Injury, Int. J. Care Injured 50 (2019) 1118–1124

Contents lists available at ScienceDirect

Injury

journal homepage: www.elsevier.com/locate/injury

External validation of a modified trauma and injury severity score model in major trauma injury

W.S. Chen[a,*], J.H. Tan[b], Y. Mohamad[b], R. Imran[b]

[a] Department of Statistics, Data Science and Epidemiology, Swinburne University of Technology, Melbourne, Australia
[b] General Surgery Department, Hospital Sultanah Aminah, Johor Bahru, Malaysia

ARTICLE INFO

Keywords:
Outcome prediction models
External validation in prognostic models
MTOS-TRISS model
NTrD-TRISS model

ABSTRACT

Background: The establishment of an accurate prognostic model in major trauma patients is important mainly because this group of patients will benefit the most. Clinical prediction models must be validated internally and externally on a regular basis to ensure the prediction is accurate and current. This study aims to externally validate two prediction models, the Trauma and Injury Severity Score model developed using the Major Trauma Outcome Study in North America (MTOS-TRISS model), and the NTrD-TRISS model, which is a refined MTOS-TRISS model with coefficients derived from the Malaysian National Trauma Database (NTrD), by regarding mortality as the outcome measurement.

Method: This retrospective study included patients with major trauma injuries reported to a trauma centre of Hospital Sultanah Aminah over a 6-year period from 2011 and 2017. Model validation was examined using the measures of discrimination and calibration. Discrimination was assessed using the area under the receiver operating characteristic curve (AUC) and 95% confidence interval (CI). The Hosmer-Lemeshow (H-L) goodness-of-fit test was used to examine calibration capabilities. The predictive validity of both MTOS-TRISS and NTrD-TRISS models were further evaluated by incorporating parameters such as the New Injury Severity Scale and the Injury Severity Score.

Results: Total patients of 3788 (3434 blunt and 354 penetrating injuries) with average age of 37 years (standard deviation of 16 years) were included in this study. All MTOS-TRISS and NTrD-TRISS models examined in this study showed adequate discriminative ability with AUCs ranged from 0.86 to 0.89 for patients with blunt trauma mechanism and 0.89 to 0.99 for patients with penetrating trauma mechanism. The H-L goodness-of-fit test indicated the NTrD-TRISS model calibrated as good as the MTOS-TRISS model for patients with blunt trauma mechanism.

Conclusion: For patients with blunt trauma mechanism, both the MTOS-TRISS and NTrD-TRISS models showed good discrimination and calibration performances. Discrimination performance for the NTrD-TRISS model was revealed to be as good as the MTOS-TRISS model specifically for patients with penetrating trauma mechanism. Overall, this validation study has ascertained the discrimination and calibration performances of the NTrD-TRISS model to be as good as the MTOS-TRISS model particularly for patients with blunt trauma mechanism.

Introduction

Development of a statistical model to predict the presence of disease and outcome occurrences of patients is a common practice in clinical research [1–3]. Generally, these models are developed using a multivariate regression approach, which resulting in a final equation for prediction purpose. Recent publications emphasised on the utilisation of three fundamental components in research involving prediction model, namely model development, external validation and impact evaluation [2–19]. Model development is the process that leads to the establishment of final prediction equation. External validation uses external data (not included in the model development) to further examine whether the model's prediction is reliable. Impact study evaluates whether the implementation of a prediction model in clinical practice actually improves the outcomes of patients [20].

In 1970, the Medical Aspect of Automotive Safety Committee coordinated the Abbreviated Injury Scale (AIS) [21,22]. The injury Severity Score (ISS) was calculated by summing up the squares of

* Corresponding author at: Department of Statistics, Data Sciences and Epidemiology, School of Health Sciences, Swinburne University of Technology, Level 9. Internal Mail H24, P.O. Box 218, Hawthorn, VIC 3122, Australia.
E-mail addresses: wchen@swin.edu.au (W.S. Chen), huei_87@hotmail.com (J.H. Tan), yuzamd72@yahoo.com (Y. Mohamad), ribajb@yahoo.com (R. Imran).

https://doi.org/10.1016/j.injury.2018.12.031
0020-1383/

the three highest AIS scores for injuries to different body regions [23,24]. The New Injury Severity Scale (NISS) was computed by summing up the patient's three largest AIS values irrespective of body region [25]. The initial Trauma and Injury Severity Score (MTOS-TRISS) was derived from both the anatomic and physiologic parameters from a population in North America [26–28]. In 2016, the coefficients of MTOS-TRISS model were refined with mortality as the outcome measurement by utilising five-year data from the National Trauma Database (NTrD) in Malaysia (called NTrD-TRISS model) [29].

