

## Clinical prediction models for the management of blunt chest trauma in the Emergency Department: a systematic review

Ceri Battle<sup>a</sup>, Elaine Cole<sup>b</sup>, Kym Carter<sup>c</sup>, Edward Baker<sup>d</sup>

<sup>a</sup> Physiotherapy Dept, Morriston Hospital, Swansea Bay University Health Board, Swansea, Wales, UK, SA6 6NL. [Ceri.battle@Wales.nhs.uk](mailto:Ceri.battle@Wales.nhs.uk) (corresponding author)

<sup>b</sup> Centre of Trauma Sciences, Blizard Institute, Queen Mary University of London, London, UK. [e.cole@qmul.ac.uk](mailto:e.cole@qmul.ac.uk)

<sup>c</sup> Swansea Trials Unit, Swansea University Medical School, Swansea University, Swansea, UK. [k.carter@swansea.ac.uk](mailto:k.carter@swansea.ac.uk)

<sup>d</sup> Emergency Dept, Kings College Hospital, London, UK. [edwardbaker@nhs.net](mailto:edwardbaker@nhs.net)

### ABSTRACT

**Introduction:** The aim of this systematic review was to investigate how clinical prediction models compare in terms of their methodological development, validation, and predictive capabilities, for patients with blunt chest trauma presenting to the Emergency Department.

**Methods:** A systematic review was conducted across databases from Jan 2000 until March 2023. Studies were categorised into three types of multivariable prediction research and data extracted regarding methodological issues and the predictive capabilities of each model. Risk of bias and applicability were assessed.

**Results:** 39 studies were included that discussed 22 different models. The most commonly observed study design was a single-centre, retrospective, chart review. The most widely externally validated clinical prediction models with moderate to good discrimination were the Thoracic Trauma Severity Score and the STUMBL Score.

**Discussion:** This review demonstrates that the predictive ability of some of the existing clinical prediction models is acceptable, but high risk of bias and lack of subsequent external validation limits the extensive application of the models. The Thoracic Trauma Severity Score and STUMBL Score demonstrate better predictive accuracy in both development and external validation studies than the other models, but require recalibration and / or update and evaluation of their clinical and cost effectiveness.

## INTRODUCTION

Patients with blunt chest trauma present an ongoing challenge for accurate triage in the Emergency Department (ED). Whilst the majority of patients with blunt chest trauma will have an uncomplicated recovery, clinical presentation at the time of ED assessment is no guarantee that a patient will be of suitable acuity for discharge to home, or for admission to award setting, as up to 10% of patients will decompensate after 48-72 hours.<sup>1-3</sup> Intensive Care Unit (ICU) referral from the ED must be carefully considered and as a result, much has been published over the last 20 years investigating the predictors of poor outcome in this patient cohort.<sup>4,5</sup> These predictors include patient age, severity of injury, number and location of rib fractures, pre-injury anticoagulants, chronic lung disease and others.<sup>4,6-8</sup>

A common aim of such primary prognostic studies is the development of clinical prediction models. The clinical prediction model is intended to estimate the individualised probability or risk that a condition, for example mortality or pulmonary complications, will occur in the future by combining multiple prognostic factors / predictors from an individual.<sup>9,10</sup> A number of different clinical prediction models have been developed for patients with blunt chest trauma, however there is still no universally accepted model in clinical practice. A recent survey study highlighted that there were 20 different clinical prediction models and pathways used when assessing whether a patient with blunt chest trauma is safe for ED home discharge.<sup>11</sup>

There is often conflicting evidence regarding the predictive capabilities of developed clinical prediction models, leading to a growing demand for evidence synthesis of external validation studies that assess model performance in a new patient cohort.<sup>10,12,13</sup> This is applicable to the range of clinical prediction models used for the management of patients with blunt chest trauma. The aim of this systematic review therefore was to investigate how clinical prediction models compare in terms of their methodological development, validation, and predictive capabilities, for clinical and healthcare utilisation outcomes for patients with blunt chest trauma presenting to the Emergency Department.

## METHODS

### Search strategy

The review was registered prospectively on the PROSPERO database ([https://www.crd.york.ac.uk/PROSPERO/display\\_record.php?RecordID=351638](https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=351638)). The CHARMS Checklist was followed for completion of this review. A broad search strategy was employed in order to capture all relevant studies. The search filter was used for PubMed and Embase Databases, the Cochrane Library, and OpenGrey from Jan 2000 until March 2023.

The search term combinations were based on Geersing et al (2012)<sup>12</sup> and used Medical Subject Heading terms, text words and word variants for blunt chest trauma. These were combined with relevant terms for both outcomes and clinical prediction model development and validation methods. The search strategy can be found in online supplementary file 1. The reference lists of all relevant studies were hand-searched in order to identify any evidence missed in the electronic search. The Annals of Emergency Medicine, Emergency Medicine Journal, Injury and the Journal of Trauma and Acute Care Surgery were hand-searched for relevant studies. Searches were international and no search limitations were used.

### **Study selection**

Studies were included that focussed on patients aged  $\geq 16$  presenting to the Emergency Department with blunt chest trauma (defined as a blunt chest injury resulting in chest wall contusion or rib fractures, with or without underlying lung injury). Prognostic multivariable prediction studies were included where the aim of the study was to predict an outcome using two or more independent variables, in order to develop a multivariable (at least two variables) weighted clinical prediction model for any outcome following blunt chest trauma. Based on the 'Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: CHARMS guidance'<sup>13</sup>, studies were categorised into three types of multivariable prediction research; 1) model development studies without external validation. 2) model development studies with external validation in independent data, and 3) external validation studies without or with model updating.

Studies were excluded which included patients presenting with: a) Penetrating trauma only, b) Multi-trauma only and no reference to chest trauma, c) Severe intra-thoracic injuries only (eg. bronchial, cardiac, oesophageal, aortic or diaphragmatic rupture) and no chest wall trauma, d) Children aged  $< 16$  years. Other exclusion criteria included, studies that investigated a single predictor (such as single prognostic marker studies), studies that investigated only causality between one or more variables and an outcome, and studies that do not contribute to patient care. For multiple publications from the same dataset, only the most relevant study to this reviews aims was included. Studies for which only an abstract was available were also excluded.

### **Data extraction**

A two-step process was used to reduce potential selection bias. Two researchers ( [REDACTED] ) analysed each title and abstract independently and then met to discuss any discrepancies. The full paper of selected studies was analysed by the reviewers. Data were extracted relating to both the reporting of and use of methods known to influence the quality

of multivariable prediction studies. A data extraction form based on CHARMS Checklist was used to record relevant information, a copy of which is available in supplementary file 2]. Study authors were contacted for any missing data and response time set at six weeks. Included studies were grouped according to the clinical prediction model under investigation for the analysis.

Data were extracted regarding the methodological issues that are considered to be important in prediction research, focussed broadly on the reporting of the domains outlined in the CHARMS Checklist. Data regarding the predictive capabilities of each model were also extracted where available, for the following outcomes; a) clinical outcomes such as mortality and any pulmonary complications, and b) healthcare utilisation outcomes such as length of stay, need for mechanical ventilation or ICU admission.

### **Quality assessment**

Risk of bias and applicability were assessed using the “Prediction model Risk Of Bias ASsessment Tool” (PROBAST)<sup>14</sup> where: “Risk of bias refers to the extent that flaws in the design, conduct, and analysis of the primary prediction modelling study lead to biased, often overly optimistic, estimates of predictive performance measures such as model calibration, discrimination, or (re)classification (usually due to over-fitted models). Applicability refers to the extent to which the primary study matches the review question, and thus is applicable for the intended use of the reviewed prediction model(s) in the target population” (Moons et al, 2014). PROBAST includes 20 signalling questions across four domains (participants, predictors, outcome, and analysis) which were scored low, high or unclear. For each included study, an overall final score for judgement of risk of bias and applicability was allocated. This process was completed independently by two reviewers (██████████), with a third reviewer (████) used to resolve any discrepancies. An outline of the PROBAST Score can be found in online supplement 3.

