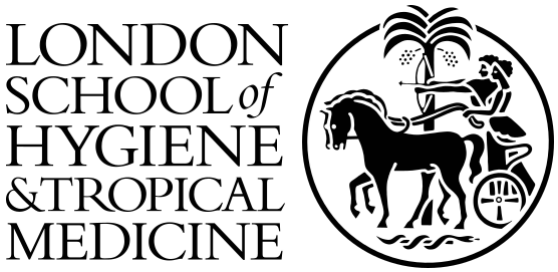


Prognosis in Traumatic Brain Injury

Pablo Andrés Perel

Thesis submitted to the University of London for the degree of Doctor of Philosophy

London School of Hygiene & Tropical Medicine

Submitted by June 2009

London School of Hygiene & Tropical Medicine
University of London

AUTHOR'S DECLARATION

I wrote the entire thesis presented hereafter. Some of the work conducted and publications written involved collaborative efforts, as mentioned below.

The study presented in Chapter 2 was published in joint names in 2007 (Perel P, Wasserberg J, Ravi RR, Shakur H, Edwards P, Roberts I. Prognosis following head injury: a survey of doctors from developing and developed countries. J Eval Clin Pract. 2007 Jun;13(3):464-5.) I designed the study with Ian Roberts and Phil Edwards. I designed and formatted the questionnaire with Phil Edwards, Haleema Shakur, John Wasserberg and R. Ravi. I conducted the analysis and drafted the manuscript

The study presented in Chapter 3 was published in joint names in 2006 (Perel P, Edwards P, Wentz R, Roberts I. Systematic review of prognostic models in traumatic brain injury BMC Med Inform Decis Mak. 2006 Nov 14;6:38.) I designed the study with Ian Roberts and Phil Edwards. I developed the search strategy with Reinhardt Wentz. I screened the records from the electronic searches with Phil Edwards and Ian Roberts. I reviewed the reports of potentially eligible studies and extracted the data. I analyzed the data and drafted the manuscript.

The study presented in Chapter 4 was published in joint names in 2008 (MRC CRASH Trial Collaborators, Perel P, Arango M, Clayton T, Edwards P, Komolafe E, Pocock S, Roberts I, Shakur H, Steyerberg E, Yutthakasemsunt S. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. BMJ. 2008 Feb 23;336(7641):425-9) I designed the study with Ian Roberts and Phil Edwards. I analyzed the data with collaboration from Stuart Pocock, Tim Clayton and Ewout Steyerberg. I designed the web page calculator with Ian Roberts which was developed by Tony Brady. I drafted the manuscript.

For the work presented in Chapter 5, I designed the study for the focus groups interviews with Ian Roberts. I conducted one of the focus group and Ian Roberts conducted the other focus group. For the survey conducted among collaborators I designed the study with Ian Roberts and Phil Edwards. I designed the questionnaire with Phil Edwards and Ian Roberts. with the help of Haleema Shakur, Lin Barnetson, Taemi Kawahara, Lisa Cook and Eni Balogun. I developed the idea for the CRASH Score Card and Maria Ramos designed it. I conducted the survey to collaborators with Taemi Kawahara, Lisa Cook and Eni Balogun. I conducted the analysis and drafted the manuscript.

The study presented in Chapter 6 was published in joint names in 2008 (Perel P, Edwards P, Shakur H, Roberts I. Use of the Oxford Handicap Scale at hospital

discharge to predict Glasgow Outcome Scale at 6 months in patients with traumatic brain injury BMC Med Res Methodol. 2008 Nov 6;8:72) I designed the study with Ian Roberts, Phil Edwards and Haleema Shakur. I analyzed the data and drafted the manuscript.

The study presented in Appendix 6 was submitted for publication to BMC Emergency Medicine in March 2009 and is currently under review. I designed the study with Ian Roberts and Fiona Lecky. Omar Bouamra conducted the analysis under my guidance. I drafted the manuscript.

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

[redacted]

Pablo Andres Perel

8/7/09

Date

Abstract

Introduction

The general purpose of this thesis was to study prognosis in traumatic brain injury (TBI) patients, with the aim of providing useful and practical information in clinical practice and clinical research. The specific objectives were: to develop and validate practical prognostic models for TBI patients and to assess the validity of the Modified Oxford Handicap Scale (mOHS) for predicting disability at six months.

Methods

A survey was first conducted to understand the importance of prognostic information among physicians. A systematic review of prognostic models for TBI patients was then carried out. Prognostic models were developed using data from a cohort of 10,008 TBI patients (CRASH trial) and validated in a cohort of 8,509 TBI patients (IMPACT study). Two focus groups and a survey were conducted to develop a paper-based prognostic score card. The correlation between the mOHS and the Glasgow Outcome Scale (GOS) was assessed, the validity of different mOHS dichotomies was assessed, and the discriminative ability of the mOHS to predict GOS was evaluated.

Results

Doctors considered prognostic information to be very important in the clinical management of TBI patients, and believed that an accurate prognostic model would change their current clinical practice. Many prognostic models for TBI have been published, but they have many methodological flaws which limit their validity. Valid prognostic models for patients from high income countries and low & middle income countries were developed and made available as a web calculator, and as a paper based score card. The mOHS was strongly correlated with and was predictive of GOS at six months.

Conclusion

The prognostic models developed are valid and practical to use in the clinical setting. The association between mOHS and GOS suggest that the mOHS could be used for interim analysis in randomised clinical trials in TBI patients, for dealing with loss to follow-up, or could be used as simple tool to inform patients and relatives about their prognosis at hospital discharge.

Acknowledgements

I am very grateful with Ian Roberts, my supervisor, for his guidance, support, and patience throughout the work for my thesis.

I am also very thankful to Phil Edwards, my co-supervisor, for his advice and constant assistance.

I would like to acknowledge the help and collaboration of all the members of the CRASH Trials Coordinating Centre, Haleema Shakur, Lin Barnetson, Eni Balogun, Lisa Cook, Taemi Kawahara, Collette Barrow, Mathew Berle, and with special to Maria Ramos.

I also wish like to make a special acknowledgment to all the CRASH trial Collaborators who participated in the survey reported in Chapter 2, provided the data for the prognostic models reported in Chapter 4, and participated in the studies reported in Chapter 5.

I am very grateful to the following LSHTM staff for advice and help throughout my thesis, Tim Clayton, Stuart Pocock, David Prieto-Merino, Val McCormack, Rhian Daniel, Mathilde Savy, Katharine Ker, Sophie Hawkesworth, Karen Blackhall, Emma Sydenham, and Branwen Hennig.

I would like to thank the IMPACT investigators, in particular Ewout Steyerberg, for making the database available for the validation of the prognostic models reported in Chapter 4.

I am also grateful to the TARN investigators, especially Fiona Lecky, for providing the necessary data for the study reported in appendix 6.

I would like to express thanks to my friends at LSHTM, Juan Pablo Casas and Jaime Miranda for their constant support.

Finally I would like to thank, my friend, Zulma Ortiz, from Argentina, who introduced me to the fascinating world of epidemiology.

Dedication

To my parents, Pommy and Jorge

Abbreviations

(alphabetical order)

AIS: Abbreviated Injury Scale

CRASH: Corticosteroids Randomisation After Significant Head Injury

CT: Computed Tomography

TCDB: Traumatic Coma Data Bank

DVT: Deep Venous Thrombosis

DMC: Data Monitoring Committee

EPH: Epidural Haemorrhage

GCS: Glasgow Coma Scale

GOS: Glasgow Outcome Scale

HIC: High Income Countries

IB: Intracranial Bleeding

ICP: Intracranial Pressure

IMPACT: International Mission for Prognosis and Analysis of Clinical Trials in TBI

IPH: Intraparenchymal Haemorrhage

ISS: Injury Severity Score

LMIC: Low & Middle Income Countries

mOHS: Modified Oxford Handicap Scale

RCT: Randomised Controlled Trial

ROC: Receiving Operator Characteristic

SDH: Subdural Haemorrhage

SAH: Subarachnoid Haemorrhage

TARN: Trauma & Audit Research Network

NFS: No Further Specified

TBI: Traumatic Brain Injury

TXA: Tranexamic Acid

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Traumatic brain injury

1.1.1 **Definition**

Traumatic brain injury (TBI) is caused by a blow or jolt to the head or a penetrating head injury that disrupts the normal function of the brain.[1] TBI is also called traumatic head injury but strictly the latter could be any traumatic injury in the head, face or skull and may or may not be associated with a TBI.

