Prognosis in moderate and severe traumatic brain injury : a systematic review of contemporary models and validation studies

Journal of Neurotrauma

DOI: 10.1089/neu.2019.6401

Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies

Running title: Prognostic Models in Moderate and Severe TBI

Simone A Dijkland, MD[1] (PhD candidate); Kelly A Foks, MD[1,2] (PhD candidate); Suzanne Polinder, PhD[1] (associate professor); Diederik WJ Dippel, MD, PhD[2] (professor); Andrew IR Maas, MD, PhD[3] (professor); Hester F Lingsma, PhD[1] (associate professor); Ewout W Steyerberg, PhD[1,4] (professor)

[1]Department of Public Health, Center for Medical Decision Making, Erasmus MC-University Medical Center Rotterdam, the Netherlands

[2]Department of Neurology, Erasmus MC-University Medical Center Rotterdam, the Netherlands

[3]Department of Neurosurgery, Antwerp University Hospital, University of Antwerp, Edegem, Belgium

[4]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

Corresponding author:

Simone Dijkland, MD, Department of Public Health, Erasmus MC-University Medical Center, PO box 2040, 3000 CA Rotterdam, the Netherlands, telephone +31 10 703 8675, fax +31 10 703 8475, e-mail: s.dijkland@erasmusmc.nl

Contact information authors:

Kelly A. Foks: Department of Public Health, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, the Netherlands
Telephone: +31107038789, e-mail: k.foks@erasmusmc.nl, fax: n/a.

Suzanne Polinder: Department of Public Health, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands
Telephone: +31107043954, e-mail: s.polinder@erasmusmc.nl, fax: n/a.

Diederik W.J. Dippel: Department of Neurology Erasmus MC, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands
Telephone: +31107030861, e-mail: d.dippel@erasmusmc.nl, fax: n/a.

Andrew I.R. Maas: University Hospital Antwerp, Department of Neurosurgery, Wilrijkstraat 10, Edegem, BE 2650, Belgium;
Telephone: 003238214537, email Andrew.Maas@uza.be, fax: n/a.

Hester F. Lingsma: Department of Public Health, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands
Telephone: +31107044269, e-mail: h.lingsma@erasmusmc.nl, fax: n/a.

Ewout W. Steyerberg: Department of Biomedical Data Sciences, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands
Telephone: +31715264574, e-mail: e.steyerberg@erasmusmc.nl, fax: n/a.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Abstract

Outcome prognostication in traumatic brain injury (TBI) is important, but challenging due to heterogeneity of the disease. The aim of this systematic review is to present the current state-of-the-art on prognostic models for outcome after moderate and severe TBI and evidence on their validity. We searched for studies reporting on the development, validation or extension of prognostic models for functional outcome after TBI with Glasgow Coma Scale (GCS) ≤ 12 published between 2006-2018. Studies with patients aged ≥14 years and evaluating a multivariable prognostic model based on admission characteristics were included. Model discrimination was expressed with the area under the receiver operating characteristic curve (AUC), and model calibration with calibration slope and intercept. We included 58 studies describing 67 different prognostic models, comprising the development of 42 models, 149 external validations of 31 models and 12 model extensions. The most common predictors were GCS (motor) score (n=55), age (n=54) and pupillary reactivity (n=48). Model discrimination varied substantially between studies. The International Mission for Prognosis and Analysis of Clinical Trials (IMPACT) and Corticoid Randomisation After Significant Head injury (CRASH) models were developed on the largest cohorts (8,509 and 10,008 patients, respectively) and were most often externally validated (n=91), yielding AUCs ranging between 0.65-0.90 and 0.66-1.00, respectively. Model calibration was reported with a calibration intercept and slope for 7 models in 53 validations, and was highly variable. In conclusion, the discriminatory validity of the IMPACT and CRASH prognostic models is supported across a range of settings. The variation in calibration, reflecting heterogeneity in reliability of predictions, motivates continuous validation and updating if clinical implementation is pursued.

PROSPERO registry number: CRD42016052100

Keywords: traumatic brain injury; prognosis; clinical prediction rules; outcome

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Introduction

Traumatic brain injury (TBI) is a major cause of injury-related death and disability.[1] It is a disease with a considerable economic impact, often affecting the working population.[2] Patients with TBI show substantial variation in injury mechanism, pathology, clinical severity and prognosis. Due to the heterogeneity of the disease, prediction of functional outcome after TBI is challenging. Outcome prognostication is important to assist clinicians in providing reliable information to patients and relatives, to guide clinical management and trial design, and to give insight in quality of care by comparing observed and expected outcomes.[3] Many prognostic models for functional outcome after moderate and severe TBI have been developed and validated, but their methodological quality was described as poor in reviews performed in 2006 and 2008.[4, 5]

Over the past decade, new prognostic models for moderate and severe TBI have been developed and existing models have been externally validated and extended in new datasets. The question remains whether the quality of the currently available models justifies further implementation in clinical practice. For instance, when informing a relative of a patient with severe TBI in the intensive care unit on prognosis, the physician might want to use a prognostic model to communicate the chance of recovery within the next six months. But can the use of this prognostic model be recommended in this setting and for this patient? The aim of this systematic review is to present the current state-of-the-art on prognostic models for outcome after moderate and severe TBI and to review their performance at internal and external validation.

Methods

This systematic review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement.[6] The protocol of this systematic review has been registered on PROSPERO (registration number 2016: CRD42016052100) and can be accessed at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42016052100.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Literature search

We performed a literature search in Embase, Medline Ovid, Web of Science, Cochrane Central, PsychInfo Ovid and Google Scholar to identify articles published between January 1st 2006 and November 12th 2018 reporting on the development, validation or extension of models predicting outcome after moderate and severe TBI. We used search terms on the following topics: brain or head injury, prediction or prognosis, model, and mortality/survival or recovery (Table S1). Studies evaluating prognostic models in moderate and severe TBI published before 2006 were already incorporated in previous systematic reviews.[4, 5] For comparison of model performance at internal versus external validation, the development studies of models published before 2006 reporting a performance measure were retrieved manually.[7-12]

Eligibility criteria

Studies were eligible if they reported on the development, validation and/or extension of multivariable prognostic models for functional outcome in patients aged ≥ 14 years with moderate and severe TBI. We included original articles that were published in English language between 2006 and 2018. Studies that enrolled both adults and children were included when >80% of the subjects was adult or when adults and children were analyzed and reported separately. Moderate or severe TBI was defined as a Glasgow Coma Scale (GCS) score ≤ 12.[13] When a study only reported inclusion of patients with moderate or severe TBI without defining this in terms of GCS, it was assumed that moderate referred to GCS 9-12 and severe referred to GCS 3-8. In case of a population including TBIs of all severities, the study was included when the data of patients with moderate and severe TBI were incorporated in the analyses (as regards the Corticoid Randomisation After Significant Head injury (CRASH) model) or analyzed separately. Studies that evaluated model performance in specific subgroups of patients (different age groups, patients that underwent neurosurgery) were also included. The predictors used in the models had to be based on patient data obtained in the first 24 hours after injury (on hospital admission), because early outcome prediction is important to provide informed expectations to relatives and to aid early inclusion of patients in clinical trials. Moreover, we wanted to enable comparison between different prognostic models within this review as well as

between this study and previous literature.[4] No limitations existed concerning outcome measurement provided that functional outcome was measured between 14 days and 24 months after injury. We excluded reviews and qualitative studies, studies confined to the rehabilitation setting, studies that focused on patients with mild TBI (defined as GCS 13-15) and studies that focused on single predictors instead of a model containing multiple predictors.

One investigator (S.A.D.) carried out the literature search and assessed studies for eligibility on title and abstract, and subsequently on full text. In case of doubt, a second investigator (K.A.F.) was involved. Disagreements were resolved by discussion or by consultation with a third (senior) investigator (H.F.L.).

Data extraction

We used a data extraction form based on the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modeling Studies (CHARMS) checklist.[14] One investigator (S.A.D.) extracted the data from the included studies, and a random check (20%) was performed by a second investigator (K.A.F.). To ensure consistency of the data extraction, the data extraction form was tested on two studies by both investigators. The random check showed no discrepancies.

For all studies, data on study design, study population and sample size, outcome measure and scale used (e.g. functional outcome according to the Glasgow Outcome Scale (Extended), GOS(E)) and timing of outcome assessment was collected. For each prognostic model described in the included studies, we extracted data on the following topics: type of model (e.g. regression analysis, decision tree), internal or external validation and model performance. Model performance can be expressed in terms of discrimination (ability of the model to distinguish between patients with good and poor outcome) and calibration (agreement between observed and predicted probabilities). A common measure for discrimination is the area under the receiver operating characteristic curve (AUC or C-statistic). The AUC ranges from 0.5 (no discriminative ability) to 1 (perfect discrimination). Calibration is often tested with the Hosmer-Lemeshow goodness-of-fit test or assessed by a calibration slope and calibration intercept.[15] The calibration slope describes the effect of

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

the predictors in the validation sample and should be equal to 1. The intercept indicates whether predictions are systematically too high or too low, and should ideally be zero.[16]

If one study reported on multiple prognostic models or multiple stages of prognostic modeling (e.g. development and validation), data extraction was performed separately for each model or stage. We classified prognostic models as separate models when they included a different set of prognostic variables. Modifications of existing prognostic models at external validation due to missing predictor data were not defined as separate models, nor were models with identical predictors but for different outcome measures (e.g. mortality and functional outcome) or outcomes measured at different time points. However, when prognostic models consisted of identical predictors but were developed on different cohorts with re-estimation of model parameters, we did consider them as independent models rather than as validation studies.