### **Data synthesis and analysis**

Narrative synthesis of included study results was conducted, grouped according to clinical prediction models. Model performance was evaluated through assessment of model discrimination, a measure of how well the model can separate those who do and those who do not have the disease of interest, and calibration, a measure of how well predicted probabilities agree with the actual observed risk. The discrimination ‘C-statistic’ (balance between negative and positive predictive value) was defined as low (below 0.70), moderate (0.70–0.79) or good (at least 0.80). Where available in the studies, the correlation between

observed and expected (calibration) outcome, as measured by the Hosmer–Lemeshow (H-L) test, was presented using a  $p > 0.050$  to indicate a good model fit.<sup>13</sup>

## RESULTS

### Study selection

The initial search strategy identified 9495 citations. Following screening titles and abstracts, we identified 172 potentially relevant studies and following full-text review, a total of 39 studies met the inclusion criteria. No additional citations were identified through the grey literature or reference list searches. Figure 1 outlines the flow diagram of study selection.

### Study characteristics

The 39 studies were categorised as; 12 model development studies without external validation, three model development studies with external validation in independent data, and 24 external validation studies without or with model updating. The most commonly observed study design was a single-centre, retrospective, chart review. A total of 22 different clinical prediction models were studied and therefore included in this review. Study design, clinical prediction model, study population (including diversity data where possible, such as age, sex, frailty and ethnicity), total sample size, outcomes and results of the included studies are outlined in Table 1.

#### Table 1: Characteristics of included studies

### Quality assessment

The quality of the included studies in this review was variable. Risk of bias was high across most of the included studies for the analysis. Selection of predictors was commonly based on univariable analysis result, handling of missing data was inadequately described and the model performance measures, in particular the model's calibration, was infrequently reported. The studies scored mostly low risk of bias in terms of the predictors included. Risk of bias for participants was variable across the studies as some used a trauma registry for their participant data. In terms of applicability, some studies scored high risk for participants, as they included paediatric patients, which this review was not investigating. The full PROBAST results are outlined in Table 2 and Figure 2.

#### Table 2: Risk of bias and applicability of included studies: PROBAST results

Figure 2 demonstrates the overall judgment of the included studies.

#### Figure 2: Risk of bias and applicability of included studies: PROBAST results

### ***Clinical prediction models***

#### *Thoracic Trauma Severity Score (TTSS)*

The TTSS was originally developed and externally validated by Pape et al (2000) to predict the risk of thoracic-trauma related complications in patients with blunt polytrauma, admitted to ICU.<sup>47</sup> Based on high risk of bias results, the c-index demonstrated good discrimination, as demonstrated by a value of 0.924 for the development set and 0.916 for the validation set, although 95% confidence intervals were not reported. Since 2002, there have been nine external validation studies of high risk of bias, that have reported various cut off values on the TTSS, with moderate to good level c-indices ranging between 0.723 and 0.848. Model calibration was not reported in any of the included studies.

#### *STUMBL Score*

The STUMBL Score was original developed and externally validated by Battle et al (2014) to predict risk of pulmonary complications in patients with isolated blunt chest wall trauma presenting to the ED.<sup>19</sup> Based on low risk of bias results, the final model demonstrated good discrimination with a reported c-index of 0.96 (95% CI: 0.93 to 0.98). The model showed good calibration when evaluated with the Hosmer Lemeshow test (9.22, P = 0.32). Since development, there have been three external validation studies completed of variable risk of bias, that have reported various cut off values on the STUMBL Score, with moderate to good level c-indices ranging between 0.73 (95% CI 0.68–0.77) and 0.90 (95% CI 0.88 - 0.93).

#### *Rib Fracture Score (RFS)*

The RFS was originally developed by Easter et al (2001), as a protocol for the management of pain, respiratory care and mobility in patients with multiple rib fractures.<sup>29</sup> The score allocated to the patient (based on number of fractures, number of sides and the patient's age), determines the treatment recommendations, rather than a risk of a particular outcome. The protocol was based on literature, rather than patient data and as a result was at high risk of bias. No predictive capabilities were reported in the original development study. Three external validation studies of high risk of bias, have been completed, demonstrating a low level of discrimination with c-indices ranging from 0.64 to 0.67 for the prediction of a number of clinical and healthcare resource outcomes. Model calibration was not reported in the included studies.

#### *Chest Trauma Score (CTS)*

The CTS, originally developed by Pressley et al (2012) for patients presenting with rib fractures, using clinical data available at the time of initial evaluation. It predicts the likelihood of mechanical ventilation and prolonged courses of care.<sup>48</sup> The development study does not report predictive capabilities of the score and was considered high risk of bias. Five external validation studies of high risk of bias have been completed, demonstrating a low to good level of discrimination with c-indices of 0.67 to 0.91. Model calibration was not reported in any of the studies.

### *RibScore*

The RibScore, originally developed by Chapman et al (2016) for blunt trauma patients with rib fractures, was based on six candidate radiographic variables, identified on CT imaging.<sup>23</sup> They reported c-indices the outcomes pneumonia, respiratory failure and tracheostomy were 0.71, 0.71, and 0.75, respectively in a high risk of bias study. Two high risk of bias external validation studies have been completed in which low and moderate c-indices of 0.66 and 0.79 (95% CI 0.69 to 0.89) were reported. Model calibration was not reported in any of the studies.

### *Other clinical prediction models*

Table 1 outlines 18 other clinical prediction models which were identified, for which only one study (all high risk of bias) per model met the inclusion criteria for this review. A number of new clinical prediction models have been developed (all high risk of bias studies) but not yet validated were included in the review. These included the Tracheostomy in Thoracic Trauma Prediction Score<sup>20</sup> (T<sub>3</sub>P-Score, c-index for tracheostomy: 0.938, 95% CI: 0.920-0.956), Sequential Clinical Assessment of Respiratory Function<sup>37</sup> (SCARF Score, c-index for pneumonia: 0.86), Rib Injury Guidelines<sup>46</sup> (RIG, c-index not reported), the Lung Organ Failure Score<sup>53</sup> (c-index not reported), and a new scoring system<sup>35</sup> (c-index: 0.82; 95% CI: 0.77-0.88).

Other models developed and validated by the original authors, but yet to be externally validated in further studies included The Rib Fracture Frailty Index<sup>25</sup> (RFFI) (c-index not reported), the Revised Intensity Battle Score<sup>21</sup> (RIBS) (c-index: 0.86), Quebec Minor Thoracic Injury Decision Rule<sup>31</sup> (c-index: 0.78; 95% CI 0.74–0.82), a single rib fracture nomogram<sup>39</sup> (c-index: 0.70), and the Trauma Induced Pulmonary Event (TIPE Score) (c-index: 0.85).<sup>40</sup>

The chest wall components of the Abbreviated Injury Scale (AIS) and Organ Injury Scale (OIS) were externally validated in a high risk of bias study by Baker et al (2020) which reported a low level of discrimination for both the OIS (c-index: 0.68; 95% CI: 0.64-0.73) and



AIS (c-index: 0.59; 95%CI: 0.55 to 0.63) for patients with rib and sternal fractures presenting to the ED.<sup>16</sup>

There were four model development studies that did not meet the inclusion criteria for this review, but subsequent validation studies were included (all high risk of bias). These included the Revised Cardiac Risk Index<sup>18</sup> (RCRI, originally developed to predict 30-day postoperative myocardial infarction, cardiac arrest, or mortality following non-cardiac surgery, c-index not reported), Pain Inspiratory Effort Cough Score<sup>17</sup> (PIC Score, c-index not reported), Revised Trauma Scale<sup>26</sup> (RTS, c-index: 0.76, 95%CI: 0.72-0.79), Lung Ultrasound Score<sup>49</sup> (LUS, c-index not reported), and the ROX Index<sup>27</sup> (which combines respiratory rate and oxygenation values, c-index: 0.88; 95%CI: 0.80–0.94).

