Development of a new score for early mortality prediction in trauma ICU patients: RETRASCORE

Luis Serviá1, Juan Antonio Llompart-Pou2, Mario Chico-Fernández3, Neus Montserrat1, Mariona Badia1, Jesús Abelardo Barea-Mendoza3, María Ángeles Ballesteros-Sanz4 and Javier Trujillano1,5*

on behalf of the Neurointensive Care and Trauma Working Group of the Spanish Society of Intensive Care Medicine (SEMICYUC)

Abstract
Background: Severity scores are commonly used for outcome adjustment and benchmarking of trauma care provided. No specific models performed only with critically ill patients are available. Our objective was to develop a new score for early mortality prediction in trauma ICU patients.

Methods: This is a retrospective study using the Spanish Trauma ICU registry (RETRAUCI) 2015–2019. Patients were divided and analysed into the derivation (2015–2017) and validation sets (2018–2019). We used as candidate variables to be associated with mortality those available in RETRAUCI that could be collected in the first 24 h after ICU admission. Using logistic regression methodology, a simple score (RETRASCORE) was created with points assigned to each selected variable. The performance of the model was carried out according to global measures, discrimination and calibration.

Results: The analysis included 9465 patients: derivation set 5976 and validation set 3489. Thirty-day mortality was 12.2%. The predicted probability of 30-day mortality was determined by the following equation: \( \frac{1}{1 + \exp(-y)} \), where \( y = 0.598 \times \text{Age 50–65} + 1.239 \times \text{Age 66–75} + 2.198 \times \text{Age > 75} + 0.349 \times \text{PRECOAG} + 0.336 \times \text{Pre-hospital intubation} + 0.662 \times \text{High-risk mechanism} + 0.950 \times \text{unilateral mydriasis} + 3.217 \times \text{bilateral mydriasis} + 0.841 \times \text{Glasgow ≤ 8} + 0.495 \times \text{MAIS-Head} - 0.271 \times \text{MAIS-Thorax} + 1.148 \times \text{Haemodynamic failure} + 0.708 \times \text{Respiratory failure} + 0.567 \times \text{Coagulopathy} + 0.580 \times \text{Mechanical ventilation} + 0.452 \times \text{Massive haemorrhage} - 5.432 \). The AUROC was 0.913 (0.903–0.923) in the derivation set and 0.929 (0.918–0.940) in the validation set.

Conclusions: The newly developed RETRASCORE is an early, easy-to-calculate and specific score to predict in-hospital mortality in trauma ICU patients. Although it has achieved adequate internal validation, it must be externally validated.

Keywords: Trauma and injury severity score, Intensive care unit, Mortality, Prognostic scoring systems

Background
Severe trauma remains the leading cause of mortality and disability in young adults [1]. In this setting, trauma registries provide relevant information in terms of benchmarking of the care provided and, therefore, constitute a relevant contribution to quality assessment and scientific...
research in an area where classical randomized trials are difficult to perform [2–4]. To this purpose, the Injury Severity Score (ISS) has been the most commonly used score to assess severity of trauma [5]. Several years later, the Trauma and Injury Severity Score (TRISS) became the most frequently used tool for outcome adjustment and benchmarking in worldwide trauma registries [6].

More recently, The TraumaRegister DGU™ (TR-DGU) developed the Revised Injury Severity Classification (RISC) score [7] and its updated version in 2014 [8] for outcome adjustment, achieving an astonishing AUC 0.95 in the derivation and validation datasets in terms of mortality prediction, with appropriate precision and calibration [8]. However, these scores can be applied to the general trauma patients and, therefore, are not specific for the trauma ICU patients, in whom the physiological consequences of trauma itself play a major role in outcomes. Indeed, intensive care units (ICUs) commonly use severity scores such as the Acute Physiology and Chronic Health Evaluation (APACHE II), the Simplified Acute Physiologic Score (SAPS II) or the Mortality Probability Models (MPM II), which take into consideration age, comorbidities and the physiological burden of critical illness rather than anatomic considerations [9].

Due to the lack of a specific trauma ICU score and by using data from the Spanish Trauma ICU registry (RETRAUCI), our objective was to develop a new score for early mortality prediction in trauma ICU patients.

