intracranial hemorrhages. Exclusion criteria included those under 15 years of age, burn victims, pregnant women, and those who were incarcerated.

Following a cross-validation approach to model prediction, patients were randomised into derivation and validation cohorts to minimise overfitting. Using a true random number generator, patients were randomly assigned zeros or ones, respectively designating selection to either the derivation or validation patient group. Randomization was subsequently assessed by examining differences in baseline characteristics via Wilcoxon rank-sum and χ^2 tests, as appropriate. In the derivation cohort, univariate analysis identified individual predictors of 28-day mortality ($P < 0.005$), and continuous variables were dichotomised via the Youden index.¹⁰ This statistical technique generates optimal threshold cut-offs by maximizing the sensitivity and specificity for each respective variable, which in practice involves finding the greatest vertical distance between the receiver operator characteristic (ROC) curve and the diagonal chance line.

Dichotomous predictors were then included in multiple logistic regression in the derivation cohort with backward variable elimination by stepwise regression ($P \leq 0.20$ was selected to minimise residual confounding). The resulting model composed of equally weighted variables was termed Trauma Early Mortality Prediction Tool and was further assessed by multiple logistic regression in the validation cohort by ROC discrimination (area under the curve (AUC)). Missingness of data was examined

using Wilcoxon rank-sum and χ^2 tests, and calibration was evaluated by Hosmer-Lemeshow goodness of fit. Performance of TEMPT was then compared in pairwise fashion to existing historical mortality and injury scores, including TRISS, RTS, Glasgow Coma Scale (GCS), and ISS by χ^2 . Lastly, bootstrapping was employed to simulate the performance of TEMPT within a larger population. This method of repetitive random sampling with replacement (replications=1000) projects estimates for the general population.

RESULTS

Of the 1427 total trauma patients enrolled in the ACIT study, 81% were male, 34% were Caucasian, and median age was 35 years (table 1).

Also, 56% suffered blunt injury and 36% experienced TBI. Median ISS was 14 and overall mortality was 18% (table 1). The group was randomly divided into a derivation (n=707) and a validation cohort (n=720 patients) using true random number generator technique. In order to assess bias in the creation of the two groups, the demographic and injury characteristics of the derivation and validation group were compared. There were no statistical difference in demographic or injury characteristics (table 1, P=NS).

Creating the TEMPT model

In univariate analysis of the derivation cohort, TBI (OR 11.52, 95% CI 7.04 to 18.83), international normalized ratio (INR) (OR 3.96, 95% CI 1.90 to 8.22), blunt injury (OR 2.46, 95% CI 1.59 to 3.80), creatinine (OR 1.75, 95% CI 1.26 to 2.43), partial thromboplastin time (PTT) (OR 1.14, 95% CI 1.10 to 1.18), age (OR 1.03, 95% CI 1.02 to 1.04), systolic blood pressure (OR 1.01, 95% CI 1.00 to 1.02), platelets (OR 0.99, 95% CI 0.99 to 1.00), base excess (OR 0.95, 95% CI 0.92 to 0.98), hemoglobin (OR 0.83, 95% CI 0.76 to 0.91), and temperature (OR 0.43, 95% CI 0.33 to 0.57) significantly predicted mortality at 28 days ($P < 0.05$) (table 2).