Model performance in terms of discrimination and calibration was summarized according to AUC, calibration intercept and calibration slope weighted for the square root of study sample size. Analyses were performed with R software version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

The literature search identified 3246 unique studies, of which 3158 were excluded based on title and abstract. Of the 88 full texts screened, 58 studies met the eligibility criteria and were included in this review (Fig 1). Data of the 58 studies were collected between 1984 and 2017 (Table S2). Sample sizes ranged from 41[17] to 10,008 patients.[18] The included studies described the development, validation or extension of 67 different prognostic models (Table S2). This comprised the development of 42 models, 149 external validations of 31 models and 12 model extensions (Fig 1). Half of the studies (n=29, 50%) evaluated multiple models in one study (Table S2). The most frequently used predictors were GCS (motor) score (n=55), age (n=54) and pupillary reactivity (n=48) (Fig 2).

Model development and internal validation

Nineteen studies described the development of 42 prognostic models (1-8 models per study).[18-36] Cohorts for model development were mostly single center and prospective,

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

with a median sample size of 700 patients (Table S2 and Table 1). Moderate or severe TBI was defined according to the GCS score in all cohorts. All models had either mortality or unfavorable outcome according to the GOS(E) as outcome measure, assessed between 14 days and one year after trauma (Table S2). For the vast majority of models, unfavorable outcome was defined as GOS 1-3 or GOSE 1-4 (Table S2). Age, GCS (motor) score and pupils were the most frequently used predictors (Fig 2). Common radiological characteristics were traumatic subarachnoid hemorrhage or intraventricular hemorrhage (19 models), presence of hematoma (14 models), compression of cisterns and third ventricle (15 models) and Marshall or Rotterdam computed tomography (CT) classification (9 models). The most often used physiological predictor was hypotension (17 models). Several laboratory predictors were studied, among which glucose, hemoglobin and coagulopathy (Fig 2). Other less frequently used predictors included sex, mechanism of injury, ethnic group, and cerebral perfusion pressure (CPP, Fig 2). Biomarkers, e.g. S100 astroglial calcium-binding protein B (S100B) and glial fibrillary acidic protein (GFAP), were only included in one newly developed model (Fig 2). Most models were developed with logistic regression (n=40, 94%) and internally validated with apparent or split-sample validation (Table 1). An AUC for internal validation was reported for 32 models (76%). The AUCs for the models for mortality ranged from 0.71 to 0.94, with a mean weighted AUC of 0.84. The models for unfavorable outcome showed AUCs ranging from 0.67 to 0.98 (mean weighted AUC 0.82).

External validation

In 49 studies, 149 external validations of newly developed (n=17) or existing (n=14) prognostic models were described (1-10 models per study).[17-19, 22, 25, 29-31, 33, 34, 36-74] The external validation cohorts had a median sample size of 409 patients, and were often multicenter (n=27, 56%) and prospective (n=37, 77%, Table 1 and Table S2). The definition of moderate and severe TBI was mostly based on GCS score, but sometimes other criteria were used (e.g. loss of consciousness and Abbreviated Injury Scale ≥ 2, Table S2). Five studies only included patients with severe TBI who underwent decompressive craniectomy.[17, 40, 48, 49, 72] The time of outcome assessment according to the GOS(E) was six months in most studies (n=36, 75%), and ranged between hospital discharge and 18

months (Table S2). The models at external validation included more physiological variables due to validation of several existing Intensive Care severity scores (e.g. Acute Physiology And Chronic Health Evaluation II, Sepsis-related Organ Failure Assessment score, Fig 2). For each external validation, at least one performance measure was reported. Model calibration was most frequently expressed with a calibration plot (54%) or the Hosmer-Lemeshow goodness-of-fit test (52%, Table 1). For 25 external validations, no measure of model calibration was reported. In 95% of the external validations, model discrimination was expressed in terms of an AUC (Table 1).

The discriminative ability of the models predicting mortality or unfavorable outcome showed substantial variation (Fig 3A and B). The AUCs at external validation ranged between 0.61-0.99 (mean weighted AUC 0.80) for the models for mortality, and between 0.66-1.00 (mean weighted AUC 0.77) for the models for unfavorable outcome. We further focused on models with a reported AUC at internal validation and one or more external validations (n=20). Discriminative ability was slightly poorer at external validation compared to internal validation, with a mean AUC difference of -0.013 (p=0.086 by paired t-test) for prediction of mortality and -0.017 (p=0.031) for unfavorable outcome.

Model calibration, reported with a calibration intercept and slope, was summarized for the models that were externally validated once or more (7 models in 53 validations, Fig 4A and B). We observed substantial variation in the agreement between observed and predicted probabilities. The mean weighted calibration intercept was -0.28 (range -3.3-0.93) for the models for mortality, and -0.019 (range -5.7-2.4) for the models for unfavorable outcome. This indicates that both mortality and unfavorable outcome were generally lower than expected. The mean weighted calibration slopes were 1.1 (range 0.42-2.3) and 0.88 (range 0.57-2.5) for mortality and unfavorable outcome respectively. The values at the extremes of the ranges for calibration slope and intercept were mainly due to selection of specific populations with moderate and severe TBI, such as patients who underwent decompressive craniectomy or TBI defined according to the Abbreviated Injury Scale.[48, 67]

The International Mission for Prognosis and Analysis of Clinical Trials (IMPACT) and CRASH models were most frequently externally validated (n=91). The mean weighted AUCs were 0.79 (mortality) and 0.77 (unfavorable outcome) for the IMPACT models (range 0.65-

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

0.90), and 0.82 (mortality) and 0.78 (unfavorable outcome) for the CRASH models (range 0.66-1.00, Fig 3A and B and Table 2). In total 51 external validations reported calibration with an intercept and slope. These 51 validations showed overestimated risks by the IMPACT and CRASH models for mortality and underestimated risks for unfavorable outcome (Fig 4A and B and Table 2). The more complex IMPACT and CRASH models, for example including CT characteristics, showed only modest improvement in discriminative ability (Fig S1 A and B), and calibration remained highly variable (Fig 4A and B). Comparison of the performance of the IMPACT and CRASH models with other models, such as Hukkelhoven and Nijmegen, was not feasible given the limited number of validations of these other models (Table 2).

Model extensions

In five studies, 12 extensions of the IMPACT and CRASH prognostic models were assessed.[41, 44, 51, 57, 58] The median sample size of the extension cohorts was 342 patients (Table 1). Moderate and severe TBI patients were selected based on GCS, except for one cohort consisting of consecutive TBI patients requiring intracranial pressure monitoring.[44] Outcomes were assessed between one week and six months (Table S2). Most studies reported model discrimination with an AUC (n=11, 92%) and calibration with the Hosmer-Lemeshow goodness-of-fit test (n=10, 83%; Table 1). The extensions included several serum and cerebrospinal fluid biomarkers, extracranial injury, coagulation parameters or dynamic predictors containing information on the first 24 hours of the clinical course (Acute Physiology And Chronic Health Evaluation II score, intracranial pressure and mean arterial pressure) (Fig 2). Performance of the extended models in terms of both discrimination and calibration improved somewhat compared to the original versions of the models. The mean AUC increase at model extension was 0.013 (p=0.18 by paired t-test) for models for mortality and 0.10 (p=0.026) for models for unfavorable outcome. Calibration was not evaluated or showed no improvement.[41, 44, 51, 57, 58] None of the extended models was externally validated.

Discussion

We systematically reviewed 58 papers describing the development, validation or extension of 67 different multivariable prognostic models for functional outcome in moderate and severe TBI. We identified 149 external validations of prognostic models. The IMPACT and CRASH models currently dominate the field of prognostic modeling in moderate and severe TBI. External validations of these models showed substantial variation in performance: overall moderate to good discrimination, but highly variable calibration.