Methods
This is a retrospective study that aims to develop and validate a mortality prognostic model with variables included in the RETRAUCI project. RETRAUCI is an observational, prospective and multicentre nationwide registry that currently includes 52 ICUs in Spain. It has the endorsement of the Neurointensive Care and Trauma Working Group of the Spanish Society of Intensive Care Medicine (SEMICYUC) and currently operates in a web-based electronic format (www.retrauci.org). Additional file 1 shows the screenshots of the web application and the list of variables collected in RETRAUCI. The records for the years 2015–2019 were used. To achieve internal validation, the total records were divided into two sets: derivation set (2015–2017) and validation set (2018–2019). It is a study with complete-case analysis with temporal validation. The development of the models was carried out following the recommendations established in the Transparency Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) initiative [10].

Ethics Committee approval for the registry was obtained (Hospital Universitario 12 de Octubre, Madrid: 12/209). Due to the retrospective analysis of de-identified collected data, informed consent was not obtained for this study.

Patients and variables included
Patients were managed according to the Advanced Trauma Life Support principles. Data on epidemiology, acute management in the pre-hospital and in-hospital settings, type and severity of injury, resources utilization, complications and outcomes were recorded. We only excluded patients with missing data about in-hospital mortality. Possible candidate variables have been selected according to clinical, bibliographic and availability criteria in the RETRAUCI database. The candidate variables must also be available within the first 24 h of admission to the ICU. The variables entered were then analysed according to different categories:

- Patient-related variables included sex, age and prior antiplatelet/anticoagulant treatment. Sex was treated as a dichotomous variable (male/female), age was distributed in four categories (less than 50 years, 50 to 65 years, 66 to 75 years and older than 75 years) and if the patient was on chronic treatment with antiplatelets or anticoagulants he/she was considered to have prior coagulation alteration (PRECOAG) [11].

- Pre-hospital care variables included pre-hospital medical attention, pre-hospital intubation and mechanism of trauma, which differentiates penetrating vs. non-penetrating types. Additionally, we coded as a high-risk mechanism trauma those mechanisms with associated mortality higher than 20%. This category included gunshot wounds, pedestrian falls, accidental falls, suicidal precipitation and those considered as unknown mechanism.

- Physiological variables: pupillary size and reactivity (normal, unilateral mydriasis and bilateral mydriasis), and score of the Glasgow Coma Scale (absolute score and percentage of patients with ≤ 8 points).

- Anatomical variables describing the severity of injuries according to the Abbreviated Injury Scale (AIS) were considered. The AIS ranges from 0 to 6, with 0 indicating no involvement and 6 indicating maximum involvement [12]. A major organic involvement was considered with a score of 3 or higher (MAIS) in any of the following six anatomical areas: head (MAIS-Head), thorax (MAIS-Thorax), abdomen (MAIS-Abdomen), upper extremity (MAIS-Ext Upper), lower extremity (MAIS-Ext Lower) and external and thermal injuries (MAIS-External).

- Organ failure-related variables were also considered: haemodynamic failure indicated by systolic blood pressure lower than 90 mmHg requiring the administration of volume, blood products and vasoactive support; respiratory failure, indicated by pO2/FiO2 below 200; and coagulopathy, indicated by the prolongation of prothrombin...
and activated partial thromboplastin times in > 1.5 times the control or by levels of fibrinogen < 150 mg/dL or thrombocytopenia < 100,000/µL in the determination of the first 24 h [13].

Treatment variables included the need of mechanical ventilation and the activation of the massive transfusion protocol because of a massive haemorrhage [14].

**Outcome definition**
The outcome variable was defined as 30-day mortality after trauma. Patients who were discharged from the hospital alive before 30 days after trauma were assumed to have survived for at least 30 days.

The ICU length of stay (LOS) was also collected to compare the derivation and validation sets. The probability of death (1-probability of survival) according to the TRISS score was used as a comparison model [15].

**Statistical analysis**
The sample size calculation helped us to verify that there were enough records for the development and validation of the model. For each possible factor, 10 deaths are needed. With a mortality of 12% and 20 variables as potential risk factors, at least \( (20 \times 10/0.12) = 1666 \) records are needed for the derivation and validation sets [16].