1.1.2 **Pathophysiology**

Primary injury occurs immediately after the impact; this is followed by a secondary injury which, some authors believe, is responsible for part of the neurological damage observed in TBI.[2] Oedema formation, both vasogenic and cytotoxic, in a rigid structure such as the skull, is followed by increased intracranial pressure (ICP) once the compensatory mechanisms are surpassed. The increased ICP reduces cerebral perfusion and the consequent ischemia generates further neurological damage.[3]

1.1.3 **Classification**

TBI can be classified by severity, mechanism or structural damage.[4]

1.1.3.1 Severity

TBI is a heterogeneous condition encompassing a wide range of manifestations, from minor symptoms to profound coma. The Glasgow Coma Scale (GCS) is the standard method used to evaluate level of consciousness and comprises three different domains (motor, verbal and ocular).[5] Adding the different components, the GCS can be rated from 3 to 15 points (table 1.1). Patients who suffer a TBI and have a GCS $\leq$ 8 are classified as having severe TBI, those with GCS between 9 and 12 are considered moderate cases, and those patients with GCS $\geq$ 13 are considered mild. The GCS was developed as a practical scale for assessing the depth and duration of impaired consciousness and enhanced communication among physicians in reporting of neurologic status after TBI. Gradually it has become accepted worldwide and has been shown to be a good predictive tool for determining outcome in TBI patients. It is not clear how classification of the GCS for mild, moderate and severe TBI was established.[6]

Table 1-1 Glasgow Coma Scale

Eye opening	
Spontaneous	4
To sound	3
To pain	2
None	1
Total eye score	1 to 4
Motor response	
Obeys commands	6
Localising	5
Normal flexion	4
Abnormal flexion	3
Extending	2
None	1
Total motor score	1 to 6
Verbal response	
Orientated	5
Confused speech	4
Words	3
Sounds	2
None	1
Total verbal score	1 to 5
Total GCS score	**3 to 15**

1.1.3.2 Mechanism

TBI can be classified as closed or penetrating. While closed TBI is related with road traffic crashes, penetrating TBI is more commonly related with gunshot wounds.[7] Some studies have shown that penetrating TBI is associated with a worse outcome. For example Peek-Asa and collaborators studied 795 patients with moderate or severe TBI. After adjusting for GCS, age, gender, and presence of multiple trauma, patients with penetrating injuries had an odds ratio of 6.6 (95% CI = 3.9-11.1) for mortality in comparison with patients with closed injuries.[8]

1.1.3.3 Structural damage

The structural damage can be assessed by neuroimaging (Computed Tomography). It can be divided into diffuse axonal injury, focal contusions and intracranial bleeding. The latter can also be classified according to the location within the cranium. Epidural haematoma refers to the collection of blood between the skull and the outer layer of the meninges (the dura mater) and is usually caused by a tear in the middle meningeal artery. Subdural haematoma refers to collection of blood between the dura mater and pia layers and has a typical crescent-shaped appearance in imaging. Subarachnoid haemorrhage is the accumulation of blood in the subarachnoid space and intraparenchymal haemorrhages occurs within brain parenchyma.

1.1.4 **Outcomes**

In-hospital mortality and disability at six months are the most frequently used outcome measures considered in TBI research. According to the International Classification of Functioning, Disability and Health (ICF), disability can be defined under a biopsychosocial model that includes biological, individual and social perspective on health.[9] Disability includes: a) impairments (problems in body functions), b) activity limitations and c) participation restrictions. Different scales have been used for assessing disability in TBI.[10] The Glasgow Outcome Scale (GOS) devised more than 30 years ago is the most widely used in TBI research.[11,12] This scale reflects disability in terms of activity limitations and participation restriction, rather than impairment.[13] It encompasses five categories: death, vegetative state, severe disability, moderate disability, and good recovery. In randomised clinical trials, it is generally dichotomized into favourable (moderate disability, good recovery) and unfavourable outcome (severe disability, vegetative state, death). GOS assesses how well patients function in their daily social interactions and is generally completed six months after hospital discharge. The fact that one of the main outcomes in clinical trials of TBI patients is measured at six months introduces a potential source of bias, as loss to follow up is a common problem in

this population. A 1998 survey of TBI trials found an average loss to follow up of 19% at six months.[14] Loss to follow up reduces study size, causing effect estimates to be less precise, and may introduce bias. Loss to follow up appears to be particularly common in TBI trials, possibly because they mostly involve young males from disadvantaged social groups who are highly geographically mobile.[15] At the moment there is no disability outcome measured at hospital discharge that predicts long term disability. Such a measure could be potentially useful when dealing with loss to follow-up, for informing interim analysis, and potentially could also be useful in clinical practice when communicating with patients and relatives.

The Disability Rating Scale (DRS) is another measurement of disability which has been used for TBI patients.[16] Neuropsychological tests and generic quality of life instruments, such as SF-36, EuroQol and the WHO-Qol, have also been recommended for assessing outcome in TBI patients.[17]

1.1.5 **Public health importance**

1.1.5.1 Global injuries

Injuries account approximately for 12% of the world's burden of disease in 2000, and it is expected that by the year 2020 road traffic crashes will become the third cause of disability worldwide.[18] More than 90% of the world's deaths from injuries occur in low and middle-income countries.[19]

1.1.5.2 Traumatic Brain injury

TBI is one of the main causes of death in patients with injuries and is a leading cause of death and disability in young people.[2] Every year approximately 1.5 million people die and at least 10 million are killed or hospitalized because of a TBI.[20,21]

The incidence, fatality and disability rates of TBI vary by region and higher rates are observed in low and middle income countries.[22] For example, according to the Global Burden of Disease Study, the Latin America region has the highest incidence of TBI related with road traffic crashes and violence at 163 per 100,000 and 67 per 100,000 in comparison with the world average of 106 per 100,000 and 42 per 100,000 respectively.[23]

Some general trends are universal, a tri-modal age specific incidence has been described with peaks in early childhood, late adolescence/early adulthood and in the elderly. Men are about twice as likely as females to experience a TBI.[20]

Falls are the main cause in children and elderly populations, while violence and road traffic crashes predominate in adolescents and young adults.[24]

Most of the patients admitted to hospitals present with mild TBI while 20% are classified as moderate or severe. In severe cases about one third of patients die in hospital.[25]

1.2 Prognosis

1.2.1 **Definition**

Prognosis, (from the Greek *"pro"* meaning *before and "gnosis "*meaning knowledge) is defined as "the result of looking forward".[26] In the context of clinical epidemiology prognosis can be defined as *"the probable course and outcome of a health condition over time"* or as *"the future risk of adverse outcomes among people with existing disease".*[27,28]

1.2.2 **Importance of prognosis in clinical practice**

Clinical practice involves three main activities: identifying diseases (diagnosis), treating diseases (therapy) and predicting diseases course and outcome (prognosis). Although the three activities are interrelated, distinctions between them are made in clinical research. Prognostic related research is considered to be the most neglected one.[28,29]

Prognosis was historically one of the most important activities of medical practice. Until the end of the nineteenth century, 10% of the content of medical textbooks was dedicated to prognosis however, by 1970 this had decreased to almost zero.[30] Predicting the future was what both priests and doctors were supposed to do for many centuries but the appearance of effective therapies has shifted the dominance of the clinical encounter to diagnosis and therapy.[31] However, in most recent years there has been an increasing interest in prognosis research.[28] Among the reasons for this resurgence, Christakis proposed:[32]

1. Interest in human terminal care and the decision of withdrawing or not life support from critical patients
2. Avoidance of futile treatment for reasons of justice or costs
3. Availability of new "technologies" (e.g. genetic tests, and biomarkers)
4. Increasing emphasis on patient autonomy
5. Increasing prevalence of chronic diseases

1.2.3 **Classification of prognosis research**

Prognostic studies can be classified into 2 according to their objective; explanatory studies or outcome prediction studies.[27]

Explanatory studies focus on the casual association between predictors and outcome. Some authors propose a further division into three stages: phase 1, identifying associations; phase 2, testing independent associations; and phase 3, understanding prognostic pathways.[27]

Outcome prediction studies, also known as prognostic models, combine different variables to obtain a probability of the outcome. According to the use of the estimated probability, these studies can be further divided into studies which are used :[33,34]

1. To inform doctors to make decisions for individual patients
2. To inform patients and relatives
3. For research purposes (for example, in the selection of patients, adjustment for baseline imbalances, or risk stratification in clinical trials)
4. To compare health services by allowing adjustment for case mix

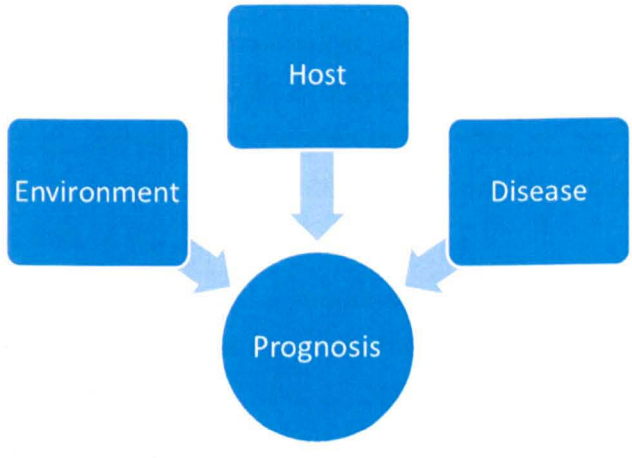
Variables influencing prognosis (predictors) can be classified into the following three categories according to their characteristics:[28]

a) Environment (e.g. country, social class, hospital care)

b) Host (e.g. age, comorbidities)

c) Disease (e.g. genes, severity)

Figure 1-1 Classification of predictors



1.2.4 **Prognostic models**

1.2.4.1 Definition

Outcome prediction studies have received different names such as prognostic models, prediction models, risk scores, prognostic indices, clinical prediction rules, clinical prediction guides, or clinical decision rules.[35 36 37 38 39]

According to some authors the term "clinical decision rules" only applies to those models that also provide a diagnostic or therapeutic recommendation.[40]

Throughout this thesis I will use the term "prognostic model" defined as the "*mathematical combination of two or more patient or disease characteristics to predict outcome*", although I acknowledge that a model could also be developed using only one variable. [41]

1.2.4.2 Performance of prognostic models

The performance of prognostic models refers to how "accurate" the model's predictions are in relation to observed data.[33] According to Rothman, "*accuracy in estimation implies that the parameter that is the object of measurement is estimated with little error*".[42] In the particular context of prognostic models, accuracy has two main components: calibration and discrimination.[43]

Calibration refers to the agreement between predicted and observed probabilities.[44] For example if, according to the model, TBI patients with certain characteristics have a probability of mortality of 30%, it would be expected that 30 out of 100 patients with those characteristics would die if the model was perfectly calibrated. Calibration can be measured in different ways; graphically by plotting observed against predicted outcomes, or through a statistical test such as the Hosmer-Lemeshow test. This test compares the observed number of people with events within risk groupings (e.g. deciles of risk) with the number predicted by the model. A small p value implies lack of fit.