Table 1 Baseline characteristics and assessment of randomization

Variable	All patients (n=1427)		Derivation (n=707)			Validation (n=720)	
	n	Median or %	n	Median or %	n	Median or %	P value
Age	1427	35 years	707	35 years	720	36 years	0.24
Sex							
Male	1160	81.2%	569	80.5%	591	81.9%	0.51
Female	269	18.8%	138	19.5%	131	18.1%	
Body mass index	1083	26.0	532	25.8	561	26.0	0.59
Race							
Caucasian/Latino	783	54.8%	385	54.5%	398	55.1%	0.97
Black	339	23.7%	166	23.5%	173	24.0%	
Asian	185	13.0%	95	13.4%	90	12.5%	
Pacific Islander	9	0.6%	5	0.7%	4	0.6%	
Native American	8	0.6%	4	0.6%	4	0.6%	
Other	36	2.5%	20	2.8%	16	2.2%	
Unknown	69	4.8%	32	4.5%	37	5.1%	
Injury Severity Score	1389	14.0	689	13.0	700	14.0	0.24
Blunt injury	801	56.3%	383	54.3%	418	58.1%	0.15
Traumatic brain injury	519	36.4%	248	35.2%	271	37.6%	0.34
Mortality at 28 days	253	17.7%	117	16.6%	136	18.8%	0.26

Table 2 Univariate analysis of the derivation cohort

Variable	n	OR	95% CI	P value
Traumatic brain injury	705	11.52	7.04 to 18.83	<0.001
INR	656	3.96	1.90 to 8.22	<0.001
Blunt injury	705	2.46	1.59 to 3.80	<0.001
Creatinine (mg/dL)	686	1.75	1.26 to 2.43	0.001
PTT (s)	657	1.14	1.10 to 1.18	<0.001
Age (years)	707	1.03	1.02 to 1.04	<0.001
Systolic blood pressure (mm Hg)	692	1.01	1.00 to 1.02	0.002
Platelets ($\times 10^3/\mu\text{L}$)	686	0.99	0.99 to 1.00	<0.001
Base excess (mmol/L)	515	0.95	0.92 to 0.98	0.001
Hemoglobin (g/dL)	689	0.83	0.76 to 0.91	<0.001
Temperature ($^{\circ}\text{C}$)	489	0.43	0.33 to 0.57	<0.001

PTT, partial thromboplastin time.

Optimal cut-offs were subsequently generated using Youden indices for all significant continuous predictors in univariate analysis ($P < 0.005$) (table 3).

These predictors were then assessed using multiple logistic regression treating the cut-off points as categorical variables. In multiple logistic regression, the presence of TBI (OR 9.7, 95% CI 3.6 to 26.1, $P < 0.001$), increased age (OR 8.0, 95% CI 3.0 to 21.4, $P < 0.001$), elevated systolic blood pressure (OR 4.6, 95% CI 1.7 to 12.3, $P = 0.003$), decreased base excess (OR 3.8, 95% CI 1.5 to 9.4, $P = 0.004$), prolonged PTT (OR 3.5, 95% CI 1.4 to 8.8, $P = 0.008$), increased INR (OR 2.8, 95% CI 1.0 to 8.0, $P = 0.049$), and decreased temperature (OR 1.9, 95% CI 0.8 to 4.6, $P = 0.17$) accurately predicted mortality at 28 days (AUC 0.93, 95% CI 0.90 to 0.96, $P < 0.001$) (table 4).

The variables that remained statistically significant in multiple logistic regression at the selected cut-off points compose the TEMPT score. The score includes presence of TBI, age (≥ 59.5 years), base excess (≥ -4.35 mmol/L), PTT (≥ 31.45 s), INR (≥ 1.25), and temperature ($\leq 36.25^{\circ}\text{C}$).

Model validation

The TEMPT score model was then tested in the validation cohort, and it performed similarly with an AUC 0.94 (95% CI 0.92 to 0.97) compared with AUC 0.93 (95% CI 0.90 to 0.96) in the derivation cohort for the prediction of mortality at 28 days (table 4). This model exhibited appropriate calibration by Hosmer-Lemeshow goodness of fit ($P = \text{NS}$) (table 5).

