

REVIEW ARTICLE

Prognostic Models in Severe Traumatic Brain Injury: A Systematic Review and Meta-analysis



Rita de Cássia Almeida Vieira^{1,2,3*} , Juliana Cristina Pereira Silveira¹, Wellingson Silva Paiva⁴, Daniel Vieira de Oliveira⁴, Camila Pedroso Estevam de Souza⁵, Eduesley Santana-Santos³ and Regina Marcia Cardoso de Sousa¹

© 2022 Springer Science+Business Media, LLC, part of Springer Nature and Neurocritical Care Society

Abstract

This review aimed to analyze the results of investigations that performed external validation or that compared prognostic models to identify the models and their variations that showed the best performance in predicting mortality, survival, and unfavorable outcome after severe traumatic brain injury. Pubmed, Embase, Scopus, Web of Science, Cumulative Index to Nursing and Allied Health Literature, Google Scholar, TROVE, and Open Grey databases were searched. A total of 1616 studies were identified and screened, and 15 studies were subsequently included for analysis after applying the selection criteria. The Corticosteroid Randomization After Significant Head Injury (CRASH) and International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury (IMPACT) models were the most externally validated among studies of severe traumatic brain injury. The results of the review showed that most publications encountered an area under the curve ≥ 0.70 . The area under the curve meta-analysis showed similarity between the CRASH and IMPACT models and their variations for predicting mortality and unfavorable outcomes. Calibration results showed that the variations of CRASH and IMPACT models demonstrated adequate calibration in most studies for both outcomes, but without a clear indication of uncertainties in the evaluations of these models. Based on the results of this meta-analysis, the choice of prognostic models for clinical application may depend on the availability of predictors, characteristics of the population, and trauma care services.

Keywords: Brain injury, Outcome assessment, Predictive modeling, Prognostic modeling

Introduction

Traumatic brain injury (TBI) is the result of any anatomical injury or functional impairment of brain segment structures caused by an external acting force [1, 2]. The pathophysiological mechanisms that start from this injury last for days to weeks, causing death and disability and mainly affecting young people of working age [1–3]. The consequences of injuries resulting from trauma remain beyond the acute phase of TBI, extending and changing for a long period after the traumatic event [1–3].

Estimating trauma severity is a constant concern of the teams who provide care to victims, as the proper definition of the severity of injuries—as well as TBI—can guide care and predict the patient's outcome. To do so, trauma severity indices were developed, beginning in 1969, to estimate the severity of victims after trauma and to assess the quality of care provided [4–9]. Trauma indices are reliable methodological tools that are essential for stratifying severity and predicting patient outcomes. The need to improve these scoring systems to assess the severity of patients with multiple injuries and to predict their outcome culminated in developing scores with anatomical, physiological, and mixed components, including the following: the Injury Severity Score, New Injury Severity Score, Revised Trauma Score, and the Trauma and Injury Severity Score (TRISS) [4–9].

*Correspondence: ritavieira@academico.ufs.br

¹ CAPES Foundation, Ministry of Education, Brasília, Brazil

Full list of author information is available at the end of the article

Prognostic models were developed in the process of improving the indices to support early clinical decision making and make predictions about the outcomes of the pathology or treatment on the basis of information collected during hospital admission [9–12]. These models are statistical tools that predict clinical outcomes on the basis of at least two predictors used to estimate outcome within a specific period [9–12]. Specific TBI models have been developed and validated in different countries in recent years to help decision making and guide the clinical care of these patients [10, 11, 13, 14].

Although the Glasgow Coma Scale (GCS) is widely used to predict TBI severity and is often associated with mortality and unfavorable outcome after this injury, new TBI-specific prognostic models have been showing good performance in predicting these outcomes and simultaneously considering imaging and laboratory test results, in addition to GCS, pupillary reactivity, and other physiological variables [9–14].

Several models were observed in the literature that have not been submitted to external evaluation and some reviews that addressed the predictive capacity of externally validated models for application after TBI [9, 11]. However, the studies that externally analyzed the models did not consider the different variations of each type of model and did not perform a meta-analysis of the predictive capacity. Given the above, the aim of this study was to analyze the results of investigations that performed external validation or compared prognostic models to identify the models and their variations which showed the best performance in predicting mortality, survival, and unfavorable outcome after severe TBI.

Method

This systematic review and meta-analysis were reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist [15] and were registered at the Prospective Register of Systematic Reviews (Center for Reviews and Dissemination, University of York and the National Institute for Health Research). The registration number is CRD42018116498.

Information Sources and Search Strategy

The inclusion criteria were based on the “PICOTS” acronym proposed by Cochrane: population, adult and older adult patients with severe TBI; index test, studies with externally invalidated prognostic models; comparator, no predefined comparator; outcome, death, survival and unfavorable outcome; timing, any moment after severe TBI; and setting, not specified [16]. We followed the guidelines proposed by the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis

or Diagnosis statement [17] to consider a study as eligible. The inclusion of studies was based on the following guiding question: which prognostic model has the best performance in predicting mortality/survival and unfavorable outcome in patients with severe TBI? Therefore, this study sought to find evidence of prognostic models and their variations used to make individualized predictions in patients with severe TBI, validated against a set of external data (meaning with different participants) to the development of the model [16].

The following databases were initially searched: National Library of Medicine National Institutes of Health (PubMed), Cumulative Index to Nursing and Allied Health Literature, Scopus and Web of Science using the keywords or descriptors, in agreement with the Health Sciences Descriptors, the Medical Subject Headings, and the Emtree of Elsevier Life Science. A partial search of the gray literature was performed using the Google Scholar, TROVE, and Open Grey databases. The Boolean operators “AND” and “OR” were used to communicate descriptors and keywords in the search strategy applied in the databases. The full list of expanded search terms is available in Supplementary Information 1. The searches were customized according to the characteristics of each database.

The initial search was performed concomitantly in all databases on December 10, 2020, and included all records found in these databases without language restrictions or time limits in the inclusion criteria. After performing the search in the databases, the records were stored in the EndNote tool (EndNote X9, Thomson Reuters) and classified in different folders, and the duplicates were excluded using the EndNote computer software program (EndNote X9, Thomson Reuters).

Eligibility Criteria

The following inclusion criteria were adopted: articles that assessed prognostic models; articles in which the studied population was patients with severe TBI; and articles that assessed the predictive capacity of externally validated prognostic models through discrimination and calibration analysis, in addition to being available in selected databases. In our review, mild and moderate TBI were not included.

The following exclusion criteria were adopted for selecting the records: experimental studies in animals and titles that did not evaluate prognostic models in TBI alone; titles that analyzed models of biomarkers to predict outcomes; titles without an abstract; or titles without the full article/text. Theses, dissertations, book chapters, case series, comments, editorials, forum abstracts, literature review, systematic reviews and meta-analyses were not included in the review.

Study Selection Data Collection Process

A selection process was conducted by two independent reviewers (RCAV and JCPS). Titles and abstracts were screened using an online software program (Rayyan, Qatar Computing Research Institute; <https://rayyan.qcri.org/welcome>) and the eligibility criteria were applied by the reviewers separately. Any discrepancies were resolved by a consensus discussion with a third reviewer (RMCS) who was involved to make a final decision, if necessary.

The publications that were retained after this review assessed the predictive ability of the tools for mortality/survival and unfavorable patient outcome after severe TBI. Thus, studies that evaluated the accuracy of the models in predicting outcomes through discrimination by applying the area of the curve (AUC) or C statistic tests, in addition to calibration as a statistical analysis, were analyzed. In our review, $AUC \geq 0.80$ was an indication of good discrimination ability, with AUCs from 0.70 to 0.80 indicating acceptable performance and an AUC below 0.70 indicating poor discrimination [18]. Calibration is a measure of agreement between predicted and observed probabilities of outcomes across the full range of probabilities. Adequate calibration (intercept and slope not statistically different than zero or one, respectively, or $p < 0.05$) indicates that outcome probabilities predicted by the models were similar to those observed [18–20].

Data Extraction

Data extraction was performed according to the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies checklist [21] by two researchers (JCPS and RCAV). In cases in which it was necessary to obtain information beyond what was published, the reviewers contacted the authors of the materials included in the analysis by email in order to clarify the question and include or exclude the study for quantitative analysis.

The EndNote program was used to sort the results of each database according to title, author, year of publication, and abstract into different folders, and it also excluded duplicate publications. The collected data were categorized and entered into a Microsoft Office Excel 2016 spreadsheet and organized into tables to facilitate interpreting the results. At this stage, the manual search for additional studies in the reference sections of the included studies was also performed by the first and second reviewer.