Discrimination is a measure of how well a model separates those who develop the outcome from those who do not.[44] It is generally measured through the area under the receiver operator curve (ROC) or the C statistic. A ROC is constructed by plotting pairs of true positive rate (sensitivity) and false positive rate (1-specificity) for several cut-off values of probability of the outcome. The area under the ROC can be interpreted as the probability that a randomly selected person with the outcome, will have a higher predicted probability than a randomly selected person without the outcome.[45] For example, if a model has an area under the ROC (or C statistic) of 0.7, this means that the model will estimate a higher probability of the

outcome for subjects with the outcome 70 out of 100 times if we choose a random pair of subjects with and without the outcome.

The relative importance of calibration and discrimination will depend on the intended application of the prognostic model.[43] For example, for counselling an individual patient calibration of the model will be more relevant, while for triage in a setting with limited resources discrimination could be more important.

In addition to the measures of discrimination and calibration we might be interested in performance measures for specific thresholds when a clinically relevant cut-off is already established. The accuracy rate (or correct classification rate) is calculated as: (*true positive +true negative)/total* and, the complement that is the error rate (misclassification rate) that is defined as (*false positive + false negative)/total*. The problem with these measures is that equal weight is given to positive and negative results whereas, in general, false negatives are more important than false positives. Furthermore the accuracy rate will be high, by definition, for a frequent or infrequent outcome. For example, if the average mortality for a condition is 7% the accuracy rate would be 93% if the model classifies all the patients as survivors.[44]

More recently new measures have been proposed, such as the net reclassification improvement (NRI). The NRI has four components: proportion of individuals with events who move up or down a category and the proportion of individuals without events who move up or down a category. The NRI is obtained by combining the four components, but they should also be reported separately.[46]

Finally there are also overall performance measures such as the $R^{2,}$, which is the amount of explained variation on the outcome explained by the model, and the Brier score which is a measure of the difference between actual outcomes and prediction. These measures do not distinguish among the different performance components, calibration and discrimination, so they are not very useful.[44]

1.2.4.3 Inaccuracy of clinical prediction

The lack of interest in prognosis has led to a weak medical training in this area and so it is not surprising that doctors feel poorly prepared and that they often disagree or are inaccurate in their predictions.[32 47]

There are numerous studies showing that physicians make errors when formulating a prognosis. In many of these studies the term accuracy is used in the more general epidemiological sense (measured with little error), and they did not necessarily use the standard specific measures of accuracy described above for evaluating prognostic models.

A systematic review compared physicians' clinical predictions of survival in terminally ill cancer patients with actual survival.[48] The authors found eight studies

(including 1563 individuals) and reported that the median clinical prediction of survival was 42 days and the actual median survival was 29, overall there was poor agreement (weighted kappa 0.36) between clinical prediction and actual survival.

A cohort study was conducted involving 16 Dutch nursing homes including 515 terminally ill non cancer patients. The authors compared physicians' predictions with actual survival. Physicians were asked to predict death in the following periods: one week (0 to 7 days), 8 to 21 days, or between 22 to 42 days. The positive predictive value of physicians' predictions was high for those patients expected to die within one week (92%), but much lower for patients who were expected to die within 8 to 21 days (16%), or within 22 to 42 days (13%).[49]

In other areas, such as cardiovascular disease, similar results have been reported. For example, Pignone and collaborators developed 12 primary prevention scenarios with a five year risk of cardiovascular heart disease events, and conducted a survey among 79 physicians to compare their predictions with values calculated from Framingham risk equations. For the analysis the authors divided the estimated risk by the Framingham estimated risk and considered results between 0.67 to 1.5 to be "accurate". They reported that only 24% of their predictions were accurate.[50] The main limitation of this study was with the use of hypothetical cases, thus the predictions could not be compared with actual survival.

In a cohort study that included 850 patients admitted for intensive care, physicians' prediction was compared with actual survival at hospital discharge and approximately 70% of the patients that were estimated to have a 30% chance of survival actually survived. But unlike for cancer patients, doctors' predictions were in general pessimistic rather than optimistic.[51]

1.2.4.4 Clinical prediction versus prognostic models

According to studies in cognitive psychology the human brain is poorly prepared for making and updating precise quantitative prediction.[52] Psychologists have been studying the question of clinical versus statistical prediction for more than 50 years.[53] Since then the results have generally shown that prognostic models are as accurate as, or more accurate than, clinical judgment.

Grove and collaborators conducted a systematic review of studies that compare statistical versus clinical prediction. Studies from the area of psychology and medicine which predicted outcomes such as human behaviour, disease diagnosis, or a disease prognosis were included.[54] They used a 25 page manual to code each study for publication variables and study design characteristics. Investigators were trained and two coders extracted the data with very high reliability (r=.97). A total of 136 studies were included. The authors used the term accuracy referring to the

error in the estimation of each of the methods in comparison with a gold standard. Different measures were reported in the studies so the authors standardized the different measures in a common metric (effect size-ES-). For this they first found a suitable transformation for each measure with a known variance and an approximate normal distribution, then they estimated the difference between the clinical and statistical prediction. Positive ES indicates superiority of statistical prediction. To conduct the meta-analysis they gave a weight to each ES that was inversely proportional to the variance. The weighted summary statistic for the ES was 0.086. This indicates that on average statistical prediction was approximately 10% more "accurate" than clinical prediction. Because there was evidence of statistical heterogeneity ($Qt=1635.2$ $p<0.0001$) the authors also reported the results using a different method. For this they considered that ES <-0.1 as substantially favouring clinical prediction, ES between -0.1 and 0.1 as being relatively equal, and those >0.1 as substantially favouring statistical prediction. With these criteria in 46% of the studies the statistical prediction was more accurate than the clinical prediction, in 48% a similar result was obtained with both methods, and in only 6% of the studies clinical prediction was superior. The authors used meta-regression to evaluate the effect in certain subgroups, such as year of publication, study design or type of setting (general medicine, mental health, education, etc.) and concluded that they did not find any exception to the general equivalence or superiority of statistical prediction. However, it is not clear from the report whether the study had enough power to evaluate the effect in these different subgroups. Another limitation of this study was that the authors did not evaluate or discuss the possibility of reporting bias.

1.2.5 **Evaluation of prognostic models**

There are two main levels to evaluate prognostic models. First we want to know if the model performance, in terms of discrimination and calibration, works satisfactorily for patients other than those from whom the data were derived. This is called "validation" of the model. The other level refers to the evaluation of the model in terms of change in behaviour of medical doctors (medical management) or changes in patient outcome. Some authors refer to this as the "impact" of the model.

Several guidelines have been proposed for the development and evaluation of prognostic models. The most recently was proposed by Reilly and collaborators, who defined five stages:

1) Derivation of the prognostic model: Identification of the predictors for multivariable model

2) Narrow validation: Assessment of the accuracy of the prognostic model in one setting

3) Broad validation: Assessment of the accuracy of the prognostic model in varied settings

4) Narrow impact analysis of prognostic model used as decision rule: Prospective demonstration that the prognostic model improves physicians' decisions in one setting

5) Broad impact analysis of the prognostic model used as decision rule: Prospective demonstration that the prognostic model improves physicians' decisions in varied settings.

According to Reilly and collaborators the two last stages (impact analysis) should be only applied to clinical decision rules (those prognostic models that recommend a diagnostic or therapeutic action according to the estimated probability), and they also consider that randomised controlled trials are the ideal study design for these two stages.[40]

During the rest of my thesis I refer, unless specified otherwise, to prognostic models that do not provide a specific course of action. Some consider that these models should also be evaluated through randomised controlled trials, while for others their evaluation could be restricted to the validation stages.[37]

To the best of my knowledge, the only randomised clinical trial evaluating the use of a prognostic model (that does not provide a course of action) was the SUPPORT study (Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments). This study enrolled 8,329 adult seriously ill patients with a 50% chance of death within six months.[55]In a first phase including 4,301 patients a prognostic model to estimate 180 days mortality was developed and, in a second phase including 4,028 patients, the investigators randomly allocated half of the physicians to receive the prognostic model estimates and patient's preferences for end life care. In this study physician's and model's discrimination were identical (area under the receiving operator curve 0.78) but physicians' predictions were worse calibrated in comparison with the prognostic model. The best discrimination was obtained when combining both physicians and the prognostic model estimates (area under the receiver operator curve 0.82). The study did not find a difference in physician's performance nor in patients' outcomes.

However, other studies have found different results. Murray and collaborators studied 1025 patients with severe TBI, with the objective of evaluating whether providing doctors with computer-based predictions influenced patient management.[56] According to their previous hypothesis there was a decrease of

39% in the use of intensive management in patients with the worst prognosis, including osmotic diuretics, ventilation and intracranial pressure monitoring. Among the limitations of this study it should be mentioned that it was a before/after design.