The NTrD-TRISS model is regarded as the first in the South East Asian region to report local data-driven coefficients. In addition, a significant improvement in discriminant properties has been revealed for the NTrD-TRISS model in comparison to the MTOS-TRISS model. Given the complex computational nature of ISS, the MTOS-TRISS and NTrD-TRISS models by incorporating NISS as one of the parameters have been recommended for practical usage [29].

In practice, a shortage of external validation studies was primarily due to lack of data available besides those data used for model development. In addition, it has increasingly been recognised that the predictive performance of the final prediction equation tends to vary across settings, populations and periods [14,30–32]. Hence, formal assessment for heterogeneity in model performance across populations, settings, and periods was rarely conducted.

Therefore, the primary objective of this study was to externally validate the predictive validity of the MTOS-TRISS and NTrD-TRISS models with mortality as the outcome measurement by using data from a trauma centre of Hospital Sultanah Aminah located in the southern part of Malaysia. In addition, this study aimed to further evaluate the predictive validity of both MTOS-TRISS and NTrD-TRISS models by incorporating NISS and ISS.

Methods

Study population

The data included in this retrospective study was collected prospectively from patients aged 16 years and above, with ISS greater than 15, and admitted to the trauma centre of Hospital Sultanah Aminah, located in the southern part of Malaysia, from 2011 to 2017. Trained trauma nurses collected patients' information, this information was further verified by trauma surgeons prior to data analysis. Data was complete for all predictors required to calculate MTOS-TRISS and NTrD-TRISS models. This study was approved by the Malaysian Research Ethics Committee (NMRR-17-1744-36885) and the Swinburne University Human Research Ethics Committee (SHR 2017/297).

MTOS-TRISS versus NTrD-TRISS models

The MTOS-TRISS model was initially developed with mortality as the outcome measurement based on the MTOS in North America and it has globally been accepted as an effective assessment tool for more than 20 years [26–28]. It is a logistic regression model of survival probability, which takes account of age, physiological derangement using the Revised Trauma Score (RTS) [33] and anatomical severity of injury using ISS plus mechanism of injury [34,35].

The predictive validity of both MTOS-TRISS and NTrD-TRISS models with mortality as the outcome measurement was previously evaluated by incorporating NISS and ISS as parameters in the models [29]. MTOS-TRISS and NTrD-TRISS models can be calculated based on coefficients summarised in Table 1. Subsequently, the survival probability for these models can be computed as $P_s = 1/(1+e^{-b})$. Prior to the computation of the models, the parameters RTS, ISS, NISS, respiratory rate (RR), systolic blood pressure (SBP) and Glasgow Coma Scale (GCS) are converted to a score ranges from 0 to 4 (Table 2). If the patient's age is between 15 to 54 years, then, the parameter AgeIndex is coded as 0. Otherwise, the parameter AgeIndex is coded as 1 for patient aged 55 years and above.

Table 1
Details in calculating survival probability for MTOS-TRISS and NTrD-TRISS models.

Blunt trauma mechanism

$b_{MTOS\text{-}TRISS\text{-}ISS} = (-1.2470) + (0.9544)*(RTS) + (-0.0768)*(ISS) + (-1.9052)*(AgeIndex)$

$b_{MTOS\text{-}TRISS\text{-}NISS} = (-1.2470) + (0.9544)*(RTS) + (-0.0768)*(NISS) + (-1.9052)*(AgeIndex)$

$b_{NTrD\text{-}TRISS\text{-}ISS} = (-3.6763) + (-0.0148)*(ISS) + (-1.1413)*(AgeIndex) + (0.2431)*(RR) + (0.8061)*(SBP) + (0.6165)*(GCS)$

$b_{NTrD\text{-}TRISS\text{-}NISS} = (-3.6167) + (-0.0160)*(NISS) + (-1.1358)*(AgeIndex) + (0.2671)*(RR) + (0.8206)*(SBP) + (0.6071)*(GCS)$

Penetrating trauma mechanism

$b_{MTOS\text{-}TRISS\text{-}ISS} = (-0.6029) + (1.1430)*(RTS) + (-0.1516)*(ISS) + (-2.6676)*(AgeIndex)$

$b_{MTOS\text{-}TRISS\text{-}NISS} = (-0.6029) + (1.1430)*(RTS) + (-0.1516)*(NISS) + (-2.6676)*(AgeIndex)$

$b_{NTrD\text{-}TRISS\text{-}ISS} = (-3.6763) + (-0.0148)*(ISS) + (-1.1413)*(AgeIndex) + (0.2431)*(RR) + (0.8061)*(SBP) + (0.6165)*(GCS)$

$b_{NTrD\text{-}TRISS\text{-}NISS} = (-3.6167) + (-0.0160)*(NISS) + (-1.1358)*(AgeIndex) + (0.2671)*(RR) + (0.8206)*(SBP) + (0.6071)*(GCS)$

AgeIndex = Patient's Age Index (0 = 15–54 years, 1 = 55 years and above); SBP = Systolic Blood Pressure; GCS = Glasgow Coma Score; ISS = Injury Severity Score; NISS = New Severity Injury Scale; RR = Respiratory Rate; RTS = Revised Trauma Score; TRISS = Trauma and Injury Severity Score; MTOS = Major Trauma Outcome Study; NTrD = National Trauma Database;
The formulae for bMTOS-TRISS-ISS and bMTOS-TRISS-NISS were sourced from Champion et al. [26], Boyd et al. [27] and Champion et al. [28].
The formulae for bNTrD-TRISS-ISS and bNTrD-TRISS-NISS were sourced from Chen et al. [29].