## DISCUSSION

This systematic review has highlighted that there are numerous clinical prediction models used for the management of patients with blunt chest trauma in various healthcare settings. These models differ widely in terms of their target patient population, included risk factors and outcomes predicted. They also differ in terms of the methods used for both their development and validation. These findings impede comparison between the models and generalisability for the patient with blunt chest wall trauma. These inherent differences also contribute to the lack of consensus in clinical practice, regarding the optimal clinical prediction model for this patient population.<sup>54, 55</sup>

This review highlights the difficulties in developing, validating and using a clinical prediction model. Instead of updating existing models and improving their predictive capabilities, most studies have developed and presented a new model. This has resulted in better performance in their population compared with existing models that were developed in another population and validated externally. Furthermore, there were no impact studies retrieved in this review that explored the clinical or cost effectiveness of any of the models. Traditional impact studies are reported to be costly to undertake and as a result, very few exist for any patient condition.<sup>55</sup> It is reasonable therefore to suggest that the ideal model does not yet exist.

Not all studies calculated a c-index to describe the discriminative abilities of the model and only one study reported a H-L analysis for calibration. Other studies may have used alternative measurements, or it must be assumed that they have compared observed with expected results, but did not report the comparison statistic. Overall, discrimination is more straightforward to calculate when compared with calibration, and the latter can be easily improved using updating methods applied to a new patient cohort.<sup>13, 55</sup> Good calibration is necessary however for calculating predictions, independent of the reported c-index.<sup>55</sup> The



clinical usefulness of a model can only be determined when both discrimination and calibration are available, and a model's cut-off value has been defined for reported sensitivity and specificity values.<sup>13, 55</sup>

The most widely externally validated clinical prediction models with moderate to good discrimination developed specifically for the management of patients with blunt chest trauma, were the TTSS<sup>47</sup> and STUMBL Score<sup>19</sup>. These models were developed for use in different healthcare settings and only the STUMBL Score had been assessed for calibration. Neither model has undergone any recalibration, updating or revision, nor have been assessed for clinical or cost effectiveness. There is limited reference to different diverse patient groups in any of the included studies, with exception to the STUMBL Score, which was the only model that was reported to have been specifically externally validated on patients of varying ethnic groups. Health inequalities across ethnic groups are reported in other disease populations<sup>56, 57</sup> but currently it isn't clear if existing blunt chest trauma clinical prediction models account for diversity-related differences.

This systematic review has a number of limitations. A large number of the included studies failed to report confidence intervals for the reported c-indices, resulting in incomplete comparisons between the models. Most of these models had been developed on Caucasian populations, and it remains unknown (other than the STUMBL Score New Zealand validation study<sup>45</sup>) whether these models would perform equally well in other ethnic groups. Frailty as a potential candidate predictor was not considered in any of the included model development studies, other than the RFFI study.<sup>25</sup> It is well-recognised that frailty identification has an important role in any clinical decision-making related in older trauma patients<sup>58, 59</sup>, therefore this needs further consideration in future studies and existing model updates. Finally, the lead author of this review is also the researcher who developed the [REDACTED], so there is the potential for interpretive bias.

In conclusion, this systematic review has examined the methodological development, validation, and predictive capabilities of the clinical prediction models, for clinical and healthcare utilisation outcomes for patients with blunt chest trauma presenting to the Emergency Department. The predictive ability of some of the existing clinical prediction models is acceptable, but high risk of bias and lack of subsequent external validation limits the extensive application of the models in the general blunt chest trauma population. TTSS and STUMBL Score demonstrate better predictive accuracy in both development and external validation studies than the other models, but both potentially still require recalibration and / or update and evaluation of their clinical and cost effectiveness.

## REFERENCES

1. Blecher GE, Mitra B, Cameron PA and Fitzgerald MC. Failed Emergency Department disposition to the ward of patients with thoracic injury. *Injury* 2008;39:586-91.
2. Baker E, Battle C. What is the optimal care pathway for patients with blunt chest wall trauma presenting to the ED? *Trauma* 2023;DOI: 10.1177/14604086221142384.
3. Plourde M, Émond M, Lavoie A, Guimont C, Le Sage N, Chauny JM, et al. [Cohort study on the prevalence and risk factors of late pulmonary complications in adults following a closed minor chest trauma]. *CJEM* 2013;15:337-44.
4. Battle C, Carter K, Newey L, Giamello JD, Melchio R and Hutchings H. Risk factors that predict mortality in patients with blunt chest wall trauma: an updated systematic review and meta-analysis. *Emerg Med J* 2022;14:14.
5. Christmas AB, Freeman E, Chisolm A, Fischer PE, Sachdev G, Jacobs DG, et al. Trauma intensive care unit 'bouncebacks': identifying risk factors for unexpected return admission to the intensive care unit. *Am Surg* 2014;80:778-82.
6. Carver TW, Milia DJ, Somberg C, Brasel K and Paul J. Vital capacity helps predict pulmonary complications after rib fractures. *J Trauma Acute Care Surg* 2015;79:413-6.
7. Choi J, Alawa J, Tennakoon L and Forrester JD. DeepBackRib: Deep learning to understand factors associated with readmissions after rib fractures. *J Trauma Acute Care Surg* 2022;93:757-61.
8. Hamilton C, Barnett L, Trop A, Leininger B, Olson A, Brooks A, et al. Emergency department management of patients with rib fracture based on a clinical practice guideline. *Trauma Surg Acute Care Open* 2017; 2: e000133.
9. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW and Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 2015;68:279-89.
10. Debray TP, Damen JA, Snell KI, Ensor J, Hooft L, Reitsma JB, et al. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;356:i6460.
11. Battle C, O'Neill C, Newey L, Barnett J, O'Neill M and Hutchings H. A survey of current practice in UK emergency department management of patients with blunt chest wall trauma not requiring admission to hospital. *Injury* 2021;52:2565-70.
12. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeftang M and Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012;7:e32844.

13. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. *PLOS Medicine* 2014;11:e1001744.
14. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess Risk of Bias and Applicability of Prediction Model Studies: Explanation and Elaboration. *Ann Intern Med* 2019;170:W1-33.
15. Aukema TS, Beenen LF, Hietbrink F and Leenen LP. Validation of the Thorax Trauma Severity Score for mortality and its value for the development of acute respiratory distress syndrome. *Open Access Emerg Med* 2011;3:49-53.
16. Baker JE, Millar DA, Heh V, Goodman MD, Pritts TA and Janowak CF. Does chest wall Organ Injury Scale (OIS) or Abbreviated Injury Scale (AIS) predict outcomes? An analysis of 16,000 consecutive rib fractures. *Surgery* 2020;168:198-204.
17. Bass GA, Stephen C, Forssten MP, Bailey JA, Mohseni S, Cao Y, et al. Admission Triage With Pain, Inspiratory Effort, Cough Score can Predict Critical Care Utilization and Length of Stay in Isolated Chest Wall Injury. *Journal of Surgical Research* 2022;277:310-8.
18. Bass GA, Duffy CC, Kaplan LJ, Sarani B, Martin ND, Ismail AM, et al. The revised cardiac risk index is associated with morbidity and mortality independent of injury severity in elderly patients with rib fractures. *Injury* 2023;54:56-62.
19. Battle C, Hutchings H, Lovett S, Bouamra O, Jones S, Sen A, et al. Predicting outcomes after blunt chest wall trauma: development and external validation of a new prognostic model. *Crit Care* 2014;18:R98.
20. Bläsius FM, Wutzler S, Störmann P, Lustenberger T, Frink M, Maegele M, et al. Predicting tracheostomy in multiple injured patients with severe thoracic injury (AIS  $\geq$  3) with the new T3P-Score: a multivariable regression prediction analysis. *Sci Rep* 2023, 24;13:3260.
21. Buchholz CJ, Jia L, Manea C, Petersen T, Wang H, Stright A, et al. Revised Intensity Battle Score (RIBS): Development of a Clinical Score for Predicting Poor Outcomes After Rib Fractures. *Am Surg* 2022:31348221123087.
22. Callisto E, Costantino G, Tabner A, Kerslake D and Reed MJ. The clinical effectiveness of the STUMBL score for the management of ED patients with blunt chest trauma compared to clinical evaluation alone. *Internal and Emergency Medicine* 2022;DOI: 10.1007/s11739-022-03001-0.
23. Chapman BC, Herbert B, Rodil M, Salotto J, Stovall RT, Biffi W, et al. RibScore: A novel radiographic score based on fracture pattern that predicts pneumonia, respiratory failure, and tracheostomy. *J Trauma Acute Care Surg* 2016;80:95-101.
24. Chen J, Jeremitsky E, Philp F, Fry W and Smith RS. A chest trauma scoring system to predict outcomes. *Surgery* 2014;156:988-93.

25. Choi J, Marafino BJ, Vendrow EB, Tennakoon L, Baiocchi M, Spain DA, et al. Rib Fracture Frailty Index: A risk stratification tool for geriatric patients with multiple rib fractures. *J Trauma Acute Care Surg* 2021;91:932-9.
26. Cinar E, Usul E, Demirtas E and Gokce A. The role of trauma scoring systems and serum lactate level in predicting prognosis in thoracic trauma. *Ulusal Travma ve Acil Cerrahi Dergisi* 2021;27:619-23.
27. Cornillon A, Balbo J, Coffinet J, Floch T, Bard M, Giordano-Orsini G, et al. The ROX index as a predictor of standard oxygen therapy outcomes in thoracic trauma. *Scand J Trauma Resusc Emerg Med* 2021;29:81.
28. Daurat A, Millet I, Roustan J-P, Maury C, Taourel P, Jaber S, et al. Thoracic Trauma Severity score on admission allows to determine the risk of delayed ARDS in trauma patients with pulmonary contusion. *Injury* 2016;47:147-53.
29. Easter A. Management of patients with multiple rib fractures. *Am J Crit Care* 2001;10:320-7.
30. El-Aziz T, El Din Abdallah A, Abdo A and El-Hag-Aly M. Thorax trauma severity score and trauma injury severity score evaluation as outcome predictors in chest trauma. *Research and Opinion in Anesthesia and Intensive Care* 2022;9:112-8.
31. Emond M, Guimont C, Chauny JM, Daoust R, Bergeron E, Vanier L, et al. Clinical prediction rule for delayed hemothorax after minor thoracic injury: a multicentre derivation and validation study. *CMAJ Open* 2017; 5:E444-53.
32. Esme H, Solak O, Yurumez Y, Yavuz Y, Terzi Y, Sezer M, et al. The prognostic importance of trauma scoring systems for blunt thoracic trauma. *Thorac Cardiovasc Surg* 2007;55:190-5.
33. Fokin A, Wycech J, Crawford M and Puente I. Quantification of rib fractures by different scoring systems. *J Surg Res* 2018;229:1-8.
34. Giamello JD, Martini G, Prato D, Santoro M, Arese Y, Melchio R, et al. A retrospective validation study of the STUMBL score for emergency department patients with blunt thoracic trauma. *Injury* 2022;DOI: 10.1016/j.injury.2022.08.028.
35. Gonzalez KW, Ghneim MH, Kang F, Jupiter DC, Davis ML and Regner JL. A pilot single-institution predictive model to guide rib fracture management in elderly patients. *J Trauma Acute Care Surg* 2015;78:970-5.
36. Harde M, Aditya G and Dave S. Prediction of outcomes in chest trauma patients using chest trauma scoring system: A prospective observational study. *Indian Journal of Anaesthesia* 2019;63:194-9.
37. Hardin KS, Leasia KN, Haenel J, Moore EE, Burlew CC and Pieracci FM. The Sequential Clinical Assessment of Respiratory Function (SCARF) score: A dynamic pulmonary

- physiologic score that predicts adverse outcomes in critically ill rib fracture patients. *J Trauma Acute Care Surg* 2019;87:1260-68.
38. Kanake V, Kale K, Mangam S and Bhalavi V. Thorax trauma severity score in patient with chest trauma: study at tertiary-level hospital. *Indian J Thorac Cardiovasc Surg* 2022;38:149-56.
39. Kishawi SK, Ho VP, Bachman KC, Alvarado CE, Worrell SG, Argote-Greene LM, et al. Are We Underestimating the Morbidity of Single Rib Fractures? *Journal of Surgical Research* 2021;268:174-80.
40. Li R, Grigorian A, Nahmias JT, Inaba K and Kuza CM. Development of a novel tool to predict pulmonary complications in trauma patients with and without chest injury. *American Journal of Surgery* 2022;224:64-8.
41. Martínez Casas I, Amador Marchante MA, Paduraru M, Fabregues Olea AI, Nolasco A and Medina JC. Thorax Trauma Severity Score: Is it reliable for Patient's Evaluation in a Secondary Level Hospital? *Bull Emerg Trauma* 2016;4:150-5.
42. Maxwell CA, Mion LC and Dietrich MS. Hospitalized injured older adults: clinical utility of a rib fracture scoring system. *J Trauma Nurs* 2012;19:168-74.
43. Mommsen P, Zeckey C, Andruszkow H, Weidemann J, Frömke C, Puljic P, et al. Comparison of Different Thoracic Trauma Scoring Systems in Regards to Prediction of Post-Traumatic Complications and Outcome in Blunt Chest Trauma. *Journal of Surgical Research* 2012;176:239-47.
44. Moon SH, Kim JW, Byun JH, Kim SH, Choi JY, Jang IS, et al. The thorax trauma severity score and the trauma and injury severity score: Do they predict in-hospital mortality in patients with severe thoracic trauma?: A retrospective cohort study. *Medicine* 2017;96:e8317.
45. Mukerji S, Tan E, May C, Micanovic C, Blakemore P, Phelps K, et al. Retrospective validation of a risk stratification tool developed for the management of patients with blunt chest trauma (the STUMBL score). *Emerg Med Australas* 2021;33:841-7.
46. Nelson A, Reina R, Northcutt A, Obaid O, Castanon L, Ditillo M, et al. Prospective validation of the Rib Injury Guidelines for traumatic rib fractures. *J Trauma Acute Care Surg* 2022; 92: 967-973.
47. Pape HC, Remmers D, Rice J, Ebisch M, Krettek C and Tscherne H. Appraisal of early evaluation of blunt chest trauma: development of a standardized scoring system for initial clinical decision making. *J Trauma Acute Care Surg* 2000;49:496-504.
48. Pressley CM, Fry WR, Philp AS, Berry SD and Smith RS. Predicting outcome of patients with chest wall injury. *American Journal of Surgery* 2012; 204:910-4.