The results of the SUPPORT study were unexpected and discouraging for those advocating the use of prognostic models. However, these results do not necessarily mean that every prognostic model would be ineffective. Other studies, as the one mentioned by Murray and collaborators, showed different results and it can be argued that the impact of prognostic models would vary according to the context in which they are applied. Their impact will be determined not only by its accuracy but by the following contextual variables:

Users: How much doctors believe in the prognostic model and incorporate its prediction into their practice is of paramount importance. There is some evidence that models which are "home grown" facilitate implementation.[57]

Setting: In settings with scarce resources doctors will need to prioritise among patients and it is plausible that accurate prognostic information could be more useful.[41]

Condition: The impact on patient outcome is related to the evidence of the effectiveness of interventions according to baseline risk. For example the evidence for interventions according to risk in primary prevention in cardiology is well established, so prognostic estimates can be easily translated into treatment recommendations.[58]

Taking into account the previous considerations, some authors argue that prognostic models should be developed to be accurate and their impact would vary according to the context where they are applied. As Kellett stated in a recent paper "*...it is unlikely that (prognostic models) worsen clinical judgment. Therefore a good physician should no more refuse use them than a good driver should refuse to use his car's headlights at night*"[31]

1.3 Prognosis in traumatic brain injury

1.3.1 Potential role of prognosis research in TBI

1.3.1.1 Potential role of prognostic models

Prognostic information could potentially be useful for decision making. Taking into account that most TBI cases occur in low and middle income countries, accurate prognostic information could be of paramount importance.

The usual difficulty for the human brain of collecting and summarizing quantitative data to make predictions, is even more extreme in emergency situations such as in the treatment of TBI patients. It has been shown that in emergency situations there are some common problems such as stress, fatigue, poor communication, interruption of thinking, which are accentuated by the need to take rapid decisions.[59] Therefore prognostic models have a potential role for the management of TBI patients.

In the context of TBI, clinical prediction models rather than clinical decision models would be more appropriate, because of the lack of evidence of effective interventions according to baseline risk. However, clinical prediction models still have the potential to influence TBI management. As mentioned before, Murray and collaborators demonstrated that the introduction of computer-based outcome prediction altered TBI patient management.[56] The potential impact of accurate prognostic information could be particularly important in low and middle income countries where resources are limited.

Furthermore, in such a critical setting, accurate and consistent prognostic information provided by a prognostic model may also be helpful in the counselling of patients and relatives. It has been shown that physicians change their own predictions when they need to communicate them to patients or relatives.[60] Therefore the use of prognostic models may not only result in more accurate predictions but may also increase the consistency when communicating with patients and relatives.

1.3.1.2 Potential role of explanatory prognosis studies

The identification of prognostic variables in causal pathways could inform the design of randomised clinical trials of potential interventions. This is particularly important in the context of TBI research where there is a lack of evidence of effective interventions.[61] There are two main pathophysiological mechanisms susceptible to be interfered after a TBI. The first is related with the inflammatory response; so far there have been numerous trials targeting inflammatory response (neuroprotection studies) but to date no evidence of a clinical effect has been found. The only trial large enough to detect a clinically plausible effect, the CRASH (Corticosteroid Randomisation After Significant Head Injury) trial reported an increase in mortality.[62,63] The second potential pathway of intervention is related with intracranial bleeding (IB). Although there is evidence that IB is associated with worse a prognosis, its relationship is not well characterized. Most of the studies evaluating this association were small and had methodological limitations. Explanatory prognosis studies could shed some light on this association, which would be useful to inform the design of future trials.[64]

1.3.2 **What is already known about prognosis in TBI**

1.3.2.1 Individual prognostic factors

The Brain Trauma Foundation conducted a systematic review of the individual early indicators of prognosis in severe TBI. [65]

The authors evaluated each of the studies included in the systematic review according to criteria intended to establish study strength. These criteria included:

1. Twenty-five or more patients in the series with complete follow-up.

2. Outcomes measured — Glasgow Outcome Scale or Mortality — at six months or more.

3. Data gathered prospectively, although retrospective examination from a database creating an ongoing cohort of patients could be used.

4. Glasgow Coma Scale score measured within 24 hours.

5. Appropriate statistics (e.g., multivariate analysis) used to include adjustment for prognostic variables.

According to these criteria they reported the evidence of prognosis as shown below:

Class I: Those papers containing all of the above characteristics.

Class II: Those papers containing four of the five characteristics, including

prospectively collected data.

Class III: Those papers containing three or fewer of the above characteristics.

They also constructed 2 X 2 tables and evaluated the positive predictive value for mortality for each of the potential predictors.

The following evidence was reported for each of the most important predictors:

Glasgow Coma Scale.: Class I evidence of an increasing probability of poor outcome with a decreasing GCS in a continuous, stepwise manner was reported.[65]

Since its introduction in 1974, the GCS has consistently showed a correlation with outcome, the lower the score the worse the prognosis. However, some controversies remain. For example, the type of relationship with the outcome (linear or non linear) and the influence of the timing of GCS measurement on its validity.[66] It has been suggested that with more intensive pre-hospital management which includes sedation and intubation, a valid GCS is more difficult to obtain. Despite these concerns GCS continues to be considered an important predictor in TBI. The motor component of the GCS has also been found to be associated with outcome.[67 68]

Age: Class I evidence of an increasing probability of poor outcome with increasing age, in a stepwise manner was reported. [65]

It is well established that an older age predicts poor outcome after TBI. It has been proposed that this could be related to an increased risk of intracranial haematomas and a decreased capacity for repair of the brain with age.[65] In addition, as in any other condition, it is possible that the worse prognosis associated with age is related with concomitant co morbidities. A meta-analysis evaluating the relationship between age and poor outcome in TBI patients, reported that this relationship was better expressed as a linear and quadratic term, and that the best fitting threshold was 40 years.[69]

Blood pressure: A systolic blood pressure less than 90 mm Hg was found to have a positive predictive value of 67% for poor outcome.[65]

Hypotension, defined as a systolic blood pressure less than 90 mmHg, is associated with a poor prognosis. Early hypotension appears to exacerbate the development of intracranial hypertension. A single episode of hypotension has been associated with a doubling in mortality.[65]

Pupillary light reflex: Class I evidence of a positive predictive value of at least 70% for poor outcome was reported for bilaterally absent pupillary light reflex.[65]

Abnormalities of the pupillary light reflex are indirect measures of herniation and brainstem injury. Dilation and fixation of one pupil signifies herniation while bilaterally dilated and fixed pupils represent brainstem injury. It has been estimated that one fixed pupil is associated with a mortality of 54% whereas those patients with bilateral fixed pupils have twice the mortality (90%).[70] Direct oculomotor trauma should be excluded when considering the prognostic information of pupil abnormalities.

CT Scan characteristics: Class I and Class II evidence of a positive predictive value of at least 70% for poor outcome was reported for: abnormalities in first CT scan, Traumatic Coma Data Bank (TCDB) CT classification, compressed or absent basal cisterns and traumatic subarachnoid haemorrhage.[65]

Different CT classifications have been proposed to be predictors of unfavourable outcome in TBI.[71] [72] The Traumatic Coma Data Bank (or Marshall) classification (table 1.2) is the most widely used and has been shown to be associated with increased mortality in most of the studies, although not in all.[72] [65] [73]

Table 1-2 Marshal CT Classification

Diffuse Injury I (no visible pathology)	No visible intracranial pathology seen on CT scan
Diffuse Injury II	Cisterns are present with midline shift 0-5mm and/or lesions densities present, no high or mixed densities lesion >25 cc, may include bone fragments and foreign bodies
Diffuse Injury III (swelling)	Cisterns compressed or absent with midline shift 0-5mm, no high or mixed densities lesion >25 cc
Diffuse Injury IV (shift)	Midline shift >5mm, no high or mixed lesion >25 cc
Evacuated mass lesion	Any lesion surgically evacuated
Non-evacuated mass lesion	High or mixed lesion > 25 cc, not surgically evacuated

The following individual CT characteristics have been shown to be associated with higher mortality: midline shift > 5mm, compression and obliteration of the basal cisterns. Intracranial bleedings (IB) are divided into extracerebral (epidural, subdural and subarachnoid) and intracerebral or parenchymal. All types of IB are associated with a worse prognosis. [65] Although there have been some studies showing an association between size of IB and prognosis, the empirical evidence is limited, most studies having small sample sizes and restricted populations.[74-77] Despite being widely accepted and biologically plausible, currently there is no clear empirical evidence showing the relationship between size of IB and poor prognosis. The further division between evacuated and non-evacuated haematomas proposed in the TCDB classification is criticized by some authors because, they argue, it could be influenced by differences in patient management between centres.[73] Recently it has been proposed that it is preferable to combine individual CT predictors rather than to use the TCDB classification for prognostic purposes.[78]

Other predictors not included in the Brain Trauma Foundation Review:

Genes: The presence of the apolipoprotein E4 (APOE4) allele has been associated with poor outcome after TBI.[79] A recent systematic review identified 14 cohort studies, including 2527 participants. Only seven studies (1,868 participants) presented dichotomous data on GOS and could be included in the meta-analysis. The APOE4 allele was significantly associated with a poor outcome of TBI (RR = 1.36; 95% CI, 1.04-1.78)[80] However, the risk of bias was high in the majority of the included studies, with only two studies having assessed outcome blinded to genotype and a follow-up larger than 80%.

Biomarkers: The protein S100B is a marker of brain damage and has been proposed as a marker of poor outcome after TBI. A review found 18 studies that evaluated this association and, reported that patients with high levels of S100B may have a higher risk of disability.[81] However, this systematic review has many limitations as it did not evaluate the risk of bias of the included studies, nor did it address the probability of reporting bias.