Strengths and limitations

This systematic review is based on a comprehensive literature search resulting in a large number of prognostic models and validation studies in the field of moderate and severe TBI. A novel feature compared to previous systematic reviews on this topic is that improvements in prognostic research in TBI now permit inclusion of a substantial number of external validation studies. However, some limitations should be considered. We did not consider models for which the outcomes (mortality or unfavorable outcome) were measured at different time points as separate models. Similarly, models with identical predictors but for different outcome measures were not defined as separate models. This may have caused an underestimation of the number of prognostic models for moderate and severe TBI. Another factor that might have unjustly reduced the number of models is the exclusion of studies that were not published in English language. Additionally, most studies in this systematic review were conducted in middle and high income countries. Therefore, our results might not be generalizable to low income countries. Finally, comparing model calibration between different models and settings was difficult due to variation in, or even absence of, calibration measures. Model calibration was reported in terms of an intercept and slope for only seven models. Our summary of model calibration might therefore not reflect the overall ability of the currently available models to provide predictions in individual patients.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Comparison with previous literature

Previous systematic reviews on prognostic models in moderate and severe TBI mainly focused on their methodological quality. Several recommendations were proposed to improve methodology and reporting of prognostic models.[4, 5] The prognostic models evaluated in the current systematic review showed advancements in reporting and statistical approaches, especially regarding external validation. Models were externally validated in independent cohorts and most validation studies reported appropriate model performance measures in terms of discrimination and calibration.[15] However, measures for discrimination are still more frequently reported than calibration measures. Moreover, although the Hosmer-Lemeshow goodness-of-fit test for model calibration is no longer recommended due to lack of power and interpretability, this was still used in more than half of the validations. The lack of adequate calibration measures is remarkable, since poor calibration implies that the predictions will be misleading when used in clinical practice. This may lead to harmful decision making.[75]

Model development and predictors

After publication of the previous systematic reviews, several new prognostic models for outcome prediction after moderate and severe TBI have been developed. Especially the introductions of the IMPACT and CRASH models have been important to confirm the core predictors for unfavorable outcome after moderate and severe TBI obtained at admission: older age, less responsive pupils and lower GCS (motor) score.[18, 31] Although these baseline predictors included in the IMPACT and CRASH models only explain around 35% of the variance in outcome, more complex models with additional predictors collected within 24 hours may not lead to substantial improvements in model performance.[3] This is supported by our observation that performance of the IMPACT and CRASH models showed only modest improvement in discriminative ability by adding CT characteristics, physiological and laboratory variables obtained within the first 24 hours, both at internal and external validation (Fig 3A and B, Fig S1 A and B). However, prognostic estimates will be refined during the course of the disease, as may be considered in dynamic prediction models.[76] Any prognostic model should only be considered an addition to clinical experience.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

In line with previous recommendations, other recently developed models introduced several new predictors (e.g. CPP, ethnic group, mechanism of injury, biomarkers).[3] However, many of these predictors were only included in a few models and not yet externally validated (Fig 2). Therefore, it remains difficult to assess the added value of these models and predictors. Further research is essential, especially external validation.

External validation

We found a large number of external validations in contemporary series. The IMPACT and CRASH models were externally validated most extensively. Model performance at external validation was on average close to performance at internal validation. Performance at external validation may best reflect the models' discriminative ability when applied in clinical practice.[77] The discriminative ability at external validation was mostly around 0.8, with one very small study even reporting an implausible AUC value of 1 for the CRASH CT model for unfavorable outcome.[74] Calibration varied highly among different models and studies. The variability in discriminative performance and calibration slopes is most likely attributable to differences in measurement of predictors or selection of the validation population.[78] For instance, a few studies investigated model performance in more homogeneous subgroups such as patients with decompressive craniectomy.[17, 40, 48, 49, 72] We also observed a substantial number of variations (i.e. differences in included predictors) on IMPACT and CRASH at external validation (Table S3), mostly due to discrepancies in predictor definitions or unavailability of predictor data.[18, 31, 38, 61, 62, 64, 68] Further, timing of outcome measurement varied substantially across different studies. Although most models were designed for outcome prediction at six months after injury, model performance was assessed in cohorts with outcome data available up to 18 months after injury.[48] Heterogeneity in baseline risk was noted according to calibration-in-the-large (intercept differences). This variability might be attributed to differences in distributions and effects of unmeasured covariates and is therefore often difficult to explain. The substantial heterogeneity in model performance across different settings indicates that models need to be recalibrated for each new setting before implementation in clinical practice is warranted.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Model extension

Highly variable model performance may be problematic when introducing the models to a specific clinical setting. Several stages have been identified in updating prognostic models, ranging from updating the intercept to addition of predictors.[16] There has been extensive research into the additional prognostic value of baseline biomarkers for TBI.[79] However, extending the IMPACT and CRASH models with markers of coagulation or serum and cerebrospinal fluid biomarkers (S100B and GFAP) barely improved model performance in the few studies that have been performed.[41, 58] Because TBI is a heterogeneous disease with a highly variable clinical course, adding new information as it becomes available over time or including factors that predict treatment response may be more promising to improve outcome prediction.[3] Extending the currently available models with such dynamic predictors has been uncommon so far, and yielded variable improvement in model performance.[44, 57] External validation of these extended models is lacking. Possibilities for updating the IMPACT and CRASH models are currently being evaluated in various studies, including the Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI) study, Transforming Research and Clinical Knowledge in TBI (TRACK-TBI) dataset and Collaborative REsearch on ACute Traumatic Brain Injury in intensiVe Care Medicine in Europe (CREACTIVE).[80-82] Given the highly variable calibration, updating of the baseline risk estimate (the intercept in the regression model) should be considered. Also, machine learning techniques are currently gaining interest and might be helpful for dynamic prediction.

Implementation in clinical practice

The availability of a large number of prognostic models for functional outcome after moderate and severe TBI suggests that outcome prediction is considered relevant for clinical practice. However, despite previous recommendations, none of the available models have been implemented in TBI guidelines. Their use in clinical practice is limited.[3] This might partly be explained by the lack of evidence-based treatment options in TBI,[83] limiting the use of prognostic models to select patients for individualized management. Previous studies evaluating the perceptions of physicians on utilization of the IMPACT calculator in clinical practice showed that approximately half of the clinicians involved in

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

TBI care was aware of its existence. Of those, only 50% occasionally used the model in clinical practice.[84, 85] Factors limiting clinical use of the IMPACT calculator comprised mistrust in the IMPACT development data, utilization for research purposes only, time needed to gather the data required to complete the online tool, and concern about misinterpretation of prognostic estimates by patients and their families.[84, 85] However, the IMPACT calculator was reported to be useful for reducing variability between physicians with different levels of clinical experience.[85]

Model discrimination, although variable, was adequate in most studies. The lack of implementation can therefore not be explained by poor discriminative ability. Moreover, models do not necessarily need high discriminative performance to be accepted in clinical practice. Examples are the Hypertension, Abnormal Renal/Liver Function, Stroke, Bleeding History or Predisposition, Labile INR, Elderly, Drugs/Alcohol Concomitantly (HAS-BLED, AUC 0.65) and Congestive heart failure, Hypertension, Age, Diabetes mellitus, Prior Stroke or TIA or Thromboembolism (CHA2DS2-VASc, AUC 0.61) models that are commonly applied in neurovascular practice, and the extensively used Gail breast cancer models (pooled AUCs between 0.55-0.75).[86-88] Compared to these widely implemented tools, the models for outcome after moderate and severe TBI perform very well (weighted mean AUCs of 0.80 and 0.76 for mortality and unfavorable outcome, respectively).

Model calibration, on the other hand, showed substantial heterogeneity between different settings. The adequate discriminative ability and highly variable calibration may indicate that the models perform well at group level, but caution is required when using them to provide predictions for individual patients in a specific clinical setting.

Based on the main findings of this systematic review, we provided a set of recommendations regarding statistical evaluation and implementation of prognostic models in moderate and severe TBI (Table 3).

Conclusion

The IMPACT and CRASH prognostic models have been developed on the largest datasets and have adequate discriminative ability across a range of settings. The reliability of predictions is highly variable. We recommend implementation of these models in clinical practice, provided that they have been validated or updated for the specific clinical setting.

Acknowledgments

None.

Author Disclosure Statement

No competing financial interests exist.

Financial Disclosure Statement

The research leading to these results was supported by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 602150 (CENTER-TBI). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

References

1. Majdan, M., Plancikova, D., Brazinova, A., Rusnak, M., Nieboer, D., Feigin, V. and Maas, A. (2016). Epidemiology of traumatic brain injuries in Europe: a cross-sectional analysis. Lancet Public Health 1, e76-e83.

2. Maas, A.I., Stocchetti, N. and Bullock, R. (2008). Moderate and severe traumatic brain injury in adults. Lancet Neurol 7, 728-741.

3. Lingsma, H.F., Roozenbeek, B., Steyerberg, E.W., Murray, G.D. and Maas, A.I. (2010). Early prognosis in traumatic brain injury: from prophecies to predictions. Lancet Neurol 9, 543-554.

4. Mushkudiani, N.A., Hukkelhoven, C.W., Hernandez, A.V., Murray, G.D., Choi, S.C., Maas, A.I. and Steyerberg, E.W. (2008). A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. J Clin Epidemiol 61, 331-343.

5. Perel, P., Edwards, P., Wentz, R. and Roberts, I. (2006). Systematic review of prognostic models in traumatic brain injury. BMC Med Inform Decis Mak 6, 38.

6. Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. and Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 62, 1006-1012.

7. Knaus, W.A., Draper, E.A., Wagner, D.P. and Zimmerman, J.E. (1985). APACHE II: a severity of disease classification system. Crit Care Med 13, 818-829.

8. Le Gall, J.R., Lemeshow, S. and Saulnier, F. (1993). A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. Jama 270, 2957-2963.

9. Lemeshow, S., Teres, D., Klar, J., Avrunin, J.S., Gehlbach, S.H. and Rapoport, J. (1993). Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. Jama 270, 2478-2486.

