External Validation and Recalibration of Risk Prediction Models for Acute Traumatic Brain Injury among Critically Ill Adult Patients in the United Kingdom

David A. Harrison,¹ Kathryn A. Griggs,¹ Gita Prabhu,¹ Manuel Gomes,² Fiona E. Lecky,³
Peter J. A. Hutchinson,⁴ David K. Menon,⁵ and Kathryn M. Rowan¹
on behalf of the Risk Adjustment In Neurocritical care (RAIN) Study Investigators

Abstract

This study validates risk prediction models for acute traumatic brain injury (TBI) in critical care units in the United Kingdom and recalibrates the models to this population. The Risk Adjustment In Neurocritical care (RAIN) Study was a prospective, observational cohort study in 67 adult critical care units. Adult patients admitted to critical care following acute TBI with a last pre-sedation Glasgow Coma Scale score of less than 15 were recruited. The primary outcomes were mortality and unfavorable outcome (death or severe disability, assessed using the Extended Glasgow Outcome Scale) at six months following TBI. Of 3626 critical care unit admissions, 2975 were analyzed. Following imputation of missing outcomes, mortality at six months was 25.7% and unfavorable outcome 57.4%. Ten risk prediction models were validated from Hukkelhoven and colleagues, the Medical Research Council (MRC) Corticosteroid Randomisation After Significant Head Injury (CRASH) Trial Collaborators, and the International Mission for Prognosis and Analysis of Clinical Trials in TBI (IMPACT) group. The model with the best discrimination was the IMPACT “Lab” model (C index, 0.779 for mortality and 0.713 for unfavorable outcome). This model was well calibrated for mortality at six months but substantially under-predicted the risk of unfavorable outcome. Recalibration of the models resulted in small improvements in discrimination and excellent calibration for all models. The risk prediction models demonstrated sufficient statistical performance to support their use in research and audit but fell below the level required to guide individual patient decision-making. The published models for unfavorable outcome at six months had poor calibration in the UK critical care setting and the models recalibrated to this setting should be used in future research.

Key words: outcome measures; prospective study; traumatic brain injury

Introduction

ACUTE TRAUMATIC BRAIN INJURY (TBI) is a leading cause of death and disability worldwide and the major cause of death under the age of 40 in developed countries.¹ While there has been improvement in outcome over the past few decades, there is a need to understand how to achieve the best results for patients with TBI, not only in terms of survival but also the quality of that survival. Risk prediction models, using information from soon after the injury to predict the risk of longer-term outcome, have a role to play in this. Risk prediction models are needed to audit the outcomes achieved from routine clinical care against a standard to underpin observational research, to enable risk stratification in randomized trials and, potentially, to support clinical communication and decision-making.²

Risk prediction models have been in established use in adult, general critical care units for over 30 years, since the publication of the original Acute Physiology And Chronic Health Evaluation (APACHE) model in 1981.³ However, there are significant limitations to using models developed and validated for general critical care for patients receiving neurocritical care. The inclusion and handling within risk prediction models for general critical care of variables with specific prognostic importance in acute TBI, for example pupil reactivity and computed tomography (CT) findings, is often limited.⁴ Further, all risk prediction models for adult, general critical care use an outcome of mortality at discharge from acute hospital, which is not considered adequate for neurocritical care patients where longer-term (e.g., six-month) morbidity and mortality are more valid.⁵

Although a large number of specific risk prediction models for TBI exists, a systematic review found that most models are limited

¹Clinical Trials Unit, Intensive Care National Audit and Research Centre, Napier House, London, United Kingdom.

²Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom.

³School of Health and Related Research, University of Sheffield, Regent Court, Sheffield, United Kingdom.

⁴Department of Clinical Neurosciences, ⁵Division of Anaesthesia, University of Cambridge, Cambridge, United Kingdom.

by being based on small samples of patients, having poor methodology, and rarely being validated on external populations.⁶ Despite the more recent development of models based on larger, more representative data sources,⁷ these models require further prospective external validation, and potentially recalibration, before they can be applied with confidence for research and audit in a different neurocritical care setting.

The primary aims of the Risk Adjustment In Neurocritical care (RAIN) Study were to externally validate risk prediction models for acute TBI and to use the best performing model(s) to evaluate the optimum location and comparative costs of neurocritical care in the UK National Health Service (NHS). This paper reports the results of the external validation and of recalibrating the models to the RAIN Study data.

Methods

Selection of risk prediction models

Perel and colleagues previously conducted a systematic review to identify and assess existing risk prediction models for TBI

published up to 2005.⁶ A total of 53 studies (reporting 102 models) were identified. Of these, the models developed by Signorini and colleagues (1999)⁸ and Hukkelhoven and colleagues (2005)⁹ were reported by the authors to be the “most clinically useful” for patients from high-income countries with moderate and severe TBI, as they fulfilled the majority of the methodological requirements and showed acceptable performance in external validation.

The first stage of the model selection process for the RAIN Study was to update the existing systematic review to identify additional risk prediction models for TBI that had been published since 2006. The systematic review was updated using the original published search strategy and inclusion criteria to identify risk prediction models that: 1) gave an overall prognostic estimation by combining the predictive information from at least two variables (studies could develop new models or evaluate previous ones); 2) used variables collected before hospital discharge; 3) included patients of any age; 4) included patients with any type or severity of TBI; and 5) predicted any outcome. To select models that were most likely to have good performance in a setting of adult patients in the UK, risk prediction models identified from the updated systematic review were additionally assessed against the following eligibility criteria to identify models for external validation in the RAIN Study that:

TABLE 1. DEFINITIONS OF PREDICTOR VARIABLES

Predictor	Definition
Age	Age in whole years at the time of TBI, calculated from the date of birth and date of TBI.
GCS	The last recorded GCS prior to sedation, which may have been either the pre-hospital GCS, the first recorded at hospital GCS, or a subsequent measurement prior to sedation and admission to the critical care unit.
Motor score	The motor score component of the last pre-sedation GCS.
Pupil reactivity	The first recorded at hospital pupil reactivity, if available, or the pre-hospital pupil reactivity if no first at hospital value was available. Pupil reactivity was recorded for the left and right eye separately and combined as either both reactive, one reactive, or neither reactive. If only one eye could be assessed, pupil reactivity was assigned as “both reactive” if this eye was reactive and “neither reactive” if this eye was not reactive.
Major extracranial injury	Defined as an injury requiring hospital admission in its own right.
Pre-hospital hypoxia	Defined as either a pre-hospital oxygen saturation less than 90%, or pre-hospital hypoxia strongly suspected if there was clinical evidence of hypoxia (e.g., tension pneumothorax) but no pre-hospital oxygen saturation was recorded.
Pre-hospital hypotension	Defined as either a pre-hospital systolic blood pressure less than 90 mm Hg, or pre-hospital hypotension strongly suspected if there was clinical evidence of hypotension (e.g., multiple trauma with massive blood loss) but no pre-hospital systolic blood pressure was recorded.
Marshall CT classification	Based on the first CT scan following the TBI, and defined as: 1) Diffuse injury I—no visible pathology seen on CT scan (first CT scan result normal); (2) Diffuse injury II—basal cisterns present, midline shift 0-5 mm, no high- or mixed-density lesion > 25 mL; 3) Diffuse injury III—basal cisterns compressed or absent, midline shift 0-5 mm, no high- or mixed-density lesion > 25 mL; 4) Diffuse injury IV—midline shift > 5 mm, no high- or mixed-density lesion > 25 mL; 5) Evacuated mass lesion—any high- or mixed-density lesion > 1 mL surgically evacuated; 6) Non-evacuated mass lesion > 25 mL—high- or mixed-density lesion > 25 mL not surgically evacuated.
Traumatic subarachnoid hemorrhage	Defined as a collection of blood between the arachnoid and pia mater either over the convexity or in the basal cisterns.
One or more small petechial hemorrhages	Includes small petechial hemorrhages at gray-white matter junction, hemispheric white matter, corpus callosum or brainstem.
Obliteration of the third ventricle or basal cisterns	Third ventricle recorded as “obliterated” or basal cisterns recorded as “absent.”
Midline shift over 5 mm	Defined as degree of displacement of the midline greater than 5 mm.
Non-evacuated hematoma	Any high- or mixed-density lesion > 1 mL not surgically evacuated.
Extradural hematoma	Defined as an accumulation of blood between the skull and dura mater.
First at hospital glucose	The first serum glucose value from blood sampled within 12 h of attendance at the first hospital for the TBI.
First at hospital hemoglobin	The first hemoglobin value from blood sampled within 12 h of attendance at the first hospital for the TBI.