1.3.2.2 Prognostic models in TBI

Since the publication of a prognostic model in TBI patients by Jennet and collaborators 30 years ago, many prognostic models have accumulated.[82] Nevertheless, there is still not one universally accepted and widely used model. A comprehensive review of all models is still lacking, but in a recent paper Hukkelhoven *et al.* described ten models in severe TBI. The authors found that most of them were based on small samples, were not well calibrated, had not been validated in external cohorts, and that hence their generalizability and utility was judged to be limited.[83] If a simple and accurate prognostic model for TBI patients could be developed then this might improve clinical practice, resource allocation, and the counselling of patient and relatives.

1.4 Aim and objectives of this thesis

The general purpose of this thesis was to study prognosis in TBI patients with the aim of providing useful and practical information in clinical practice and clinical research.

The two main objectives were

1) To develop accurate and practical prognostic models which could be used at the bedside in the early management of TBI patients. In order to achieve this objective I:
    a. Conducted a survey among doctors to describe the importance and use of prognostic information in the context of the management of TBI patients (Chapter 2)
    b. Evaluated existing prognostic models for TBI patients and critically appraised them (Chapter 3)
    c. Developed and validated prognostic models for the initial management of TBI patients using data from a large international cohort of TBI patients from the CRASH trial (Chapter 4)
    d. Developed a user friendly interface (paper based score card) for the prognostic models (Chapter 5)

2) To evaluate the validity of a simple disability scale (Modified Oxford Handicap Scale) completed at hospital discharge, to predict disability at six months, which could inform interim analysis in randomised clinical trials, could be used to help tackle the problem of loss to follow-up, or could be used to inform patients and relatives at hospital discharge. (Chapter 6)

Chapter 2 Doctors' perceptions of the importance of prognosis for TBI

2.1 Introduction

In Chapter 1 I discussed the major features of TBI, prognosis in clinical practice and research, characteristics of prognostic models and, in particular, the potential role of prognosis (and prognostic models) in the context of managing TBI patients.

In this chapter I will explore physicians' perceptions in relation to different aspects related to prognostic information in the context of the management of TBI patients.

Although prognosis has been always an essential part of the practice of medicine, its practice is often informal and non explicit.[84] Prognostic information is used when making treatment decisions and when communicating with patients and relatives.[29] However, the role of prognostic information from the perspective of physicians who routinely treat TBI patients is currently unknown.

An excellent opportunity to explore this issue occurred with the final meeting of the CRASH trial. The CRASH trial was a randomised controlled clinical trial that allocated adults patients (aged 16 years or older) with a TBI and a GCS of 14 or less to either a 48 hour infusion of methylprednisolone or a matching placebo within eight hours of injury.[63] A total of 10,008 patients from 239 hospitals in 48 countries were randomised. Primary outcomes were all cause mortality within 14 days and death and disability six months after injury.

I conducted a survey among the collaborators attending the final meeting of the CRASH trial to assess their perception in relation to the importance of prognostic information in the management of TBI patients and the potential use of a prognostic model in this context. Furthermore, I inquired their preferences in terms of predictors to be included, outcomes to be predicted, and ways of expressing prognosis for a prognostic model for TBI patients.

2.2 Methods

2.2.1 Sample

The sampling frame for the survey was all doctors participating in the final results meeting of the CRASH trial.[63] The principal investigators from centres that recruited at least 30 patients were invited to the final CRASH trial meeting in April 2005. A session about prognostic models in TBI was planned where all the collaborators were to be given a self-completed questionnaire about prognosis in TBI.

2.2.2 **Questionnaire**

Before the meeting a questionnaire was developed. *(Appendix 2.1)*

Four domains were defined:

- Attributes: Physicians' socio-demographic characteristics
- Behaviours: What do physicians do in relation to prognosis in TBI?
- Beliefs: What do physicians think about prognosis in TBI?
- Attitudes: What do physicians want in relation to a prognostic model in TBI?

Items were developed for each of the domains and revised five times in iterative procedures by an expert in questionnaire design (Phil Edwards) until a final version was obtained. Fourteen closed-ended questions (five of them also included an open-ended option) and two open-ended questions were developed. After developing the first draft, an evaluation of the questionnaire was made by three reviewers: one neurosurgeon (Jonathan Wasserberg), one expert in clinical epidemiology (Ian Roberts) and one epidemiologist expert in qualitative research (Caroline Free). The revised questionnaire was then tested on a convenience sample (7 respondents) and written comments were obtained regarding the instructions and the face validity of the questionnaire. Finally the revised version was checked by the expert in questionnaires. As one third of the CRASH collaborators that attended the final meeting were Spanish speaking, the final version was translated into Spanish.

2.3 Results

2.3.1 **Attributes**

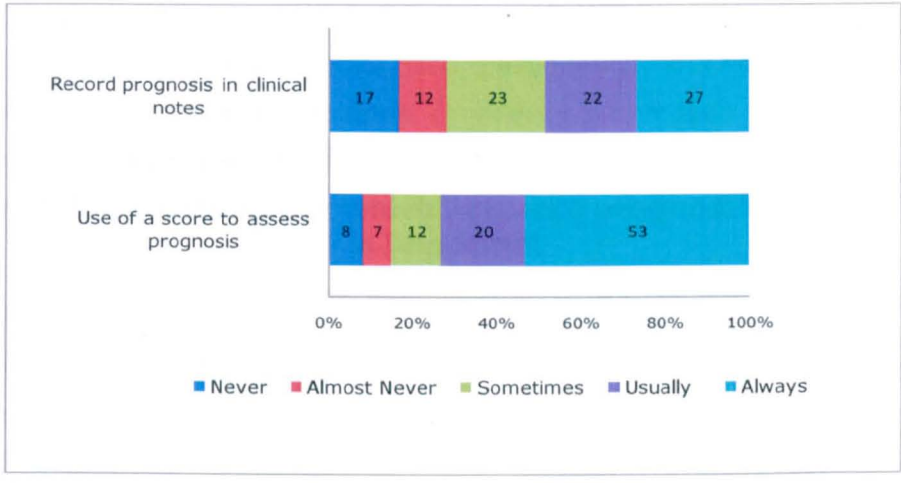
A total of 91 collaborators were invited to attend the final CRASH trial meeting of which 67 were able to attend. The session about prognostic models in TBI was attended by 60 (90%) of the collaborators.

The mean age of the 60 respondents was 44.9 years (*SD* 8.3 years) and 48 (81%) were males. Regarding their specialities, 20 (34%) were neurosurgeons, 19 (33%) intensive care specialists, 12 (20%) emergency care specialists, 4 (7%) anaesthesiologists, 2 (3%) general surgeons and 2 (3%) had other speciality. Most doctors worked in low and middle income countries: 15 (25%) in Latin America & the Caribbean, 12 (20%) in Africa, 11 (19%) in Asia and 5 (9%) in Eastern Europe & Central Asia; while 16 (27%) were from Western Europe. Most worked in hospitals with intensive care (100%), computed tomography (97%), and neurosurgical facilities (92%).

2.3.2 **Behaviours**

A total of 51 respondents (85%) reported using a score to assess prognosis, of which the summated Glasgow Coma Scale (GCS) was the most popular, used by 41 (80%) of them. A total of 43 (72%) reported recording prognosis in clinical notes. (Figure 2.1)

Figure 2-1 Behaviour in relation to prognosis in TBI patients



2.3.3 **Beliefs**

A total of 22 (37%) of the respondents believed that they currently assess prognosis accurately. Forty respondents (67%) said that a more accurate prognostic model could change the way that they manage patients and for 52 (88%) it would also change the way that they currently tell the prognosis to a patient or relative (table 2.1).

Table 2-1 Beliefs about prognosis in TBI patients

	Totally disagree %	Disagree %	Unsure %	Agree %	Totally agree %
Currently assess prognosis accurately (n:59)	8	27	27	25	12
An accurate prognostic model would change the way they manage patients (n: 60)	2	10	22	35	32
An accurate prognostic model would change communication with patient or relative (n: 59)	3	3	5	46	42

2.3.4 **Attitudes**

Accurate prognostic information was considered to be very important for a number of clinical decisions. The need to undertake a decompressive craniotomy, who should receive intensive care, and the decision on treatments were the three categories for which accurate prognosis was more frequently considered very important. On the other hand accurate prognostic information was deemed as not important in approximately one third of the respondents to decide in which patients to conduct a CT scan or which patients need rehabilitation. It is noteworthy that 50 (86%) respondents considered accurate prognostic information as important, or very important, to withdrawn treatment (table 2.2).

Table 2-2 Situations for which accurate prognostic information is important

	Very important %	Important %	Not important %
To decide which patients need decompressive craniotomy (n:59)	61	27	12
To decide which patients need Intensive Care Unit (n:58)	60	26	14
To decide which patients should receive treatment(e.g. hyperventilation, barbiturates, mannitol) (n:58)	55	26	19
To give counselling to patients and/or relatives (n:57)	54	37	9
To decide in which patients treatment should be withdrawn (n:58)	52	34	14
To decide in which patients intracranial pressure should be monitored (n:58)	50	40	10
To decide which patients need surgery (n:59)	49	32	19
To decide in which patients CT scan should be done (n:59)	19	49	32.
To decide which patients need rehabilitation (n:57)	14	53	33

The outcomes considered most important to predict were in-hospital death, need for intensive care, need for surgery and major disability (Table 2.3).