10. Signorini, D.F., Andrews, P.J., Jones, P.A., Wardlaw, J.M. and Miller, J.D. (1999). Predicting survival using simple clinical variables: a case study in traumatic brain injury. J Neurol Neurosurg Psychiatry 66, 20-25.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

11. Hukkelhoven, C.W., Steyerberg, E.W., Habbema, J.D., Farace, E., Marmarou, A., Murray, G.D., Marshall, L.F. and Maas, A.I. (2005). Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. J Neurotrauma 22, 1025-1039.

12. Maas, A.I., Hukkelhoven, C.W., Marshall, L.F. and Steyerberg, E.W. (2005). Prediction of outcome in traumatic brain injury with computed tomographic characteristics: a comparison between the computed tomographic classification and combinations of computed tomographic predictors. Neurosurgery 57, 1173-1182; discussion 1173-1182.

13. Teasdale, G. and Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. Lancet 2, 81-84.

14. Moons, K.G., de Groot, J.A., Bouwmeester, W., Vergouwe, Y., Mallett, S., Altman, D.G., Reitsma, J.B. and Collins, G.S. (2014). Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med 11, e1001744.

15. Steyerberg, E.W., Vickers, A.J., Cook, N.R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M.J. and Kattan, M.W. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21, 128-138.

16. Steyerberg, E.W. In: *Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. New York: Springer, 2009.*

17. Honeybul, S., Ho, K.M., Lind, C.R.P., Corcoran, T. and Gillett, G.R. (2009). The retrospective application of a prediction model to patients who have had a decompressive craniectomy for trauma. Journal of Neurotrauma 26, 2179-2183.

18. Perel, P.A., Olldashi, F., Muzha, I., Filipi, N., Lede, R., Copertari, P., Traverso, C., Copertari, A., Vergara, E.A., Montenegro, C., De Huidobro, R.R., Saladino, P., Surt, K., Cialzeta, J., Lazzeri, S., Piñero, G., Ciccioli, F., Videtta, W., Barboza, M.F., Svampa, S., Sciuto, V., Domeniconi, G., Bustamante, M., Waschbusch, M., Gullo, M.P., Drago, D.A., Linares, J.C.A., Camputaro, L., Tróccoli, G., Galimberti, H., Tallott, M., Eybner, C., Buchinger, W., Fitzal, S., Mazairac, G., Oleffe, V., Grollinger, T., Delvaux, P., Carlier, L., Braet, V., Jacques, J.M., De Knoop, D., Nasi, L., Choi, H.K., Schmitt, M., Gentil, A., Nacul, F., Barrios, P.B., Xinkang, C., Hua, L.S., Tian, H.H., Xiaodong, C., Gualteros, W., Otero, A.A., Arango, M., Ciro, J., Jaramillo, H., Gonzalez, I., Gomez, C., Arias, A., Fonseca, M., Mora, C., Cabrera, E.G.L., Betancurth, J.L., Muñoz, P., Quiñónez, J.A., Castillo, M.E.G., Lopez, O., Yepes, R.P., Cuellar,

D.L., Paez, G., Chaves, H.D., Ordoñez, P.E., Plata, R., Pineda, M., Pulido, L.E., Jaramillo, J.S.V., Rebolledo, C., Palma, O., Soler, C., Pastrana, I., Falero, R., Perera, M.D., García, A.A., Oliva, R., Delgado, H.L., Carnero, A.M., López, B.L., Gallardo, A.L., Morales, A.O., Lezcano, H., Ferrer, M.I., Bess, I.Z., Canino, G.R., Ruiz, E.M.P., Cruz, O.G., Svoboda, P., Kantorová, I., Ochmann, J., Scheer, P., Kozumplík, L., Maršová, J., Edelmann, K., Chytra, I., Bosman, R., Andrejsová, H., Pachl, J., Bürger, J., Kramar, F., Ulloa, M.I., Gonzalez, L., Daccach, A., Ortega, A., Cevallos, S., Cueva, B.Z., Ochoa, M., Tapia, J.V., Hurtado, J., Wong, M.C.S., Santos, R., Khamis, H., Abaza, A.H., Fekry, A., El Kordy, S., Shawky, T., El-Sayed, H., Khalil, N., Negm, N., Fisal, S., Alamin, M., Shokry, H., Elhusseny, A.Y., Radwan, A., Rashid, M., Gogichaisvili, T., Ingorokva, G., Gongadze, N., Otarashvili, A., Kleist, W., Kalkum, M., Ulrich, P., Andrews, N., Nakos, G., Karavelis, A., Archontakis, G., Myrianthefs, P., Yadav, Y., Yadav, S., Khatri, R., Baghel, A., Husain, M., Jha, D., Chhang, W.H., Dhandhania, M., Fonning, C., Iyengar, S.N., Gupta, S., Ravi, R.R., Bopiah, K.S., Herur, A., Venkataramana, N.K., Satish, A., Bhavadasan, K., Morris, R., Ramesh, S., Satish, A., Dewan, Y., Singh, Y., Bhagchandani, R., Bhagchandani, S., Sethurayar, V.U., Ipe, S., Sreekumar, G., Panigrahi, M., Reddy, A., Khosla, V., Gupta, S., Pillay, H., Thomas, N., Sridhar, K., Jose, B., Kurian, N., Praharaj, S., Pillai, S., Ramana, Gupta, S., Kiyawat, D., Maheshwari, K., Panikar, D., Chawla, J., Shenoy, S., Raja, A., Rupayana, Y., Reddy, S., Mohan, N., Kelkar, S., Johri, M., Golden, N., Maliawan, S., Fauzi, A., Farouk, U., Fakharian, E., Aramesh, A., Eghtedari, M., Ahmadzadeh, F., Gholami, A., Plunkett, P., Redican, C., McMahon, G., Annetta, M.G., Mouchaty, H., Bruzzone, E., Harding, B., Qureshi, M., Idris, Z., Jafri Abdullah, N.C., Ghazali, G., Ghani, A.R.I., Cheah, F., Cabrera, A., González, J.L.M., Loría-Castellanos, J., Jackson, S., Hutchinson, R., Komolafe, E., Adeolu, A., Komolafe, M., Adeyemi-Doro, O., Bankole, F., Shehu, B., Danlami, V., Odebode, O., Oluwadiya, K., Sanni, A., Giebel, H., Kumar, S., Jooma, R., Mezquita, J.E., Ovelar, C.O., Portillo, M.G., Rodriguez, D., Balica, L., Oprita, B., Sklerniacof, M., Steflea, L., Bandut, L., Danil, A., Iliescu, R., Ciurea, J., El-Dawlatly, A., Alwatidy, S., Al-Yafi, W., El-Dawlatly, M., Krunic-Protic, R., Janosevic, V., Tan, J., Seah, C., Trenkler, S., Humenansky, M., Stajančová, T., Schwendt, I., Laincz, A., Julius, Z., Maros, S., Firment, J., Cifraničova, M., Sániová, B., Kalig, K., Medekova, S., Wiszt, R., Mačuga, I., Hartzenberg, B., Du Plessis, G., Houlie, Z., Nathoo, N., Khumalo, S., Tracey, R., Muñoz-Sánchez, A., Francisco Murillo-Cabezas, N.C., Flores-Cordero, J., Rincón-Ferrari, D., Rubi, M., Caler, L., Del Campo, M.M., Laguna, L.B., Nava, J.M., Minguillón, M.A.,

Lopez, A.M., Ramos-Gómez, L., De La Torre-Prados, V., Pellejero, R., Laloë, V., Mandrella, B., Suganthan, Perera, S., Mahendran, K., Stocker, R., Ludwig, S., Zimmermann, H., Denzler, U., Yutthakasemsunt, S., Kittiwattanagul, W., Piyavechvirat, P., Tapsai, P., Namuang-Jan, A., Chantapimpa, U., Watanachai, C., Subsompon, P., Pussanakawatin, W., Khunjan, P., Tangchitvittaya, S., Nilapong, S., Klangsang, T., Taechakosol, W., Srinat, A., Jerbi, Z., Borsali-Falfoul, N., Rezgui, M., Cakar, N., Ssenyonjo, H., Kobusingye, O., Lomas, G., Yates, D., Lecky, F., Bleetman, A., Baldwin, A., Jenkinson, E., Pantrini, S., Stewart, J., Contractor, N., Roberts, T., Butler, J., Pinto, A., Lee, D., Brayley, N., Robbshaw, K., Dix, C., Graham, S., Pye, S., Green, M., Kellins, A., Moulton, C., Fogg, B., Cottingham, R., Funnell, S., Shanker, U., Summers, C., Malek, L., Ashcroft, C., Powell, J., Moore, S., Buckley, S., Grocutt, M., Chambers, S., Morrice, A., Marshall, H., Harris, J., Matthews, W., Tippet, J., Mardell, S., MacMillan, F., Shaw, A., Luthra, P., Dixon, G., Ahmed, M., Butler, J., Young, M., Mason, S., Loveday, I., Clark, C., Taylor, S., Wilson, P., Ali, K., Greenwood, S., White, M., Perez, R., Eljamel, S., Wasserberg, J., Shale, H., Read, C., McCarron, J., Pennell, A., Ray, G., Thurston, J., Brown, E., Jaffey, L., Graves, M., Bailey, R., Loveridge, N., Evans, G., Hughes, S., Ahmed, M.K., Richardson, J., Gallagher, C., Odedun, T., Lees, K., Foley, D., Payne, N., Pennycook, A., Griffiths, C., Moore, D., Byrne, D., Dasan, S., Banerjee, A., McGuinness, S., Chikhani, C., Zoltie, N., Barlow, I., Stell, I., Hulse, W., Crossley, J., Watkins, L., Dorani, B. and Van Viet, T. (2008). Predicting outcome after traumatic brain injury: Practical prognostic models based on large cohort of international patients. BMJ 336, 425-429.