In order to compare our new AIS-independent score to commonly used scores, we selected the subset of patients from

Table 3 Youden index cut-offs of the derivation cohort

Variable	Cut-off
Age (years)	≥ 59.50
Systolic blood pressure (mm Hg)	≥ 163.50
Creatinine (mg/dL)	≥ 1.35
INR	≥ 1.25
PTT (s)	≥ 31.40
Hemoglobin (g/dL)	≤ 12.75
Platelets ($10^3/\mu\text{L}$)	≤ 224.50
Base excess (mmol/L)	≤ -4.35
Temperature ($^{\circ}\text{C}$)	≤ 36.25

INR, international normalised ratio; PTT, partial thromboplastin time.

the validation group that had complete data needed for calculation of each of the comparison scores of TRISS, RTS, GCS, and ISS scores. These data were available for 305 patients of the validation cohort (42.4%); the data were largely missing at random, with missing patients exhibiting a greater percentage of females (21.6% vs 13.4%, $P = 0.005$) but no significant difference in age, race, BMI, ISS, TBI, or mechanism of injury compared with their non-missing counterparts ($P > 0.40$). In pairwise comparisons of TEMPT to TRISS, RTS, GCS, and ISS in the validation cohort, TEMPT (AUC 0.92, 95% CI 0.88 to 0.95) performed similarly to 1995 updated TRISS (AUC 0.93, 95% CI 0.89 to 0.97, $P = 0.58$) and 2009 updated TRISS (AUC 0.89, 95% CI 0.84 to 0.95, $P = 0.47$), but significantly outperformed RTS (AUC 0.83, 95% CI 0.77 to 0.89, $P = 0.003$), GCS (AUC 0.84, 95% CI 0.78 to 0.89, $P = 0.003$), and ISS (AUC 0.84, 95% CI 0.79 to 0.90, $P = 0.001$).

In mild to moderately injured patients (ISS < 16), TEMPT performed comparably to 2009 updated TRISS (AUC 0.69, 95% CI 0.22 to 1.00, $P = 0.20$) and significantly outperformed 1995 updated TRISS (AUC 0.78, 95% CI 0.59 to 0.97, $P = 0.02$), RTS (AUC 0.71, 95% CI 0.62 to 0.79, $P < 0.001$), GCS (AUC 0.72, 95% CI 0.64 to 0.81, $P < 0.001$), and ISS (AUC 0.69, 95% CI 0.48 to 0.90, $P = 0.03$) (table 6).

In severely injured patients (ISS ≥ 16), TEMPT performed comparably to 1995 updated TRISS (AUC 0.89, 95% CI 0.83 to 0.94, $P = 0.15$), 2009 updated TRISS (AUC 0.83, 95% CI 0.75 to 0.90, $P = 0.75$), RTS (AUC 0.80, 95% CI 0.77 to 0.89, $P = 0.38$), and GCS (AUC 0.83, 95% CI 0.78-0.89, $P = 0.52$), and significantly outperformed ISS (AUC 0.68, 95% CI 0.78 to 0.90, $P = 0.002$) (table 6). Bootstrapping projected similar 95% CIs for all patient groups (table 6).

DISCUSSION

In 1983, the advent of TRISS provided a common mechanism for capturing trauma mortality. TRISS included both anatomic and physiological data as a comprehensive approach to injury evaluation. For decades, TRISS was widely accepted as the standard method for mortality prediction.⁵⁻¹¹ However, Rogers *et al* recently found that TRISS significantly overestimates trauma mortality and does so by an increasing margin annually.⁸ This likely reflects advances in trauma and critical care as well as changing population demographics that have affected outcomes.

Between 2002 and 2010, Sise *et al* identified a substantial increase in mortality due to falls (46%) and decrease due to motor vehicle traffic (27%).¹² In addition, McGwin *et al*, Gunst *et al*, and Evans *et al* noted a shift away from the classic trimodal distribution of immediate (< 1 hour after injury), early (several hours after injury), and late (days to weeks after injury) deaths and toward a bimodal profile with fewer deaths beyond 1 week.¹³⁻¹⁵ These shifts have occurred simultaneous with advances in resuscitation. Historically, crystalloid was the initial fluid replacement strategy for hemorrhaging patients and often patients would have exposure to large volumes. In the past decade, there has been a diffusion of the adoption of balanced resuscitation with physiological ratios of packed red blood cells (PRBCs), platelets, and plasma while limiting crystalloid. The net result has been an improvement in mortality.¹⁶⁻¹⁷ In PROPPR, a recently reported randomised control trial of balanced resuscitation ratios of platelets, plasma, and PRBCs, the major difference in mortality was attributable to a decrease in exsanguinating deaths with balanced resuscitation.¹⁸