Table 2-3 Outcomes considered important to predict

	Very important %	Important %	Not important %
In-hospital death (n:60)	73	20	7
6 month death (n:58)	40	48	12
Need for surgery (n:59)	57	34	9
Need for Intensive Care Unit (n:59)	68	25	7
Days of stay in hospital (n:58)	9	60	31
Major disability (n:59)	56	41	3
Minor disability(n:58)	21	60	19
Need for rehabilitation (n:58)	31	52	17

The most favoured way of expressing the prognosis was as a percentage followed by a qualitative scale and as survival time (Table 2.4).

Table 2-4 Preferences on ways for expressing prognosis

	Very Useful %	Useful %	Not useful %
As a percentage (n:60)	62	35	3
Qualitatively (n:60)	22	60	18
As survival time (n:60)	20	57	23

In relation to prognostic variables, age, presence of extracranial injury, GCS and pupil reactions were considered among the most important predictors; abnormal CT scan results, in particular midline shift and obliteration of the third ventricle, were also considered as very important predictors by the majority of the respondents. (Table 2.5)

Table 2-5 Variables considered important predictors

	Very important %	Important %	Not important %
Demographics			
Age (n:60)	57	38	5
Gender (n:60)	2	13	85
Injury related			
Cause of injury (n:59)	34	42	24
Presence of major extracranial injury (n:60)	60	35	5
Time since injury to hospital arrival (n:60)	58	38	3
Clinical Examination			
Total GCS (n:60)	82	15	3
Eye component of the GCS (n:59)	36	44	20
Motor component of the GCS (n:60)	58	33	8
Verbal component of the GCS (n:59)	29	53	19
Pupil reactions (n:59)	80	20	0
Complications			
Presence of complications (n:60)	48	47	5
Wound infection (n:59)	8	61	31
Gastrointestinal bleeding (n:59)	17	66	17
Seizures (n:59)	51	34	15
Pneumonia (n:60)	37	52	12
CT scan results			
Abnormal CT scan (n:59)	69	29	2
One or more petechial haemorrhages within the brain (n:60)	30	63	7
Obliteration of the third ventricle or basal cisterns (n:60)	73	27	0
Subarachnoid bleed (n:59)	36	58	7
Midline shift over 5mm (n:60)	88	12	0
Non evacuated haematoma (n:57)	72	28	0
Evacuated haematoma (n:58)	36	55	9

2.4 Discussion

2.4.1 **Principal findings**

In a survey of doctors who routinely treat TBI patients I found that many believe that they make important decisions about the care of patients with TBI, including the decision to withdraw care, based on judgments about prognosis. Nevertheless, most of the doctors believe their prognosis is not accurate and think that a more accurate method of assessing prognosis could change clinical management in TBI patients, as well as communication with patients and relatives. Age, GCS, pupil reactivity, major extracranial injury, time to injury and abnormal CT scans were among the predictors considered most important. Gender, wound infection and cause of injury were among those considered least important. In-hospital mortality and major disability were among the outcomes considered most important to predict accurately and days of stay in hospitals among the least important.

2.4.2 **Comparison with other studies**

In a survey about attitudes and practice regarding prognostication among internists, Christakis and collaborators found similar results.[32] More than 80% of the respondents face several situations that require formulation of a prognosis and almost 60% reported that "they find it difficult to make predictions".

To the best of my knowledge the only study about the role of prognosis when treating TBI patients was conducted more than 20 years ago.[85] It included 59 neurosurgeons who also attended a conference, but unlike our study most of the participants were from high income countries. They also reported that assessment of prognosis was a frequent practice when treating a patient with TBI. Almost 90% of the respondents estimated prognosis preoperatively and, they thought that statistical prediction would improve their prognosis. Almost two thirds of the neurosurgeons considered that "computer predictions" would be more, or as reliable as, those predictions of an experienced physician.

2.4.3 **Strengths and weaknesses**

A key strength of the survey is that it includes physicians treating TBI patients from different specialties and from diverse regions of the world, mainly from developing countries, where the major burden of TBI occurs. Furthermore the response rate of those attending the conference was high (90%).

This survey is not free of limitations. It should be mentioned that this was a convenience sample and it was not intended to represent doctors worldwide. Furthermore, it could be argued that it was a biased sample as the respondents

were high recruiters from a clinical trial. This could be the explanation for the high rate of availability of facilities reported (intensive care, CT scans and neurosurgery). The extent to which the results may be generalized to other settings is therefore a matter for judgment. However, it might be argued that the need for prognostic information in settings with fewer resources could be even more important.

Another limitation is that doctors "reported" their current practice and this could be a source of information bias. The explicit use of a prognosis score and the recording of prognosis in clinical notes were quite prevalent and it might be likely that observation of actual practices of these doctors could show different results. Another potential type of bias could be "desirability" bias as collaborators were aware of the intention to develop a prognostic model using the CRASH trial data, however, we limited this source of bias because it was a self-completed and anonymized questionnaire.

2.4.4 **Implications of the findings and future research**

More than 20 years have passed since the previous survey on this topic and the results are remarkably similar. Doctors use prognostic information frequently, prognosis influences important decisions on patient management, and the majority of doctors would welcome a more accurate way of estimating prognosis such as by a prognostic model. The question is why are doctors still demanding prognostic models although many have been published? To answer this question is necessary to conduct a systematic review of existing prognostic models for TBI.

Chapter 3 Systematic review of prognostic models in TBI

3.1 Introduction

In the previous chapter I showed that some doctors believe that a more accurate way of predicting outcome would be clinically useful in the management of TBI patients and the counselling of patients and relatives.

A systematic and critical appraisal of existing prognostic models would enable doctors to know which of the available models are accurate and clinically useful. It would also help to critically appraise existing models, and to inform the development of future studies.

So far, there has not been any comprehensive systematic review of prognostic models in traumatic brain injury. Previous reviews of prognostic studies in TBI have only focused on individual predictors, or have been restricted to prognostic models of some type of traumatic brain injury or outcome.[65,83] It has, therefore, become increasingly important to identify and evaluate prognostic models in TBI patients.

In this chapter a systematic review of prognostic models in TBI patients is presented.

3.2 Objectives

The objectives were to:
(a) identify prognostic models in TBI
(b) describe the characteristics of prognostic models in TBI
(c) investigate the quality of prognostic models in TBI, and
(d) describe the models that were validated in an external population.

3.3 Methods

3.3.1 **Eligibility criteria**

3.3.1.1 Type of studies

Studies that gave an overall estimation of prognosis after TBI combining the predictive information from at least two variables were included. Studies could develop new prognostic models (derivation studies) or evaluate previous ones (validation studies). Studies published prior to 1990 were excluded because patient management and diagnostic techniques may have changed since this time. Studies

that investigated more than one variable but did not combine them for obtaining a prediction of outcome were excluded.

3.3.1.2 Type of exposures

Only variables that were collected before hospital discharge were considered as predictors. The Glasgow Coma Scale (GCS) was considered as one predictor variable.

3.3.1.3 Type of participants

Patients could be of any age, with any type, or severity of TBI.

3.3.1.4 Type of outcome measures

Studies that predicted any outcome in a TBI patient (e.g. neurological impairment, disability, survival, etc.) were included. There was no time restriction for the evaluation of the outcomes.

3.3.2 **Search strategy for identification of studies**

I developed jointly with a librarian (Reinhardt Wentz) the search strategy which is presented in appendix 3.1. I used the terms related with TBI and search strategies recommended by Altman for prognostic studies.[34] I conducted the search in Medline (Pubmed version), and Embase databases. The reference lists of included studies were inspected for further possible studies meeting the inclusion criteria. A forward search (citing references in the Web of Knowledge) was conducted with selected seminal papers, and some of the citing papers, not found by the database search, were inspected for relevance and possible inclusion. All records were imported into an Endnote database.

3.3.3 **Trial identification and selection**

I examined titles, abstracts and keywords of records from electronic databases, for eligibility. The full text of all potentially relevant records was obtained and I assessed whether each met the pre-defined inclusion criteria. This procedure was also done independently by another reviewer (Phil Edwards). Disagreement between me and the other reviewer was resolved by a third reviewer (Ian Roberts).

3.3.4 **Quality assessment**

Quality assessment scores for controlled clinical trials and diagnostic studies have been criticized. [86] [87] The main problem with quality scores is to determine the weight that each item should contribute to the overall score. The abundance of quality scores shows that there is no consensus on this issue. Instead, a component approach appraisal allows one to evaluate each methodological aspect. Depending

on the question and on the study design, some components may be more relevant than others. Recently it has been proposed that the term "risk of bias" should be used instead of quality assessment, as one study could have been conducted according to the highest possible standards yet still have an important risk of bias (e.g. with a surgical intervention blinding of the patient and caregiver would be unachievable).[88] However, because of its wide acceptance, I will use the term "quality assessment" throughout my thesis.

In studies of prognostic models in particular, although diverse quality assessment criteria have been proposed, none is widely accepted. [35,36,38] [89] I analyzed the quality of the prognostic models included in this systematic review according to two main domains:

3.3.4.1 Internal validity

This refers to the systematic error of the study and is related to study design, variables and analysis strategy.

3.3.4.2 External validity or generalizability

This refers to the extrapolation of the study to other settings. For making judgments about generalizability it is important to consider the characteristics of the sample from which the model was derived, how clearly the results were presented , and ideally the model should be evaluated (validated) in a different sample from the original.[43] Justice and collaborators distinguishes two components of "generalizability".[43]

a) Reproducibility or internal validation means that the model should replicate its accuracy in the setting where the model was developed. There are different methods for evaluating reproducibility such as re-sampling techniques (bootstrapping) or cross validation. These techniques are particularly important when the sample used to develop the model is small.

b) Transportability or external validation means that the model should replicate its accuracy in patients from a different but plausible related population

Finally, for a model to be used in clinical practice Wyatt and Altman argued that prognostic models should be "clinically credible"; by this they mean that it should be clearly explained how to use the model and ideally it should be simple to calculate a prediction.[37]

3.3.4.3 Questions assessing quality of prognostic models

Taking into account these two domains, 18 questions were considered for each of the models included. The quality assessment was restricted to the reports that included derivation studies.