Risk of Bias and Quality Assessment

The selected studies were assessed by using the Prediction Model Risk of Bias Assessment Tool [22] to assess risk of bias and applicability of the models. Two

reviewers (JCPS and RCAV) independently evaluated the included studies and all decisions about the scoring system were agreed on by all reviewers prior to critical appraisal assessments. The tool comprises four domains: participants, predictors, outcome, and analysis, each appraised separately and answered with “no,” “probably no,” “probably yes,” “yes,” or “no information.” If one or more questions in a domain were answered with “probably no” or “no,” the study was considered at high risk of bias regarding that domain. Studies were assumed to be at a low risk of bias if it scored low in all domains, high risk if anyone domain had a high risk of bias, and unclear risk of bias for any other rating.

There is no Grading of Recommendations, Assessment, Development, and Evaluation system guidance available to judge the quality of evidence across prognostics models; therefore, this type of assessment was not performed.

Data Synthesis

The quantitative analysis of results was performed by differences between AUC and calibration results. The estimated AUC and its 95% confidence interval (95% CI) were evaluated separately for each outcome (mortality/survival and unfavorable outcome).

A meta-analysis was conducted to evaluate the discrimination of externally validated models using the random-effects model [16, 23]. A prognosis meta-analysis was performed on the AUC of each model in each study. A logistic transformation was applied in the AUC and its variance was computed as suggested in Ruscio [24]. Meta-regression models were also conducted using mixed-effects models, using the AUC as a dependent variable and the CRASH model, IMPACT model, risk of bias, and age were used as independent variables. A sensitivity analysis was conducted to evaluate studies with low and unclear risk of bias and observational studies. All analyses were conducted using the R version 4.1.3 software program using the metafor package.

Summarizing the estimates of calibration mortality and unfavorable outcomes in the external validation of each model was not performed because the predicted outcome frequency per group at a specific time point and the studies reported different types of summary statistics calibration. This made it impossible to extract and compare the information from different models on the total number of observed and expected events ratio of the overall model calibration [23].

Results

The search in the databases resulted in 1616 records, with 1400 from the five databases, 214 from the gray literature,

and 2 were added through analysis of the reference list of previously eligible publications. Using the EndNote tool, 927 duplicate records were identified that were excluded, leaving 689 records selected for the analysis of titles and abstracts. Of the 689 records, 655 were excluded according to exclusion criteria, as can be seen in Fig. 1.

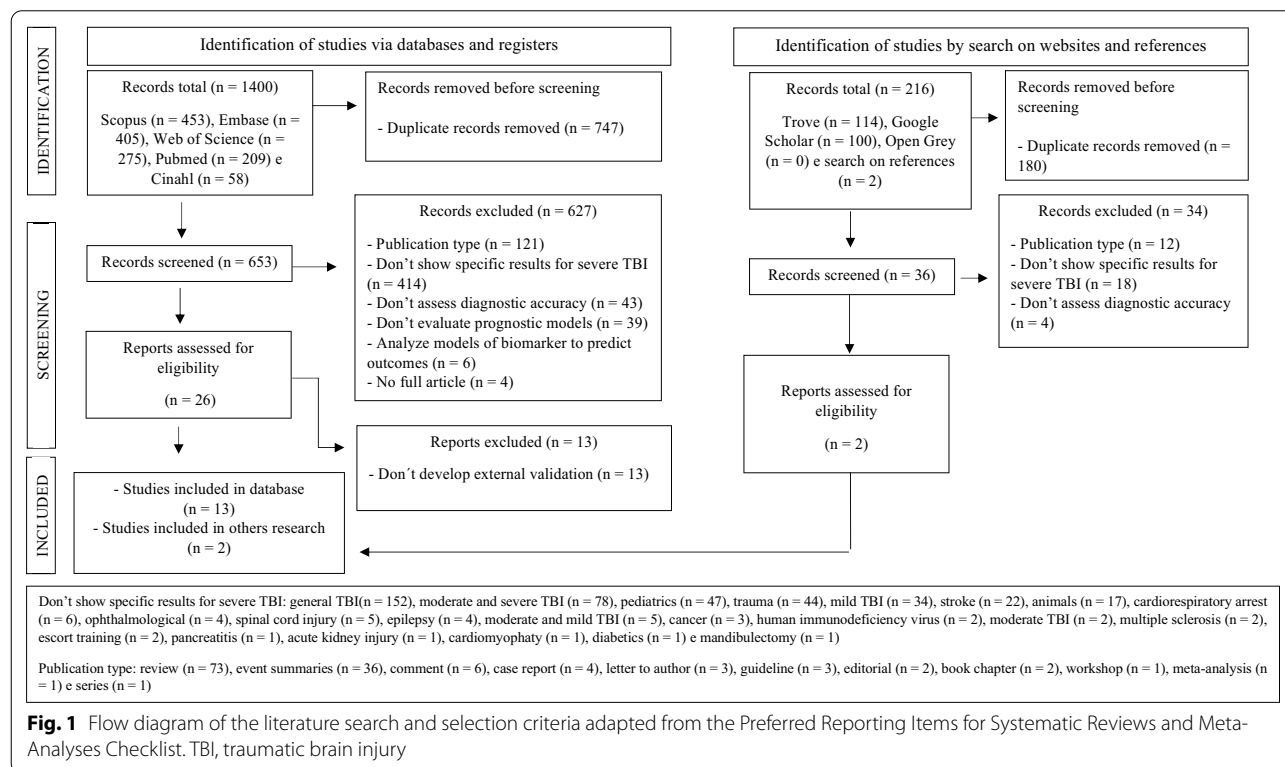
Of the total of 15 eligible scientific articles included in our study [25–39], 8 (53.33%) studies externally validated only one model and 7 (46.66%) studies compared the performance of different prognostic predictors. Among the publications that analyzed one model, five analyzed the International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury (IMPACT), one analyzed the Corticosteroid Randomization After Significant Head Injury (CRASH), and two publications developed and externally validated other models. The seven articles that compared the performance of prognostic predictors used the CRASH and IMPACT ($n=4$), CRASH, IMPACT, and TRISS ($n=1$), Mortality Probability Model (MPM II), Simplified Acute Physiology Score II (SAPS II) ($n=1$), and Rotterdam Computer Tomography (CT) and Helsinki CT score system ($n=1$).

Studies have been published in Europe, Singapore, China, India, Japan, United States of America, and South America, of which 26.66% ($n=4$) were conducted in the United States and 13.33% ($n=2$) in Sweden. The publication period of the 15 articles included in the review was

from 2006 to 2020. The number of events per variable for mortality mean was 29.20 (standard deviation of 29.96) in 33 studies, and for unfavorable outcome the events per variable mean was 32.37 (standard deviation of 30.44) in 31 studies, and the age range ranged from 8 to 98 years. The studies were originally cohort in 73.33% and derived from randomized clinical trials in 26.67% of the articles (Supplementary Information 2).

AUC of External Validation Studies

The data in Table 2 show the eight external validation studies [25, 27, 28, 30, 31, 33, 34, 36] included in the review. The eight selected studies had an AUC for mortality from 0.67 to 0.89 and an AUC for an unfavorable outcome from 0.68 to 0.81. In the mortality analyses, one of the models proposed and externally validated by the authors reached $AUC \geq 0.80$, in addition to studies that evaluated the IMPACT in the extended and laboratory proposal. The IMPACT also reached $AUC \geq 0.80$ for unfavorable outcome in a study in the core modality and in another for laboratory, as well as the model developed by Cremer et al. [25]. On the other hand, Sun et al. [34] showed $AUC < 0.70$ for the IMPACT core in the analysis of the two outcomes and for the extended and laboratory IMPACT when mortality was the studied outcome. Olivecrona et al. [31] observed $AUC = 0.69$ for



an unfavorable outcome when they analyzed the CRASH CT (Table 1).

AUC of the Comparison Studies

Table 2 shows the seven studies [26, 29, 32, 35, 37–39] that compared the prognostic models, in which five studies [29, 32, 35, 37, 39] used IMPACT and CRASH as comparisons. The AUC results ranged from 0.64 to 0.90 for mortality and 0.60–0.89 for an unfavorable outcome. The models that showed good results ($AUC \geq 0.80$) in predicting mortality were: MPM II (AUC between 0.81 and 0.90), SAPS II (AUC between 0.84 and 0.87), Helsinki CT score system ($AUC = 0.82$), CRASH CT and basic (AUC 0.83 and 0.80, respectively), as well as IMPACT core, extended and laboratory (AUC 0.80, 0.81 and 0.80, respectively). The models that presented $AUC \geq 0.80$ for an unfavorable outcome were: CRASH basic and CT (AUC = 0.86 and 0.89, respectively), IMPACT core, extended and laboratory (AUC = 0.84, 0.88 and 0.87, respectively). The study by Roozenbeek et al. [29], in which CRASH basic and IMPACT core were applied in three large databases, showed $AUC < 0.70$ for both outcomes when applying the models in the Phamos data set and $AUC = 0.69$ when applying the CRASH basic in the National Acute Brain Injury Study data set in the unfavorable outcome analysis. Charry et al. [35] found weak discrimination capacity for unfavorable outcome in CRASH CT and IMPACT lab, whereas Wongchareon et al. [39] observed $AUC < 0.70$ for mortality when

applying CRASH basic and CT and in IMPACT core (Table 2).