With advances in critical care and trauma resuscitation, it is likely that the TRISS score has demonstrated poorer performance

Table 4 Multiple logistic regression of the derivation and validation cohorts

Variables	Derivation (n=334)			Validation (n=327)		
	OR	95% CI	P value	OR	95% CI	P value
Traumatic brain injury	9.7	3.6 to 26.1	<0.001	10.6	3.7 to 30.9	<0.001
Age (years)	8.0	3.0 to 21.4	<0.001	9.6	3.6 to 25.4	<0.001
Systolic blood pressure(mm Hg)	4.6	1.7 to 12.3	0.003	5.3	1.9 to 14.6	0.001
Base excess (mmol/L)	3.8	1.5 to 9.4	0.004	3.0	0.8 to 4.8	0.16
PTT (s)	3.5	1.4 to 8.8	0.008	5.8	2.1 to 15.7	0.001
INR	2.8	1.0 to 8.0	0.049	2.2	1.1 to 8.5	0.04
Temperature (°C)	1.9	0.8 to 4.6	0.17	1.9	0.9 to 4.8	0.09

recently because a proportion of patients have been converted from those with a high probability of death due to complications of hemorrhage to a lower risk group. Although TRISS has undergone several iterations in the form of coefficient revisions since its first inception, its real-time clinical utility remains limited by the need to have fully defined anatomic injury for calculation. This has restricted the use of TRISS to essentially a retrospective tool designed to benchmark performance after all care is complete. Generating a novel mortality prediction score in real time that could augment early clinical management to mitigate poor outcome is the next step toward achieving precision health solutions for trauma care. To achieve this, an AIS-independent prediction tool would be required.

In this study, we generated a novel AIS-independent trauma mortality prediction score (TEMPT) that accurately assesses survival after injury and can be practically adopted in an acute setting. All predictor variables selected for TEMPT are often available on ED admission and are similarly well documented in trauma and critical care literature to be important for trauma mortality.^{2, 19–26} Further, despite its AIS independence, the tool has equivalent performance to the most recent revision of TRISS for all patients and exceeds its performance in the minor injured population. One of the most important advantages of TEMPT compared with TRISS is that TRISS cannot be calculated until the AIS scores are known for individual patients. In most trauma centers, these values are obtained after the patient care episode has been concluded and all the injuries are known. The purpose of the TEMPT score is that it can be used early in care prospectively before the full spectrum of injuries are known.

This is an important distinction. Prediction tools calculated at the end of care are useful for benchmarking and research purposes; however, they provide little prospective benefit at the point of care for the bedside practitioner considering treatment options. As we move toward the goal of precision medicine in trauma care, we will require the ability to discern those more likely to do poorly earlier in care. Earlier identification of these patients creates opportunity to apply precision medicine techniques in an effort to personalise care pathways to mitigate poor

outcome. Although most practitioners recognise obviously ill or not ill persons, the value in these types of prediction rules is the ability to discern between the large group in the middle. The TEMPT score performs particularly well in this patient group. Considering the present push toward precision medicine, we believe the accuracy and practical simplicity of TEMPT as an AIS-independent prediction algorithm offers considerable advantages toward directing personalised care in an acute setting.