Statistical analysis

Validation was examined using the measures of discrimination and calibration [36]. Discrimination was assessed using the area under the receiver operating characteristic curve (AUC) and 95% confidence interval (CI). By convention, an AUC $\leq$ 0.5 is regarded to represent a non-discriminative model, whereas an AUC $\geq$ 0.8 is considered to represent adequate discriminative ability by the model.

On the other hand, calibration was examined using the Hosmer-Lemeshow (H-L) goodness-of-fit test, which tests the null hypothesis that the model's estimates fit the observed data. In theory, The H-L test is a chi-square test conducted by sorting the n

Table 2
Conversion of parameters in MTOS-TRISS and NTrD-TRISS models from original value to a score between 0 and 4.

Parameter	Value	Score
Respiratory rate (breaths/min) [RR]	10–29	4
	> 29	3
	6–9	2
	1–5	1
	0	0
Systolic blood pressure (mmHg) [SBP]	> 89	4
	76–89	3
	50–75	2
	1–49	1
	0	0
Glasgow Coma Scale [GCS]	13–15	4
	9–12	3
	6–8	2
	4–5	1
	3	0

records in the dataset by estimated probability of success dividing the sorted set into g equal-sized groups, and evaluating the H-L statistic [37]. Since the power increases with sample size, therefore, it can be undesirable for the H-L goodness-of-fit tests because in very large dataset, small departures from the proposed model will be considered significant. In order to achieve a reliable estimate, it is desirable for g to fulfil the following conditions: 1) $g < n/5$ (at least 5 observations per group); 2) $g \geq 6$ in order for the H-L statistic to be distributed approximately as chi-square with (g-2) degree of freedom; 3) chi-square approximation fails when the event rate is small and g is large. Therefore, the value of g was calculated based on the following equation where n denotes sample size [37].

$$g = 2 + 8\left(\frac{n}{1000}\right)^2$$

A p-value is computed based on a chi-square distribution, to test the fit of the prognostic model. The H-L goodness-of-fit test was applied directly to the original predictive formulae summarised in Table 1. If the p-value of H-L goodness-of-fit test is larger than 0.05, then, we would fail to reject the null hypothesis. A p-value of less than 0.05 (two-sided) was deemed to be statistically significant. All statistical analyses were performed using STATA Intercool version 13 (Stata Corp, College Station, TX).

Results

Study population

This study included 3788 major trauma cases reported during 2011 to 2017 (Fig. 1). The age ranged from 15 to 91 years old, with a median age of 33 years (interquartile range of 24–53 years). Included patients were predominantly males (89%) with injuries sustained from a blunt mechanism (91%), with majority of these patients (89%) were alive when discharged. Overall, deceased patients had almost 2 folds higher median ISS and NISS scores but lower median RTS value in comparison to alive patients (Table 3).

NTrD-TRISS model's performance

Table 4 summarises the discrimination and calibration results for both MTOS-TRISS and NTrD-TRISS models by the mechanism of injury. Both models performed well with AUCs ranging from 0.86 to 0.89 for patients with blunt trauma mechanism and 0.89 to 0.99 for patients with penetrating trauma mechanism (Figs. 2a, b, 3).

For blunt mechanism, p-values from all models indicated these models' estimates fit the observed data well (p-value = 0.524 for MTOS-TRISS-ISS; p-value = 0.242 for MTOS-TRISS-NISS; p-value = 0.719 for NTrD-TRISS-ISS; p-value = 0.765 for NTrD-TRISS-NISS) (Table 4).

Due to only 352 patients reported with penetrating trauma mechanism, the estimated g value was 3, which was below the minimum value of 6 in order for the chi-square approximation to be valid [37]. In addition, only 8 (2%) deaths were observed out of the 352 patients, therefore, chi-square approximation was not valid for such a small mortality rate [37]. Hence, no H-L statistics and p-values were reported for patients with penetrating trauma mechanism in Table 4.