Categorical variables were described as percentage and continuous variables as median (interquartile range), as they did not follow a normal distribution (Kolmogorov–Smirnov test). For the comparison between the groups derivation–validation and survivors–non-survivors, the Mann–Whitney test was used for continuous variables and the Chi-square test for categorical variables. A \( p < 0.05 \) was considered statistically significant.

In the derivation set, a multivariable logistic regression model was used to determine predictors for 30-day mortality. We use the LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression algorithm in order to obtain a subset of predictor variables from the 20 candidate variables. The LASSO algorithm can select from the set of candidate variables that achieve greater importance once regularized. The LASSO algorithm finds the variables that contribute the least in the logistic regression model and forces them to have coefficients equal to zero. In this way, only the significant variables will be part of the final model [17]. This subset of predictor variables was used to carry out the logistic regression model and the odds ratios with their 95% confidence intervals, and the \( \beta \)-coefficients of each factor were calculated.

To check the stability in the selection of variables and in the calculation of the coefficients of the logistic regression, an analysis with random partitions of the derivation and validation sets was carried out. One hundred sets of derivation have been created randomly with the same number of records as those used in the temporal validation (\( n = 5976 \)). The LASSO variable selection methodology has been applied to each of these sets, and the corresponding logistic regression coefficients have been calculated. These results were compared with those obtained in the temporal validation.

Internal validity and adjustment for overfitting of the model were performed with a bootstrapping procedure. One thousand bootstrap samples were drawn from the derivation set. A shrinkage factor that multiplied the \( \beta \)-coefficients of the predictive factors and made them adjusted was calculated. These \( \beta \)-coefficients were used to calculate the individual probability of death in the derivation and validation sets.

A simple score (RETRASCORE) was developed based on predictors that were associated with 30-day mortality in the multivariate analysis. Score points were defined multiplying the regressions \( \beta \)-coefficients by 2 and rounded them to the nearest integer. The sum of the points is the value of the final score.

The performance of the models, logistic regression (LR), TRISS and RETRASCORE was determined in the derivation and validation sets. Global measure was used with the Brier score, discrimination measures using the area under the ROC curve (AUROC) with 95% CI, and calibration with calibration plots (the mean of the predicted probabilities was computed for each risk decile) and the calculation of the values of the fitted lines with intercept and slope with 95% CI [18].

The calculations were performed using STATA software, version 15.0 (Stata Corporation, College Station, Texas, USA) and R statistics 4.0.3 with the “glmnet” package (R Foundation for Statistical Computing, Vienna, Austria) [19].

**Results**
The study group included a total of 9465 patients with complete data who were divided into a derivation set with 5976 patients and a validation set with 3489 patients. Additional file 2: Figure S1 shows the flow-chart of the study. Thirty-day mortality reached 12.2%. Table 1 shows the variables potentially associated with mortality selected from the RETRAUCI database, differentiating between the derivation and validation sets. The only significant differences found in the validation set were a higher percentage of pre-hospital medical attentions (80.1 vs. 76.1%), a higher proportion of assaults (6.0 vs. 4.3%) and a lower need of mechanical ventilation (49.8 vs. 45.9%).
Table 2 shows the twenty candidate variables according to 30-day mortality in the derivation set. The LASSO regression analysis determined that 13 variables, with β-coefficients other than zero, were included in the logistic regression model (Additional file 2: Figure S2).
Table 3 shows the final multivariate logistic regression (LR) model with internal validation (1000 bootstrap samples) with the 13 selected variables. β-coefficients multiplied by a shrinkage factor of 0.98.

The predicted probability of 30-day mortality was determined by the following equation:

\[
\text{Prob 30-day mortality} = \frac{1}{1 + \exp(-y)},
\]

where

- \( y = 0.598 \, \text{(Age 50–65)} + 1.239 \, \text{(Age 66–75)} + 2.198 \, \text{(Age > 75)} + 0.349 \, \text{(PRECOAG)} + 0.336 \, \text{(Pre-hospital intubation)} + 0.662 \, \text{(High risk mechanism)} + 0.950 \, \text{(Unilateral mydriasis)} + 3.217 \, \text{(Bilateral mydriasis)} + 0.841 \, \text{(GCS \leq 8)} + 0.495 \, \text{(MAIS-Head)} - 0.271 \, \text{(MAIS-Thorax)} + 1.148 \, \text{(Haemodynamic failure)} + 0.708 \, \text{(Respiratory failure)} + 0.567 \, \text{(Coagulopathy)} + 0.580 \, \text{(Mechanical ventilation)} + 0.452 \, \text{(Massive haemorrhage)} - 5.432 \).