19. Cremer, O.L., Moons, K.G.M., Van Dijk, G.W., Van Balen, P. and Kalkman, C.J. (2006). Prognosis following severe head injury: Development and validation of a model for prediction of death, disability, and functional recovery. Journal of Trauma - Injury, Infection and Critical Care 61, 1484-1491.

20. Demetriades, D., Kuncir, E., Brown, C.V.R., Martin, M., Salim, A., Rhee, P. and Chan, L.S. (2006). Early prediction of mortality in isolated head injury patients: A new predictive model. Journal of Trauma - Injury, Infection and Critical Care 61, 868-872.

21. Fabbri, A., Servadei, F., Marchesini, G., Stein, S.C. and Vandelli, A. (2008). Early predictors of unfavourable outcome in subjects with moderate head injury in the emergency department. Journal of Neurology, Neurosurgery and Psychiatry 79, 567-573.

22. Gradisek, P., Osredkar, J., Korsic, M. and Kremzar, B. (2012). Multiple indicators model of long-term mortality in traumatic brain injury. Brain Injury 26, 1472-1481.

23. Han, J.X., See, A.A.Q., Gandhi, M. and King, N.K.K. (2017). Models of Mortality and Morbidity in Severe Traumatic Brain Injury: An Analysis of a Singapore Neurotrauma Database. World Neurosurgery 108, 885-893.e881.

24. Iba, J., Tasaki, O., Hirao, T., Mohri, T., Yoshiya, K., Hayakawa, K., Shiozaki, T., Hamasaki, T., Nakamori, Y., Fujimi, S., Ogura, H., Kuwagata, Y. and Shimazu, T. (2014). Outcome prediction model for severe traumatic brain injury. Acute Medicine and Surgery 1, 31-36.

25. Jacobs, B., Beems, T., Van Der Vliet, T.M., Van Vugt, A.B., Hoedemaekers, C., Horn, J., Franschman, G., Haitsma, I., Van Der Naalt, J., Andriessen, T.M.J.C., Borm, G.F. and Vos, P.E. (2013). Outcome prediction in moderate and severe traumatic brain injury: A focus on computed tomography variables. Neurocritical Care 19, 79-89.

26. Kamal, V.K., Agrawal, D. and Pandey, R.M. (2016). Prognostic models for prediction of outcomes after traumatic brain injury based on patients admission characteristics. Brain Injury 30, 393-406.

27. Pang, B.C., Kuralmani, V., Joshi, R., Hongli, Y., Kah, K.L., Beng, T.A., Li, J., Tze, Y.L. and Ng, I. (2007). Hybrid outcome prediction model for severe traumatic brain injury. Journal of Neurotrauma 24, 136-146.

28. Petroni, G., Quaglino, M., Lujan, S., Kovalevski, L., Rondina, C., Videtta, W., Carney, N., Temkin, N. and Chesnut, R. (2010). Early Prognosis of severe traumatic brain injury in an urban argentinian trauma center. Journal of Trauma - Injury, Infection and Critical Care 68, 564-570.

29. Raj, R., Skrifvars, M.B., Bendel, S., Selander, T., Kivisaari, R., Siironen, J. and Reinikainen, M. (2014). Predicting six-month mortality of patients with traumatic brain injury: Usefulness of common intensive care severity scores. Critical Care 18, R60.

30. Rizoli, S., Petersen, A., Bulger, E., Coimbra, R., Kerby, J.D., Minei, J., Morrison, L., Nathens, A., Schreiber, M. and de Oliveira Manoel, A.L. (2016). Early prediction of outcome after severe traumatic brain injury: A simple and practical model. BMC Emergency Medicine 16, 32.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

31. Steyerberg, E.W., Mushkudiani, N., Perel, P., Butcher, I., Lu, J., McHugh, G.S., Murray, G.D., Marmarou, A., Roberts, I., Habbema, J.D.F. and Maas, A.I.R. (2008). Predicting outcome after traumatic brain injury: Development and international validation of prognostic scores based on admission characteristics. PLoS Medicine 5, 1251-1261.

32. Tasaki, O., Shiozaki, T., Hamasaki, T., Kajino, K., Nakae, H., Tanaka, H., Shimazu, T. and Sugimoto, H. (2009). Prognostic indicators and outcome prediction model for severe traumatic brain injury. Journal of Trauma - Injury, Infection and Critical Care 66, 304-308.

33. Yuan, F., Ding, J., Chen, H., Guo, Y., Wang, G., Gao, W.W., Chen, S.W. and Tian, H.L. (2012). Predicting outcomes after traumatic brain injury: The development and validation of prognostic models based on admission characteristics. Journal of Trauma and Acute Care Surgery 73, 137-145.

34. Mahadewa, T.G.B., Golden, N., Saputra, A. and Ryalino, C. (2018). Modified Revised Trauma-Marshall score as a proposed tool in predicting the outcome of moderate and severe traumatic brain injury. Open access emerg. med. 10, 135-139.

35. Pannatier, M., Delhumeau, C. and Walder, B. (2018). Comparison of two prehospital predictive models for mortality and impaired consciousness after severe traumatic brain injury. Acta Anaesthesiologica Scandinavica 63, 74-85.

36. Rached, M.A.K.B., Gaudet, J.G., Delhumeau, C. and Walder, B. (2018). Comparison of two simple models for prediction of short term mortality in patients after severe traumatic brain injury. Injury 50, 65-72.

37. Bonds, B.W., Dhanda, A., Wade, C., Massetti, J., Diaz, C. and Stein, D.M. (2015). Prognostication of Mortality and Long term Functional Outcomes Following Traumatic Brain Injury: Can We Do Better? Journal of Neurotrauma (Epub ahead of print).

38. Castaño-Leon, A.M., Lora, D., Munarriz, P.M., Cepeda, S., Paredes, I., De La Cruz, J., Gómez Lopez, P.A. and Lagares, A. (2016). Predicting Outcomes after Severe and Moderate Traumatic Brain Injury: An External Validation of Impact and Crash Prognostic Models in a Large Spanish Cohort. Journal of Neurotrauma 33, 1598-1606.

39. Charry, J., Falla, J., Ochoa, J., Pinzón, M., Tejada, J., Henriquez, M., Solano, J. and Calvache, C. (2017). External validation of the rotterdam computed tomography score in the prediction of mortality in severe traumatic brain injury. Journal of Neurosciences in Rural Practice 8, 23-26.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

40. Charry, J.D., Tejada, J.H., Pinzon, M.A., Tejada, W.A., Ochoa, J.D., Falla, M., Tovar, J.H., Cuellar-Bahamón, A.M. and Solano, J.P. (2017). 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 Brain Injury Managed with Early Decompressive Craniectomy. World Neurosurgery 101, 554-558.

41. Czeiter, E., Mondello, S., Kovacs, N., Sandor, J., Gabrielli, A., Schmid, K., Tortella, F., Wang, K.K.W., Hayes, R.L., Barzo, P., Ezer, E., Doczi, T. and Buki, A. (2012). Brain injury biomarkers may improve the predictive power of the IMPACT outcome calculator. Journal of Neurotrauma 29, 1770-1778.

42. Egea-Guerrero, J.J., Rodríguez-Rodríguez, A., Gordillo-Escobar, E., Fernández-Delgado, E., Martínez-Roldán, Á., Roldán-Reina, Á., Durán-Martínez, P., de Vega-Ríos, E., Freire-Aragón, M.D., Vilches-Arenas, Á., Murillo-Cabezas, F. and Quintana-Díaz, M. (2017). IMPACT Score for Traumatic Brain Injury: Validation of the Prognostic Tool in a Spanish Cohort. Journal of Head Trauma Rehabilitation 33, E46-E52.

43. Fischler, L., Lelais, F., Young, J., Buchmann, B., Pargger, H. and Kaufmann, M. (2007). Assessment of three different mortality prediction models in four well-defined critical care patient groups at two points in time: A prospective cohort study. European Journal of Anaesthesiology 24, 676-683.

44. Güiza, F., Depreitere, B., Piper, I., Van Den Berghe, G. and Meyfroidt, G. (2013). Novel methods to predict increased intracranial pressure during intensive care and long-term neurologic outcome after traumatic brain injury: Development and validation in a multicenter dataset. Critical Care Medicine 41, 554-564.

45. Han, J., King, N.K.K., Neilson, S.J., Gandhi, M.P. and Ng, I. (2014). External validation of the CRASH and IMPACT prognostic models in severe traumatic brain injury. Journal of Neurotrauma 31, 1146-1152.