Risk of Bias and Applicability

In evaluating the 15 studies by using the Prediction Model Risk of Bias Assessment Tool [22], it was observed that six studies had low risk of bias [25, 26, 32–34, 39], four studies had an unclear risk of bias [28–30, 38], and five studies had a high risk of bias [27, 31, 35–37]. For the judgment of applicability, only two studies were classified as high applicability [36, 38] and five as unclear [26, 27, 29, 31, 34] applicability (Fig. 2).

Meta-analysis

The results in Figs. 3 and 4 describe the mixed-effects models for meta-regression estimated by restricted maximum likelihood for the studies included in the meta-analysis. To increase the accuracy of the estimated effect, studies with less than 100 participants with the outcome of interest were excluded from the meta-analysis [27, 31] and a model in which the authors applied CRASH at a different time to that proposed in their validation [35]. The remaining 12 studies analyzed nine models, of which seven (two developed models for authors, MPMII, SAPS II, Rotterdam CT, Helsinki CT, and TRISS) were excluded from the meta-analysis because they were tested in a single data set. After these exclusions, CRASH and IMPACT remained with

Table 1 Discrimination results of the eight studies that performed external validation

Author	Outcome	Model	AUC or C statistic mortality	AUC or C statistic unfavorable outcome
Cremer et al. [25]	Functional recovery or disability after 12 months	Developed model	–	0.81
Olivecrona et al. [27]	6-month mortality and unfavorable outcome	IMPACT	Core: 0.74 Extended: 0.75 Lab: 0.78	Core: 0.73 Extended: 0.72 Lab: 0.81
Roozenbeek et al. [28]	14-day mortality	IMPACT	Core: 0.79 (95% CI 0.77–0.81) Extended: 0.83 (95% CI 0.81–0.85)	–
Panczykowski et al. [30]	6-month mortality and unfavorable outcome	IMPACT	Core: 0.78 Extended: 0.83 Lab: 0.83	Core: 0.76 Extended: 0.78 Lab: 0.76
Olivecrona et al. [31]	14-day mortality and 6-month unfavorable outcome	CRASH	Do not have this statistical analysis for mortality	CT: 0.69
Gómez et al. [33]	Early death (within 48 h)	Developed model	0.89 (95% CI 0.84–0.93)	–
Sun et al. [34]	6-month mortality and unfavorable outcome	IMPACT	Core: 0.67 Extended: 0.69 Lab: 0.69	Core: 0.68 Extended: 0.70 Lab: 0.71
Wan et al. [36]	6-month mortality and unfavorable outcome	IMPACT	Core: 0.76 Extended: 0.76 Lab: 0.73	Core: 0.80 Extended: 0.79 Lab: 0.77

AUC, area under the curve; CI, confidence interval; CRASH, Corticosteroid Randomization After Significant Head Injury; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; Lab, laboratory

Table 2 Discrimination results of the seven selected comparison studies

Author	Outcome	Model	AUC mortality	AUC unfavorable outcome
Fischler et al. [26]	Hospital- mortality and 1-year mortality	Mortality Probability Model II (MPM II) and Simplified Acute Physiology Score II (SAPS II)	MPM II Hospital mortality: Zero hour: 0.83 (95% CI 0.77–0.88) and 24 h: 0.90 (95% CI 0.86–0.94) One-year mortality Zero hour: 0.81 (95% CI 0.75–0.87) and 24 h: 0.89 (95% CI 0.85–0.93) SAPS II Hospital mortality: 0.87 (95% CI 0.82–0.92) One-year mortality: 0.84 (95% CI 0.78–0.89)	–
Roozenbeek et al. [29] ^a	14-day and 6-month mortality and 6-month unfavorable outcome	CRASH and IMPACT	CRASH Basic: 0.69 ^b Basic: 0.74 ^c Basic: 0.65 ^d IMPACT Core: 0.70 ^b Core: 0.75 ^c Core: 0.65 ^d	CRASH Basic: 0.74 ^b Basic: 0.75 ^c Basic: 0.68 ^d IMPACT Core: 0.73 ^b Core: 0.71 ^c Core: 0.66 ^d
Han et al. [32]	14-day mortality, 6-month mortality and 6-month unfavorable outcome	CRASH and IMPACT	CRASH Basic: 0.80 (95% CI 0.75–0.85) CT: 0.83 (95% CI 0.78–0.87) IMPACT Core: 0.80 (95% CI 0.75–0.85) Extended: 0.81 (95% CI 0.76–0.86) Lab: 0.80 (95% CI 0.75–0.86)	CRASH Basic: 0.86 (95% CI 0.81–0.90) CT: 0.89 (95% CI 0.84–0.93) IMPACT Core: 0.84 (95% CI 0.80–0.89) Extended: 0.88 (95% CI 0.83–0.92) Lab: 0.87 (95% CI 0.82–0.92)
Charry et al. [35]	14-day and 6-month mortality, and 6-month unfavorable outcome	CRASH and IMPACT	CRASH CT: 14-day death: 0.71 (CI 95% 0.59–0.82); 6-month death: 0.70 (95% CI 0.58–0.81) IMPACT Lab: 6-month death: 0.75 (95% CI 0.65–0.84)	CRASH CT: 0.60 (IC 95% 0.50–0.71) IMPACT Lab: 0.67 (95% CI 0.57–0.76)
Maeda et al. [37]	Hospital mortality and 6-month unfavorable	TRISS, CRASH and IMPACT	TRISS 0.75 (95% CI 0.72–0.79)	CRASH Basic: 0.86 (95% CI 0.82–0.90) CT: 0.86 (95% CI 0.82–0.89) IMPACT Core: 0.81 (95% CI 0.77–0.85) Extended: 0.85 (95% CI 0.80–0.89)
Pargaonkar et al. [38]	3-month mortality	Rotterdam CT and Helsinki CT score system	Rotterdam: 0.75 Helsinki: 0.82	–
Wongchareon et al. [39]	14-day and 6-month mortality and 6-month unfavorable outcome	CRASH and IMPACT	CRASH Basic: 0.64 (95% CI 0.59–0.68) CT: 0.66 (95% CI 0.62–0.70) IMPACT Core: 0.68 (95% CI 0.63–0.73) Extended: 0.73 (95% CI 0.67–0.77) Lab: 0.73 (95% CI 0.68–0.77)	CRASH Basic: 0.72 (95% CI 0.68–0.75) CT: 0.73 (95% CI 0.69–0.76) IMPACT Core: 0.76 (95% CI 0.71–0.79) Extended: 0.77 (95% CI 0.73–0.81) Lab: 0.77 (CI 95% 0.73–0.81)

AUC, area under the curve; CI, confidence interval; CRASH, Corticosteroid Randomization After Significant Head Injury; CT, computerized tomography; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; MPM II, Mortality Probability Model; NABIS, National Acute Brain Injury Study; Lab, laboratory; SAPS II, Simplified Acute Physiology Score II; TBI, traumatic brain injury; TRISS, Trauma and Injury Severity Score

^a In the study by Roozenbeek et al. [29], only severe TBI data related to the prediction of mortality and unfavorable outcome were extracted

^b NABIS data set is a clinical trial conducted in the United States with 392 patients investigating the effect of hypothermia on severe TBI

^c Cerestat is an unpublished randomized controlled multicenter study in a combined group of North American and European centers investigating the effect of Aptiganel HCl, a noncompetitive *N*-methyl-*D*-aspartate receptor antagonist in patients with TBI severe (*n* = 547)

^d Pharmos is a multicenter, randomized, placebo-controlled, phase III study that investigated the safety and efficacy of dexanabol in patients with severe TBI

Table 2 (continued)