49. Sayed MS, Elmeslmany KA, Elsayy AS and Mohamed NA. The Validity of Quantifying Pulmonary Contusion Extent by Lung Ultrasound Score for Predicting ARDS in Blunt Thoracic Trauma. *Crit Care Res Pract* 2022;3124966.
50. Schmoekel N, Berguson J, Stassinopoulos J, Karamanos E, Patton J and Johnson JL. Rib fractures in the elderly: physiology trumps anatomy. *Trauma Surg Acute Care Open* 2019;4:e000257.
51. Seok J, Cho HM, Kim HH, Kim JH, Huh U, Kim HB, et al. Chest Trauma Scoring Systems for Predicting Respiratory Complications in Isolated Rib Fracture. *Journal of Surgical Research* 2019;244:84-90.
52. Ujjaneswari B, Kumari VM, Madhusudhan KV, Lokesh K, Lakshmipathi Naik KS. A study of role of chest trauma scoring system in blunt chest trauma. *European Journal of Molecular and Clinical Medicine* 2023;10:42-66.
53. Wutzler S, Wafaisade A, Maegele M, Laurer H, Geiger EV, Walcher F, et al. Lung Organ Failure Score (LOFS): probability of severe pulmonary organ failure after multiple injuries including chest trauma. *Injury* 2012;43:1507-12.
54. Birse F, Williams H, Shipway D and Carlton E. Blunt chest trauma in the elderly: an expert practice review. *Emergency Medicine Journal* 2020;37:73-8.
55. Bouwmeester W, Zuithoff NP, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012;9:1-12.
56. Bhopal R, Steiner MF, Cezard G, Bansal N, Fischbacher C, Simpson CR, et al. Risk of respiratory hospitalization and death, readmission and subsequent mortality: scottish health and ethnicity linkage study. *Eur J Public Health* 2015;25:769-74.
57. Jennison T, Kulenkampff C, Lee J and Mahmood A. Is ethnicity a risk factor for mortality in major trauma? A single-centred cohort study. *Ann R Coll Surg Engl* 2023:racsann20220097.
58. Hamidi M, Haddadin Z, Zeeshan M, Saljuqi AT, Hanna K, Tang A, et al. Prospective evaluation and comparison of the predictive ability of different frailty scores to predict outcomes in geriatric trauma patients. *Journal of Trauma and Acute Care Surgery* 2019;87:1172-80.
59. Braude P, Short R, Bouamra O, Shipway D, Lecky F, Carlton E, et al. A national study of 23 major trauma centres to investigate the effect of a geriatrician assessment on clinical outcomes in older people admitted with serious injury in England (FiTR 2): a multicentre observational cohort study. *The Lancet Healthy Longevity* 2022;3:e549-e57.

**Figures:**

**Figure 1: PRISMA Flow Diagram**

**Figure 2: Risk of bias and applicability of included studies: PROBAST results**



**Table 1: Characteristics of included studies**

Author / year	Risk score	Study type	Study design	Participants	Number	Outcomes	Results
Aukema 2011 <sup>15</sup>	TTSS	External validation	Single centre, retrospective, trauma database	Patients with a score of 1+ on the AISThorax admitted to ED	516	Mortality, pneumonia, second PTX, persistent HTX, ARDS, empyema	AUROC mortality: 0.844. TTSS was significant higher in patients who died of thorax-related complications than in patients who died because of non thorax-related complications (p<0.001).
Baker 2020 <sup>16</sup>	OIS & AIS	External validation	Single centre, retrospective, trauma database	Adult patients with rib / sternal #s admitted to ED	3033	Mortality, tracheostomy, cardiopulmonary complications, readmissions within 30 days	OIS AUROCs: 0.679 for mortality and 0.667 for tracheostomy. TTSS and CTS outperformed both OIS and AIS for all outcomes except for readmissions.
Bass 2022 <sup>17</sup>	PIC Score	External validation	Single centre, retrospective, trauma database	Patients with isolated chest wall injuries (excluded AIS >2 in head or abdomen)	194	ICU admission, mechanical ventilation and length of stay.	A cut-off PIC score of ≤7 was associated with ICU admission OR: 8.19. 95%CI: 3.39-22.55, p<0.001 and with ICU admission for >48 hours OR: 26.9 95%CI: 5.5-43.96, p<0.001.
Bass 2023 <sup>18</sup>	RCRI	External validation	Multi-centre, retrospective, trauma database	Patients aged ≥65 with ≥1 rib fracture. Exclusion: managed operatively	96,750	In-hospital mortality, myocardial infarction, cardiac arrest with CPR, stroke, ARDS	Compared to RCRI 0, an RCRI score of 1 had a 16% increased risk of in-hospital mortality: adj-IRR: 1.16 95%CI: 1.02-1.32, p=0.020; RCRI score of 2: adj-IRR: 1.72 95%CI: 1.44-2.06, p<0.001
Battle 2014 <sup>19</sup>	STUMBL	Development / External validation	Single centre, retrospective chart review (development study). Multi-centre prospective observational (external validation)	Patients with primary diagnosis of blunt chest-wall trauma. Exclusion: <18 yrs, any immediate life-threatening injury.	274 237	Composite outcome: in-hospital mortality, morbidity including all pulmonary complications, ICU admission, or a prolonged LOS 7+ days	Final model reported AUROC of 0.96 (95% confidence intervals: 0.93 to 0.98), sensitivity was 80%, specificity was 96%, positive predictive value was 93% and negative predictive value was 86%.
Blasius 2023 <sup>20</sup>	T <sub>3</sub> P-Score	Development / Internal validation	Multi-centre, retrospective, trauma database	Adult patients with multi-trauma and severe thoracic trauma, requiring MV	1019	Tracheostomy, multi-organ failure, sepsis	The T3P-Score had high predictive validity for tracheostomy (AUROC: 0.938, 95% CI: 0.920, 0.956; Nagelkerke's R2 was 0.601). Specificity was 0.68, and the sensitivity was 0.96
Buchholz 2022 <sup>21</sup>	RIBS	Development / Internal validation	Single centre, retrospective, trauma database	Patients admitted with at least one rib fracture	838	Composite outcome: >7 days ventilated, tracheostomy, pneumonia, upgrade to ICU, unplanned intubation, mortality.	Final model AUROC of .858. Sensitivity is 72%, specificity is 84%, positive predictive value is 48.4%, and negative predictive value is 93.5%
Callisto 2022 <sup>22</sup>	STUMBL	External validation	Single centre, retrospective, chart review	Adult patients with ED diagnosis of blunt chest trauma. Exclusion: any immediate life-threatening injury, ICU admission.	369	Lower respiratory tract infection, pulmonary consolidation, empyema, pneumothorax, haemothorax, splenic or hepatic injury and 30-day mortality.	ED clinician decision to admit had a sensitivity of 83.9% and specificity of 86.0% for predicting complications. STUMBL score ≥11 had a sensitivity of 79.0% and specificity of 77.9%. AUROC of STUMBL score and ED clinician decision to admit was 0.84 (95% CI 0.78–0.90) and 0.85 (95% CI 0.79–0.91).
Chapman 2016 <sup>23</sup>	RibScore	Development	Single centre, retrospective, trauma database	Patient with blunt trauma with one or more rib fractures visualized on CT	385	Pneumonia, respiratory failure, and tracheostomy	RibScore was linearly associated with pneumonia (p<0.01), ARF (p<0.01), tracheostomy (p<0.01). AUROC for the outcomes were 0.71, 0.71, and 0.75, respectively.
Chen 2014 <sup>24</sup>	CTS	External validation	Single centre, retrospective, trauma database	Patients with blunt torso trauma	1361	Mortality, acute pneumonia and respiratory failure	CTS 5+ had nearly 4-fold increased odds of mortality (OR; 3.99, 95%CI: 1.92–8.31, p=0.001) compared with CTS < 5.
Choi 2021 <sup>25</sup>	RRFI	Development / External validation	Multi-centre, retrospective, trauma database.	Geriatric patients admitted with multiple rib fractures	55,540 77,710	Mortality, pneumonia, mechanical ventilation, hospital length of stay, discharge disposition	Among external validation cohort, increasing frailty risk was associated with stepwise worsening OR of mortality (1.5 [1.2–1.7], 3.5 [3.0–4.0]), intubation (2.4 [1.5–3.9]), 4.7 [3.1–7.5])
Cinar 2021 <sup>26</sup>	RTS, ISS and NISS	External validation	Single centre retrospective, chart review	Patients with isolated thoracic trauma. Exclusions: <18 years, major injury,	683	Mortality	NISS: AUROC: 0.876 (cut off score: >27), sensitivity: 85.3%, specificity: 80.7%, 95%CI: 0.848–0.899, P=0.000.
Cornillon 2021 <sup>27</sup>	ROX Index	External validation	Single centre, retrospective, chart review	All patients admitted to the ICU with AIS thorax.	171	Standard oxygen therapy failure	AUROC: 0.88 with a 95% CI [0.80–0.94]. ROX cut-off: 12.8: sensitivity: 81.7, 95%CI 0.7–0.9, specificity: 88.5, 95%CI 0.8–0.9
Daurat 2016 <sup>28</sup>	TTSS	External validation	Single centre retrospective, chart review	All blunt thoracic trauma with pulmonary contusion	329	Delayed ARDS	AUROC for TTSS for ARDS: 0.82 (95% CI 0.78–0.86). A TTSS of 13–25: risk factor for ARDS (OR 25.8 [95% CI 6.7–99.6] P < 0.001)
Easter 2001 <sup>29</sup>	RFS	Development	Based on literature only	Not stated	n/a	ICU Length of stay	Not stated