TBI, traumatic brain injury; GCS, Glasgow Coma Scale; CT, computed tomography.

1) were based on predominantly adult (>16 years) populations; 2) had a sample size of greater than 500 patients in either the development and/or validation dataset; 3) aimed to evaluate outcome regardless of care received during the hospital stay; and 4) were multicenter (or, if single center, UK based).

Eligible models identified from the updated systematic review, plus the two most clinically useful models from the original systematic review, were reviewed by the RAIN Study Steering Group (see Acknowledgements) to determine if there were any relevant models that had been missed and to select the most appropriate models for external validation in the RAIN Study.

Prospective cohort study

In order to externally validate the selected risk prediction models, a prospective cohort study was undertaken in adult critical care units (including dedicated neurocritical care units, combined neuro/general critical care units within a neuroscience center and general critical care units outside a neuroscience center). All critical care units in the 32 UK neuroscience centers were invited to participate and also critical care units outside neuroscience centers that participate in the Intensive Care National Audit and Research Centre Case Mix Programme (ICNARC CMP) – the national clinical audit for adult critical care in England, Wales, and Northern Ireland. All adult patients (ages 16 years or older) admitted to participating critical care units following an actual or suspected TBI, and with a Glasgow Coma Scale (GCS) score of less than 15 following resuscitation were identified. Patients who were subsequently found to not have a TBI were excluded. The RAIN Study was approved by the NHS Research Ethics Committee for Wales (09/MRE09/10) and the Scotland A Research Ethics Committee (09/MRE00/15). Patients were enrolled without informed consent under the provisions of the Mental Capacity Act 2005 and the Adults with Incapacity (Scotland) Act 2000. The Ethics and Confidentiality Committee of the National Information Governance Board gave support under Section 251 of the NHS Act 2006 for permission to hold sufficient patient identifiable data to contact the patient and gain their consent (approval number ECC 2-06(d)/2009).

The RAIN Study dataset was developed from the original publications reporting each risk prediction model, any other associated publications and, where available, original study documentation relating to the data sources. Data were entered onto a dedicated secure web portal and were validated centrally for incomplete, inaccurate and inconsistent data. To avoid duplication of data collection, data for the RAIN Study were linked with data for the same patients from the ICNARC CMP and from the Scottish Intensive Care Society Audit Group—the equivalent national clinical audit in Scotland.

Patients were followed up to six months for mortality by linkage with death registration and for unfavorable outcome (death or severe disability) with the Extended Glasgow Outcome Scale administered by postal or telephone questionnaire, using methods based on those employed in previous research studies.^{10,11}

Although a GCS score of less than 15 was an inclusion criterion for the RAIN Study, some patients with suspected TBI were recruited with no GCS recorded, and others that had an initial GCS score of less than 15, either pre-hospital or on presentation at hospital, subsequently improved prior to sedation. Patients from the RAIN Study database were included in the analysis if their last GCS prior to sedation/admission to critical care was less than 15.

Predictor variables were calculated from the raw data in the RAIN Study database; definitions are provided in Table 1.

Handling of missing data

Missing data in the RAIN Study database were addressed with multiple imputation.^{12,13} As the Extended Glasgow Outcome Scale is conditional on survival status, the imputation was conducted in two stages. In the first stage, imputation models were specified for predictors and for mortality at six months, considering other fully-observed covariates (such as age and GCS score). In the second stage, for each of the previously imputed datasets, imputation models were specified for the Extended Glasgow Outcome Scale for those patients who were either known to be alive at six months or were predicted to be alive by the first stage imputation model, considering fully-observed covariates and intermediate endpoints

TABLE 2. OVERVIEW OF RISK PREDICTION MODELS FOR TRAUMATIC BRAIN INJURY INCLUDED IN THE EXTERNAL VALIDATION

	<i>Hukkelhoven</i>	<i>CRASH</i>	<i>IMPACT</i>
Data	Two multicenter RCTs (<i>n</i> = 2269)	One multicenter RCT (<i>n</i> = 2482 – high income)	Eight RCTs and three observational studies (<i>n</i> = 8509)
Predictors in model	Age, GCS motor score, pupil reactivity, pre-hospital hypoxia, pre-hospital hypotension, Marshall CT class, traumatic subarachnoid hemorrhage	<i>CRASH Basic</i> : Age, GCS, pupil reactivity, major extracranial injury <i>CRASH CT</i> : CRASH Basic predictors plus traumatic subarachnoid hemorrhage, one or more small petechial hemorrhages, obliteration of the third ventricle or basal cisterns, midline shift over 5 mm, non-evacuated hematoma	<i>IMPACT Core</i> : Age, GCS motor score, pupil reactivity <i>IMPACT Extended</i> : IMPACT Core predictors plus pre-hospital hypoxia, pre-hospital hypotension, Marshall CT class, traumatic subarachnoid hemorrhage, extradural hematoma <i>IMPACT Lab</i> : IMPACT Extended predictors plus glucose, hemoglobin
Outcomes	Mortality at six months Unfavorable outcome at six months	Unfavorable outcome at six months*	Mortality at six months Unfavorable outcome at six months
Validation	European Brain Injury Consortium (<i>n</i> = 796)	IMPACT dataset (<i>n</i> = 8509)	CRASH dataset (<i>n</i> = 6681)

*Models were also developed for mortality at 14 d—not included in this external validation.

CRASH, Corticosteroids After Significant Head Injury Trial; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in TBI (traumatic brain injury); RCT, randomized controlled trial; GCS, Glasgow Coma Scale; CT, computed tomography.

TABLE 3. CASE MIX AND OUTCOMES FOR THE HUKKELHOVEN, CRASH, AND IMPACT MODEL DEVELOPMENT SAMPLES, COMPARED WITH THE RAIN STUDY DATABASE