(n = 861) conducted in Europe, Israel, Australia, and the United States

Study	Model	Prediction Model Risk of Bias Assessment Tool (PROBAST)				
		Overall judgement of risk of bias				
		Domain 1: Participants	Domain 2: Predictors	Domain 3: Outcome	Domain 4: Analysis	Overall judgement
Charry et al 2017 (cohort)	CRASH and IMPACT	Low	Low	Unclear	High	High
Maeda et al 2019 (cohort)	TRISS, CRASH and IMPACT	Low	Unclear	Low	High	High
Wongchareon et al 2020 (cohort)	CRASH and IMPACT	Low	Low	Low	Low	Low
Wan et al 2017 (cohort)	IMPACT	Low	High	Low	Unclear	High
Olivecrona et al 2012 (randomized controlled trial)	IMPACT	Low	High	Low	High	High
Cremer et al 2006 (cohort)	Model developed	Low	Low	Low	Low	Low
Fischler et al 2007 (cohort)	SAPS II and MPM II	Low	Low	Low	Low	Low
Gómez et al 2014 (cohort)	Model developed	Low	Low	Low	Low	Low
Han et al 2014 (cohort)	CRASH and IMPACT	Low	Low	Low	Low	Low
Roozenbeek et al 2012 (randomized controlled trial)	CRASH and IMPACT	Low	Low	Low	Unclear	Unclear
Panczykowski et al 2012 (cohort)	IMPACT	Low	Low	Low	Unclear	Unclear
Roozenbeek et al 2012 (cohort)	IMPACT	Low	Low	Low	Unclear	Unclear
Sun et al 2016 (randomized controlled trial)	IMPACT	Low	Low	Low	Low	Low
Olivecrona et al 2012 (randomized trial)	CRASH	Low	Low	Low	High	High
Pargaonkar et at 2019 (cohort)	Rotterdam and Helsinki	Low	Low	Low	Unclear	Unclear

Study	Model	Prediction Model Risk of Bias Assessment Tool (PROBAST)			
		Overall judgement of applicability			
		Domain 1: Participants	Domain 2: Predictors	Domain 3: Outcome	Overall judgement
Charry et al 2017 (cohort)	CRASH and IMPACT	Low	Low	Low	Low
Maeda et al 2019 (cohort)	TRISS, CRASH and IMPACT	Low	Low	Low	Low
Wongchareon et al 2020 (cohort)	CRASH and IMPACT	Low	Low	Low	Low
Wan et al 2017 (cohort)	IMPACT	High	Low	Low	High
Olivecrona et al 2012 (randomized controlled trial)	IMPACT	Unclear	Low	Low	Unclear
Cremer et al 2006 (cohort)	Model developed	Low	Low	Low	Low
Fischler et al 2007 (cohort)	SAPS II and MPM II	Unclear	Low	Low	Unclear
Gómez et al 2014 (cohort)	Model developed	Low	Low	Low	Low
Han et al 2014 (cohort)	CRASH and IMPACT	Low	Low	Low	Low
Roozenbeek et al 2012 (randomized controlled trial)	CRASH and IMPACT	Unclear	Low	Low	Unclear
Panczykowski et al 2012 (cohort)	IMPACT	Low	Low	Low	Low
Roozenbeek et al 2012 (cohort)	IMPACT	Low	Low	Low	Low
Sun et al 2016 (randomized controlled trial)	IMPACT	Unclear	Low	Low	Unclear
Olivecrona et al 2012 (randomized trial)	CRASH	Unclear	Low	Low	Unclear
Pargaonkar et at 2019 (cohort)	Rotterdam and Helsinki	High	Low	Low	High

Fig. 2 Risk of bias and applicability (PROBAST) assessment of the prediction models after severe TBI. CRASH, Corticosteroid Randomization After Significant Head Injury; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; ISS, Injury Severity Score; MPM II, Mortality Probability Model; NABIS, National Acute Brain Injury Study; PROBAST, Prediction Model Risk of Bias Assessment Tool; TBI, traumatic brain injury; TRISS, Trauma and Injury Severity Score; SAPS II, Simplified Acute Physiology Score II

their variations, which were evaluated in nine studies [28–30, 32, 34–37, 39].

In Figs. 3 and 4, we can see the result of pooled AUC ROC for mortality and unfavorable outcome of the studies included in the meta-analysis [28–30, 32, 34–37, 39]. The studies for mortality included n=7012 participants, and it was observed that all models had an acceptable pooled AUC (>0.70) with substantial heterogeneity (tau²=0.08; standard error [SE]=0.03; I²=78.81%) between studies.

A total of 5134 participants were analyzed for unfavorable outcome at 6 months, and all models presented pooled AUC>0.70 (Fig. 4); however, the meta-analysis for unfavorable outcome exhibited considerable heterogeneity (tau²=0.11; SE 0.04; I²=82.96%). The analysis of the results of Figs. 3 and 4 did not show the superiority of any model for predicting mortality and unfavorable outcome when considering the 95% CI of the pooled AUC [28–30, 32, 34–37, 39].

Meta-regression

The meta-regression analysis using to examine if the models (CRASH and IMPACT), risk of bias and age (mean/median) of a study can be used to predict its effect size, mortality, and unfavorable outcome exhibited the effect size of the studies have increased over time (Table 3). For mortality (tau²=0.0011; SE 0.0006; I²=59.7%; R²=61.14%; p<0.001 [test of moderators]), IMPACT extended had AUC 0.073 times greater than CRASH basic, unclear risk of bias had AUC 0.130 times greater than high risk of bias, and each additional year in the age (mean/median) increase the AUC by 0.004. Unfavorable outcome (tau²=0.0012; SE 0.0007; I²=61.65%; R²=61.73%; p<0.001 [test of moderators]) showed low risk of bias had AUC 0.134 times greater than high risk of bias, and unclear risk of bias had AUC 0.108 times greater than high risk of bias.

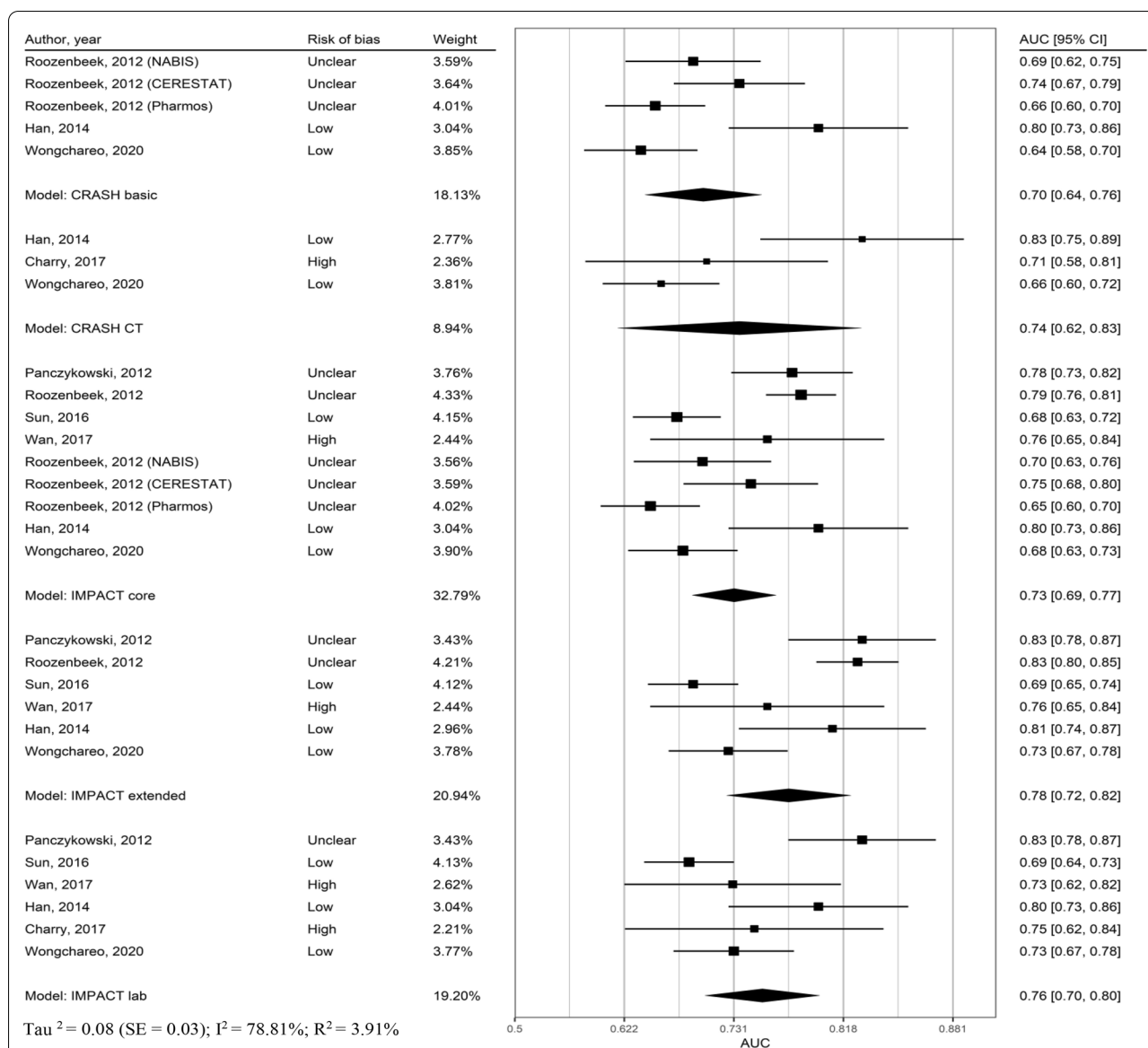


Fig. 3 Meta-analysis of the AUC for prediction models of mortality after severe TBI. Weights are from random-effects analysis. AUC, area under the curve; CI, confidence interval; CERESTAT, Cerestat data set; CT, computerized tomography; CRASH, Corticosteroid Randomization After Significant Head Injury; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; Lab, laboratory; NABIS, National Acute Brain Injury Study; Pharmos, Pharmos data set; SE, standard error; TBI, traumatic brain injury

Sensitivity Analysis

Studies with low or unclear risk of bias [28–30, 32, 34, 39] selected for the meta-analysis were analyzed separately to assess the model performance robustness, as these were observational studies [28, 30, 32, 35, 36, 39] (Supplementary Information 3 and 4) Randomized clinical trials were excluded in this last analysis because they may not be representative for all individuals with severe TBI [29, 34].