Some results directly parallel findings in the literature. Similar to our own age cut-off of ≥ 59.5 years, Campbell-Furtick *et al* recommended age 60 to be considered the new ‘elderly’ in the setting of trauma as it marked the start of the most accelerated rise in mortality rate in relation to advancing age.²⁷ A base excess threshold of ≤ -4.35 mmol/L aligns with classifications for class II mild shock (<2 to -6 mmol/L), and in an analysis of $>16\,000$ patients of the TraumaRegister DGU database, this cohort exhibited an average aPTT of 32.1 s, thus, mirroring our own PTT cut-off of ≥ 31.4 s.²⁸ Mutschler and colleagues found individuals presenting in class II shock or greater (base excess < -2 mmol/L) displayed a mortality rate more than twice that of non-shock patients (19% vs 7.4%).²⁸ Neville *et al* identified that among elderly trauma patients this rate becomes even more pronounced as those with an initial base excess of ≤ -4 mmol/L were more than five times more likely to die than patients with a base excess of > -4 mmol/L.²⁹ Lastly, systolic blood pressure as a predominant upper limit cut-off (≥ 163.5 mm Hg) likely represents the Cushing’s reflex in the setting of severe brain injury, thus signifying the overwhelming detriment of central nervous system injury. TBI is the leading cause of death immediately after trauma and is also associated with a threefold increased odds of mortality.²¹

In contrast, other variable cut-offs that we have identified on first pass would not raise concern in most providers; however, we found them to be predictive of poor outcome. This underscores the utility of analytic solutions and precision approaches that have potential to enhance clinician diagnostic abilities especially in patients that initially appear stable. A trauma patient presenting with a hemoglobin of 12.75 g/dL or temperature of 36.25°C does not raise red flags, but these results suggest perhaps they should. It may be time to revise the classic paradigm ‘sick or not sick’ as seemingly innocuous thresholds may now be required for sufficient sensitivity to identify those at risk of higher mortality. The findings suggest that there is inherent value in seemingly benign cut-offs to detect individuals at risk of poor outcome despite initially appearing to be less injured patient. Detecting these patients early may allow for more rigorous care or stratification along a different care pathway that may serve to alter outcomes in a more favorable direction.

Table 5 Calibration of models

Model	n	P value
Trauma Early Mortality Prediction Tool	305	0.45
Trauma Injury Severity Score (2009)	305	0.11
Trauma Injury Severity Score (1995)	305	0.18
Revised Trauma Score	305	0.04
Glasgow Coma Scale	305	0.045
Injury Severity Score	305	0.26

Table 6 Performance of the Trauma Early Mortality Prediction Tool compared with previously published scores

Model	Mild to moderately injured (n=157)				Severely injured (n=148)			
	Area under the curve	95% CI	95% CI*	P	Area under the curve	95% CI	95% CI*	P value
Trauma Early Mortality Prediction Tool	1.00	0.99 to 1.00	0.99 to 1.00		0.84	0.82 to 0.93	0.76 to 0.90	
Trauma Injury Severity Score (2009)	0.69	0.22 to 1.00	0.38 to 0.95	0.20	0.83	0.75 to 0.90	0.75 to 0.89	0.70
Trauma Injury Severity Score (1995)	0.78	0.59 to 0.97	0.63 to 0.90	0.02	0.89	0.83 to 0.94	0.82 to 0.93	0.15
Revised Trauma Score	0.71	0.77 to 0.89	0.62 to 0.78	<0.001	0.80	0.77 to 0.89	0.72 to 0.87	0.38
Glasgow Coma Scale	0.72	0.78 to 0.89	0.63 to 0.79	<0.001	0.83	0.78 to 0.89	0.74 to 0.88	0.52
Injury Severity Score	0.69	0.79 to 0.90	0.51 to 0.87	0.03	0.68	0.79 to 0.90	0.58 to 0.76	0.002

*Bootstrap result.

Nearly 30 years have passed since the original development of TRISS (1983), RTS (1989), GCS (1974), and ISS (1974), and these findings insinuate their limitations may lie in the assessment of mild to moderately injured patients.^{3-5 11 30} In cases of severe injury, irreparable damage often dictates outcomes; within this subset, these scores still perform accurately as even the most aggressive care is minimally beneficial for the patient with an anatomically non-survivable injury. In contrast, the historic prediction algorithms are not as effective at discerning outcome in those less severely injured and the TEMPT score appears to offer an advantage in this patient group.