<b>El-Aziz 2022</b> <sup>30</sup>	TTSS & TRISS	External validation	Single centre, prospective cohort	Patients with chest trauma either penetrating or blunt trauma	100	Hospital mortality, need for oxygenation, ventilator, hospital length of stay	TTSS (cut-off value 4.5): AUROC: 0.88, P>0.001, sensitivity: 84.6%, specificity: 80.5%, 95%CI: 0.788–0.972. TRISS (cut off value: 24.55): AUROC: 0.892, P>0.001, sensitivity: 92.3%, specificity: 81.6%, 95%CI: 0.828–0.956.
<b>Emond 2017</b> <sup>31</sup>	Quebec Decision Rule	Development / Internal validation	Multi-centre, prospective cohort	Adult patients with a minor thoracic injury	830 552	Delayed haemothorax at 7, 14, 30 and 90 days	AUROC: 0.78 (95% CI 0.74–0.82) for the derivation cohort and 0.74 (95% CI 0.67–0.81) for the validation cohort
<b>Esme 2007</b> <sup>32</sup>	RTS, TRISS, ISS, LIS, CWIS	External validation	Single centre, retrospective, chart review	Patients with blunt chest trauma	152	Mechanical ventilation, thoracotomy, tube thoracostomy duration, LOS hospital and ICU stay, morbid conditions, mortality	TRISS was a predictor of mortality, LIS was an predictor of morbidity, the need for thoracotomy. CWIS, and LIS were independent predictors of the need for mechanical support. RTS, TRISS, ISS and LIS were predictors of the LOS
<b>Fokin 2018</b> <sup>33</sup>	RFS, CTS & RibScore	External validation	Single centre, retrospective, chart review	Patients with radiologically confirmed rib fractures	1089	Mortality, hospital and ICU length of stay, mechanical ventilation, pneumonia, tracheostomy, epidural analgesia.	RFS: AUROCs (mortality): all patients: 0.636, non-geriatric: 0.642, geriatric: 0.614. CTS: AUROCs (mortality): all patients: 0.669, non-geriatric: 0.687, geriatric: 0.646. RS: AUROCs (mortality): all patients: 0.654, non-geriatric: 0.656, geriatric: 0.656.
<b>Giamello 2022</b> <sup>34</sup>	STUMBL	External validation	Single centre, retrospective, chart review	Adult patients with isolated blunt thoracic trauma. Exclusion: immediately life-threatening lesion.	745	Composite outcome: in-hospital mortality, pulmonary complications, need for ICU, hospital length of stay 7+days	Primary outcome c-index: 0.90 (95% CI 0.88 - 0.93), and the result of the H-L test was 9.01 (p=0.34). STUMBL score = 16 has a sensitivity: 0.8 (95% CI 0.75-0.85), specificity: 0.87 (95% CI 0.84-0.90), PPV: 0.7 (95%CI 0.64-0.76), NPV: 0.92 (95% CI 0.90-0.94). AUROC: 0.82 (95% confidence interval [95% CI], 0.77-0.88). In cross-validation, sensitivity: mean of 70.43%. Specificity mean of 78.3%, NPV: mean of 93.1%.
<b>Gonzalez 2015</b> <sup>35</sup>	Trauma Scoring System	Development	Single centre, retrospective, chart review	Patients aged ≥55 with rib fractures	400	Intubation, pneumonia	AUROC: 0.82 (95% confidence interval [95% CI], 0.77-0.88). In cross-validation, sensitivity: mean of 70.43%. Specificity mean of 78.3%, NPV: mean of 93.1%.
<b>Harde 2019</b> <sup>36</sup>	CTS	External validation	Single centre, prospective cohort	Adult patients with chest trauma. Exclusion: significant injury.	30	Mortality, pneumonia and need for ventilator support	AUROC: 0.75. A CTS score 5.5: maximum sensitivity is 87.5% and specificity is 68%
<b>Hardin 2019</b> <sup>37</sup>	SCARF score	Development	Single centre, prospective cohort	Adult patients with rib fractures admitted to the surgical ICU	100	Pneumonia, FIO2 requirement >50%, respiratory failure, empyema, tracheostomy, ICU LOS, ICU re-admission, and mortality.	AUROC: the maximum SCARF score for these outcomes were 0.86, 0.76, and 0.79, respectively.
<b>Kanake 2022</b> <sup>38</sup>	TTSS	External validation	Single centre, prospective cohort	All patients chest trauma, with associated minor head injury	284	Mortality (hospitalised and non-hospitalised)	AUROC for the TTSS of 7.5: 0.9
<b>Kishawi 2021</b> <sup>39</sup>	Single rib fracture nomogram	Development / internal validation	Multi-centre, retrospective, trauma database	Adult patients with a single rib fracture associated with blunt trauma	2398	Composite outcome: mortality, pneumonia, tracheostomy, and hospital LOS >12 days	Among the training set, the AUROC: 0.700. When applied to the validation set, the model demonstrated AUROC: 0.672.
<b>Li 2022</b> <sup>40</sup>	TIPE score	Development / Internal validation	Multi-centre, retrospective, trauma database	Adult trauma patients	311,608 312,751	Pulmonary complications	AUROC for the TIPE score was 0.844 for both the derivation and validation-set
<b>Martinez-Casas 2016</b> <sup>41</sup>	TTSS	External validation	Single centre, retrospective, chart review	All patients with thoracic trauma	238	Length of hospital and ICU stay; need for mechanical ventilation; admission; complications and mortality	AUROC for TTSS was significant for predicting complications (0.848) and mortality (0.856) values. TTSS with a cut off value of 8: sensitivity: 66%, specificity: 94% to predict complications and 80% sensitivity and 94% specificity for predicting mortality
<b>Maxwell 2012</b> <sup>42</sup>	RFS	External validation	Single centre, retrospective, trauma database	Patients aged 50 years or older with rib fracture(s)	81	Hospital and ICU length of stay, discharge disposition	Correlation between hospital LOS with the RFS score: 0.29 (P=0.10). Correlation between RFS and ICU length of stay: 0.29 (P=0.009) No association of RFS with discharge disposition
<b>Mommsen 2012</b> <sup>43</sup>	PCS, AISchest, TTSS,	External validation	Single centre, retrospective chart review	Adult patients with polytrauma with severe thoracic trauma (AIS chest > 3)	278	ICU length of stay, mechanical ventilation, mortality	TTSS had the best prediction power for ARDS, MODS, and mortality among the examined thoracic trauma scores. No association between the TTSS and the development of SIRS and sepsis could be observed.
<b>Moon 2017</b> <sup>44</sup>	TTSS & TRISS	External validation	Single centre, retrospective, chart review	Patients with severe thoracic injury (ISS>18) who required ICU	228	In-hospital mortality	AUROC: 0.787 for the TRISS. At a cut-off value of 25.9%, the TRISS had a sensitivity of 83.6% and specificity of 73.5% to predict in-hospital mortality.