Predictor value is one when present and zero when absent.

Additional file 3 shows the stability of the model in the selection of variables and in the coefficients calculated using the comparison between the temporal validation and the use of random partitions.

RETRASCORE model

All factors associated with 30-day mortality in the LR model were used to develop the RETRASCORE model. Table 3 shows the points assigned to each factor. The sum of the different points constitutes the total score of the RETRASCORE. Figure 1 shows the probability of death, in the derivation and validation sets, according to the total score. Score values above 13 have been grouped by the limited number of records.

Performance analysis of models

The LR model and the RETRASCORE models were analysed, and the TRISS model was used as a comparison. Additional file 2: Figures S3, S4 and S5 show the values of global performance (Brier score) in the three models, which was similar in the derivation and validation sets and amongst all of them; discrimination with the AUROC with higher values (from the DeLong test with \( p < 0.001 \)) of the LR model and RETRASCORE (higher than 0.9) compared to the TRISS. The calibration is shown as a calibration curve in risk deciles and intercept and slope fit values, which were acceptable in the LR and RETRASCORE models and overestimated in the TRISS.

Figure 2 shows the summary sheet for the RETRASCORE model. The predicted probability of death was grouped into six categories: very low (less than 5%) with less than 5 points, low (10–15%) with 5–6 points, medium (20–30%) with 7–8 points, high (40–60%) between 9 and 11 points, very high (70–80%) with 12–13 points, and extreme (higher than 90%) with > 13 points. Figure 3 shows these separate probability categories in the derivation and validation sets.

Discussion

The appropriate use of trauma scoring is of paramount importance in terms of outcome adjustment and benchmarking of the care provided [20]. The ideal score for stratifying the risk of death in the trauma ICU should be performed early, must be easy-to-calculate and include specific factors derived from the intensive care management of critical trauma patients. However, classical trauma-specific scoring systems (ISS, RTS, TRISS) are still used with a poor calibration and discrimination capacity, and their performance is lower than that obtained by general ICU scores where anatomical injuries caused by trauma are not considered [21, 22], raising
the need of specific trauma ICU scores that take into consideration early anatomic consideration but also pre-hospital care, physiological derangements, organ failures and treatments provided. We believe that the model presented meets these expectations [23].

Different types of variables were finally included in our model, according to its relationship with mortality in conventional models.

Non-modifiable factors depending on the patient were age and the prior treatment of antiplatelet/anticoagulant therapy. Age is a critical factor in the outcomes of trauma patients, as mortality increases with age [8, 24]. On the other side, prior treatment with antiplatelets/anticoagulants contributes to the size and burden of traumatic injuries, especially in patients with traumatic brain injury (TBI) [11].

In our model, two variables related to the pre-hospital care were included, the type of attention with trained medical attention specialized in advanced resuscitation techniques and the need of pre-hospital intubation [25, 26]. These factors are closely related to the quality of a mature trauma system in each country. We also considered type and intentionality of trauma. We grouped different mechanisms into a single one variable associated with high mortality including pedestrian falls, suicidal high-energy falls, firearms injuries and unknown mechanisms [27, 28]. These are well-known mechanisms associated with higher mortality. For the purpose of an external validation of our score, it would be especially important to check how this variable performs.

TBI is the leading cause of mortality in the trauma ICU in our environment. Therefore, it is not surprising that three neurological factors were finally included in our model. All of them are well-known clinical factors related to outcomes: pupillary reactivity, which is easy to determine and carries an important weight in TBI [8, 29], GCS $\leq$ 8 and the anatomical burden of head injury evaluated by the AIS [8].

Additionally, haemodynamic instability, respiratory failure and trauma-induced coagulopathy in patients with critical bleeding are associated with a worst prognosis and were also present in our model [13, 14, 30, 31].