Internal validity

Study

1) Did the patients have an adequate follow-up? *In the case of prognosis in TBI all the studies are conceptually cohort studies, although the method of data collection could be either prospective or retrospective (i.e. use of databases). In cohort studies a large loss to follow-up could lead to attrition bias. To minimize attrition bias the follow-up should be at least 90% of the original cohort.*

Variables

2) Was a discussion included about rationale to include the predictors? *The variables included should be important predictors reported in previous studies (e.g. for early indicators in severe traumatic brain injury the systematic review conducted by the Brain Trauma Foundation identified five: GCS, age, pupillary reflex, hypotension, CT scan features).*

3) Were the predictive variables clearly defined? *Variable definition and measurement should be clearly described in the method section of the report.*

4) Were the outcomes predicted valid? *Validity, for outcomes other than mortality, should be reported. GOS is the most frequent outcome considered in TBI studies and its validity has been reported previously.*

5) Were missing data adequately managed? *Imputation strategies are preferable to complete case analysis when the amount of missing data is large.*

Analysis

6) Was an adequate strategy performed to build the multivariable model? *Multivariable analysis strategy should consider clinical criteria when entering variables in the model and not only automatic selection strategies such as stepwise. In some cases important clinical predictors could be "forced" into the model.*

7) Were interactions between the variables examined? *When a multivariable analysis is performed interactions between variables should be explored.*

8) Were continuous variables handled appropriately? *It is preferable to keep continuous variables as originally recorded because they can give more information and are more powerful to detect an association. Categorization of a variable into groups assumes a constant risk in each group created, which is often not true. In the case where variables are categorized, the rationale for the cut-point should be clearly explained.*

9) Were more than 10 events per variable included? *The estimates may be unreliable if the data contain less than 10 outcome events relative to the number of parameters.*

External validity or generalizability

10) Was the description of the sample reported? *For making judgments about generalizability it is important to know the characteristics of the sample from which the model was derived, so it is very important that studies include information*

about the population included: e.g. time of inclusion of the patients in relation to the injury, time of the measurement of the variables, treatment received.

11) Was it clearly explained how to estimate the prognosis? *For a prognostic model to be clinically useful it should be clearly explained how to estimate the prognosis in a clinical setting. Probability of the outcome could be obtained through simple scores, nomograms, or simple figures. Reporting just the coefficients of the multivariable model is not enough to be clinically practical in the emergency setting.*

12) Were measures of discrimination reported? *To evaluate a model's performance its discrimination and calibration should be assessed: Discrimination refers to the ability to rank in the correct order individuals with different prognosis. It is usually measured with the area under the Receiving Operator Curve (R.O.C) or c statistic.*

13) Were measures of calibration reported? *For assessing the usefulness of a model the calibration should be reported. Calibration refers to the ability to predict correctly the prognosis (not too high or too low). It could be measured graphically or with the Hosmer-Lemeshow test.*

14) Were confidence intervals presented? *Clinicians that will use the model should know the precision of the estimates derived from the model.*

15) Was the model validated? *For a prognostic model to be generalized to a population different from the one from which the model was derived, it should be evaluated (validated) in a different set of patients.*

16) Was the model internally validated? *Internal validation assesses the validity of the model for the setting where the development data originated from. It is evaluated through different techniques, such as bootstrapping or cross validation.*

17) Was the model externally validated? *Refers to the evaluation of the model in a different population (e.g. different geographical region, historical periods or different methods of data collection).*

18) Was the clinical credibility of the model evaluated? *For a model to be used it should be well accepted by physicians. Ideally the "acceptability" and "practicality" of a model should be evaluated.*

3.3.5 Data extraction

I extracted the information from each study for assessing the quality of reporting in each of the questions into an excel file.

3.3.6 Description of models externally validated

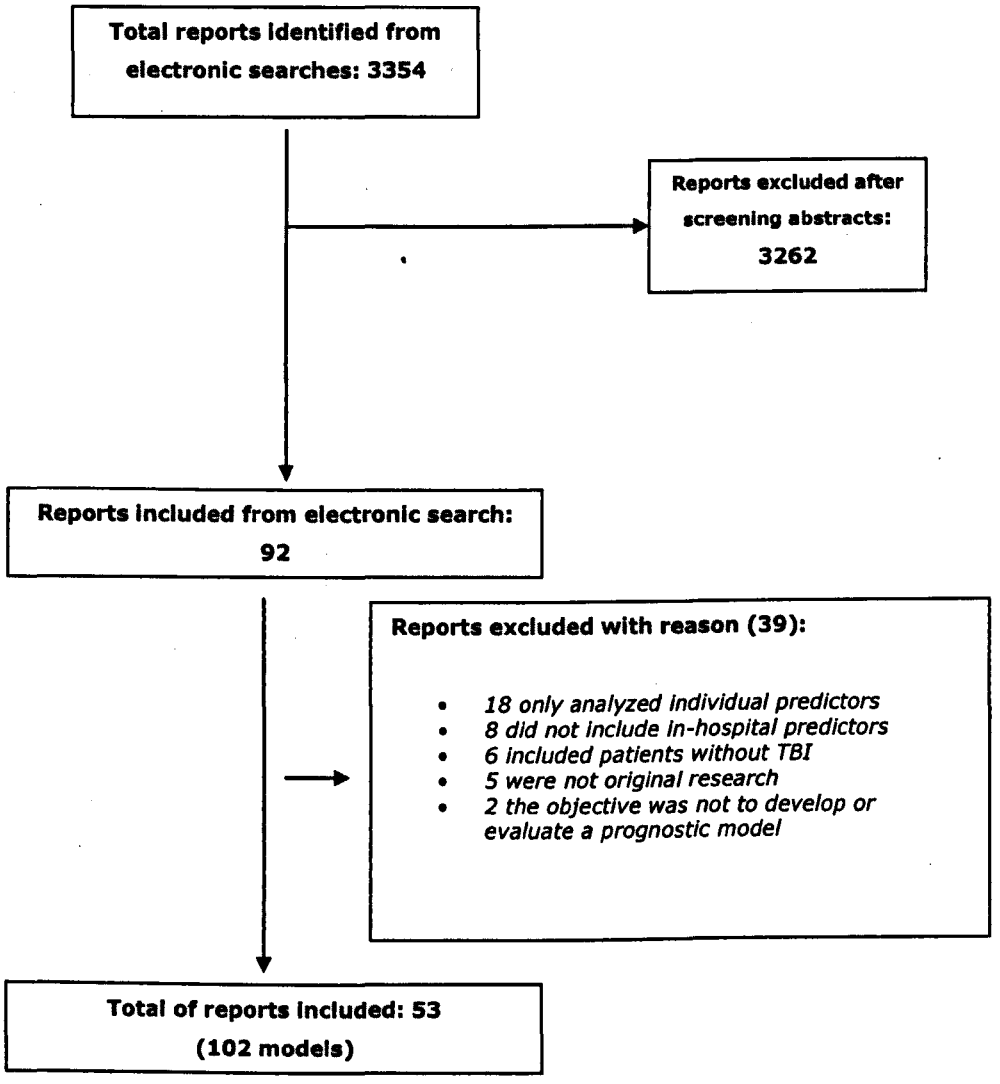
The characteristics and performance of models that were validated in external samples were reported. Those models that were reported by the authors as evaluated in a different cohort of TBI patients from the derivation set were

considered as externally validated. Only TBI specific models, as opposed to general trauma scores, were fully described.

3.4 Results

A total of 3354 records were identified. After reading all the records, 92 reports were identified and read in full. Thirty nine were excluded for the following reasons: 18 analyzed individual predictors but did not combine them in a single score, eight did not include in-hospital predictors, six included patients without TBI, five were not original research (e.g. discussion, letter), and in two the objective was not to evaluate prognosis in TBI patients. The remaining 53 reports described 102 prognostic models (Figure 3.1).

Figure 3-1 Study selection process for the systematic review



3.4.1 **General characteristics of the prognostic models**

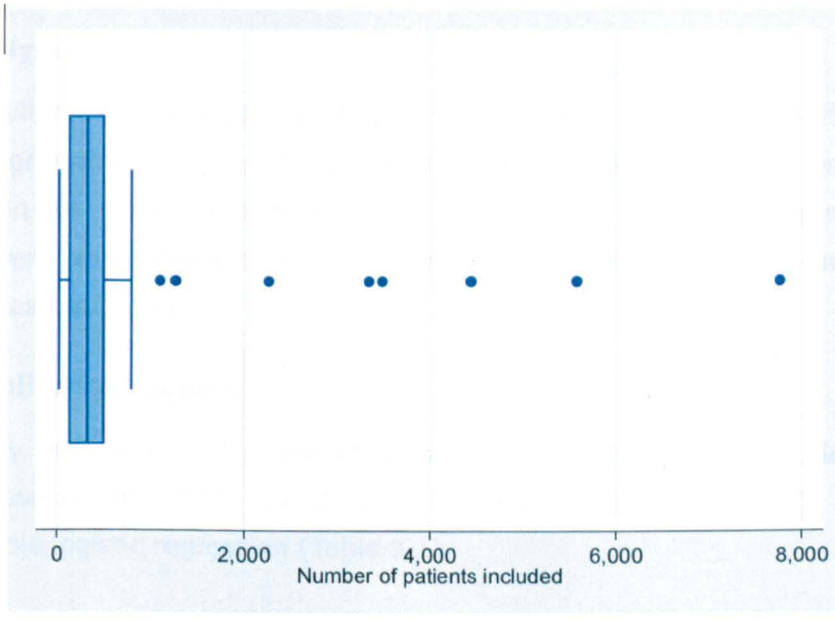
3.4.1.1 Population included

A total of 48 of the models were derived from an adult population, 12 were derived from a child population, while 21 were derived from a population that included both adults and children. In 21 of the models it was not clearly reported from which population they were derived. For details see appendix 3.2 (list of included studies), and appendix 3.3 (characteristics of each study included).