	RAIN (all)	RAIN (meeting Hukkelhoven inclusion)	Hukkelhoven development sample	RAIN (meeting CRASH inclusion)	CRASH development sample ^a	RAIN (meeting IMPACT inclusion)	IMPACT development sample
Number of patients	2975	1806	2269	2710	2482	2528	8509
Age, mean (SD)	44.7 (18.9)	38.0 (14.4)	33.2 (13.5)	43.9 (18.8)	40.6 (19.4)	-	-
Age, median [IQR]	44 [28, 59]	-	-	-	-	44 [28, 59]	30 [21, 45]
Male, <i>n</i> (%)	2263 (76.1)	1414 (78.3)	1757 (77.4)	2069 (76.4)	(78.9)	-	-
GCS, <i>n</i> (%)							
13–14 (Mild TBI)	447 (15.0)	-	-	398 (14.7)	(32.6)	-	-
9–12 (Moderate TBI)	765 (25.7)	-	-	677 (25.0)	(23.6)	-	-
3–8 (Severe TBI)	1763 (59.3)	-	-	1635 (60.3)	(43.8)	-	-
GCS motor score, <i>n</i> (%)							
5/6 (Localizes/obeys)	1409 (48.5)	672 (37.2)	933 (41.1)	-	-	974 (39.5)	2591 (32.2)
4 (Normal flexion)	349 (12.0)	242 (13.4)	659 (29.0)	-	-	345 (14.0)	1940 (24.1)
3 (Abnormal flexion)	209 (7.2)	140 (7.8)	374 (16.5)	-	-	209 (8.5)	1085 (13.5)
1/2 (None/extension)	940 (32.3)	752 (41.6)	303 (13.4)	-	-	940 (38.1)	2437 (30.3)
Pupil reactivity, <i>n</i> (%)							
Both reactive	2215 (80.3)	1275 (75.9)	1401 (70.5)	2016 (80.0)	(84.0)	1828 (77.9)	4486 (63.0)
One reactive	167 (6.1)	122 (7.3)	279 (14.0)	154 (6.1)	(6.6)	152 (6.5)	886 (12.4)
Neither reactive	378 (13.7)	283 (16.8)	308 (15.5)	350 (13.9)	(9.5)	367 (15.6)	1754 (24.6)
Major extracranial injury, <i>n</i> (%) [N]	1213 (40.8) [2974]	-	-	1163 (42.9) [2709]	(22.5)	-	-
Pre-hospital hypoxia, <i>n</i> (%) [N]	354 (12.6) [2800]	260 (15.2) [1709]	429 (21.4) [2006]	-	-	327 (13.8) [2379]	1116 (20.5) [5452]
Pre-hospital hypotension, <i>n</i> (%) [N]	190 (6.8) [2791]	141 (8.3) [1703]	395 (17.9) [2208]	-	-	169 (7.1) [2370]	1171 (18.2) [6420]
Marshall CT classification, <i>n</i> (%)							
1/2 (Diffuse injury I/II)	1051 (37.9)	674 (39.4)	1006 (44.7)	-	-	864 (36.7)	2198 (42.3)
3 (Diffuse injury III)	248 (8.9)	193 (11.3)	426 (18.9)	-	-	214 (9.1)	863 (16.6)
4 (Diffuse injury IV)	78 (2.8)	51 (3.0)	88 (3.9)	-	-	64 (2.7)	187 (3.6)
5/6 (Evacuated mass lesion/ Non-evacuated mass lesion > 25 ml)	1396 (50.3)	792 (46.3)	729 (32.4)	-	-	1213 (51.5)	1944 (37.4)
Traumatic subarachnoid hemorrhage, <i>n</i> (%) [N]	1565 (56.0) [2797]	1006 (58.4) [1724]	1090 (49.0) [2223]	1456 (56.8) [2563]	(26.4)	1365 (57.5) [2375]	3313 (44.8) [7393]
Laboratory results, median (IQR) [N]							
Glucose (mmol L ⁻¹)	7.6 (6.3, 9.5) [2496]	-	-	-	-	7.7 (6.3, 9.5) [2142]	8.2 (6.7, 10.4) [4830]
Hemoglobin (g dL ⁻¹)	13.4 (11.9, 14.6) [2683]	-	-	-	-	13.4 (11.8, 14.6) [2280]	12.7 (10.8, 14.3) [4376]
Six month outcomes, <i>n</i> (%) [N]							
Mortality	748 (26.0) [2881]	397 (22.7) [1746]	500 (22.0) [2269]	-	-	685 (27.9) [2455]	2396 (28.2) [8509]
Unfavorable outcome	1481 (61.3) [2420]	867 (60.0) [1444]	862 (40.1) [2149]	1340 (61.0) [2198]	(38.5)	1323 (64.3) [2059]	4082 (48.0) [8509]

^aAbsolute numbers were not reported, only percentages.CRASH, Corticosteroids After Significant Head Injury Trial; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in TBI (traumatic brain injury); RAIN, Risk Adjustment In Neurocritical Care Study; SD, standard deviation; IQR, interquartile range; *N*, number of patients with complete data for variable; GCS, Glasgow Coma Scale; CT, computed tomography.

(such as hospital length of stay). Each multiple imputation model assumed that the data were missing at random, conditional on the variables included in the imputation models. Five imputed datasets were generated at each stage of the imputation,¹³ giving 25 imputed datasets in total. All multiple imputation models were implemented in R (www.r-project.org) using Multivariate Imputation by Chained Equations.¹⁴ As a sensitivity analysis, all analyses were repeated using the original (non-imputed) data.

Validation datasets

For the primary analysis, each model was validated in the multiply imputed datasets, using: 1) all patients in the RAIN Study with a last pre-sedation GCS score of less than 15; and 2) only those patients meeting the original inclusion criteria for each model. For the validation of risk prediction models for mortality at six months, the five datasets from the first stage imputation were used. For the validation of risk prediction models for unfavorable outcome at six months, the 25 datasets from the second stage imputation were used.

As a sensitivity analysis, the analysis for models predicting unfavorable outcome in all patients in the RAIN Study database—analysis 1) above—was repeated excluding those patients assigned to the Extended Glasgow Outcome Scale category of severe disability that was either pre-existing or not due to the TBI.

Statistical analysis

The case mix and outcomes of patients reported for the original development sample for each risk prediction model were compared with those in the RAIN Study database that met the original inclusion criteria for the respective model and with all patients in the RAIN Study database.

Each risk prediction model was validated in each of the five RAIN Study validation datasets using measures of calibration, discrimination and overall fit, as described below. For multiply imputed data, measures of model performance were calculated in

each imputed dataset and combined using Rubin's rules.¹² Chi-square test statistics were combined to calculate an F statistic using the method of Li and colleagues.¹⁵

Discrimination was assessed on the C index,¹⁶ which is equivalent to the area under the receiver operating characteristic curve.¹⁷ Calibration was assessed by graphical plots of observed against expected risk both continuously, using non-parametric locally-weighted sum of squares (LOWESS) smoothers,¹⁸ and by categories, in 10 equal-sized groups by predicted risk. The magnitude and direction of miscalibration was assessed using Cox's calibration regression.¹⁹ Overall fit was assessed using Brier's score, the mean squared error between the outcome and prediction.²⁰

The risk prediction models were recalibrated by estimating new coefficients for each model in the full validation dataset. The recalibrated models were assessed using the C index and Brier's score, adjusted for potential optimism from fitting and evaluating the models in the same dataset using Efron's bootstrap method.²¹

Analyses were conducted using Stata/SE Version 13.0 (Stata-Corp LP, College Station, TX).

Results

Selection of risk prediction models

The electronic database searches identified a total of 1832 citations (date of search, December 22, 2008). After screening of titles and abstracts, 23 potentially eligible papers were identified for full text review.^{22–44} In addition, the electronic database searches identified seven review articles,^{1,45–50} the references lists of which did not identify any further potentially eligible papers. Of the 23 full text papers reviewed, 13 were excluded either because they were not studies that had developed or evaluated risk prediction models for TBI or because the models included predictors measured outside of the hospital, leaving 10 studies published since 2006 (reporting 78 models) meeting the Perel and colleagues inclusion criteria.^{22,27,30,35–37,39,40,42,43}

Of these 10 studies, eight did not fulfil the RAIN Study eligibility criteria—because the models had been developed in pediatric populations ($n=3$), were based on samples of fewer than 500 patients ($n=6$), adjusted for care received within hospital ($n=1$), or had been conducted in a single-center, non-UK setting ($n=6$)—resulting in two eligible studies (reporting 14 models).^{35,40} Additionally including the two studies reporting the most clinically useful models from the original systematic review resulted in four studies (reporting 17 models) for review by the RAIN Study Steering Group—the models of Signorini and colleagues (1999),⁸ Hukkelhoven and colleagues (2005),⁹ the MRC CRASH Trial Collaborators (2008),³⁵ and Steyerberg and colleagues for the IMPACT project (2008).⁴⁰

The RAIN Study Steering Group did not identify any further studies, either published or ongoing, that would be potentially eligible for the RAIN Study (date of meeting, April 17, 2009). Of the 17 risk prediction models reported by the four eligible studies, 11 were appropriate for validation in the RAIN Study (models developed by the MRC CRASH Trial Collaborators³⁵ for low-income countries and predicting mortality at 14d were excluded). Following review by the RAIN Study Steering Group, 10 risk prediction models were selected for external validation in the RAIN Study. The model developed by Signorini and colleagues (1999)⁸ was rejected because the model was developed using data from a relatively small single-center study, there was substantial additional data burden associated with its collection, and the outcome (mortality at one year) did not incorporate functional status and was

TABLE 4. DISCRIMINATION, CALIBRATION, AND OVERALL FIT—MORTALITY AT SIX MONTHS (PUBLISHED MODELS)