The need for MV is an indicator of severity in critical trauma patients. It does not only potentially express respiratory failure, but also neurological or haemodynamic involvement or the need of procedures that require sedation.

We observed that severe chest trauma was a protective factor of trauma ICU patients. This intriguing result merits further investigation since the real contribution of

| Table 3  Multivariable analysis of factors associated with 30-day mortality |
|----------------------------------------------------------|
| Variable | $\beta$-coefficients$^a$ | OR (95% CI) | $p$ value | Points |
| Age groups | | | | |
| $< 50$ | Reference | | | 0 |
| 50–65 | 0.598 | 1.83 (1.34–2.47) | $< 0.001$ | 1 |
| 65–75 | 1.239 | 4.78 (3.43–6.66) | $< 0.001$ | 2 |
| $> 75$ | 2.198 | 10.35 (7.47–14.36) | $< 0.001$ | 4 |
| PRECOAG | 0.349 | 1.43 (1.10–1.85) | 0.007 | 1 |
| Pre-hospital intubation | 0.336 | 1.41 (1.08–1.83) | 0.012 | 1 |
| High Risk mechanism | 0.662 | 1.96 (1.56–2.47) | $< 0.001$ | 1 |
| PUPILS | | | | |
| Normal | Reference | | | 0 |
| Unilateral mydriasis | 0.950 | 2.63 (1.97–3.51) | $< 0.001$ | 2 |
| Bilateral mydriasis | 3.217 | 26.38 (17.51–39.74) | $< 0.001$ | 6 |
| GCS $\leq$ 8 | 0.841 | 2.35 (1.79–3.10) | $< 0.001$ | 2 |
| MAIS-Head | 0.495 | 1.66 (1.30–2.12) | $< 0.001$ | 1 |
| MAIS-Thorax | $-$0.271 | 0.76 (0.60–0.96) | 0.026 | $-$1 |
| Haemodynamic failure | 1.148 | 3.21 (2.51–4.12) | $< 0.001$ | 2 |
| Respiratory failure | 0.708 | 2.06 (1.58–2.67) | $< 0.001$ | 1 |
| Coagulopathy | 0.567 | 1.78 (1.36–2.33) | $< 0.001$ | 1 |
| Mechanical ventilation | 0.580 | 1.80 (1.39–2.34) | $< 0.001$ | 1 |
| Massive haemorrhage | 0.452 | 1.58 (1.10–2.27) | $< 0.001$ | 1 |

Logistic regression. Derivation set (n = 5976)

PRECOAG, prior treatment with antiplatelets or anticoagulants; GCS, Glasgow Coma Score; AIS, Abbreviated Injury Scale; MAIS, AIS $\geq$ 3

$^a$ Regression coefficient multiplied with a shrinkage factor (bootstrapping procedure) of 0.98
chest trauma to the mortality of severe trauma patients is still a matter of debate [32]. Approximately 600 patients were admitted to the participating ICUs with severe chest trauma and without other major injuries, and most of these patients received pain control, had early physiotherapy and had a good outcome. Likely, this is the reason why severe chest trauma received a score of $-1$ in the RETRASCORE model.

Taking together these variables, our easy-to-calculate and specific trauma ICU score achieves an excellent performance in terms of mortality prediction with an AUROC of almost 0.93 in the validation dataset. This is very close to the astonishing predictive ability of the RISC-II updated score, but ours is specific to the trauma ICU patients and is easier to use since it does not include data on pre-trauma ASA status, base deficit, haemoglobin and cardiopulmonary resuscitation [8]. Instead, it uses common clinical conditions such as respiratory failure or trauma-induced coagulopathy that reflect the physiological response to trauma based on predefined definitions. In this line, our score should be also confronted with general ICU scores. Magee et al. recently compared general ICU scores with traditional trauma scores and observed an improved performance of general ICU scores evaluating physiological derangements in the initial 24 h over anatomical scores [20]. Due to the increasing age of trauma populations who carry on additional comorbidities, we believe that scores like ours combining anatomical injuries and physiological derangements in the initial 24 h will improve our early prediction ability, especially in different subgroups of trauma patients [21].

We expect our score to be further externally validated in different databases allowing additional comparison with general ICU scores.