Discussion

Our findings showed that both the MTOS-TRISS and NTrD-TRISS models performed adequately when used to predict outcomes in

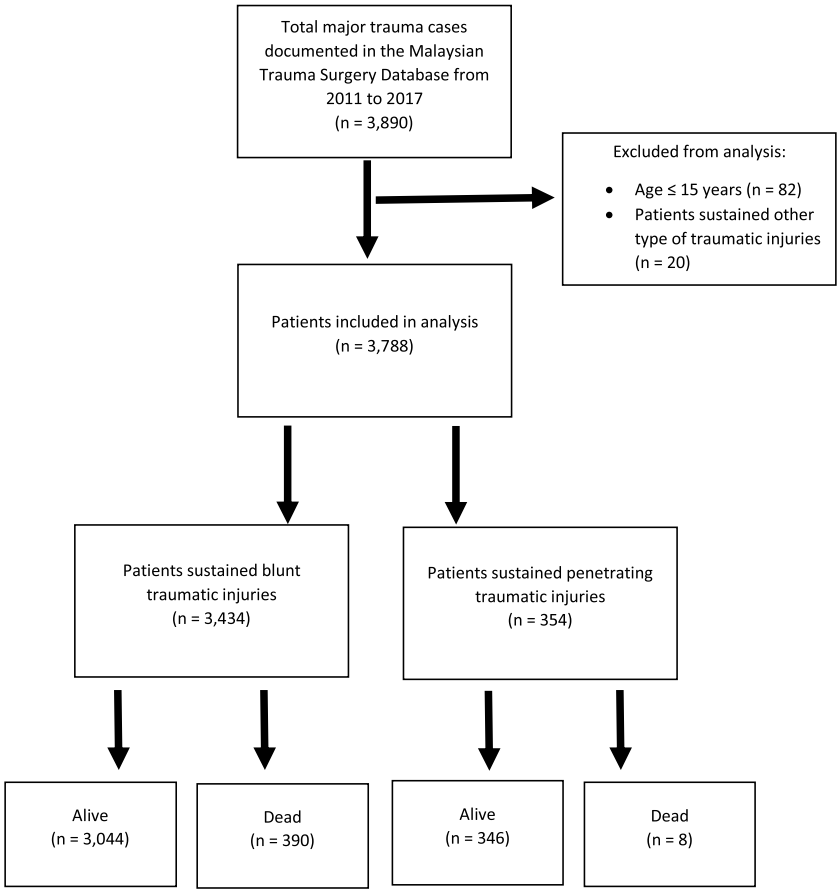


Fig. 1. Flow of patients.

Table 3
General characteristics of patients.

Variable	Alive (n = 3390)	Dead (n = 398)	All (n = 3788)
Male sex [n (%)]	3001 (89%)	355 (89%)	3356 (89%)
Age (years)	36.3 (15.6)	39.5 (18.5)	36.6 (16.0)
• Mean (SD)	33 (23–47)	36 (24–53)	33 (23–48)
• Median (IQR)			
Blunt trauma [n (%)]	3044 (90%)	390 (98%)	3434 (91%)
SBP (mmHg)	126.6 (24.0)	120.5 (36.8)	125.9 (25.7)
• Mean (SD)	125 (111–140)	117 (95–140)	124 (110–140)
• Median (IQR)			
GCS	13.8 (2.8)	9.0 (4.8)	13.3 (3.4)
• Mean (SD)	15 (15–15)	8 (4–15)	15 (14–15)
• Median (IQR)			
ISS	15.8 (9.8)	30.8 (12.2)	17.4 (11.0)
• Mean (SD)	16 (9–24)	32 (25–41)	16 (9–25)
• Median (IQR)			
NISS	16.4 (11.3)	35.5 (13.3)	18.4 (13.0)
• Mean (SD)	13 (9–22)	34 (25–43)	16 (9–27)
• Median (IQR)			
RTS	7.6 (0.7)	6.0 (1.6)	7.4 (0.10)
• Mean (SD)	7.8 (7.8–7.8)	6.0 (4.7–7.8)	7.8 (7.8–7.8)
• Median (IQR)			

IQR = Interquartile Range; SD = Standard Deviation; SBP = Systolic Blood Pressure; GCS = Glasgow Coma Score; ISS = Injury Severity Score; NISS = New Severity Injury Scale; RTS = Revised Trauma Score.

Table 4
Logistic regression models with discrimination and calibration.