	<i>All patients</i>	<i>Eligible for model</i>
C index (95% CI)		
Hukkelhoven	0.784 (0.764 to 0.805)	0.773 (0.745 to 0.801)
IMPACT Core	0.754 (0.733 to 0.774)	0.750 (0.728 to 0.772)
IMPACT Extended	0.780 (0.760 to 0.799)	0.780 (0.760 to 0.801)
IMPACT Lab	0.777 (0.756 to 0.798)	0.779 (0.758 to 0.800)
Cox calibration regression, α (95% CI) β (95% CI) [p value]		
Hukkelhoven	−0.01 (−0.13 to 0.11)	−0.13 (−0.29 to 0.04)
	1.06 (0.96 to 1.16)	1.06 (0.92 to 1.20)
	[0.27]	[0.022]
IMPACT Core	−0.50 (−0.60 to −0.40)	−0.51 (−0.61 to −0.40)
	0.95 (0.86 to 1.05)	0.94 (0.84 to 1.05)
	[<0.001]	[<0.001]
IMPACT Extended	−0.37 (−0.48 to −0.27)	−0.39 (−0.50 to −0.28)
	1.02 (0.92 to 1.11)	1.03 (0.93 to 1.13)
	[<0.001]	[<0.001]
IMPACT Lab	−0.11 (−0.23 to 0.01)	−0.11 (−0.23 to 0.01)
	1.00 (0.90 to 1.11)	1.01 (0.90 to 1.12)
	[0.080]	[0.076]
Brier's score (95% CI)		
Hukkelhoven	0.150 (0.143 to 0.158)	0.141 (0.132 to 0.151)
IMPACT Core	0.166 (0.159 to 0.173)	0.175 (0.167 to 0.182)
IMPACT Extended	0.156 (0.149 to 0.163)	0.163 (0.155 to 0.170)
IMPACT Lab	0.153 (0.145 to 0.161)	0.159 (0.151 to 0.167)

IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in TBI (traumatic brain injury); CI, confidence interval.

FIG. 1. Calibration plots—mortality at six months (published models)

Observed mortality is plotted against predicted mortality in 10 equal sized groups. The gray dashed line is from a locally-weighted sum of squares smoother. The distribution of the predicted probability of mortality is shown at the foot of each figure. Calibration plots are shown for all patients; results were similar when restricted to patients eligible for each model.

at a different time-point from all the other models being considered. The remaining three studies were large multicenter studies that considered functional outcome, as well as mortality, at six months.

Consequently, three families of risk prediction models were selected for external validation in the RAIN Study: the models of: Hukkelhoven and colleagues (hereafter termed “Hukkelhoven

models”)⁹; the MRC CRASH Trial Collaborators (“CRASH models”)³⁵; and Steyerberg and colleagues (“IMPACT models”)⁴⁰. A total of 10 models were included in the RAIN Study validation: four predicting mortality at six months (one Hukkelhoven model and three IMPACT models) and six predicting unfavorable outcome at six months (one Hukkelhoven model, two

TABLE 5. DISCRIMINATION, CALIBRATION AND OVERALL FIT—UNFAVORABLE OUTCOME AT SIX MONTHS (PUBLISHED MODELS)

	<i>All patients</i>	<i>Eligible for model</i>
C index (95% CI)		
Hukkelhoven	0.708 (0.686 to 0.730)	0.688 (0.660 to 0.716)
CRASH Basic	0.699 (0.677 to 0.721)	0.697 (0.675 to 0.720)
CRASH CT	0.708 (0.686 to 0.729)	0.710 (0.687 to 0.733)
IMPACT Core	0.693 (0.671 to 0.716)	0.689 (0.665 to 0.713)
IMPACT Extended	0.708 (0.686 to 0.730)	0.706 (0.683 to 0.730)
IMPACT Lab	0.714 (0.692 to 0.736)	0.713 (0.689 to 0.737)
Cox calibration regression, α (95% CI) β (95% CI) [p value]		
Hukkelhoven	0.44 (0.35 to 0.53)	0.39 (0.28 to 0.50)
	0.60 (0.52 to 0.68)	0.57 (0.47 to 0.66)
	[<0.001]	[<0.001]
CRASH Basic	0.74 (0.64 to 0.85)	0.73 (0.62 to 0.84)
	0.59 (0.52 to 0.67)	0.60 (0.52 to 0.67)
	[<0.001]	[<0.001]
CRASH CT	0.40 (0.31 to 0.49)	0.38 (0.28 to 0.47)
	0.56 (0.49 to 0.63)	0.57 (0.50 to 0.64)
	[<0.001]	[<0.001]
IMPACT Core	0.41 (0.32 to 0.50)	0.43 (0.34 to 0.52)
	0.66 (0.58 to 0.75)	0.65 (0.55 to 0.74)
	[<0.001]	[<0.001]
IMPACT Extended	0.47 (0.38 to 0.56)	0.48 (0.38 to 0.58)
	0.66 (0.58 to 0.75)	0.66 (0.57 to 0.75)
	[<0.001]	[<0.001]
IMPACT Lab	0.58 (0.48 to 0.67)	0.59 (0.49 to 0.74)
	0.65 (0.57 to 0.73)	0.65 (0.56 to 0.74)
	[<0.001]	[<0.001]
Brier's score (95% CI)		
Hukkelhoven	0.232 (0.222 to 0.242)	0.239 (0.227 to 0.252)
CRASH Basic	0.262 (0.251 to 0.273)	0.262 (0.251 to 0.273)
CRASH CT	0.232 (0.222 to 0.242)	0.230 (0.220 to 0.241)
IMPACT Core	0.232 (0.223 to 0.241)	0.230 (0.220 to 0.239)
IMPACT Extended	0.230 (0.221 to 0.240)	0.227 (0.217 to 0.237)
IMPACT Lab	0.236 (0.227 to 0.246)	0.233 (0.222 to 0.243)

CI, confidence interval; CRASH, Corticosteroids After Significant Head Injury Trial; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in TBI (traumatic brain injury).

CRASH models, and three IMPACT models). Table 2 describes these models.

Prospective cohort study

A total of 67 critical care units participated in the RAIN Study—31 within a neuroscience center and 36 general critical care units outside a neuroscience center. Between August 2009 and March 2011, a total of 3626 critical care unit admissions were recruited. After excluding subsequent admissions of the same patient ($n=227$; 6.3%), patients who did not prove finally to have a TBI ($n=189$; 5.2%), patients with no GCS score recorded ($n=112$; 3.1%), and patients with a last pre-sedation GCS score of 15 ($n=134$; 3.7%), 2975 patients were included in the analysis.

Of the 2975 patients included, 26 were homeless, 59 were non-UK residents, and for nine patients (military) the address details were withheld, resulting in a cohort of 2881 patients (97%) that

were followed up for six month survival by data linkage with death registrations. Of 2133 patients not reported to have died by six months, 225 (11%) declined follow-up and 220 (10%) were lost to follow-up. This left 1688 (79%) who were successfully contacted and completed either a paper ($n=1152$; 68%) or telephone ($n=536$; 32%) questionnaire. Fourteen questionnaires (0.8%) were not sufficiently complete to assign an Extended Glasgow Outcome Scale category. When combined with the 748 patients known to have died by six months, this resulted in an overall follow-up rate for unfavorable outcome of 81% (2422/2975).

Table 3 describes, for each model, the case mix and outcomes of the patients from the RAIN Study, both those meeting the original inclusion criteria for the respective model and all patients in the RAIN Study database, compared with the original model development sample.

Data completeness

Age and GCS score were 100% complete (as a requirement for meeting RAIN Study inclusion criteria), and major extracranial injury was missing for only one patient (and assumed absent). All other predictors and outcomes were included in the multiple imputation. The core predictors were available for between 92.8% (pupil reactivity) and 97.7% (GCS motor score) of patients. The first computed tomography (CT) scan was available for 2817 patients (94.7%) and the majority of these had all CT fields recorded. Laboratory variables were the least complete with 83.9% completeness for glucose and 90.2% for hemoglobin. Missing predictors, together with the outcomes for the 94 patients lost to follow-up for six-month survival (see above) were imputed in the first stage of the two-stage imputation. Outcomes for the 553 patients lost to follow-up for unfavorable outcome at six months were imputed in the second stage of the two-stage imputation.

The patients lost to follow-up for unfavorable outcomes at six months were, on average, slightly younger (mean 42 vs. 45 years) with a slightly higher proportion of males (81% vs. 75%) but with similar severity of the TBI (57%, 27%, and 15% vs. 60%, 25%, and 15% with GCS scores of 3–8, 9–12, and 13–14, respectively).

Following multiple imputation, the overall mortality at six months for the RAIN Study cohort was 25.7% and the rate of unfavorable outcomes 57.4%.