Although observational studies showed higher pooled AUC values and lower I² values for mortality and an

unfavorable outcome at 6 months, it was not possible to state that the sensitivity analyses showed better performance when compared with the meta-analyses that included all studies (Figs. 3 and 4). Similar to previous findings (Figs. 3 and 4), there was no evidence of superiority of any prognostic model in the sensitivity analyses for both outcomes according to 95% CI of pooled AUC.

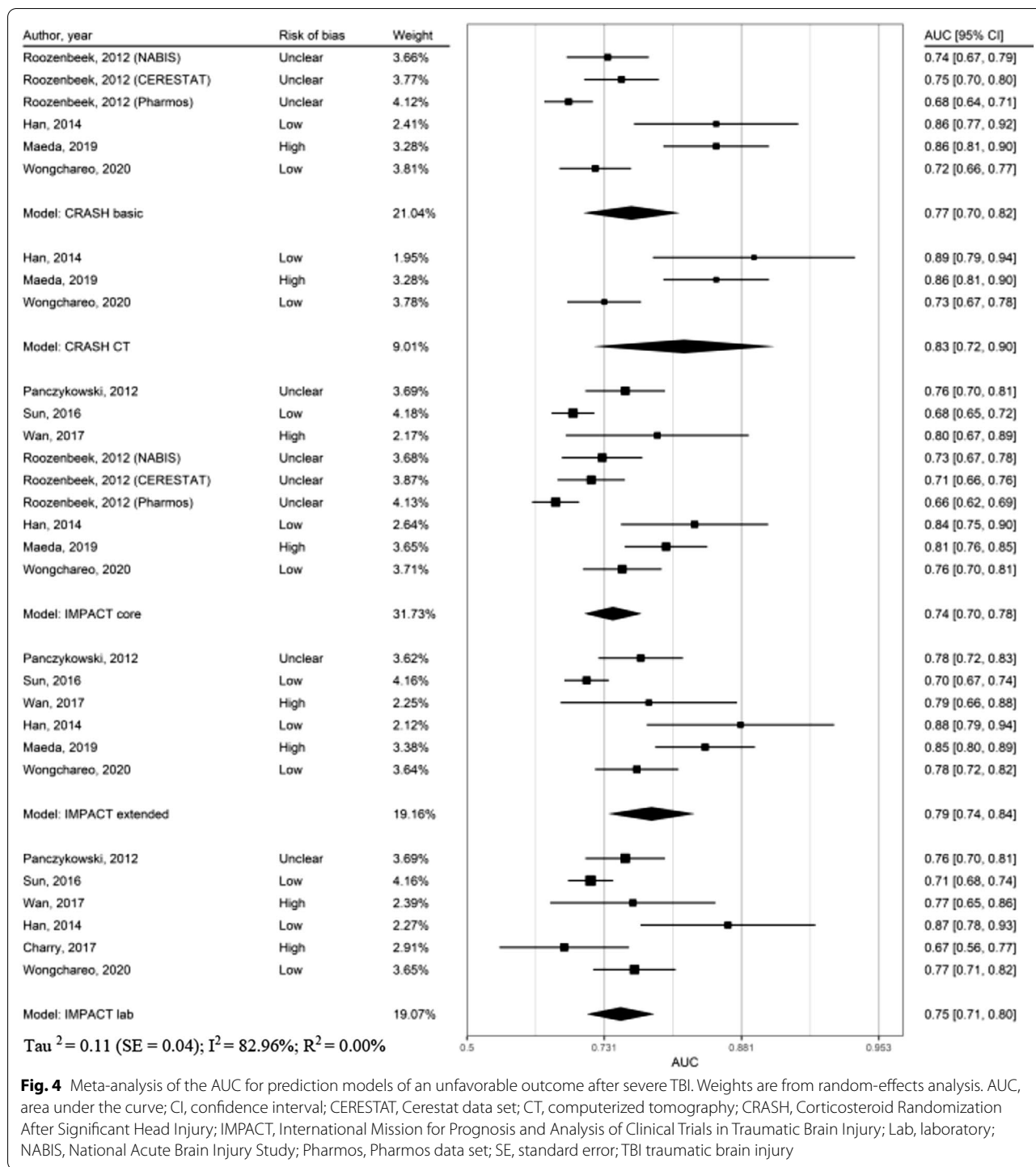


Fig. 4 Meta-analysis of the AUC for prediction models of an unfavorable outcome after severe TBI. Weights are from random-effects analysis. AUC, area under the curve; CI, confidence interval; CERESTAT, Cerestat data set; CT, computerized tomography; CRASH, Corticosteroid Randomization After Significant Head Injury; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; Lab, laboratory; NABIS, National Acute Brain Injury Study; Pharmos, Pharmos data set; SE, standard error; TBI traumatic brain injury

Calibration of Studies

Tables 4 and 5 show the calibration results of the models, in which 11 studies analyzed mortality and 7 articles analyzed unfavorable outcomes. IMPACT was the most evaluated model in terms of calibration for these two outcomes.

For mortality (Table 4), the Hosmer–Lemeshow (HL) test showed adequate fit for the proposed models and externally validated by the authors themselves for MPM II/24 h (hospital mortality), MPM II/initial (mortality one year after TBI), and SAPS II (in-hospital and 1-year mortality). The HL test and the logistic fit indicated

Table 3 AUC Meta-regression analysis^a with means AUC, and CRASH model, IMPACT model, risk of bias, and age

	Mortality			Unfavorable outcome		
	Models estimate	Z value	p value	Models estimate	Z value	p value
CRASH CT	0.065	1.732	0.083	0.024	0.580	0.562
IMPACT core	0.025	1.011	0.312	-0.016	-0.644	0.520
IMPACT extended	0.073	2.697	0.007	0.011	0.363	0.717
IMPACT lab	0.068	2.342	0.019	0.008	0.263	0.793
Low risk of bias	0.079	2.192	0.028	0.134	3.091	0.002
Unclear risk of bias	0.130	3.365	0.001	0.108	2.372	0.018
Age (mean/median) ^b	0.004	4.424	<0.001	0.004	4.956	<0.001

AUC, area under the curve; CRASH, Corticosteroid Randomization After Significant Head Injury; CT, computerized tomography; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury

Significant values are highlighted in bold

^a This is the expected effect size when predictor reference is CRASH basic, high risk of bias, and age is zero

^b Age analysis included ten studies

adequate fitting of IMPACT and CRASH for mortality in four studies. On the other hand, Wongchareon et al. [39] observed poor calibration in all variations of CRASH and IMPACT in their analysis.

The slope and intercept values differed in the studies when IMPACT and CRASH were analyzed for the mortality outcome (Table 4). The values of these parameters for the IMPACT core ranged from 0.73 to 1.14 for slope and -0.77 to +0.81 for intercept [29, 32]; the slope in the extended IMPACT varied between 0.76 and 1.15 and the intercept between -0.66 to +0.59 [32, 34]; and the laboratory IMPACT varied between 0.76 and 1.17 for slope and -0.59 and +0.81 for intercept [32, 34]. The CRASH basic presented a slope variation between 0.64 and 0.95, in the intercept between -1.00 and +0.51 [29, 32], and the CT, which was only evaluated in two samples, presented results of 0.85 and 1.05 for slope and +0.03 and +0.07 for intercept [32, 39].

HL and logistic fit tests for unfavorable outcome showed adequate fit for the IMPACT core and extended models in four out of five samples [27, 30, 32, 36, 39] and in all analyses performed with IMPACT laboratory. CRASH basic was evaluated in two studies [32, 39], with a poor calibration result for one of them [39], whereas the CRASH CT presented poor results in one of the three analyses carried out with this model variation [31, 32, 39].