TEMPT statistically outperformed 1995 updated TRISS, RTS, and GCS in the less severely injured patient group. While differences in performance with the most recent version of TRISS did not achieve statistical significance, TRISS had a low AUC and wide CI raising concern about precision, accuracy, and overfitting. Simulation in a larger population through the use of bootstrapping further supports these conclusions and suggests the score would be beneficial within a larger generalised trauma population. Although bootstrapping enhances confidence in the results, it is important to acknowledge that the derivation and validation groups were derived from a population from a single center. The trauma population this sample was drawn from is similar to those previously studied in multi-center studies with our cohort having similar ages, percentage of males, race and blunt mechanism prevalence.¹⁸ The patients in the current study are slightly less injured compared with these studies, which reflect a selection bias in the comparison studies that were designed to select only patients requiring transfusion. Although we believe our ACIT cohort to represent a consistent demographic compared with a typical trauma center population, TEMPT should next be studied in a true larger cohort before we can recommend universal adoption.

Finally, despite its role as the gold standard of injury assessment, ISS, derived from AIS scores, exhibited the worst predictive ability of all scoring systems. These results compliment the many studies documenting the limitations of ISS. Narrow inclusion of single injuries from the three most severely damaged body regions diminishes ability to account for polytrauma. Heavy reliance on AIS predisposes to substantial inter-coder variability.⁷ Gupta *et al* found the increased utilization of CT pan scanning promotes ISS inflation due to greater detection of minor injuries that increase ISS but do not alter clinical decision-making.³¹ Joesse *et al* even identified specific injury profiles of low ISS patients associated with mortality.³² TEMPT provides benefit early in the course of care and does so independently of retrospective ISS.

This study is not without its limitations. First, given this was an observational, non-interventional study, we could not mandate specific laboratory values or clinical data (temperature) be

obtained on patients enrolled. Therefore, data for each variable had some missingness, which was 0%–7% for all variables except base excess (28%) and temperature (31%). To assess the possibility that bias may have arisen from missingness, we evaluated differences between the patients with missing and non-missing data. There were no significant differences with regard to injury characteristics (ISS, TBI, or mechanism of injury) and the data appear to be missing at random. Additionally, in order to be certain that the TBI variable collected prospectively in the ED was accurate, TBI status was verified using physician adjudicators blinded to the conduct of the ACIT study. This independent evaluation was done to ensure that patients classified as having TBI did have TBI known at the time of hospital admission from the ED. We specifically used information only known at the time of hospital admission with regards to TBI status. This was important as the goal of developing TEMPT was to be an AIS-independent score able to be used at the time of admission from the ED.

The use of dichotomous treatment of independent variables in place of continuous variables is also a potential limitation. This has the net effect of potentially weakening the predictive ability of regression models. However, as it was our aim to develop a clinically practical score, we believe the inclusion of simple yes/no variables greatly encourages adoption and hospital implementation. All predictor variables were dichotomised and equally weighted to increase TEMPT's clinical utility. While these amendments do indeed detract from TEMPT's predictive ability, we argue the added simplicity provides a considerable advantage for adoption within the hospital. Further, we surmise these statistically derived cut-offs themselves pose exciting implications for the current state of injury assessment.

CONCLUSION

In conclusion, this study provides a novel, prospectively derived, AIS-independent trauma mortality prediction score that can be used on ED presentation and improves detection of poor outcome in mildly injured patients. Further validation is needed in a larger patient cohort, but TEMPT may provide an alternate assessment tool that is more reflective of current resuscitation practices and better aligns with precision medicine goals.

Contributors Literature search, data analysis and writing: RCK and RAC. Study design: RCK, MJC and RAC. Data collection and data interpretation: LZK, MJC, RCK and RAC. Critical revision: LZK, MJC and RAC.