Model	AUC (95% CI)	H-L statistic	p-value
(a) Blunt trauma mechanism (n = 3434, g = 96)			
• NTrD-TRISS-ISS	0.86 (0.83, 0.89)	41.96	0.719
• NTrD-TRISS-NISS	0.87 (0.84, 0.89)	44.36	0.765
• MTOS-TRISS-ISS	0.88 (0.86, 0.91)	46.76	0.524
• MTOS-TRISS-NISS	0.89 (0.86, 0.91)	58.75	0.242
(b) Penetrating trauma mechanism (n = 354, g = 3*)			
• NTrD-TRISS-ISS	0.90 (0.82, 0.98)	NA+	NA+
• NTrD-TRISS-NISS	0.89 (0.82, 0.97)	NA+	NA+
• MTOS-TRISS-ISS	0.98 (0.95, 1.00)	NA+	NA+
• MTOS-TRISS-NISS	0.99 (0.97, 1.00)	NA+	NA+

AUC = Area under the receiver operating characteristic curve; H-L statistic = Hosmer-Lemeshow statistic; p-value = p-value was calculated based on a chi-square distribution to test the fit of the model; NA = Not Applicable;

* The initial estimated value of g was 3 due to small sample size. The smallest value of g has to be at least 6 in order for the chi-square approximation to be valid [37].

+ Since only 8 deaths (2%) were recorded for penetrating trauma mechanism, therefore, chi-square approximation was not valid. Hence, no H-L statistic and p-value were reported [37].

the trauma centre of Hospital Sultanah Aminah. The calibration capability for penetrating trauma mechanism was not reported due to invalid chi-square approximation in such a small sample size and low mortality rate. Overall, the NTrD-TRISS model performance seems to be as good as the performance of the MTOS-TRISS model in this external validation study particularly for patients with blunt trauma mechanism. This finding was consistent with a previous study [29].

Comparing this dataset with the original NTrD dataset that was used in developing the NTrD-TRISS model, generally patients in the current dataset were predominantly younger males, with lower average ISS and average NISS scores, similar average SBP, and slightly higher average GCS score. Specifically, in the current dataset, the average age was 36.6 years (standard deviation of 16.0 years), with overall mortality rate of 11%, approximately 91% blunt trauma mechanism and 9% penetrating trauma mechanism. In contrast, in the original NTrD dataset used to develop the NTrD-TRISS model, the average age was 34.6 years (standard deviation of 16.2 years), with overall mortality rate of 27%, approximately 97% blunt trauma mechanism and 3% penetrating trauma mechanism. This study emphasised on major trauma patients mainly because this group of patients will benefit the most from an accurate prognostic model.

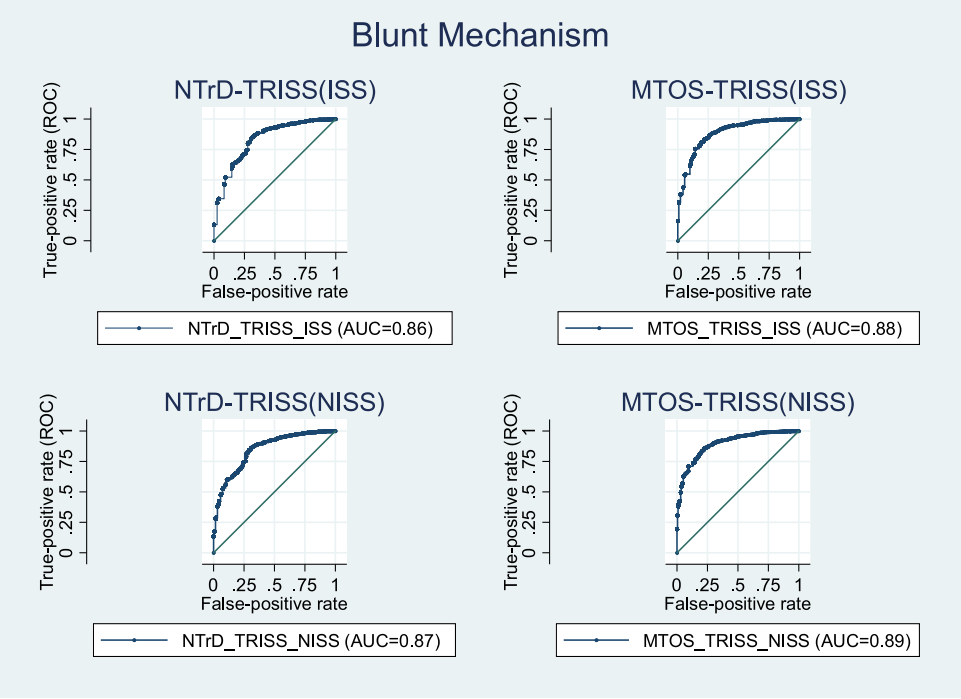
Differences between the current dataset and the original NTrD dataset may likely to influence the validation of this prognostic model [38]. In addition, the mortality rate in the current dataset was approximately 2.5 folds lower than the original NTrD dataset. Despite the differences in these datasets, the NTrD-TRISS model was consistently found to have good discrimination for both blunt and penetrating trauma mechanisms while only good calibration was found for patients with blunt trauma mechanism.