The slope and intercept in the unfavorable outcome analyses (Table 5) also showed important differences in the values found. The variation for the IMPACT core was from 0.64 to 1.37 for slope [29, 32] and -0.36 to +1.07 for intercept [29, 32]; in the extended IMPACT, values between 0.72 and 1.58 were observed for slope [32, 34] and -0.01 and 0.88 for intercept [32, 39]; the IMPACT laboratory presented values between 0.71 and 1.46 for

slope [32, 34] and +0.11 to +0.94 [32, 39] for intercept. CRASH basic ranged from 0.64 to 1.34 for slope [29, 32] and -0.01 to 2.39 for intercept [29, 32], whereas the slope values in CRASH CT were 0.98 and 1.22 and the intercept values were +0.17 and +1.78 in two analyses performed [32, 39].

Discussion

In the present systematic review and meta-analysis, we evaluated the effectiveness of prognostic models that went external validation in predicting mortality/survival and unfavorable outcome after severe TBI. This review is, to our knowledge, the first to identify published studies attempting to evaluate prognostic models specifically for patients with severe TBI and systematically analyze the CRASH and IMPACT variation. In a recent systematic review, Dijkland et al. [11] found more than 67 internally and externally validated prognostic models in patients with moderate and severe TBI in the literature. Although the brilliant review by Dijkland et al. [11] brought several contributions, the authors did not compare CRASH and IMPACT in a meta-analysis with other evaluated models (e.g., Helsinki, Hukkelhoven and Nijmegen) due to the limited number of validations found. Though similarities between our results and the results of Dijkland et al. [11] were observed, the variations of CRASH and IMPACT models (e.g., CRASH basic, CRASH CT, IMPACT core, IMPACT extended and IMPACT lab) were not included in Dijkland et al. [11], and patients with severe TBI were not analyzed separately, but these aspects were explored in our current review.

CRASH and IMPACT are regression-based models derived from large databases of randomized, observational studies which have tested specific treatments in patients with TBI [10, 13]. As a result, the eligibility

Table 4 Calibration results of the prognostic models for mortality

Calibration test	Model	Results	Author
Hosmer–Lemeshow	Developed model	$p=0.28$	Cremer et al. [25]
	Developed model	$p=0.32$	Gómez et al. [33]
	CRASH basic	$p>0.05$	Han et al. [32]
	CRASH CT	$p>0.05$	
	IMPACT core	$p>0.05$	
	IMPACT extended	$p>0.05$	
	IMPACT lab	$p>0.05$	
	IMPACT core	$p=0.70$	Panczykowski et al. [30]
	IMPACT extended	$p=0.89$	
	IMPACT lab	$p=0.80$	
	IMPACT core	$p=0.28$	Wan et al. [36]
	IMPACT extended	$p=0.25$	
	IMPACT lab	$p=0.58$	
	CRASH basic	$p=0.051$	Wongchareon et al. [39]
	CRASH CT	$p=0.001$	
	IMPACT core	$p=0.040$	
	IMPACT extended	$p=0.049$	
	IMPACT lab	$p=0.002$	
	MPM II		Hospital: Zero hour: 17.83; $p=0.02$ 24 h: 10.68; $p=0.22$ One year: Zero hour: 14.21; $p=0.08$ 24 h: 18.14; $p=0.02$
SAPS II		Hospital: 7.15; $p=0.52$ One year: 5.67; $p=0.68$	
Logistic fit	IMPACT core	$p=0.03$	Olivecrona et al. [27]
	IMPACT extended	$p=0.03$	
	IMPACT lab	$p=0.02$	
Goodness-of-fit test	Developed model	Intercept: +0.03 Slope: 0.83	Gómez et al. [33]
Calibration plots and calibration-in-the-large or Cox calibration	IMPACT core	Intercept: -0.61 Slope: 1.06	Roozenbeek et al. [28]
	IMPACT extended	Intercept: +0.03 Slope: 1.12	
	CRASH basic	Intercept: -0.65 Slope: 0.67 ^a Intercept: -0.65 Slope: 0.89 ^b Intercept: -1.00 Slope: 0.64 ^c	Roozenbeek et al. [29] ^d
	IMPACT core	Intercept: -0.40 Slope: 0.84 ^a Intercept: -0.22 Slope: 1.05 ^b Intercept: -0.77 Slope: 0.73 ^c	
	CRASH basic	Intercept: +0.51 (95% CI 0.20–0.82) Slope: 0.95 (95% CI 0.71–1.18)	Han et al. [32]
	CRASH CT	Intercept: +0.03 (95%CI -0.24 to 0.31) Slope: 0.85 (95% CI 0.64–1.05)	
	IMPACT core	Intercept: +0.81 (95% CI 0.48–1.13) Slope: 1.14 (95% CI 0.85–1.43)	
	IMPACT extended	Intercept: +0.59 (95% CI 0.29–0.89) Slope: 1.15 (95% CI 0.86–1.43)	
	IMPACT lab	Intercept: +0.81 (95% CI 0.48–1.15) Slope: 1.17 (95% CI 0.86–1.49)	
	IMPACT core	Intercept: -0.70 Slope: 0.81	Sun et al. [34]
	IMPACT extended	Intercept: -0.66 Slope: 0.76	
	IMPACT lab	Intercept: -0.59 Slope: 0.76	
	CRASH basic	Intercept: +0.10 Slope: 0.76	Wongchareon et al. [39]
	CRASH CT	Intercept: +0.07 Slope: 1.05	
	IMPACT core	Intercept: +0.07 Slope: 0.98	
IMPACT extended	Intercept: +0.03 Slope: 1.08		

Table 4 (continued)

Calibration test	Model	Results	Author
	IMPACT lab	Intercept: + 0.05 Slope: 1.05	

CI, confidence interval; CRASH, Corticosteroid Randomization After Significant Head Injury; CT, computerized tomography; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; lab, laboratory; MPM II, Mortality Probability Model; NABIS, National Acute Brain Injury Study; SAPS II, Simplified Acute Physiology Score II; TBI, traumatic brain injury

^a NABIS Dataset

^b Cerestat Dataset

^c Pharmos Dataset

^d In the study by Roozenbeck et al. [29], data related to the prediction of mortality and unfavorable outcome were considered only after severe TBI

criteria of the participants of these databases can interfere with the efficiency of the generated prognostic models, which will fit better for certain case mixes. Our meta-analysis results that analyzed the discriminative capacity showed similar performance of the prognostic models and their variations. The studies [29, 32, 35, 37, 39] that compared the CRASH and IMPACT models showed that both have good performance in predicting mortality and unfavorable outcomes in their populations. However, some studies warn that a discrimination comparison between CRASH and IMPACT models may be influenced by the inclusion criteria regarding patient age, countries, trauma care organization, treatment, predictor effects, outcome distribution and lack of available data [34, 39–43]. Therefore, the difference in case-mix distribution between the databases of the models submitted for validation and development can influence the predictive capacity. Therefore, some researchers [29, 34, 37] recommend fitting the coefficients of the models for the target population before its implementation in clinical practice to correct these interferences [11, 34].

Although contemporary clinical studies indicate the continuing relevance of CRASH and IMPACT, in a recent study, Wongchareon et al. [39] propose updating prognostic models, including new predictors, to maintain currency and generalizability, as these models were derived from a database that included patients from 1985, and advances in the treatment and diagnostic resources of patients with TBI may have altered the performance of the models. Understanding the role of prognostic models could be helpful for risk stratification; however, external validation is recommended to assess prediction in a new setting. CRASH and IMPACT showed moderate to good discrimination results; however, the calibration analysis showed overestimated risks for mortality and underestimated risks for unfavorable results [11]. Although, in this review, IMPACT extended and CRASH CT led to high pooled AUC values when predicting mortality and unfavorable outcomes, the meta-analysis did not show the superiority of CRASH and IMPACT variations in predicting outcomes. The present systematic review confirmed that heterogeneous models that used blood

biomarkers, imagining biomarkers and clinical characteristics could be used as a caution for global prediction outcome, being recommended to update the model (external validation) per setting.

Over the past decades, CT scans, GCS, pupillary changes, and age have been frequently used as predictors to estimate the outcome after TBI in prognostic models. Dijkland et al. [11] observed that age, GCS motor score, and pupils were the most common components of the models and that CRASH and IMPACT were the most frequently submitted to external validation. By identifying the need to predict the outcome of patients with severe TBI readily, some authors have externally validated models using only variables that can be reliably investigated after trauma. The findings of these studies showed that the models proposed and externally validated by Gómez et al. [33] and Cremer et al. [25] performed well in predicting mortality at 48 h (AUC=0.89) [33] and unfavorable outcomes after 12 months (AUC=0.81) [25]. In the past decade, CRASH and IMPACT models have been used in patients with TBI (GCS 3–14). Although CRASH and IMPACT models were development in large cohorts with different severity, it is recommended that these models are adjusted according to the local and clinical characteristics population.