External validation of the risk prediction models

The performance of the four models predicting mortality at six months is reported in Table 4. Discrimination was similar for the Hukkelhoven, IMPACT Extended and IMPACT Lab models (C index, 0.777 to 0.784, compared with an ideal value of 1) and higher than that for the IMPACT Core model (0.754). Inspection of the calibration plots (Fig. 1) showed good calibration for the Hukkelhoven and IMPACT Lab models, whereas the IMPACT Core and Extended models tended to overestimate the risk of mortality at six months (i.e., observed mortality was lower than predicted). This visual interpretation is supported by the Cox calibration regression (Table 4). Brier's score, summarizing the overall fit of the predictions, was best (closest to zero) for the Hukkelhoven model and

FIG. 2. Calibration plots—unfavorable outcome at six months (published models)

Observed unfavorable outcome is plotted against predicted unfavorable outcome in 10 equal sized groups. The gray dashed line is from a locally-weighted sum of squares smoother. The distribution of the predicted probability of unfavorable outcome is shown at the foot of each figure. Calibration plots are shown for all patients; results were similar when restricted to patients eligible for each model.

TABLE 6. DISCRIMINATION AND OVERALL FIT—MORTALITY AND UNFAVORABLE OUTCOME AT SIX MONTHS (RECALIBRATED MODELS—OPTIMISM ADJUSTED ESTIMATES)

	<i>Mortality at six months</i>	<i>Unfavorable outcome at six months</i>
C index (95% CI)		
Hukkelhoven	0.790 (0.771 to 0.809)	0.715 (0.693 to 0.737)
CRASH Basic	0.775 (0.756 to 0.795)	0.702 (0.681 to 0.724)
CRASH CT	0.795 (0.776 to 0.815)	0.710 (0.688 to 0.731)
IMPACT Core	0.766 (0.746 to 0.785)	0.698 (0.675 to 0.720)
IMPACT Extended	0.790 (0.771 to 0.809)	0.715 (0.694 to 0.737)
IMPACT Lab	0.796 (0.776 to 0.815)	0.723 (0.701 to 0.745)
Brier's score (95% CI)		
Hukkelhoven	0.150 (0.142 to 0.157)	0.212 (0.204 to 0.219)
CRASH Basic	0.152 (0.144 to 0.160)	0.215 (0.209 to 0.222)
CRASH CT	0.145 (0.137 to 0.153)	0.213 (0.206 to 0.220)
IMPACT Core	0.155 (0.148 to 0.163)	0.217 (0.210 to 0.223)
IMPACT Extended	0.150 (0.142 to 0.157)	0.212 (0.204 to 0.219)
IMPACT Lab	0.148 (0.140 to 0.156)	0.209 (0.202 to 0.217)

CI, confidence interval; CRASH, Corticosteroids After Significant Head Injury Trial; CT, computed tomography; IMPACT, International Mission for Prognosis and Analysis of Clinical Trials in TBI (traumatic brain injury).

worst for the IMPACT Core model. There was little difference in any measures of model performance between assessment among all patients and among those eligible for each model.

The discrimination of the models for predicting unfavorable outcome at six months was worse than for mortality at six months, with C index values among all patients ranging from 0.693 to 0.714 (Table 5). The IMPACT Lab model had slightly better discrimination (C index, 0.714), followed by the models making use of CT findings but not laboratory variables (Hukkelhoven, CRASH CT, and IMPACT Extended, all 0.708) with the models using only core predictors having the worst discrimination (CRASH Basic, 0.699 and IMPACT Core, 0.693). However, these differences were extremely small. All models were poorly calibrated, substantially underestimating the risk of unfavorable outcome at six months, particularly at low predicted risk (Fig. 2); Cox calibration regression confirmed very significant departures from perfect calibration. The lack of calibration resulted in very poor values for Brier's score—little better than (and in the case of the CRASH Basic model, worse than) the value of 0.25 that would be observed from predicting a constant risk of 0.5 for all patients (Table 5).

The sensitivity analysis excluding patients with severe disability that was either pre-existing or not due to the TBI resulted in very small improvements in discrimination and no discernible change to calibration or overall fit. Results were similar when repeated in the original (non-imputed) data.

Recalibration of the risk prediction models

Recalibrating the risk prediction models resulted in small improvements in discrimination, with C index values ranging from 0.766 to 0.796 for mortality and 0.698 to 0.723 for unfavorable outcome at six months (Table 6). All recalibrated models showed excellent calibration (Fig. 3 and Fig. 4). There was little difference in performance between the recalibrated Hukkelhoven, CRASH CT, and IMPACT Extended/Lab models. The full coefficients of

the recalibrated models are presented in Tables 7, 8, and 9, compared with the original published coefficients.

Discussion

The RAIN Study has demonstrated that the Hukkelhoven, CRASH, and IMPACT risk prediction models for mortality and unfavorable outcome at six months have acceptable levels of discrimination when applied to a UK cohort of adult patients admitted to critical care units following acute TBI. While the Hukkelhoven and IMPACT Lab models for mortality at six months had good calibration, the other IMPACT models for mortality at six months over-predicted risk and all models for unfavorable outcome at six months under-predicted risk, particularly at the lower end of the risk spectrum. Recalibration of the models resulted in small improvements in discrimination and excellent calibration for all models.

The main strengths of the RAIN Study are the large, representative sample with high levels of data completeness and the rigorous methodological approach. The RAIN Study included 84% of all neuroscience centers in the UK plus representation from the important but often overlooked cohort of critically ill patients with acute TBI managed outside neuroscience centers. While the completeness of follow-up for unfavorable outcome at six months (81%) was lower than ideal, this reflects the difficulties inherent in following up such a population. Further, consent was obtained at the point of follow-up, rather than prior to enrolment (as would be the case in a randomized controlled trial), so patients that declined follow-up were still included within the initial study population. Loss to follow-up was addressed by implementing a two-stage multiple imputation process—first imputing predictors and mortality, and subsequently, for those alive or imputed to be alive at six months, imputing the Extended Glasgow Outcome Scale.

The RAIN Study is the first prospective external validation of these risk prediction models, with all previous validations having been undertaken using existing data sources from either randomized controlled trials or trauma registries.^{51–54} Consequently, while previous validation studies have been restricted to validating modified and refitted versions of the models, due to differences between the available data for validation and the data used for model development, the RAIN Study was able to define variables in advance to accurately implement all risk models as originally reported and provide a fair, direct, head-to-head comparison of the alternative models. The discrimination of the risk prediction models in the RAIN Study was generally within the range observed in previous development and validation populations.

In terms of the statistical assessment of model performance, there was very little to choose between models of similar complexity from Hukkelhoven, CRASH, and IMPACT. The best discrimination overall was from the IMPACT Lab model – the only one of the models to include laboratory parameters; however, the improvement in performance over the models of the next level down in complexity (Hukkelhoven, CRASH CT, IMPACT Extended) was small. There was a larger difference in performance between these models and the simplest models using core data only (CRASH Basic and IMPACT Core), suggesting that there is important prognostic information within the CT scan and the presence/absence of pre-hospital hypoxia/hypotension. The CRASH models included two predictors—major extracranial injury and small petechial hemorrhages—that did not demonstrate any association with outcomes in the RAIN Study. This may be seen as a reason to prefer the Hukkelhoven and IMPACT Extended models over the CRASH CT model; however, despite the inclusion of these apparently non-prognostic variables, the overall

FIG. 3. Calibration plots—mortality at six months (recalibrated models)
 Observed mortality is plotted against predicted mortality in 10 equal sized groups. The gray dashed line is from a locally-weighted sum of squares smoother. The distribution of the predicted probability of mortality is shown at the foot of each figure.