Funding RCK is supported by the National Center for Advancing Translational Sciences, National Institutes of Health via the University of California San Francisco Clinical and Translational Institute (TL1 TR000144). RAC is supported with a career development award from the National Institute of Environmental Health Sciences, National Institute of Health Big Data to Knowledge Initiative (K01ES026834).

Disclaimer The contents of this paper are solely the responsibility of the authors

and do not represent the official views of the NIH.

Competing interests None declared.

Ethics approval This investigation was approved by the Institutional Review Board (IRB) at the University of California, San Francisco School of Medicine and Zuckerberg San Francisco General Hospital.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Rhee P, Joseph B, Pandit V, Aziz H, Verccruysse G, Kulvatunoy N, Friese RS. Increasing trauma deaths in the United States. *Ann Surg* 2014;260:13–21.
- Grossman MD, Miller D, Scaff DW, Arcona S. When is an elder old? Effect of preexisting conditions on mortality in geriatric trauma. *J Trauma* 2002;52:242–6.
- Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974;2:81–4.
- Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the trauma score. *J Trauma* 1989;29:623–9.
- Champion HR, Sacco WJ, Hunt TK. Trauma severity scoring to predict mortality. *World J Surg* 1983;7:4–11.
- Palmer CS, Gabbe BJ, Cameron PA. Defining major trauma using the 2008 abbreviated injury scale. *Injury* 2016;47:109–15.
- Ringdal KG, Skaga NO, Hestnes M, Steen PA, Røislien J, Rehn M, Røise O, Krüger AJ, Lossius HM. Abbreviated injury scale: not a reliable basis for summation of injury severity in trauma facilities? *Injury* 2013;44:691–9.
- Rogers FB, Osler T, Krasne M, Rogers A, Bradburn EH, Lee JC, Wu D, McWilliams N, Horst MA. Has TRISS become an anachronism? A comparison of mortality between the National Trauma Data Bank and major trauma outcome study databases. *J Trauma Acute Care Surg* 2012;73:326–31.
- Schluter PJ. Trauma and Injury Severity Score (TRISS): is it time for variable re-categorisations and re-characterisations? *Injury* 2011;42:83–9.
- Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology* 2005;16:73–81.
- Schluter PJ, Nathens A, Neal ML, Goble S, Cameron CM, Davey TM, McClure RJ. Trauma and Injury Severity Score (TRISS) coefficients 2009 revision. *J Trauma* 2010;68:761–70.
- Sise RG, Calvo RY, Spain DA, Weiser TG, Staudenmayer KL. The epidemiology of trauma-related mortality in the United States from 2002 to 2010. *J Trauma Acute Care Surg* 2014;76:913–20.
- Evans JA, van Wessem KJ, McDougall D, Lee KA, Lyons T, Balogh ZJ. Epidemiology of traumatic deaths: comprehensive population-based assessment. *World J Surg* 2010;34:158–63.
- Gunst M, Ghaemmaghami V, Gruszecki A, Urban J, Frankel H, Shafi S. Changing epidemiology of trauma deaths leads to a bimodal distribution. *Proc* 2010;23:349–54.
- McGwin G, Nunn AM, Mann JC, Griffin R, Davis GG, MacLennan PA, Kerby JD, Acker JE, Rue LW. Reassessment of the tri-modal mortality distribution in the presence of a regional trauma system. *J Trauma* 2009;66:526–30.
- Holcomb JB, Wade CE, Michalek JE, Chisholm GB, Zarzabal LA, Schreiber MA, Gonzalez EA, Pomper GJ, Perkins JG, Williams KL *et al*. Increased plasma and platelet to red blood cell ratios improves outcome in 466 massively transfused civilian trauma patients. *Ann Surg* 2008;248:97–108.