In the descriptive analysis, it was observed that all the studies except for the study by Wongchareon et al. [39] indicated adequate calibration of IMPACT and CRASH for both outcomes by the HL test, but without a clear indication regarding the type and extent of the calibration error identified by the slope and intercept of the Cox calibration, which had widely varied values in the studies included in this review. Using the HL test, the study by Wongchareon et al. [39] showed that the CRASH basic and IMPACT laboratory models presented adequate fit ($p > 0.05$) for mortality and unfavorable outcome, respectively. However, from the calibration plots and intercept, CRASH and IMPACT generally predicted lower probabilities of both outcomes (mortality and unfavorable outcome) than what was observed, except for the IMPACT CT model, which was more accurate in predicting an unfavorable outcome at six months, only if the expected

Table 5 Calibration results of the prognostic models for unfavorable outcome

Calibration test	Model	Results	Author
	IMPACT core	$p=0.39$	Panczykowski et al. [30]
	IMPACT extended	$p=0.17$	
	IMPACT lab	$p=0.78$	
	CRASH basic	$p>0.05$	Han et al. [32]
	CRASH CT	$p>0.05$	
	IMPACT core	$p>0.05$	
	IMPACT extended	$p>0.05$	
	IMPACT lab	$p>0.05$	
	IMPACT core	$p=0.35$	Wan et al. [36]
	IMPACT extended	$p=0.23$	
	IMPACT lab	$p=0.55$	
	CRASH basic	$p=0.00$	Wongchareon et al. [39]
	CRASH CT	$p=0.04$	
	IMPACT core	$p=0.00$	
	IMPACT extended	$p=0.002$	
	IMPACT lab	$p=0.16$	
Logistic fit	IMPACT core	$p=0.003$	Olivecrona et al. [27]
	IMPACT extended	$p=0.007$	
	IMPACT lab	$p=0.0005$	
	CRASH CT	$p<0.01$	Olivecrona et al. [31]
Calibration plots or calibration-in-the-large or Cox calibration	CRASH Basic	Intercept: +0.04 Slope: 0.80 ^a Intercept: -0.72 Slope: 0.95 ^b Intercept: -0.01 Slope: 0.64 ^c	Roozenbeek et al. [29] ^d
	IMPACT core	Intercept: +0.42 Slope: 0.89 ^a Intercept: -0.36 Slope: 0.81 ^b Intercept: +0.29 Slope: 0.63 ^c	
	CRASH basic	Intercept: +2.39 (95% CI 1.84–2.95) Slope: 1.34 (95% CI 1.01–1.67)	Han et al. [32]
	CRASH CT	Intercept: +1.78 (95% CI 1.33–2.22) Slope: 1.12 (95% CI 0.85–1.39)	
	IMPACT Core	Intercept: +1.07 (95% CI 0.74–1.39) Slope: 1.37 (95% CI 1.02–1.71)	
	IMPACT extended	Intercept: +0.88 (95% CI 0.55–1.21) Slope: 1.58 (95% CI 1.20–1.97)	
	IMPACT lab	Intercept: +0.94 (95% CI 0.59–1.30) Slope: 1.46 (95% CI 1.09–1.83)	
	IMPACT Core	Intercept: +0.23 Slope: 0.72	Sun et al. [34]
	IMPACT extended	Intercept +0.34 Slope: 0.72	
	IMPACT lab	Intercept: +0.42 Slope: 0.71	
	CRASH basic	Intercept: +0.17 Slope: 0.88	Wongchareon et al. [39]
	CRASH CT	Intercept: +0.17 Slope: 0.98	
	IMPACT Core	Intercept: +0.07 Slope: 1.01	
	IMPACT extended	Intercept: -0.01 Slope: 1.15	
	IMPACT lab	Intercept: +0.11 Slope: 1.04	

CI, confidence interval; CRASH, Corticosteroid Randomization After Significant Head Injury; CT, computerized tomography; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury; lab, laboratory; MPM II, Mortality Probability Model; NABIS, National Acute Brain Injury Study; SAPS II, Simplified Acute Physiology Score II; TBI, traumatic brain injury

^a NABIS Dataset

^b Cerestat Dataset

^c Pharmos Dataset

^d In the study by Roozenbeek et al. [29], data related to the prediction of mortality and unfavorable outcome were considered only after severe TBI

outcome is less severe (probability from 0 to approximately 20%) [39].

The meta-regression analysis of the CRASH and IMPACT models showed higher AUC for studies with a low risk of bias when compared with an increased risk of bias in both outcomes (mortality and unfavorable outcome). This may be attributable to the studies assessing discrimination but not providing the model calibration analysis, thereby increasing the risk of bias. In this case, the model's performance in providing accurate individual probabilities is completely unknown [22]. Moreover, we observed that studies with a higher risk of bias analyzed in this review had not complete information about the predictors available, which may result in the overestimation of discrimination. Thus, using specific guidelines to assess the risk of bias is essential to prevent overestimation discrimination and identify if the results are distorted estimates of model performance [22]. Another interesting observation is about the increase of AUC values with age. This may be related to the good ability of the models to predict the outcomes in older patients. It is known that increased age is related to diminished functional capacity and ability to recover [40].

It is important to emphasize that studies included in this review applied different types of statistical analyses to assess the calibration of the models (11 studies have conducted calibration), with the HL test and Cox calibration being more frequent. Furthermore, some studies did not report the 95% CI values and did not report estimated probabilities at each calibration point, making it impossible to perform the meta-analysis of the calibration for the models in our study.

The use of prognostic models after severe TBI can guide the team in predicting resources, informing the family about the patient's prognosis, and assessing the quality of care provided. However, their use should be cautioned because these models predict group estimates of patients, and they cannot accurately predict individual outcomes or are not recommended for clinical decision making [32]. It is known that the consequences and outcomes caused by severe TBI are different in patients with mild and moderate TBI, so it is important that studies carry out validation of models according to the severity of the injury, identifying whether this is an interfering variable in its accuracy. Therefore, the authors recommend taking into account the inclusion criteria adopted in the studies of the model used and the application of the AUC case-mix in the studies in order to make the results generalizable [34, 39, 41–43].

Some difficulties occurred in the analysis of the studies included in the review, with a discrepancy in the inclusion criteria observed in the studies and a lack of

similarity in the information presented in the results of the articles standing out. Nevertheless, it was notable that the discrimination and calibration results present adequate levels when the inclusion criteria of the studies are the same or very similar to those of the validation study of the developed model, indicating an accurate estimation of the outcomes by the models [11, 34, 39]. We also noted that a lack of clear information on the methodology used for model validation, identification of complete AUC and calibration results, and information on the number of deaths were also difficulties faced during the analyses.

Conclusions

The results of this review showed that the IMPACT and CRASH models were the most validated and externally compared in patients with severe TBI and generally presented acceptable performance in predicting the studied outcomes. Among these two models and their variations, a similarity in the ability to discriminate for mortality and unfavorable outcome was observed, indicating that both models can contribute to estimate prognosis in clinical practice in patients with severe TBI. The results also showed that the variations of the CRASH and IMPACT models had adequate calibration for both outcomes in most studies, but without a clear indication of uncertainties in the evaluations of these models. Overall, the analyses did not show a model that performed better than the others.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1007/s12028-022-01547-7>.

Author details

¹ CAPES Foundation, Ministry of Education, Brasília, Brazil. ² School of Nursing, University of Sao Paulo, São Paulo, Brazil. ³ Nursing Postgraduate Program, University of Sergipe, Sao Cristovao, Sergipe, Brazil. ⁴ School of Medicine, University of Sao Paulo, São Paulo, Brazil. ⁵ Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Canada.

Acknowledgements

The authors thank Bernardo dos Santos for his help in statistical analysis.

Author contributions

The study concept, design, data acquisition, analysis, and interpretation were performed by RCAV, JCPS, WSP, DVO, and RMCS. Data were acquired, analyzed, and interpreted by RCAV, JCPS, WSP, DVO, CPES, ESS, and RMCS. Statistical expertise was provided by BS and CPES. The article was drafted by all authors. Supervision was provided by RCAV, CPES, ESS, and RMCS. The final manuscript was approved by all authors.