46. Harrison, D.A., Griggs, K.A., Prabhu, G., Gomes, M., Lecky, F.E., Hutchinson, P.J.A., Menon, D.K. and Rowan, K.M. (2015). External validation and recalibration of risk prediction models for acute traumatic brain injury among critically ill adult patients in the United Kingdom. Journal of Neurotrauma 32, 1522-1537.

47. Hashemi, B., Amanat, M., Baratloo, A., Forouzanfar, M.M., Rahmati, F., Motamedi, M. and Safari, S. (2016). Validation of CRASH model in prediction of 14-day mortality and 6-month unfavorable outcome of head trauma patients. Emergency 4, 196-201.

48. Honeybul, S. and Ho, K.M. (2016). Predicting long-term neurological outcomes after severe traumatic brain injury requiring decompressive craniectomy: A comparison of the CRASH and IMPACT prognostic models. Injury 47, 1886-1892.

49. Honeybul, S., Ho, K.M., Lind, C.R.P. and Gillett, G.R. (2014). Validation of the CRASH model in the prediction of 18-month mortality and unfavorable outcome in severe traumatic brain injury requiring decompressive craniectomy. Journal of Neurosurgery 120, 1131-1137.

50. Hukkelhoven, C.W.P.M., Rampen, A.J.J., Maas, A.I.R., Farace, E., Habbema, J.D.F., Marmarou, A., Marshall, L.F., Murray, G.D. and Steyerberg, E.W. (2006). Some prognostic models for traumatic brain injury were not valid. Journal of Clinical Epidemiology 59, 132-143.

51. Lingsma, H., Andriessen, T.M.J.C., Haitsema, I., Horn, J., Van Der Naalt, J., Franschman, G., Maas, A.I.R., Vos, P.E. and Steyerberg, E.W. (2013). Prognosis in moderate and severe traumatic brain injury: External validation of the IMPACT models and the role of extracranial injuries. Journal of Trauma and Acute Care Surgery 74, 639-646.

52. Majdan, M., Brazinova, A., Rusnak, M. and Leitgeb, J. (2017). Outcome prediction after traumatic brain injury: Comparison of the performance of routinely used severity scores and multivariable prognostic models. Journal of Neurosciences in Rural Practice 8, 20-29.

53. Majdan, M., Lingsma, H.F., Nieboer, D., Mauritz, W., Rusnak, M. and Steyerberg, E.W. (2014). Performance of IMPACT, CRASH and Nijmegen models in predicting six month outcome of patients with severe or moderate TBI: an external validation study. Scandinavian journal of trauma, resuscitation and emergency medicine 22, 68.

54. Olivecrona, M. and Koskinen, L.O.D. (2012). The IMPACT prognosis calculator used in patients with severe traumatic brain injury treated with an ICP-targeted therapy. Acta Neurochirurgica 154, 1567-1573.

55. Olivecrona, M. and Olivecrona, Z. (2013). Use of the CRASH study prognosis calculator in patients with severe traumatic brain injury treated with an intracranial pressure-targeted therapy. Journal of Clinical Neuroscience 20, 996-1001.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

56. Panczykowski, D.M., Puccio, A.M., Scruggs, B.J., Bauer, J.S., Hricik, A.J., Beers, S.R. and Okonkwo, D.O. (2012). Prospective independent validation of IMPACT modeling as a prognostic tool in severe traumatic brain injury. Journal of Neurotrauma 29, 47-52.

57. Raj, R., Siironen, J., Kivisaari, R., Hernesniemi, J. and Skrifvars, M.B. (2014). Predicting outcome after traumatic brain injury: Development of prognostic scores based on the IMPACT and the APACHE II. Journal of Neurotrauma 31, 1721-1732.

58. Raj, R., Siironen, J., Kivisaari, R., Hernesniemi, J., Tanskanen, P., Handolin, L. and Skrifvars, M.B. (2013). External validation of the international mission for prognosis and analysis of clinical trials model and the role of markers of coagulation. Neurosurgery 73, 305-311.

59. Røe, C., Skandsen, T., Manskow, U., Ader, T. and Anke, A. (2015). Mortality and One-Year Functional Outcome in Elderly and Very Old Patients with Severe Traumatic Brain Injuries: Observed and Predicted. Behavioural Neurology 2015.

60. Rønning, P.A., Pedersen, T., Skaga, N.O., Helseth, E., Langmoen, I.A. and Stavem, K. (2011). External validation of a prognostic model for early mortality after traumatic brain injury. Journal of Trauma - Injury, Infection and Critical Care 70, E56-E61.

61. Roozenbeek, B., Chiu, Y.L., Lingsma, H.F., Gerber, L.M., Steyerberg, E.W., Ghajar, J. and Maas, A.I.R. (2012). Predicting 14-day mortality after severe traumatic brain injury: Application of the IMPACT models in the brain trauma foundation TBI-trac® New York state database. Journal of Neurotrauma 29, 1306-1312.

62. Roozenbeek, B., Lingsma, H.F., Lecky, F.E., Lu, J., Weir, J., Butcher, I., McHugh, G.S., Murray, G.D., Perel, P., Maas, A.I. and Steyerberg, E.W. (2012). 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. Critical Care Medicine 40, 1609-1617.

63. Staples, J.A., Wang, J., Mills, B., Temkin, N., Zaros, M.C., Jurkovich, G.J. and Rivara, F.P. (2016). The application of the CRASH-CT prognostic model for older adults with traumatic brain injury: A population-based observational cohort study. Journal of Head Trauma Rehabilitation 31, E8-E14.

64. Staples, J.A., Wang, J., Zaros, M.C., Jurkovich, G.J. and Rivara, F.P. (2016). The application of IMPACT prognostic models to elderly adults with traumatic brain injury: A population-based observational cohort study. Brain Injury 30, 899-907.

65. Sun, H., Lingsma, H.F., Steyerberg, E.W. and Maas, A.I.R. (2016). 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. Journal of Neurotrauma 33, 1535-1543.

66. Wan, X., Zhao, K., Wang, S., Zhang, H., Zeng, L., Wang, Y., Han, L., Beejadhursing, R., Shu, K. and Lei, T. (2017). Is It Reliable to Predict the Outcome of Elderly Patients with Severe Traumatic Brain Injury Using the IMPACT Prognostic Calculator? World Neurosurgery 103, 584-590.

67. Wong, G.K.C., Teoh, J., Yeung, J., Chan, E., Siu, E., Woo, P., Rainer, T. and Poon, W.S. (2013). Outcomes of traumatic brain injury in Hong Kong: Validation with the TRISS, CRASH, and IMPACT models. Journal of Clinical Neuroscience 20, 1693-1696.

68. Yeoman, P., Pattani, H., Silcocks, P., Owen, V. and Fuller, G. (2011). Validation of the IMPACT outcome prediction score using the Nottingham head injury register dataset. Journal of Trauma - Injury, Infection and Critical Care 71, 387-392.

69. Kandil, A., Kenawi, M., Samir, A. and Hussein, K. (2017). Traumatic brain injury predictive value of common intensive care severity scores. Research and Opinion in Anesthesia and Intensive Care 4, 124.

70. Cicuendez, M., Castano-Leon, A., Ramos, A., Hilario, A., Gomez, P.A. and Lagares, A. (2018). The added prognostic value of magnetic resonance imaging in traumatic brain injury: The importance of traumatic axonal injury when performing ordinal logistic regression. Journal of Neuroradiology (Epub ahead of print).

71. Faried, A., Satriawan, F.C. and Arifin, M.Z. (2018). Feasibility of Online Traumatic Brain Injury Prognostic Corticosteroids Randomisation After Significant Head Injury (CRASH) Model as a Predictor of Mortality. World neurosurgery 116, e239-e245.

72. Ho, K.M., Honeybul, S. and Ambati, R. (2018). Prognostic Significance of Magnetic Resonance Imaging in Patients with Severe Nonpenetrating Traumatic Brain Injury Requiring Decompressive Craniectomy. World Neurosurgery 112, 277-283.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

73. Sadaka, F., Doctors, N., Pearson, T., Snyders, B. and O'Brien, J. (2018). Does red cell distribution width predict outcome in traumatic brain injury: comparison to corticosteroid randomization after significant head injury. Journal of clinical medicine research 10, 9.

74. Sadaka, F., Jadhav, A., Miller, M., Saifo, A., O'Brien, J. and Trottier, S. (2018). Is it possible to recover from traumatic brain injury and a Glasgow coma scale score of 3 at emergency department presentation? The American journal of emergency medicine.

75. Van Calster, B. and Vickers, A.J. (2015). Calibration of risk prediction models: impact on decision-analytic performance. Med Decis Making 35, 162-169.

76. van Houwelingen H.C., P.H. In: *Dynamic Prediction in Clinical Survival Analysis. Taylor & Francis Inc, 2011*.

77. Siontis, G.C., Tzoulaki, I., Castaldi, P.J. and Ioannidis, J.P. (2015). External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. J Clin Epidemiol 68, 25-34.

78. Vergouwe, Y., Moons, K.G. and Steyerberg, E.W. (2010). External validity of risk models: Use of benchmark values to disentangle a case-mix effect from incorrect coefficients. Am J Epidemiol 172, 971-980.