- Sperry JL, Ochoa JB, Gunn SR, Alarcon LH, Minei JP, Cuschieri J, Rosengart MR, Maier RV, Billiar TR, Peitzman AB, *et al*. An FFP:PRBC transfusion ratio $\geq 1:1.5$ is associated with a lower risk of mortality after massive transfusion. *J Trauma* 2008;65:986–93.
- Holcomb JB, Tilley BC, Baraniuk S, Fox EE, Wade CE, Podbielski JM, del Junco DJ, Brasel KJ, Bulger EM, Callcut RA, *et al*. Transfusion of plasma, platelets, and red blood cells in a 1:1:1 vs a 1:1:2 ratio and mortality in patients with severe trauma: the PROPPR randomized clinical trial. *JAMA* 2015;313:471–82.
- DeLa'O CM, Kashuk J, Rodriguez A, Zipf J, Dumire RD. The Geriatric Trauma Institute: reducing the increasing burden of senior trauma care. *Am J Surg* 2014;208:988–94.
- Cosgriff N, Moore EE, Sauaia A, Kenny-Moynihan M, Burch JM, Galloway B. Predicting life-threatening coagulopathy in the massively transfused trauma patient: hypothermia and acidosis revisited. *J Trauma* 1997;42:857–61.
- Fazel S, Wolf A, Pillas D, Lichtenstein P, Långström N. Suicide, fatal injuries, and other causes of premature mortality in patients with traumatic brain injury: a 41-year Swedish population study. *JAMA Psychiatry* 2014;71:326–33.
- Callcut RA, Cotton BA, Muskat P, Fox EE, Wade CE, Holcomb JB, Schreiber MA, Rahbar MH, Cohen MJ, Knudson MM, *et al*. Defining when to initiate massive transfusion: a validation study of individual massive transfusion triggers in PROMTTT patients. *J Trauma Acute Care Surg* 2013;74:59–65.
- Callcut RA, Cripps MW, Nelson MF, Conroy AS, Robinson BB, Cohen MJ. The massive transfusion score as a decision aid for resuscitation: learning when to turn the massive transfusion protocol on and off. *J Trauma Acute Care Surg* 2016;80:450–6.
- Nunez TC, Voskresensky IV, Dossett LA, Shinal R, Dutton WD, Cotton BA. Early prediction of massive transfusion in trauma: simple as ABC (assessment of blood consumption)? *J Trauma* 2009;66:346–52.
- Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, Iapichino G, Edbrooke D, Capuzzo M, Le Gall JR. SAPS 3-From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. *Intensive Care Med* 2005;31:1345–55.
- Knaus WA, Wagner DP, Draper EA, Zimmerman JE, Bergner M, Bastos PG, Sirio CA, Murphy DJ, Lotring T, Damiano A. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619–36.
- Campbell-Furtick M, Moore BJ, Overton TL, Laureano Phillips J, Simon KJ, Gandhi RR, Duane TM, Shafi S. Post-trauma mortality increase at age 60: a cutoff for defining elderly? *Am J Surg* 2016;212:781–5.
- Mutschler M, Nienaber U, Brockamp T, Wafaisade A, Fabian T, Paffrath T, Bouillon B, Maegele M. TraumaRegister DGU. Renaissance of base deficit for the initial assessment of trauma patients: a base deficit-based classification for hypovolemic shock developed on data from 16,305 patients derived from the TraumaRegister DGU®. *Crit Care* 2013;17:R42.
- Neville AL, Nemtsev D, Manasrah R, Bricker SD, Putnam BA. Mortality risk stratification in elderly trauma patients based on initial arterial lactate and base deficit levels. *Am Surg* 2011;77:1337–41.
- Baker SP, O'Neill B, Haddon W, Long WB. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma* 1974;14:187–96.
- Gupta M, Gertz M, Schriger DL. Injury severity score inflation resulting from pan-computed tomography in patients with blunt trauma. *Ann Emerg Med* 2016;67:71–5.
- Joose P, Schep NW, Goslings JC. Injury profiles related to mortality in patients with a low injury severity score: a case-mix issue? *J Trauma Acute Care Surg* 2012;73:179–85.