<b>Mukerji 2021<sup>45</sup></b>	STUMBL	External validation	Multi-centre, retrospective, chart review	Adult patients aged with isolated blunt chest trauma. Exclusion: penetrating chest trauma, immediate life-threatening injuries or multi-trauma	445	Composite outcome: in-hospital mortality, morbidity including all pulmonary complications, ICU admission, hospital length of stay 7+ days	AUROC for all complications composite were (0.73, 95% CI 0.68–0.77), mortality (0.92, 95% CI 0.89–0.94), ICU admissions (0.78, 95% CI 0.73–0.81) and prolonged LOS (0.80, 95% CI 0.76–0.83)
<b>Nelson 2022<sup>46</sup></b>	RIG	Development	Single centre, prospective cohort	Adult patients with blunt trauma with at least one rib fracture on CT	1100	Readmission, unplanned ICU admission, in-hospital mortality	Predictive capabilities not stated
<b>Pape 2000<sup>47</sup></b>	TTSS	Development / External validation	Single centre, retrospective, chart review (development study). Multi-centre, retrospective database (validation study)	Patients with a thoracic injury admitted to ICU	1495	Morbidity and mortality	AUROC demonstrated an adequate discrimination, as demonstrated by a value of 0.924 for the development set and 0.916 for the validation set. The score was also superior to the ISS (0.881) or the thorax Abbreviated Injury Score (0.693)
<b>Pressley 2012<sup>48</sup></b>	CTS	Development	Single centre, retrospective, trauma database	Patients with rib fractures	649	Mortality, ICU admission, mechanical ventilation, LOS	Predictive capabilities not stated
<b>Sayed 2022<sup>49</sup></b>	LUS	External validation	Single centre, prospective cohort	Patients with polytrauma with blunt chest trauma admitted to ICU	50	ARDS	A LUS of 4 was defined as a cut-off value for predicting ARDS development within 72 hours of trauma with sensitivity and specificity (91.67% and 84.21%), respectively
<b>Schmoekel 2019<sup>50</sup></b>	RibScore, MFI, PaCO2	External validation	Single centre, retrospective, chart review	Patients aged $\geq 55$ with blunt trauma and $\geq 1$ rib fracture identified by CT	263	Pneumonia, respiratory failure and tracheostomy	AUROC: RibScore: 0.79 (95% CI 0.69 to 0.89); mFI: 0.83 (95% CI 0.75 to 0.91) and PaCO2: 0.88 (95% CI 0.80 to 0.95). The PaCO2 had the highest discriminative ability of the three models.
<b>Soek 2019<sup>51</sup></b>	AIS, TTSS, RFS, CTS	External validation	Single centre, retrospective, chart review	Adult patients with sustained blunt trauma and isolated rib fractures (AIS $<2$ except in the chest area).	177	Pulmonary complications	Highest AUROC was TTSS (0.723, 95%CI 0.651-0.788). In patients with pulmonary contusion, TTSS also showed the highest AUROC (0.704, 95% CI 0.613-0.784 and without pulmonary contusion, RFS showed the highest AUROC (0.759, 95% CI 0.630-0.861).
<b>Ujjaneswari 2023<sup>52</sup></b>	CTS	External validation	Single centre, retrospective, chart review	Adult patients with $\geq 1$ rib fracture. Exclusion: associated injuries, COPD		Morbidity and mortality	There was a highly significant association between CTS score and mortality. (AUROC: 0.905, p<0.0001)
<b>Wutzler 2012<sup>53</sup></b>	LOFS	Development	Multi-centre, retrospective, trauma database	Adult patients admitted to the ICU with lung contusion/ lacerations	5892	Pulmonary organ failure	Predictive capabilities not stated

AUROC: Area under the receiver operator curve; H-L: Hosmer-Lemeshow; OR: Odds ratio; CI: Confidence Interval; adj-IRR: adjusted incidence risk ratio; LOS: length of stay; ICU: Intensive Care Unit; ARDS: Acute Respiratory Distress Syndrome; PCS: Pulmonary Contusion Score; RCRI: Revised Cardiac Risk Index; T3P-Score: Tracheostomy in Thoracic Trauma Prediction Score; ISS: Injury Severity Score, MFI: Modified Five-item Frailty Index; AIS: Abbreviated Injury Scale; TTSS: Thoracic Trauma Severity Score; TRISS: Trauma Score Injury Severity Score; RFS: Rib Fracture Score; CTS: Chest Trauma Score; LUS: RIG: Rib Injury Guidance; TIPE: Trauma Induced Pulmonary Event; LIS: Lung Injury Score; CWIS: Chest Wall Injury Score; OIS: Organ Injury Score; SCARF: Sequential Clinical Assessment of Respiratory Function Score; NISS: New Injury Severity Score; RRFI: Rib Fracture Frailty Index; RIBS: Revised Intensity Battle Score; PIC: Pain, Inspiratory Effort, Cough Score.