TABLE 7. COEFFICIENTS FOR THE ORIGINAL AND RECALIBRATED HUKKELHOVEN MODELS

Variable	Mortality at six months		Unfavorable outcome at six months	
	Original	Recalibrated	Original	Recalibrated
Constant	−3.267	−3.452	−2.842	−1.507
Age (per year)	−0.020	−0.018	0.0011	0.010
Age squared	0.00053	0.00063	0.00039	0.00019
GCS motor score				
4	0.494	0.194	0.675	0.357
3	0.918	0.527	1.275	0.497
½	1.126	0.868	1.690	0.728
Pupil reactivity				
One reactive	0.364	0.067	0.440	0.253
Neither reactive	0.648	1.176	0.938	0.952
Pre-hospital hypoxia	0.377	0.451	0.449	0.112
Pre-hospital hypotension	0.745	0.654	0.740	0.626
Marshall CT classification				
3 (Diffuse injury III)	0.808	1.008	0.637	0.502
4 (Diffuse injury IV)	1.354	1.284	0.628	0.731
5/6 (Evacuated mass lesion/ non-evacuated mass lesion >25 mL)	0.860	0.984	0.619	0.624
Traumatic subarachnoid hemorrhage	0.694	0.338	0.657	0.134

GCS, Glasgow Coma Scale; CT, computed tomography.

performance of the CRASH CT model was similar. For the planned subsequent analyses around optimum location and costs of care in the RAIN Study, we therefore selected the IMPACT Lab model as the primary model for risk adjustment in the base case analyses, with the CRASH CT model used for sensitivity analyses.

The risk prediction models evaluated in the RAIN Study demonstrated sufficient statistical performance to support their use for comparative audit and in research studies, for example, as a basis for risk adjustment in observational studies and risk stratification in randomized controlled trials. Given the poor calibration of the original published models, particularly for unfavorable outcome, we would recommend that the recalibrated models be used for studies in a UK critical care setting. The recalibrated risk prediction models have the potential to be clinically useful at the individual patient level both to aid communication between health care professionals and in communicating risks to patients and their loved ones to enable shared decision-making.⁵⁵ We would recommend that the models be used as an adjunct to clinical judgement and not to directly guide individual patient clinical decision-making, particularly where this is likely to be of a life-or-death nature, for example withdrawal of life-sustaining therapy.⁵⁶

Conclusions

While the existing risk prediction models for acute TBI demonstrated acceptable discrimination for mortality at six months and less so for unfavorable outcome at six months, there appears to be considerable scope for further improvement. Recalibrating the models to the RAIN Study data resulted in substantially improved calibration and we recommend that these recalibrated models be used in a UK critical care setting. Additional areas that may be

explored to further improve the risk models include: consideration of mechanism of injury; incorporation of pupil size in addition to reactivity; addressing the severity and/or site(s) of major extracranial injury; incorporating additional data on the physiological response to injury; and consideration of genetic factors in the host response to injury.

Acknowledgments

This project was funded by the National Institute for Health Research Health Technology Assessment (NIHR HTA) Programme (project number 07/37/29) and has been published in full in Health Technology Assessment. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HTA programme, NIHR, NHS or the Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. PIAH was supported by the NIHR Cambridge Biomedical Research Centre.

The authors wish to thank all the participants for taking part in the study, GPs and health care professionals for their assistance in patient follow-up, the Scottish Intensive Care Society Audit Group for providing access to their database, the Trauma Audit and Research Network for providing data for verification of recruitment, and the Neurocritical Care Network for adopting and supporting the study.

RAIN Study Investigators

Dr. David Harrison (Chief Investigator); Dr. Richard Grieve; Professor Peter Hutchinson; Professor Fiona Lecky; Professor David Menon; Professor Kathy Rowan; Professor Martin Smith; Dr. Paddy Yeoman.

FIG. 4. Calibration plots—unfavorable outcome at six months (recalibrated models)

Observed unfavorable outcome is plotted against predicted unfavorable outcome in 10 equal sized groups. The gray dashed line is from a locally-weighted sum of squares smoother. The distribution of the predicted probability of unfavorable outcome is shown at the foot of each figure.

TABLE 8. COEFFICIENTS FOR THE ORIGINAL AND RECALIBRATED CRASH MODELS

Variable	Mortality at six months*		Unfavorable outcome at six months			
			Basic		CT	
	Recalibrated	CT	Original	Recalibrated	Original	Recalibrated
Constant	−3.041	−3.824	−3.062	−0.982	−3.214	−1.268
Age (per year)	0.063	0.062	0.073	0.043	0.070	0.042
GCS (per point decrease)	0.141	0.126	0.196	0.111	0.164	0.102
Pupil reactivity						
One reactive	0.199	0.002	0.942	0.308	0.767	0.185
Neither reactive	1.363	1.055	1.193	1.023	1.036	0.856
Major extracranial injury	0.009	0.057	0.485	0.084	0.433	0.137
One or more small petechial hemorrhages	-	0.068	-	-	0.179	0.047
Obliteration of the third ventricle	-	0.809	-	-	0.813	0.446
or basal cisterns						
Traumatic subarachnoid hemorrhage	-	0.293	-	-	0.490	0.161
Midline shift over 5 mm	-	0.466	-	-	0.751	0.269
Non-evacuated hematoma	-	0.635	-	-	0.500	0.149

*Note that original model coefficients were not reported for mortality at six months; original coefficients for unfavorable outcome at six months were those reported for high income countries.

CRASH, Corticosteroids After Significant Head Injury Trial; CT, computed tomography; GCS, Glasgow Coma Scale.

RAIN Study Management Group

Dr. Gita Prabhu (Study Coordinator); Andrew Craven (previous Research Administrator); Kathryn Griggs (Statistical Research Assistant); Dr. David Harrison (Chief Investigator); Dr. Sheila Harvey (CTU Manager/Senior Research Fellow); Phil Restarick (previous Research Coordinator); Professor Kathy Rowan (Co-investigator); Andrew Stenson (previous Research Manager); Emma Walmsley (Research Assistant).

RAIN Study Steering Group

Professor Monty Mythen (Independent chair); Julie Bridgewater (Independent); Dr. Richard Grieve; Dr. David Harrison; Professor Peter Hutchinson; Mr. Jonathan Hyam (Independent); Professor Fiona Lecky; Professor David Menon; Professor Kathy Rowan; Professor Martin Smith; Dr. Ian Tweedie (Independent); Dr. Paddy Yeoman.

RAIN Study Research Team

D. Menon, L. Moore, J. Outtrim (Addenbrooke's Hospital); H. Black, R. Jacob, C. Smalley (Arrowe Park Hospital); S. Chau, A. Bowry, C. Denniss, D. Raw, M. Reid (Barnsley District Hospital); M. Sun Wai, E. Bilton, H. Chitty (Basildon Hospital); J. Cupitt, S. Baddeley (Blackpool Victoria Hospital); R. Meacher, M. Templeton (Charing Cross Hospital); R. Wroth, A. Jarvis, J. Toms (Chesterfield Royal Hospital); J. Bleasdale, J. Edwards, A. McCann (City Hospital and Sandwell General Hospital); N. Robin, H. Jeffrey (Countess of Chester Hospital); V. Prasad, P. Wakefield (Darent Valley Hospital); E. Thomas, H. McMillan, T. Quintrell (Derriford Hospital); A. Asumang, M. Wain (Diana, Princess of Wales Hospital); A. Manara, S. Grier (Frenchay Hospital); A. Gratrix, N. Smith (Hull Royal Infirmary); I. Littlejohn, J. Powell, G. Spurling (Hurstwood Park Neurological Centre); H. Madder, J. Smith, J. Titchell (John Radcliffe Hospital); P. Hopkins, J. Dawson, D. Hadfield (King's College Hospital); D. Holden, M. Coggon, S. Fleming (King's Mill Hospital); J. Adams, Z. Beardow, S. Elliot (Leeds General Infirmary); N. Flint, K. Clarkson, A. Smith