Source of support

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Brasil), Finance Code 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Conflicts of interest

The authors report no conflicts of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 14 February 2022 Accepted: 4 June 2022

Published: 9 August 2022

References

- Najem D, Rennie K, Ribecco-Lutkiewicz M, Ly D, Haukenfrers J, Liu Q, Nzau M, Fraser DD, Bani-Yaghoob M. Traumatic brain injury: classification, models, and markers. *Biochem Cell Biol*. 2018;96:391–406.
- Capizzi A, Woo J, Verduzco-Gutierrez M. Traumatic brain injury: an overview of epidemiology, pathophysiology, and medical management. *Med Clin N Am*. 2020;104:213–38.
- Sandsmark DK. Clinical outcomes after traumatic brain injury. *Curr Neurol Neurosci Rep*. 2016;16:52.
- Tohira H, Jacobs I, Mountain D, Gibson N, Yeo A. Systematic review of predictive performance of injury severity scoring tools. *Scand J Trauma Resuscit Emerg Med*. 2012;20:63.
- Champion HR, Copes WS, Sacco WJ, Lawnick MM, Keast SL, Bain JL, Frey CF. The Major Trauma Outcome Study: establishing national norms for trauma care. *J Trauma*. 1990;30:1356–65.
- Boyd C, Tolson M, Copes W. Evaluating trauma care: the TRISS method. *J Trauma*. 1987;27:370–8.
- Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the trauma score. *J Trauma*. 1989;29:623–9.
- Domingues CA, Sousa RM, Nogueira LS, Poggetti RS, Fontes B, Muñoz D. The role of the new trauma and injury severity score (NTRISS) for survival prediction. *Revista Escola de Enfermagem USP*. 2011;45:1353–8.
- Perel P, Edwards P, Wentz R, Roberts I. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak*. 2006;6:38.
- Steyerberg EW, Mushkudiani N, Perel P, Butcher I, Lu J, McHugh GS, Murray GD, Marmarou A, Roberts I, Habbema JDF, Maas AIR. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med*. 2008;5:e165.
- Dijkland SA, Foks K, Polinder S, Dippel DWJ, Maas AIR, Lingsma HF, Steyerberg EW. Prognosis in moderate and severe traumatic brain injury: a systematic review of contemporary models and validation studies. *J Neurotrauma*. 2020;37:1–13.
- Helmrich IRAR, Lingsma HF, Turgeon AF, Yamal JM, Steyerberg EW. Prognostic research in traumatic brain injury: markers, modeling, and methodological principles. *J Neurotrauma*. 2020;38:2502–13.
- Perel P, Arango M, Clayton T, Edwards P, Komolafe E, Pocock S, Roberts I, Shajur H, Steyerberg E, Yuthakasemsunt S. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ*. 2008;336:425–9.
- Dijkland SA, Helmrich IRAR, Nieboer D, van der Jagt M, Dippel DWJ, Menon DK, Stocchetti N, Maas AIR, Lingsma HF, Steyerberg EW, CENTER-TBI Participants and Investigators. Outcome prediction after moderate and severe traumatic brain injury: external validation of two established prognostic models in 1742 European patients. *J Neurotrauma*. 2021;38:1377–88.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Wilson-Mayo E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2020;372:n71.
- Debray TPA, Damen JAAG, Snell KIE, Ensor J, Hooft LH, Reitsma JB, Riley RD, Moons KGM. A guide to systematic review and meta-analysis of prediction model performance. *BMJ*. 2017;356:i6460.
- Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2014;350:g7594.
- Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, McGinn T, Guyatt G. Discrimination and calibration of clinical prediction models. *JAMA*. 2017;318:1377–84.
- Vergouwe Y, Steyerberg EW, Eijkemans MJC, Habbema JDF. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol*. 2005;58:475–83.
- Calster VB, McLernon DJ, Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17:230.
- Moons KGM, Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, Reitsma JB, Collins GS. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11:e1001744.
- Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Internal Med*. 2019;170:51–8.
- Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, Collins GS, Moons KGM. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res*. 2018;28:2768–86.
- Ruscio J. A probability-based measure of effect size: robustness to base rates and other factors. *Psychol Methods*. 2008;13:19–30.
- Cremer OL, Moons KG, Dijk GW, Balen P, Kalkman CJ. Prognosis following severe head injury: development and validation of a model for prediction of death, disability, and functional recovery. *J Trauma*. 2006;61:1484–91.
- Fischler L, Lelais F, Young J, Buchmann B, Pargger H, Kaugmann M. Assessment of three different mortality prediction models in four well-defined critical care patient groups at two points in time: a prospective cohort study. *Eur J Anaesthesiol*. 2007;24:676–83.
- Olivecrona M, Koskinen LO. The IMPACT prognosis calculator used in patients with severe traumatic brain injury treated with an ICP-targeted therapy. *Acta Neurochirurg*. 2012;154:1567–73.
- Roozenbeek B, Chiu YL, Lingsma HF, Gerber LM, Steyerberg EW, Ghajar J, Maas AIR. Predicting 14-day mortality after severe traumatic brain injury: application of the IMPACT models in the brain trauma foundation TBI-trac (R) New York State Database. *J Neurotrauma*. 2012;29:1306–12.
- Roozenbeek B, Lingsma HF, Lecky FE, Lu J, Weir J, Butcher I, McHugh GS, Murray GD, Perel P, Maas AI, Steyerberg EW. Prediction of outcome after moderate and severe traumatic brain injury: external validation of the International Mission on Prognosis and Analysis of Clinical Trials (IMPACT) and Corticoid Randomisation After Significant Head injury (CRASH) prognostic models. *Crit Care Med*. 2012;40:1609–17.
- Panczykowski DM, Puccio AM, Scruggs BJ, Bauer JS, Hricik AJ, Beers SR, Okonkwo DO. Prospective independent validation of IMPACT modeling as a prognostic tool in severe traumatic brain injury. *J Neurotrauma*. 2012;29:47–52.
- Olivecrona M, Olivecrona Z. Use of the CRASH study prognosis calculator in patients with severe traumatic brain injury treated with an intracranial pressure-targeted therapy. *J Clin Neurosci*. 2013;20:996–1101.
- Han J, King NKK, Neilson SJ, Gandhi MP, Ng I. External validation of the CRASH and IMPACT prognostic models in severe traumatic brain injury. *J Neurotrauma*. 2014;31:1146–52.
- Gómez PA, Cruz J, Lora D, Jiménez-Roldán L, Rodríguez-Boto G, Sarabia R, Sahuquillo J, Lastra R, Morera J, Lazo E, Dominguez J, Ibañez J, Brell M, Lama A, Lobato RD, Lagares A. Validation of a prognostic score for early mortality in severe head injury cases. *J Neurosurg*. 2014;121:1314–22.
- Sun H, Lingsma HF, Steyerberg EW, Maas AIR. External validation of the international mission for prognosis and analysis of clinical trials in traumatic brain injury: prognostic models for traumatic brain injury on the study of the neuroprotective activity of progesterone in severe traumatic brain injuries trial. *J Neurotrauma*. 2016;33:1535–43.
- Charry JD, Tejada JH, Pinzon MA, Tejada WA, Ochoa JD, Falla M, Tovar JH, Cuellar-Bahamón AM, Solano JP. Predicted unfavorable neurologic outcome is overestimated by the marshall computed tomography score, corticosteroid randomization after significant head injury (CRASH), and international mission for prognosis and analysis of clinical trials in

- traumatic brain injury (IMPACT) models in patients with severe traumatic. *World Neurosurg.* 2017;101:554–8.
36. Wan X, Zhao K, Wang S, Zhang H, Zeng L, Wang Y, Han L, Beejadhursing R, Shu K, Lei T. Is it reliable to predict the outcome of elderly patients with severe traumatic brain injury using the IMPACT prognostic calculator? *World Neurosurg.* 2017;103:584–90.
 37. Maeda Y, Ichikawa R, Misawa J, Shibuya A, Hishiki T, Maeda T, Yoshino A, Kondo Y. External validation of the TRISS, CRASH, and IMPACT prognostic models in severe traumatic brain injury in Japan. *PLoS ONE.* 2019;14:e0221791.
 38. Pargaonkar R, Kumar V, Menon G, Hedge A. Comparative study of computed tomographic scoring systems and predictors of early mortality in severe traumatic brain injury. *J Clin Neurosci.* 2019;66:100–6.
 39. Wongchareon K, Thompson HJ, Mitchell PH, Barber J, Temki N. IMPACT and CRASH prognostic models for traumatic brain injury: external validation in a South American cohort. *Inj Prev.* 2020;26:546–54.
 40. Baztán JJ, De la Puente M, Socorro A. Frailty, functional decline and mortality in hospitalized older adults. *Geriatr Gerontol Int.* 2017;17:664–6.
 41. Dullaert M, Oerlemans J, Hallaert G, Paepe P, Kalala JPO. Comparison of the CRASH score-predicted and real outcome of traumatic brain injury in a retrospective analysis of 417 patients. *J Clin Neurosci.* 2020;137:e159–65.
 42. Letsinger J, Rommel C, Hirschi R, Nirula R, Hawryluk GWJ. The aggressiveness of neurotrauma practitioners and the influence of the IMPACT prognostic calculator. *PLoS ONE.* 2017;12:e0183552.
 43. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res.* 2016;25:1692–706.