Our study presents some limitations:
First, the most important limitation is the lack of an external validation. This must be performed before the score can be applied in daily clinical practice. Second, the inclusion criteria of being admitted to the participating ICUs may not appropriately reflect the critical trauma population due to differences in admission criteria and bed and staffing availability. Third, unless patients were managed following the Advanced Trauma Life Support principles, we cannot rule out deviations so this could affect patients’ management and therefore outcomes [22]. Fourth, we have only used the TRISS as a comparison model; RETRASCORE should

![Fig. 1 Probability of death according to the total score. Derivation and validation sets. S: survivors, NS: non-survivors](image-url)
be compared with other scores, both specific to trauma patients and of a general type. Fifth, if a random partition had been made of the derivation and validation groups, no differences would have been found between the sets; we chose a temporary partition since we had a sufficiently large sample that could provide us with a temporal validation [10]. Finally, machine learning techniques have been used in our environment to predict outcome [33], but in this study, our objective was to develop a simple and early, easy-to-calculate and specific trauma ICU score rather than using complex methodologies.

**Conclusions**

In conclusion, the newly developed RETRASCORE is an early, easy-to-calculate and specific score to predict in-hospital mortality in trauma ICU patients. This is the first trauma score specifically designed for the trauma ICU population. Although it has achieved adequate internal validation, it must be externally validated.
Authors' contributions
LS, JALL, MCF, NM, MB, JAB, MAB and JT contributed to the study design. JT, LS, NM, JALL and MB performed data analysis and drafted and revised the article. LS, JALL and JT gave the final approval of the version to be submitted. All authors read and approved the final manuscript.

Funding
RETRAUCI was supported by a grant for the development of an electronic web-based system awarded to Dr. Chico-Fernández (Fundación Mutua Madrileña, reference number API 17892013).

Availability of data and materials
The data are available for other investigators' use under reasonable request.

Declarations

Acknowledgements
None.

Additional file 1: Screenshots of the web application and the list of variables collected in RETRAUCI.

Additional file 2: Figure S2-1. Flow diagram for selection of patients.
Figure S2-2: Selection of variables using LASSO (Least Absolute Shrinkage and Selection Operator) binary logistic regression model. Figure S2-3: Performance evaluation of the logistic regression (LR) model. Figure S2-4: Performance evaluation of the TRISS model. Figure S2-5: Performance evaluation of the RETRASCORE model.

Additional file 3: Comparative analysis between temporal and random partitions in the selection of variables and calculated coefficients.

References
1. Ray JG, Guttmann A, Silveira J, Park AL. Mortality in a cohort of 3.1 million children, adolescents and young adults. J Epidemiol Community Health. 2020;74:260–8. https://doi.org/10.1136/jech-2019-213365.
2. De Munter L, Polinder S, Lansink KW, Cnossen MC, Steyerberg EW, de Jongh MA. Mortality prediction models in the general trauma population: a systematic review. Injury. 2017;48:221–9. https://doi.org/10.1016/j.injury.2016.12.009.
3. Lefering R. Trauma scoring systems. Curr Opin Crit Care. 2012;18:637–40.
4. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. J Trauma. 1987;27:370–8.
5. López F, Grupo de Trabajo de Trauma y Neurointensivismo de SEMICYUC. Quality of trauma care and trauma registries. Med Intensiva. 2015;39:114–23. https://doi.org/10.1016/j.medint.2014.06.008.
6. Boyd CR, Telson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. J Trauma. 1987;27:370–8.
7. Lefering R. Development and validation of the Revised Injury Severity Classification (RISC) score for severely injured patients. Eur J Trauma Emerg Surg. 2009;35:437–47. https://doi.org/10.1007/s00068-009-9122-0.
8. Lefering R, Huber-Wagner S, Nienaber U, Maegele M, Bouillon B. Update of the trauma risk adjustment model of the TraumaRegister DGU™: the Revised Injury Severity Classification, version II. Crit Care. 2014:18:476. https://doi.org/10.1186/s13054-014-0476-2.
9. Vincent JL, Moreno R. Clinical review: scoring systems in the critically ill. Crit Care. 2010;14:207. https://doi.org/10.1186/cc8204.
10. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis.