(Leicester Royal Infirmary); D. Simpson, N. Pinto (Medway Maritime Hospital); H. Jones (Morrison Hospital); M. Smith, B. Boyd, H. Burgess, F. Chong (National Hospital for Neurology and Neurosurgery); S. Edwards, C. Hughes (Nevill Hall Hospital); S. Fletcher, M. Rosbergen (Norfolk and Norwich University Hospital); A. Walder, K. Jones (North Devon District Hospital); J. Cuesta, M. Dlamini (North Middlesex University Hospital); J. Wilkinson, R. Marsh, K. Brough (Northampton General Hospital); T. Gallacher, M. Lycett, T. Matthews, T. Thornhill (Queen Elizabeth Hospital Birmingham and Selly Oak Hospital); R. Jain, K. Reid (Queen's Hospital, Romford and Royal Berkshire Hospital); P. Yeoman, A. Jarvis, J. Litchfield, H. Wright (Queen's Medical Centre); S. Thornton (Royal Bolton Hospital); J. Paddle, K. Burt (Royal Cornwall Hospital); J. Andrzejewski, A. Pritchard, K. Bouakline (Royal Hallamshire Hospital); A. Vincent, C. Higham, P. Laws, H. McConnell (Royal Victoria Infirmary—previously Newcastle General Hospital); R. Protheroe (Salford Royal Hospital); R. Sharawi, A. Day (Scunthorpe General Hospital); R. Lightfoot, K. Linford, J. Mitchell, R. Oram, S. Tollerfield (Southampton General Hospital); D. Higgins, S. Andrews (Southend University Hospital); SP Young, L. Stewart (Southern General Hospital); P. Razis, V. Frazer (St George's Hospital); M. Watters, K. Challis, K. Mayell (The Great Western Hospital); R. Lewis, S. Bell, A. Kong (The Ipswich Hospital); J. Wright, K. Hugill (The James Cook University Hospital); K. Blenk, L. Everett (The James Paget Hospital); D. Dutta, C. McClements, J. Power (The Princess Alexandra Hospital); A. Guha, J. Nolan, A. Walker, K. Williams (The Royal Liverpool University Hospital); V. Verma, K. Maitland, G. Marshall (The Royal London Hospital); D. Watson, J. Bradshaw, K. Flahive, R. Jackson, G. Ward (University Hospital Coventry); M. Mostert, V. Todman (University Hospital Lewisham); V. Gupta (University Hospital of Hartlepool and University Hospital of North Tees); S. Krueper, R. Ahern, D. Cartledge (University Hospital of North Staffordshire); G. Scholey, S. Fernandez, N. Haskins, S. Shah, A. Williams (University Hospital of Wales); C. Whitehead, J. Cater, D. Davies, C. Owen, D. Watling (Walton Centre for Neurology and Neurosurgery); J. Cardy, S. Humphreys, C. Swanevelde (West Suffolk Hospital); P. Andrews,

TABLE 9. COEFFICIENTS FOR THE ORIGINAL AND RECALIBRATED IMPACT MODELS

Variable	Mortality at six months						Unfavorable outcome at six months					
	Core			Extended			Lab			Core		
	Orig.	Recal.		Orig.	Recal.		Orig.	Recal.		Orig.	Recal.	
Constant	-3.109	-3.861		-3.787	-4.671		-3.184	-3.834		-2.644	-1.493	
Age (per year)	0.034	0.044		0.032	0.044		0.020	0.042		0.038	0.030	
GCS motor score												
4	0.390	0.199		0.313	0.169		0.310	0.163		0.632	0.355	
3	0.797	0.545		0.746	0.531		0.778	0.489		1.266	0.496	
2	1.397	0.784		1.207	0.745		1.109	0.711		2.078	0.910	
1	1.447	0.983		1.205	0.885		0.965	0.866		1.393	0.744	
Pupil reactivity												
One reactive	0.514	0.188		0.334	0.019		0.090	-0.039		0.592	0.332	
Neither reactive	1.239	1.462		0.970	1.198		0.533	1.156		1.216	1.134	
Pre-hospital hypoxia	-	-		0.237	0.445		0.360	0.414		-	-	
Pre-hospital hypotension	-	-		0.667	0.669		0.366	0.557		-	-	
Marshall CT classification	-	-								-	-	
1 (Diffuse injury I)				-0.298	-0.349		-0.150	-0.324				
3/4 (Diffuse injury III/IV)				0.774	1.083		0.715	1.064				
5/6 (Evacuated mass lesion/non-evacuated mass lesion > 25 mL)				0.651	0.939		0.807	0.956				
Traumatic subarachnoid hemorrhage	-	-		0.606	0.283		0.740	0.291		-	-	
Extradural hematoma	-	-		-0.379	-0.216		-0.510	-0.224		-	-	
Laboratory results	-	-		-	-					-	-	
Glucose (per 1 mmol l ⁻¹)							0.097	0.040				
Hemoglobin (per 1 g dL ⁻¹)							-0.086	-0.081				

GCS, Glasgow Coma Scale; CT, computed tomography.

E. Grant (Western General Hospital); F. Keane, I. Bird (Whipps Cross University Hospital); C. Jones, S. Dowling, A. McCairn (Whiston Hospital); D. Southern, V. Cunningham, C. Hirst (Wrexham Maelor Hospital).

Author Disclosure Statement

DKM is a paid consultant or member of a Data Monitoring Committee for Solvay Ltd, GlaxoSmithKline Ltd, Brainscope Ltd., Ornim Medical, Shire Medical, and Neurovive Ltd. and his institution receives payment for a registered patent for a new positron emission tomography ligand assessing mitochondrial function. For the remaining authors, no competing financial interests exist.