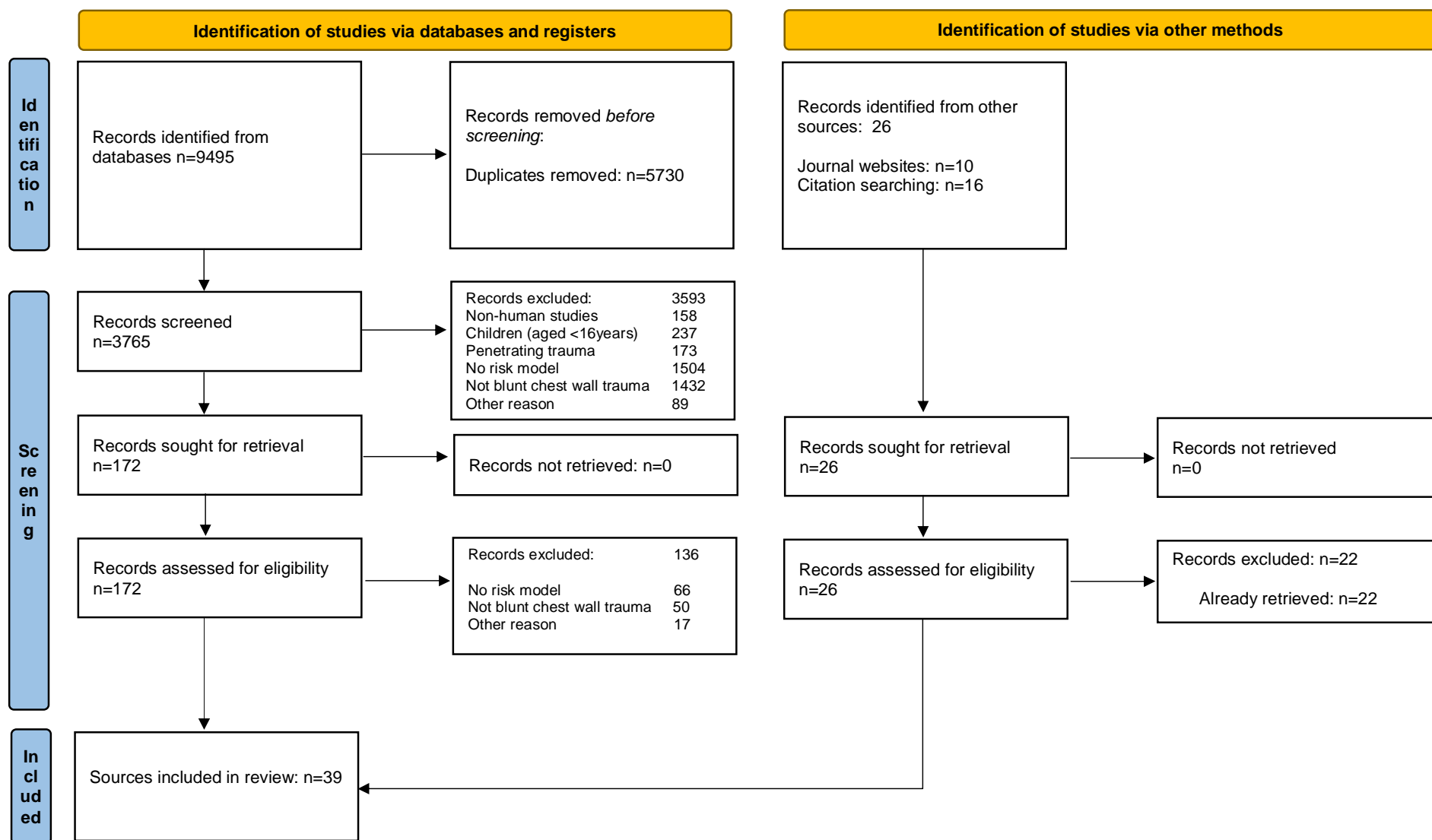
**Table 2: Risk of bias and applicability of included studies: PROBAST results**

Study	ROB				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Aukema 2011	-	+	-	-	-	+	+	-	-
Baker 2020	-	+	?	-	+	+	+	-	+
Bass 2022	-	+	-	-	?	+	+	-	+
Bass 2023	-	+	+	-	+	+	+	-	+
Battle 2014 Development	+	+	+	-	+	+	+	-	+
Battle 2014 Validation	+	+	+	+	+	+	+	+	+
Blasius 2023	-	+	+	-	+	+	+	-	+
Buchholz 2022	-	+	+	-	+	+	+	-	+
Callisto 2022	+	+	+	-	+	+	+	-	+
Chapman 2016	+	+	?	-	-	+	+	-	-
Chen 2014	-	+	+	-	-	+	+	-	-
Choi 2021 Development	-	+	?	-	+	+	+	-	+
Choi 2021 Validation	-	+	?	-	+	+	+	-	+
Cinar 2021	+	+	+	-	+	+	+	-	+
Cornillon 2021	+	+	+	-	-	+	+	-	+
Daurat 2016	+	+	+	-	?	+	+	-	?
Easter 2001	-	+	-	-	+	+	+	-	+
El-Aziz 2022	+	+	+	-	-	+	+	-	+
Emond 2017 Development	+	+	+	-	+	+	+	-	+
Emond 2017 Validation	+	+	+	-	+	+	+	-	+
Esme 2007	+	+	+	-	+	+	+	-	+
Fokin 2018	-	+	+	-	+	+	+	-	+
Giamello 2022	+	+	+	+	+	+	+	+	+
Gonzalez 2015	+	?	+	-	+	+	+	-	+
Harde 2019	+	+	+	-	?	+	+	-	?
Hardin 2019	+	+	+	-	+	+	+	-	+
Kanake 2022	+	+	+	-	-	+	+	-	-
Kishawi 2021	-	+	+	-	+	+	+	-	+
Li 2022 Development	-	+	?	-	-	-	+	-	-
Li 2022 Validation	-	+	?	-	-	-	+	-	-
Martinez 2016	+	+	?	-	-	+	+	-	-
Maxwell 2012	-	+	+	-	+	+	+	-	+
Mommsen 2012	?	+	+	-	+	+	+	-	+

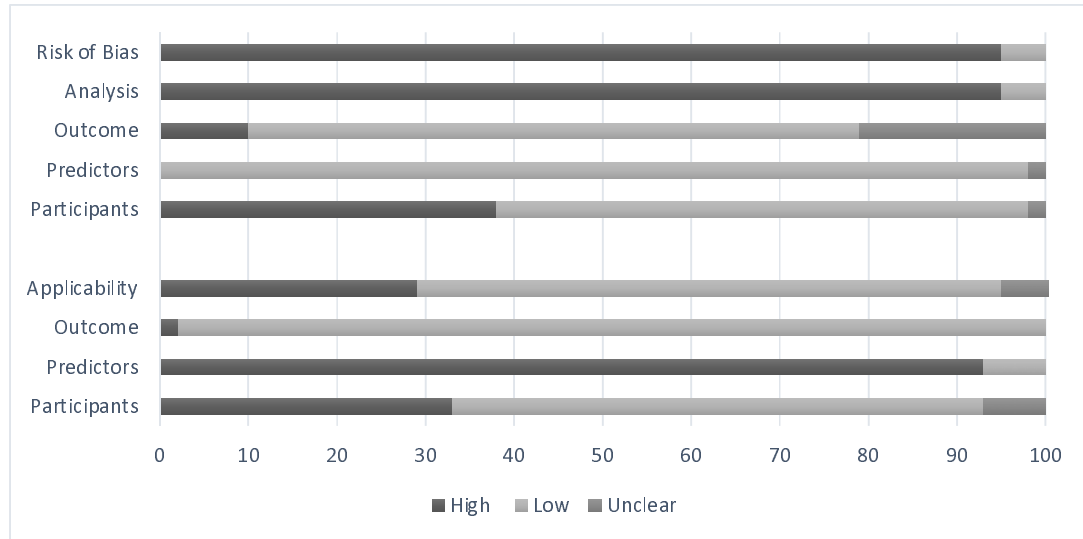
Moon 2017	+	+	+	-	-	+	+	-	-
Mukerji 2021	+	+	+	-	+	+	+	-	+
Nelson 2022	+	+	-	-	-	+	+	-	-
Pape 2000 Development	+	+	+	-	+	+	+	-	+
Pape 2000 Validation	-	+	+	-	+	+	+	-	+
Pressley 2012	-	+	?	-	+	+	+	-	+
Sayed 2022	+	+	+	-	-	-	-	-	-
Schmoekel 2019	+	+	+	-	+	+	+	-	+
Soek 2019	+	+	+	-	+	+	+	-	+
Ujjansewari	-	+	+	-	+	+	+	-	+
Wutzler 2012	-	+	+	-	-	+	+	-	-

+ low risk, ? unclear risk, - high risk

Figure 1: PRISMA flow diagram



From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>



**Figure 2: Risk of bias and applicability of included studies: PROBABST results**