79. Maas, A.I.R., Menon, D.K., Adelson, P.D., Andelic, N., Bell, M.J., Belli, A., Bragge, P., Brazinova, A., Buki, A., Chesnut, R.M., Citerio, G., Coburn, M., Cooper, D.J., Crowder, A.T., Czeiter, E., Czosnyka, M., Diaz-Arrastia, R., Dreier, J.P., Duhaime, A.C., Ercole, A., van Essen, T.A., Feigin, V.L., Gao, G., Giacino, J., Gonzalez-Lara, L.E., Gruen, R.L., Gupta, D., Hartings, J.A., Hill, S., Jiang, J.Y., Ketharanathan, N., Kompanje, E.J.O., Lanyon, L., Laureys, S., Lecky, F., Levin, H., Lingsma, H.F., Maegele, M., Majdan, M., Manley, G., Marsteller, J., Mascia, L., McFadyen, C., Mondello, S., Newcombe, V., Palotie, A., Parizel, P.M., Peul, W., Piercy, J., Polinder, S., Puybasset, L., Rasmussen, T.E., Rossaint, R., Smielewski, P., Soderberg, J., Stanworth, S.J., Stein, M.B., von Steinbuchel, N., Stewart, W., Steyerberg, E.W., Stocchetti, N., Synnot, A., Te Ao, B., Tenovuo, O., Theadom, A., Tibboel, D., Videtta, W., Wang, K.K.W., Williams, W.H., Wilson, L., Yaffe, K., In, T.P. and Investigators (2017). Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. Lancet Neurol 16, 987-1048.

80. Maas, A.I., Menon, D.K., Steyerberg, E.W., Citerio, G., Lecky, F., Manley, G.T., Hill, S., Legrand, V., Sorgner, A., Participants, C.-T. and Investigators (2015). Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI): a prospective longitudinal observational study. Neurosurgery 76, 67-80.

81. Yue, J.K., Vassar, M.J., Lingsma, H.F., Cooper, S.R., Okonkwo, D.O., Valadka, A.B., Gordon, W.A., Maas, A.I., Mukherjee, P., Yuh, E.L., Puccio, A.M., Schnyer, D.M., Manley, G.T. and Investigators, T.-T. (2013). Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury. J Neurotrauma 30, 1831-1844.

82. CREACTIVE - Collaborative REsearch on ACute Traumatic Brain Injury in intensiVe Care Medicine in Europe (CREACTIVE). ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT02004080. Accessed December 3, 2018.

83. Carney, N., Totten, A.M., O'Reilly, C., Ullman, J.S., Hawryluk, G.W., Bell, M.J., Bratton, S.L., Chesnut, R., Harris, O.A., Kissoon, N., Rubiano, A.M., Shutter, L., Tasker, R.C., Vavilala, M.S., Wilberger, J., Wright, D.W. and Ghajar, J. (2017). Guidelines for the Management of Severe Traumatic Brain Injury, Fourth Edition. Neurosurgery 80, 6-15.

84. Letsinger, J., Rommel, C., Hirschi, R., Nirula, R. and Hawryluk, G.W.J. (2017). The aggressiveness of neurotrauma practitioners and the influence of the IMPACT prognostic calculator. PloS one 12, e0183552.

85. Moskowitz, J., Quinn, T., Khan, M.W., Shutter, L., Goldberg, R., Col, N., Mazor, K.M. and Muehlschlegel, S. (2018). Should We Use the IMPACT-Model for the Outcome Prognostication of TBI Patients? A Qualitative Study Assessing Physicians' Perceptions. MDM Policy Pract 3, 2381468318757987.

86. Wang, X., Huang, Y., Li, L., Dai, H., Song, F. and Chen, K. (2018). Assessment of performance of the Gail model for predicting breast cancer risk: a systematic review and meta-analysis with trial sequential analysis. Breast Cancer Res 20.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

87. Lip, G.Y., Frison, L., Halperin, J.L. and Lane, D.A. (2011). Comparative validation of a novel risk score for predicting bleeding risk in anticoagulated patients with atrial fibrillation: the HAS-BLED (Hypertension, Abnormal Renal/Liver Function, Stroke, Bleeding History or Predisposition, Labile INR, Elderly, Drugs/Alcohol Concomitantly) score. J Am Coll Cardiol 57, 173-180.

88. Lip, G.Y., Nieuwlaat, R., Pisters, R., Lane, D.A. and Crijns, H.J. (2010). Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 137, 263-272.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 1. Summary of characteristics of development, validation and extension of models for moderate and severe traumatic brain injury.

Characteristics	Development (n=42)	External validation (n=149)	Extension (n=12)
No. of models	42	31	12
Median number of patients (IQR)	700 (381-1466)	409 (290-890)	342 (160-534)
Type of model			
Regression analysis	40 (94)	142 (95)	12 (67)
Classification tree	1 (3)	7 (5)	-
Other[ab]	1 (3)	-	4 (33)
Internal validation		NA	
Apparent	15 (36)		4 (33)
Cross-validation	6 (14)		-
Bootstrapping	11 (26)		8 (67)
Split sample	13 (31)		3 (25)
Performance measures			
Calibration			
Plot	15 (36)	80 (54)	1 (8)
Goodness of fit	36 (86)	77 (52)	10 (83)
Slope	2 (5)	53 (36)	5 (42)
Intercept	2 (5)	53 (36)	5 (42)
Other[c]	2 (5)	7 (5)	3 (25)
Discrimination			
Accuracy rate	1 (8)	6 (4)	-
	2 (5)	4 (3)	-
Sensitivity/specificity	32 (76)	142 (95)	11 (92)
ROC/AUC	13 (31)	39 (26)	8 (67)
Other[d]			

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

IQR = interquartile range, ROC = receiver operating characteristic curve, AUC = area under the receiver operating characteristic curve

[a]E.g. Bayesian methods, discriminant analysis, machine learning

[b]One study compared five different statistical approaches on the same cohort: logistic regression, decision tree, neural network, Bayesian methods and discriminant analysis.[27]

[c]E.g. Calibration belt

[d]E.g. Nagelkerke R^2, Brier score

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 2. Summary model performance IMPACT and CRASH models versus other models at external validation

Performance measure	**IMPACT models**		**CRASH models**		**Nijmegen clinical model 1**		**Hukkelhoven**	
Mortality	Discrimination: 56 validations Calibration: 31 validations		Discrimination: 23 validations Calibration: 16 validations		Discrimination: 2 validations Calibration: 1 validation		Discrimination: 4 validations Calibration: 1 validation	
	Mean[a]	Range	Mean[a]	Range	Mean[a]	Range	Mean[a]	Range
AUC	0.79	0.65-0.90	0.82	0.66-0.99	0.84	0.82-0.86	0.81	0.74-0.89
Calibration slope	1.1	0.42-2.3	1.1	0.64-1.9	0.98	-	1.1	-
Calibration intercept	-0.22	-3.3-0.93	-0.41	-3.2-0.51	-0.29	-	-0.13	-
Unfavorable outcome	Discrimination: 55 validations Calibration: 26 validations		Discrimination: 24 validations Calibration: 17 validations		Discrimination: 2 validations Calibration: 1 validation		Discrimination: 3 validations Calibration: 1 validation	
	Mean[a]	Range	Mean[a]	Range	Mean[a]	Range	Mean[a]	Range
AUC	0.77	0.66-0.92	0.78	0.66-1.00	0.82	0.81-0.82	0.74	0.69-0.83
Calibration slope	0.90	0.63-2.1	0.89	0.57-2.5	0.87	-	0.57	-
Calibration intercept	0.044	-4.2-1.1	-0.13	-5.7-2.4	-0.74	-	0.39	-

[a]Weighted for the square root of sample size

IMPACT = International Mission for Prognosis and Analysis of Clinical Trials; CRASH = Corticoid Randomisation After Significant Head injury; AUC = area under the receiver operating characteristic curve

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Table 3. Recommendations on (statistical) evaluation and implementation of prognostic models for moderate and severe TBI

- Continuous validation and updating of prognostic models is required to judge generalizability and transportability to other TBI populations.
- Calibration reflects the ability of the prognostic model to provide reliable predictions and should thus be reported at every external validation.
- The currently available prognostic models for moderate and severe TBI discriminate well between low risk and high risk patients.
- Caution is required when providing predictions for patients in a specific clinical setting.
- Prognostic models for moderate and severe TBI may need to be recalibrated for each new setting before implementation in clinical practice is warranted.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Figure Legends

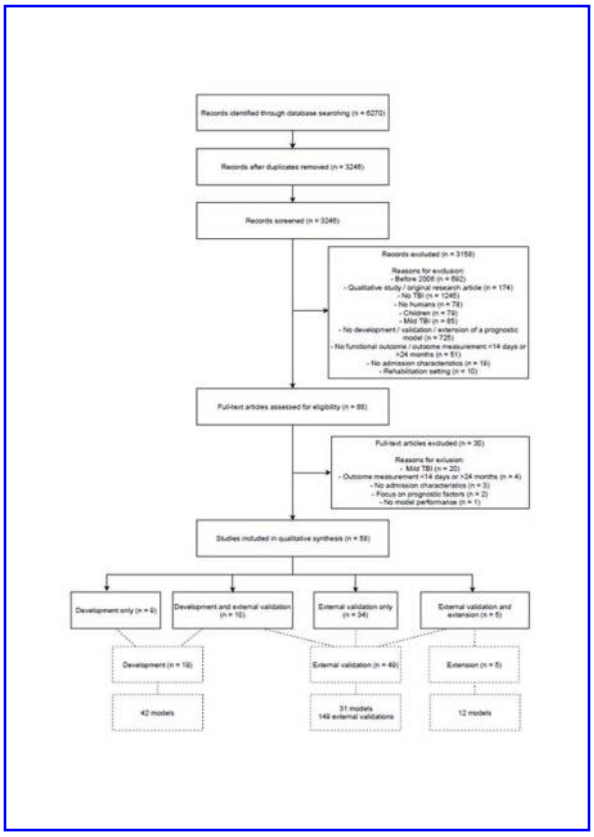

Figure 1. PRISMA flow diagram of selected articles.