References

- Maas, A.I., Stocchetti, N., and Bullock, R. (2008). Moderate and severe traumatic brain injury in adults. *Lancet Neurol.* 7, 728–741.
- Steyerberg, E.W., Moons, K.G.M., van der Windt, D.A., Hayden, J.A., Perel, P., Schroter, S., Riley, R.D., Hemingway, H., and Altman, D.G.; PROGRESS Group. (2013). Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med.* 10, e1001381.
- Knaus, W.A., Zimmerman, J.E., Wagner, D.P., Draper, E.A., and Lawrence, D.E. (1981). APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit. Care Med.* 9, 591–597.
- Hyam, J.A., Welch, C.A., Harrison, D.A., and Menon, D.K. (2006). Case mix, outcomes and comparison of risk prediction models for admissions to adult, general and specialist critical care units for head injury: a secondary analysis of the ICNARC Case Mix Programme Database. *Crit. Care* 10 Suppl 2, S2.
- Hayes, J.A., Black, N.A., Jenkinson, C., Young, J.D., Rowan, K.M., Daly, K., and Ridley, S. (2000). Outcome measures for adult critical care: a systematic review. *Health Technol. Assess.* 4, 1–111.
- 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.
- Menon, D.K. and Zahed, C. (2009). Prediction of outcome in severe traumatic brain injury. *Curr. Opin. Crit. Care* 15, 437–441.
- 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.
- 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.
- Hutchinson, P.J., Corteen, E., Czosnyka, M., Mendelow, A.D., Menon, D.K., Mitchell, P., Murray, G., Pickard, J.D., Rickels, E., Sahuquillo, J., Servadei, F., Teasdale, G.M., Timofeev, I., Unterberg, A., and Kirkpatrick, P.J. (2006). Decompressive craniectomy in traumatic brain injury: the randomized multicenter RESCUEicp study (www.RESCUEicp.com). *Acta Neurochir. Suppl* 96, 17–20.
- The CRASH Trial management group. (2001). The CRASH trial protocol (Corticosteroid randomisation after significant head injury) [ISRCTN74459797]. *BMC Emerg. Med.* 1, 1.
- Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys. J Wiley & Sons: New York.
- Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G., Wood, A.M., and Carpenter, J.R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 338, b2393.
- van Buuren, S. and Groothuis-Oudshoorn, K. (2011). MICE: Multi-variate Imputation by Chained Equations in R. *Journal of Statistical Software* 45, 1–67.
- Li, K.H., Meng, X.L., Raghunathan, T.E., and Rubin, D.B. (1991). Significance levels from repeated p-values with multiply-imputed data. *Statistica Sinica* 1, 65–92.
- Harrell, F.E., Jr., Califf, R.M., Pryor, D.B., Lee, K.L., and Rosati, R.A. (1982). Evaluating the yield of medical tests. *JAMA* 247, 2543–2546.
- Hanley, J.A. and McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36.
- Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74, 829–836.
- Cox, D.R. (1958). Two further applications of a model for binary regression. *Biometrika* 45, 562–565.
- Brier, G.W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Rev.* 75, 1–3.
- Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall: New York.
- Catroppa, C., Anderson, V.A., Morse, S.A., Haritou, F., and Rosenfeld, J.V. (2008). Outcome and predictors of functional recovery 5 years following pediatric traumatic brain injury (TBI). *J. Pediatr. Psychol.* 33, 707–718.
- Colantonio, A., Escobar, M.D., Chipman, M., McLellan, B., Austin, P.C., Mirabella, G., and Ratcliff, G. (2008). Predictors of postacute mortality following traumatic brain injury in a seriously injured population. *J. Trauma* 64, 876–882.
- Corrigan, J.D., Selassie, A.W., Lineberry, L.A., Millis, S.R., Wood, K.D., Pickelsimer, E.E., and Rosenthal, M. (2007). Comparison of the Traumatic Brain Injury (TBI) Model Systems national dataset to a population-based cohort of TBI hospitalizations. *Arch. Phys. Med. Rehabil.* 88, 418–426.
- Dawson, D.R., Schwartz, M.L., Winocur, G., and Stuss, D.T. (2007). Return to productivity following traumatic brain injury: cognitive, psychological, physical, spiritual, and environmental correlates. *Disabil. Rehabil.* 29, 301–313.
- Devitt, R., Colantonio, A., Dawson, D., Teare, G., Ratcliff, G., and Chase, S. (2006). Prediction of long-term occupational performance outcomes for adults after moderate to severe traumatic brain injury. *Disabil. Rehabil.* 28, 547–559.
- Ducrocq, S.C., Meyer, P.G., Orliaguet, G.A., Blanot, S., Laurent-Vannier, A., Renier, D., and Carli, P.A. (2006). Epidemiology and early predictive factors of mortality and outcome in children with traumatic severe brain injury: experience of a French pediatric trauma center. *Pediatr. Crit. Care Med.* 7, 461–467.
- Feiz-Erfan, I., Horn, E.M., Theodore, N., Zabramski, J.M., Klopstein, J.D., Lekovic, G.P., Albuquerque, F.C., Partovi, S., Goslar, P.W., and Petersen, S.R. (2007). Incidence and pattern of direct blunt neurovascular injury associated with trauma to the skull base. *J. Neurosurg.* 107, 364–369.
- Hesdorffer, D.C. and Ghajar, J. (2007). Marked improvement in adherence to traumatic brain injury guidelines in United States trauma centers. *J. Trauma* 63, 841–847.
- Hsiao, K.Y., Hsiao, C.T., Weng, H.H., Chen, K.H., Lin, L.J., and Huang, Y.M. (2008). Factors predicting mortality in victims of blunt trauma brain injury in emergency department settings. *Emerg. Med. J.* 25, 670–673.
- Jagannathan, J., Okonkwo, D.O., Yeoh, H.K., Dumont, A.S., Saulle, D., Haizlip, J., Barth, J.T., Jane, J.A. Sr., and Jane, J.A. Jr. (2008). Long-term outcomes and prognostic factors in pediatric patients with severe traumatic brain injury and elevated intracranial pressure. *J. Neurosurg. Pediatr.* 2, 240–249.
- Johnson, P., Thomas-Stonell, N., Rumney, P., and Oddson, B. (2006). Long-term outcomes of pediatric acquired brain injury. *Brain Cogn.* 60, 205–206.
- Liao, C.C., Hsu, Y.P., Chen, C.T., and Tseng, Y.Y. (2007). Transarterial embolization for intractable oronasal hemorrhage associated with craniofacial trauma: evaluation of prognostic factors. *J. Trauma* 63, 827–830.
- Mauritz, W., Janciak, I., Wilbacher, I., and Rusnak, M. (2007). Severe traumatic brain injury in Austria IV: intensive care management. *Wien Klin. Wochenschr.* 119, 46–55.
- MRC CRASH Trial Collaborators. (2008). Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 336, 425–429.
- Oh, H.S., Seo, W.S., Lee, S., and Song, H. (2006). Comparisons of the prognostic predictors of traumatic brain injury according to admission Glasgow Coma Scale scores-based on 1- and 6-month assessments. *Taehan Kanho Hakhoe Chi* 36, 621–629.
- Pang, B.C., Kuralmani, V., Joshi, R., Hongli, Y., Lee, K.K., Ang, B.T., Li, J., Leong, T.Y., and Ng, I. (2007). Hybrid outcome prediction model for severe traumatic brain injury. *J. Neurotrauma* 24, 136–146.
- Pape, T.L., Lundgren, S., Heinemann, A.W., Guernon, A., Giobbie-Hurder, A., Wang, J., Roth, H., Blahnik, M., and Williams, V. (2006). Establishing a prognosis for functional outcome during coma recovery. *Brain Inj.* 20, 743–758.

39. Salorio, C.F., Slomine, B.S., Guerguerian, A.M., Christensen, J.R., White, J.R., Natale, J.E., Shaffner, D.H., Grados, M.A., Vasa, R.A., and Gerring, J.P. (2008). Intensive care unit variables and outcome after pediatric traumatic brain injury: a retrospective study of survivors. *Pediatr. Crit. Care Med.* 9, 47–53.
40. Steyerberg, E.W., Mushkudiani, N., Perel, P., Butcher, I., Lu, J., McHugh, G.S., Murray, G.D., Marmarou, A., Roberts, I., Habbema, J.D., and Maas, A.I. (2008). Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med.* 5, e165.
41. Stulemeijer, M., van der, W.S., Borm, G.F., and Vos, P.E. (2008). Early prediction of favourable recovery 6 months after mild traumatic brain injury. *J. Neurol. Neurosurg. Psychiatry* 79, 936–942.
42. Thompson, H.J., Rivara, F.P., Jurkovich, G.J., Wang, J., Nathens, A.B., and MacKenzie EJ. (2008). Evaluation of the effect of intensity of care on mortality after traumatic brain injury. *Crit. Care Med.* 36, 282–290.
43. Xu, X.Y., Liu, W.G., Yang, X.F., and Li, L.Q. (2007). Evaluation of models that predict short-term outcome after traumatic brain injury. *Brain Inj.* 21, 575–582.
44. Zheng, W.B., Liu, G.R., Li, L.P., and Wu, R.H. (2007). Prediction of recovery from a post-traumatic coma state by diffusion-weighted imaging (DWI) in patients with diffuse axonal injury. *Neuroradiology* 49, 271–279.
45. Hollingworth, W., Vavilala, M.S., Jarvik, J.G., Chaudhry, S., Johnston, B.D., Layman, S., Tontisirin, N., Muangman, S.L., and Wang, M.C. (2007). The use of repeated head computed tomography in pediatric blunt head trauma: factors predicting new and worsening brain injury. *Pediatr. Crit. Care Med.* 8, 348–356.
46. Nampiaparampil, D.E. (2008). Prevalence of chronic pain after traumatic brain injury: a systematic review. *JAMA* 300, 711–719.
47. Ricker, G.M., Cook, D.J., and Shemie, S.D. (2006). Brief review: Practice variation in end of life care in the ICU: implications for patients with severe brain injury. *Can. J. Anaesth.* 53, 814–819.
48. Saatman, K.E., Duhaime, A.C., Bullock, R., Maas, A.I., Valadka, A., and Manley, G.T. (2008). Classification of traumatic brain injury for targeted therapies. *J. Neurotrauma* 25, 719–738.
49. Tagliaferri, F., Compagnone, C., Korsic, M., Servadei, F., and Kraus, J. (2006). A systematic review of brain injury epidemiology in Europe. *Acta Neurochir. (Wien)* 148, 255–268.
50. Willemse-van Son, A.H., Ribbers, G.M., Verhagen, A.P., and Stam, H.J. (2007). Prognostic factors of long-term functioning and productivity after traumatic brain injury: a systematic review of prospective cohort studies. *Clin. Rehabil.* 21, 1024–1037.
51. 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. *J. Trauma* 71, 387–392.
52. 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. *J. Neurotrauma* 29, 47–52.
53. 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. *Crit. Care. Med.* 40, 1609–1617.
54. Lingsma H, Andriessen TM, Haitsema I, Horn J, van der Naalt J, Franschman G, Maas, A.I., 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. *J. Trauma Acute Care. Surg.* 74, 639–646.
55. Spiegelhalter, D., Pearson, M., and Short, I. (2011). Visualizing uncertainty about the future. *Science* 333, 1393–1400.
56. Skrobik, Y. and Kavanagh, B.P. (2006). Scoring systems for the critically ill: use, misuse and abuse. *Can. J. Anaesth.* 53, 432–436.

Address correspondence to:

David A Harrison, PhD

Clinical Trials Unit

Intensive Care National Audit and Research Centre

Napier House

24 High Holborn

London WC1V 6AZ, United Kingdom

E-mail: david.harrison@icnarc.org