Strength and weakness of study

This study offers several strengths; it is the first in the South East Asian region which reported local data-driven coefficients in the calculation of TRISS (NTrD-TRISS model). Furthermore, this study has ascertained the discrimination and calibration of the NTrD-TRISS model to be as good as the MTOS-TRISS model particularly for patients with blunt trauma mechanism. The computation of the NISS was much simpler compared to the ISS, this study revealed the overall performance of the NTrD-TRISS model with NISS was similar to the NTrD-TRISS model with ISS. Currently, only MTOS-TRISS model is utilised via online calculator in the clinical care of patients in the trauma centre of Hospital Sultanah Aminah, Malaysia. Therefore, the promising performance of the NTrD-TRISS model suggested it could easily be adopted and applied as an audit tool to monitor and compare the performance of trauma centres in Malaysia. Additionally, the NTrD-TRISS model will help to predict in-hospital mortality and potentially guide the resource allocation as there is currently limited bed specifically in the intensive care unit. Precise prediction may help to inform prognosis to family and clinician to decide on early withdrawal of severely injured cases, which is likely to be irreversible

Several limitations existed in this study. Firstly, the dataset utilised in this validation study was captured from a single trauma centre. Therefore, this study unable to investigate the possible centre level effects in the outcome prediction of models examined. Nevertheless, the trauma centre of Hospital Sultanah Aminah is regarded as the next largest centre managing traumatic injuries after the NTrD in Malaysia [39]. Therefore, this representative sample enabled the NTrD-TRISS model to be validated externally. In order to generalise the NTrD-TRISS model, it is essential for the NTrD-TRISS model to be examined in numerous and diverse settings [40]. In addition, it is equally important to validate prognostic model externally on regular basis to ensure the model remains up-to-date. The second limitation presence in this study was the low number of patients reported with penetrating trauma mechanism, which has subsequently led to low mortality rate that made chi-square approximation impossible. This limitation has ultimately restricted the assessment of the calibration capability of the MTOS-TRISS and NTrD-TRISS models for patients with penetration trauma mechanism.

(a)



(b)

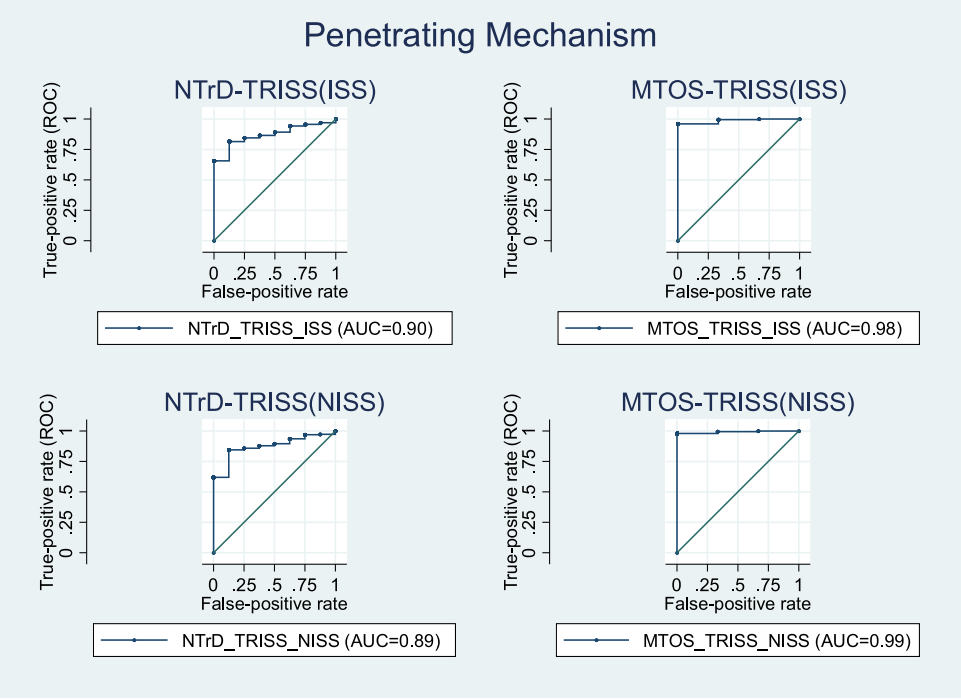


Fig. 2. a) AUC for NTrD-TRISS and MTOS-TRISS models for patients sustained blunt traumatic injuries. b) AUC for NTrD-TRISS and MTOS-TRISS models for patients sustained penetrating traumatic injuries.