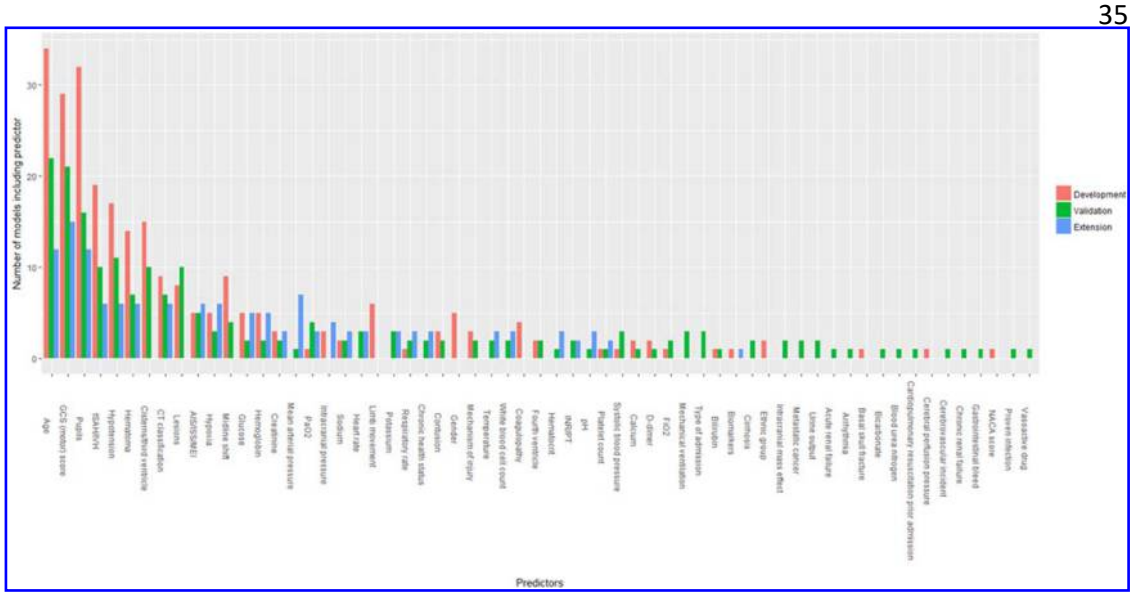

Figure 2. Overview of predictors included in 42 models for development, 31 models for external validation and 12 models for extension.

GCS = Glasgow Coma Scale; tSAH = traumatic subarachnoid hemorrhage; IVH = intraventricular hemorrhage; CT = computed tomography; AIS = Abbreviated Injury Scale; ISS = Injury Severity Score; MEI = major extracranial injury; PaO2 = partial arterial pressure of oxygen; INR = international normalized rate; PT = prothrombin time; pH = potential hydrogen; FiO2 = fraction of inspired oxygen; NACA = National Advisory Committee for Aeronautics

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.

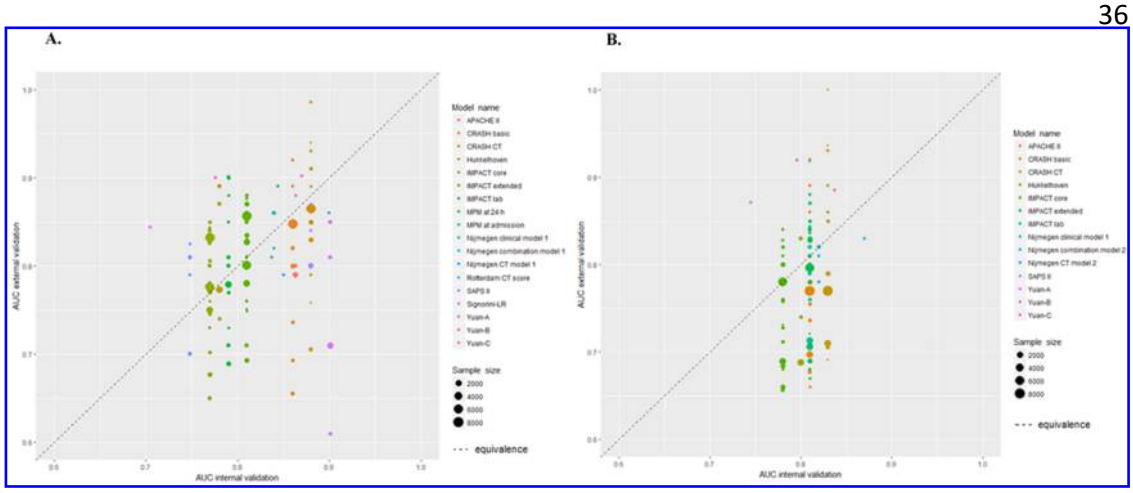

Figure 3. Correlation of discriminative ability (area under the receiver operating characteristic curve, AUC) of models for (**A**) mortality and (**B**) unfavorable outcome (defined as GOS 1-3 or GOSE 1-4) between internal and external validation. The colors of the dots represent the different prognostic models that have been externally validated once or more, and the dot size refers to the sample size of the different validation cohorts. The diagonal line indicates equivalence between model discrimination at internal and external validation. Above this line, model discrimination at external validation was better than at internal validation. The dots below the line indicate a higher AUC at internal validation compared to external validation.

APACHE = Acute Physiology And Chronic Health Evaluation; CRASH = Corticoid Randomisation After Significant Head injury; CT = computed tomography; IMPACT = International Mission for Prognosis and Analysis of Clinical Trials; MPM = mortality probability models; SAPS = Simplified Acute Physhiology Score; LR = logistic regression; AUC = area under the receiver operating characteristic curve

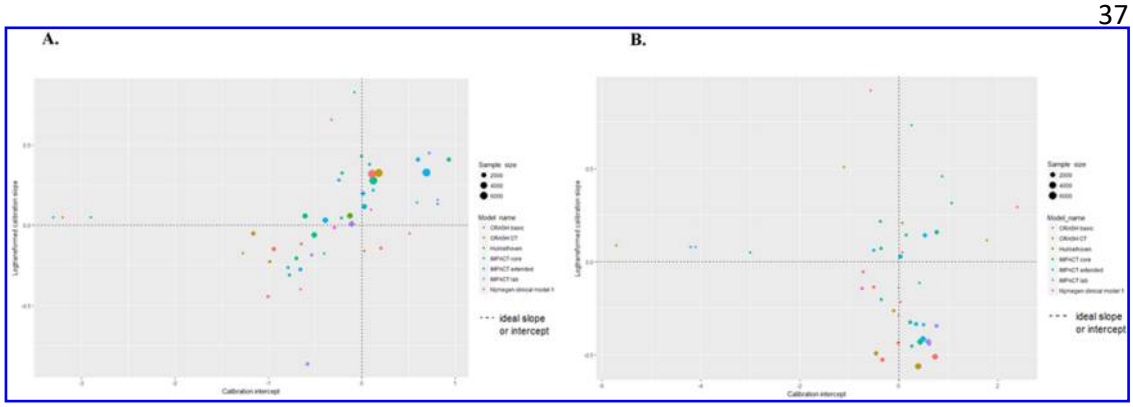

Figure 4. Calibration intercept and slope reported for models for (**A**) mortality and (**B**) unfavorable outcome (defined as GOS 1-3 or GOSE 1-4) at external validation. The colors of the dots represent the different prognostic models that have been externally validated once or more, and the dot size refers to the sample size of the different validation cohorts. The vertical line indicates the ideal calibration intercept and the horizontal line shows the perfect calibration slope. A calibration intercept > 0 indicates systematic underestimation of mortality or unfavorable outcome, and an intercept < 0 refers to systematic overestimation of outcome risk. A calibration slope below or above the horizontal line indicates that predictions were too extreme: low predictions too low, and high predictions too high. In short, the closer the dots are to the intersection of these lines, the better the ability of the model to provide reliable predictions for patients in that specific setting.

CRASH = Corticoid Randomisation After Significant Head injury; CT = computed tomography; IMPACT = International Mission for Prognosis and Analysis of Clinical Trials

Journal of Neurotrauma
Prognosis in Moderate and Severe Traumatic Brain Injury: A Systematic Review of Contemporary Models and Validation Studies (DOI: 10.1089/neu.2019.6401)
This paper has been peer-reviewed and accepted for publication, but has yet to undergo copyediting and proof correction. The final published version may differ from this proof.