Future research

Currently, more than 90% of patients are reported with blunt trauma mechanism. Therefore, it is extremely important to have an accurate prognostic model that could assist physicians in their real-time assessment in accurately predicting patient's outcome to ensure efficient triaging of patients. This study failed to assess the calibration capability of both NTrD-TRISS and MTOS-TRISS models for patients with penetrating trauma mechanism due to low number of cases reported. Thus, future research is warranted to ascertain the overall performance of the NTrD-TRISS model especially for patients with penetrating trauma mechanism. In addition, assessment of

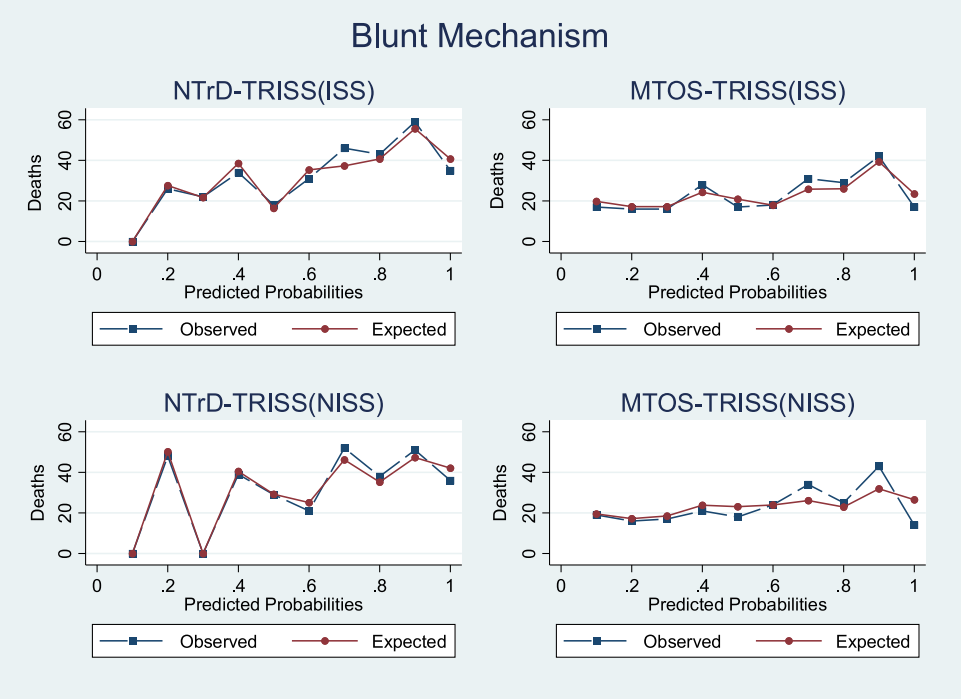


Fig. 3. Observed versus Expected Number of Deaths According to Predicted Probabilities.

clinical utility of the NTrD-TRISS model in a mobile application format will be the essential component of model evaluation in the next phase of study. The clinical implications of the findings from the current study enable physicians to provide accurate real-time assessments to patients in their clinical care and potentially improve the outcomes of patients.

Conclusion

This study has successfully validated the MTOS-TRISS and NTrD-TRISS models using a dataset collected from a single trauma centre. The overall discrimination and calibration performances of these models by incorporating NISS as a parameter were found to be consistent with the same models by incorporating ISS as a parameter. Overall, this validation study has ascertained the discrimination and calibration performances of the NTrD-TRISS model to be as good as the MTOS-TRISS model particularly for patients with blunt trauma mechanism.

Conflict of interest and disclosure

There is no conflict of interest to disclose.

Acknowledgements

The authors would like to thank Swinburne University of Technology for financially supporting a trauma nurse through a seed grant to complete data entry for this study. In addition, the authors thank the Director General of Health Malaysia for his approval to publish this data.

References

[1] Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer; 2009, doi:http://dx.doi.org/10.1007/978-0-387-77244-8.
[2] Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009;338:b604, doi:http://dx.doi.org/10.1136/bmj.b604.
[3] Steyerberg EW, Moons KG, van der Windt DA, et al. PROGRESS group. Prognosis research strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381, doi:http://dx.doi.org/10.1371/journal.pmed.1001381.
[4] Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605, doi:http://dx.doi.org/10.1136/bmj.b605.
[5] Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606, doi:http://dx.doi.org/10.1136/bmj.b606.
[6] Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ 2009;338:b375, doi:http://dx.doi.org/10.1136/bmj.b375.
[7] Hemingway H, Croft P, Perel P, et al. PROGRESS Group. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ 2013;346:e5595, doi:http://dx.doi.org/10.1136/bmj.e5595.
[8] Riley RD, Hayden JA, Steyerberg EW, et al. PROGRESS group. Prognosis research strategy (PROGRESS) 2: prognostic factor research. PLoS Med 2013;10:e1001380, doi:http://dx.doi.org/10.1371/journal.pmed.1001380.
[9] Hingorani AD, Windt DA, Riley RD, et al. PROGRESS Group. Prognosis research strategy (PROGRESS) 4: stratified medicine research. BMJ 2013;346:e5793, doi:http://dx.doi.org/10.1136/bmj.e5793.
[10] Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events [correction in BMJ 2016;353:i3235]. BMJ 2015(351)h3868, doi:http://dx.doi.org/10.1136/bmj.h3868.
[11] Moons KG, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart 2012;98:683–90, doi:http://dx.doi.org/10.1136/heartjnl-2011-301246.
[12] Harrell FE. Regression modeling strategies, with applications to linear models, logistic regression, and survival analysis. Springer; 2001.
[13] Bleeker SE, Moll HA, Steyerberg EW, et al. External validation is necessary in prediction research: a clinical example. J Clin Epidemiol 2003;56:826–32, doi:http://dx.doi.org/10.1016/S0895-4356(03)00207-5.
[14] Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015;68:279–89, doi:http://dx.doi.org/10.1016/j.jclinepi.2014.06.018.
[15] Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. BMC Med 2010;8:21, doi:http://dx.doi.org/10.1186/1741-7015-8-21.
[16] Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann Intern Med 2006;144:201–9, doi:http://dx.doi.org/10.7326/0003-4819-144-3-200602070-00009.
[17] Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med 2012;9:e1001221, doi:http://dx.doi.org/10.1371/journal.pmed.1001221.
[18] Wyatt J, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? BMJ 1995;311:1539–41, doi:http://dx.doi.org/10.1136/bmj.311.7019.1539.

[19] Collins GS, Michaëlsson K. Fracture risk assessment: state of the art, methodologically unsound, or poorly reported? Curr Osteoporos Rep 2012;10:199–207, doi:http://dx.doi.org/10.1007/s11914-012-0108-1.
[20] Riley RD, Ensor J, Snell KIE, Debray TPA, Altman DG, Moons KGM, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ 2016;22(June (353))i3140.
[21] Committee on Medical Aspects of Automotive Safety. Rating the severity of tissue damage: I. The abbreviated scale. JAMA 1972;215:277–80.
[22] Committee on Medical Aspects of Automotive Safety. Rating the severity of tissue damage: II. The comprehensive scale. JAMA 1972;220:717–20.
[23] Baker SP, o'Neill B, Haddon Jr. W, et al. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma Acute Care Surg 1974;14:187–96.
[24] Baker SP, O'Neill B. The injury severity score: an update. J Trauma 1976;16:882–5.
[25] Osler T, Baker SP, Long W. A modification of the injury severity score that both improves accuracy and simplifies scoring. J Trauma 1997;43:922–6.
[26] Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. Trauma score. Crit Care Med 1981;9:672–6.
[27] Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. J Trauma 1987;27(370–378):6.
[28] Champion HR, Copes WS, Sacco WJ, et al. The Major Trauma Outcome Study: establishing national norms for trauma care. J Trauma 1990;30:1356–65.
[29] Chen WS, Lee SWH, Jamaluddin S, Wong CP. Comparison of Trauma and Injury Severity Score model with alternative approach in outcome prediction in trauma using National Trauma Database in Malaysia. Trauma 2016;19(April (2)):103–12.
[30] Debray TPA, Riley RD, Rovers MM, Reitsma JB, Moons KG. Cochrane IPD Meta-analysis Methods group. Individual participant data (IPD) meta-analyses of diagnostic and prognostic modeling studies: guidance on their use. PLoS Med 2015;12:e1001886, doi:http://dx.doi.org/10.1371/ journal.pmed.1001886.
[31] Royston P, Parmar MKB, Sylvester R. Construction and validation of a prognostic model across several studies, with an application in superficial bladder cancer. Stat Med 2004;23:907–26, doi:http://dx.doi.org/10.1002/sim.1691.
[32] Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: use of benchmark values to disentangle a case-mix effect from incorrect coefficients. Am J Epidemiol 2010;172:971–80, doi:http://dx.doi.org/10.1093/aje/kwq223.
[33] Moore L, Lavoie A, LeSage N, et al. Statistical validation of the revised trauma score. J Trauma Acute Care Surg 2006;60:305–11.
[34] Baker SP, o'Neill B, Haddon Jr. W, et al. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma Acute Care Surg 1974;14:187–96.
[35] Baker SP, O'Neill B. The injury severity score: an update. J Trauma 1976;16:882–5.
[36] Mushkudiani NA, Hukkelhoven CW, Hernandez AV, Murray GD, Choi SC, Maas AI, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. J Clin Epidemiol 2008;61:331–43.
[37] Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. Statist Med 2013;32:67–80, doi:http://dx.doi.org/10.1002/sim.5525.
[38] Vergouwe Y, Moons KGM, Steyerberg EW. External validity of risk models: use of benchmark values to disentangle a casemix effect from incorrect coefficients. Am J Epidemiol 2010;172(971):980.
[39] Tan JH, Tan HCL, Md Noh NA, Mohamad Y, Alwi RI. Validation of the trauma mortality prediction scores from a Malaysian population. Burns Trauma 2017;5(37):1–6, doi:http://dx.doi.org/10.1186/s41038-017-0102-z.
[40] Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999;130:515–24.