In relation to the severity of the TBI studied, 45 models included all grades of severity, 31 included severe TBI, nine moderate or severe TBI, nine mild TBI and in eight the severity of TBI was not clearly reported.

A median of 319 patients (range 22–7764 patients) were included per model. Three quarters included less than 500 patients (Figure 3.2).

Figure 3-2 Number of patients included per model



A total of 95 models included populations from high income countries, five included populations from middle income countries and in two the population was from a low income country.

3.4.1.2 Objectives

A total of 66 models reported were derived for the first time (derivation models) while 35 were validating pre-existing models (validation models). The objective of one model was unclear.

3.4.1.3 Variables included as predictors in the model

A total of 89 variables were included in the prognostic models. A mean of 5 variables were included in each model (range 2 to 13). GCS was the most common predictor included in the models (50%), followed by age (46%) and pupil reactivity (26%). Overall clinical variables were included in 66% of the models, demographic variables were included as predictors in 50% of the models, CT scan predictors were used in 19% of the models and 7% included variables related to characteristics of the injury. In 7% of the models other predictors were included (e.g. other complementary tests or existing scores).

3.4.1.4 Outcomes

Mortality was the main outcome in 30% of the models and GOS in 28%. Other functional outcomes were reported in 31% of the models. The presence of a CT scan lesion was the main outcome in 7%, the need of neurosurgical intervention in 2%, and raised intracranial pressure in 1%.

3.4.2 **Analysis**

In the multivariable analysis for the development of prognostic models (n=66) logistic regression was used in 31 (47%) models. Regression tree analysis was reported in 14 (21%) and neural networks in nine (13%). Other methods of analysis were performed in nine (14%) models while in three (4%) no multivariable analysis was performed.

3.4.3 **Quality assessment**

The quality assessment was restricted to the 66 derivation models. Some of the quality assessment items could only be applied to the models that used multivariable logistic regression (Table 3.1).

Table 3-1 Quality assessment of prognostic models

INTERNAL VALIDITY	*All models*	*Logistic regression*	*Other*
	N=66	*N=31*	*N=35*
STUDY			
Loss to follow-up			
<10%	10 (15%)	5 (16%)	5 (14%)
>10%	19 (29%)	7 (23%)	12 (34%)
Not reported	37 (56%)	19 (61%)	18 (52%)
VARIABLES			
Discussion about predictors			
Yes	21 (32%)	11 (35%)	10 (29%)
No	45 (68%)	20 (65%)	25 (71%)
Description of measurement of predictors			
Yes	12 (18%)	8 (26%)	3 (9%)
No	54 (82%)	23 (74%)	32 (91%)
Validity of outcome reported			
Yes	31 (47%)	14 (45%)	17 (49%)
No	20 (30%)	7 (23%)	13 (37%)
Not applicable	15 (23%)	10 (32%)	5 (14%)
Handling of missing data			
Estimated statistically	4 (6%)	4 (13%)	0
Excluded	36 (55%)	16 (52%)	20 (57%)
Not reported	26 (39%)	11 (35%)	15 (43%)
ANALYSIS			
Multivariable analysis			
Stepwise backwards		12 (39%)	
Stepwise forwards		3 (10%)	
Stepwise not specified		10 (32%)	
Not reported		5 (16%)	
Other		1 (3%)	
Interactions examined			
Yes		4 (13%)	
Not reported		27 (87%)	NA
Handling of predictors variables			
Continuous		6 (19%)	
Categorical		16 (52%)	
Unclear		9 (29%)	
More than 10 events per variable			
Yes		9 (29%)	
No		16 (52%)	
Not reported		6 (19%)	

Continuation table 3.1

EXTERNAL VALIDITY	*All models*	*Logistic regression*	*Other*
	N=66	*N=31*	*N=35*
Description of the sample			
Yes	55 (83%)	28 (90%)	27 (77%)
No	11 (17%)	3 (10%)	8 (23%)
Presentation of the prognostic model			
Nomogram	1 (1%)	1 (3%)	0
Simplified score	8 (12%)	4 (13%)	4 (11%)
Figure	13 (20%)	1 (3%)	12 (34%)
Regression formula	15 (23%)	12(39%)	3 (9%)
Not explained	29 (44%)	13(42%)	16 (46%)
Performance reported			
Area under ROC (Discrimination)			
Yes		18 (58%)	
No		13 (42%)	
C.I. presented		8 out of 18 (44%)	
H-L (Calibration)			NA
Yes		7 (23%)	
No		23 (74%)	
Other		1 (3%)	
Overall accuracy			
Yes	37 (56%)	15 (48%)	22 (63%)
No	29 (44%)	16 (52%)	13 (17%)
Validation			
Yes	25 (38%)	17 (55%)	8 (23%)
External	*7 (11%)*	*7 (23%)*	*0*
No	41 (62%)	14 (45%)	27 (77%)

3.4.3.1 Internal validity

In over half of the models loss to follow-up was not reported. Ten models (15%) reported less than 10% loss to follow-up.

Most models (68%) did not include a discussion about the rationale for including the predictors in the model. A detailed description of the measurement of the predictors was absent in 82% of the models. In one third of the models the validity of the outcome measures was not reported.

In relation to the analysis of the models that used multivariate logistic regression, stepwise was the most common approach (81%). Interactions were examined in 13% of the models. Predictor variables were analyzed as continuous in 19% of the models. A third (29%) of the models included at least 10 events per variable analyzed as predictor. The most common strategy to handle missing data was exclusion of observations (55%).

3.4.3.2 External Validity

The sample was described in almost all the models (83%). The procedure to obtain the score was explained in approximately half of the models (56%). However, in those that used logistic regression only 19% included a user-friendly presentation.

In relation to the performance of the models, discrimination was reported in 58% of the models through the area under the ROC, 44% of which included the respective confidence interval. Calibration was reported with the Homer-Lemeshow test in 27% of the models. Almost half the models (56%) reported their overall accuracy.

Less than half of the models (38%) were validated, of which 11% were validated in an external population. None of the models evaluated their clinical credibility.

3.4.4 **Description of externally validated models**

Seven models were developed that also reported an external validation (Table 3.2).

Table 3-2 Characteristics of the models externally validated

Author	Derivation sample	Validation sample	Predictors	Outcomes	Performance in the validation sample	Presentation of a simplified score
Pillai et al.	289 patients from India with severe TBI	26 patients from the same centre	1-oculocephalic reflex 2-motor GCS 3-midline shift	Death or vegetative state	Sensitivity (75%) Specificity (67%) PPV 50%	No
Signorini et al	372 patients from Scotland with moderate and severe TBI	520 patients from the same centre	1-GCS 2-ISS 3-pupils reactivity 4-haematoma (CT scan)	Survival at 1 year	ROC (0.835) Error rate (15.2%) Brier score (0.1160) Hosmer-Lemeshow ($p<0.001$)	Nomogram
Signorini et al	110 patients from Scotland with moderate and severe TBI	140 patients from the same centre	1-GCS 2-ISS 3-pupils reactivity 4-haematoma (CT scan) 5-ICP measures	Survival at 1 year	Not reported	No
Hukkelhoven et al.	275 patients from Netherlands with moderate and severe TBI	205 patients from the same centre	1-age 2-motor GCS 3-pupils reactivity 4-pupillary size 5-hypotension 6-ISS	Raised ICP	ROC (0.50) Hosmer-Lemeshow ($p=0.18$)	No

Continuation table 3.2

Hukkelhoven et al.	275 patients from Netherlands with moderate and severe TBI	205 patients from the same centre	1-age 2-cause of injury 3-pupils reactivity 4-pupillary size 5-hypotension 6-ISS	Surgical removable lesions	ROC (0.67) Hosmer-Lemeshow (p=0.01)	No
Hukkelhoven et al.	2269 patients from 2 trials in high income countries with moderate and severe TBI	796 patients from Europe	1-age 2-motor GCS 3-pupils reactivity 4-hypoxia 5-hypotension 6-CT classification 7-subarachnoid haemorrhage	Death or disability at six months	ROC (0.83) Hosmer-Lemeshow (p=0.05)	Score chart
Hukkelhoven et al.	2269 patients from 2 trials in high income countries with moderate and severe TBI	796 patients from Europe and 746 from the United States	1-age 2-motor GCS 3-pupils reactivity 4-hypoxia 5-hypotension 6-CT classification 7-subarachnoid haemorrhage	Death at 6 months	ROC (0.87/0.89) Hosmer-Lemeshow (p=0.42/<0.001)	Score chart

Pillai and collaborators developed a prognostic model to predict unfavourable outcome (death or vegetative state) at one month.[90] They developed the model in a cohort of 289 patients and validated the model in 26 patients from the same centre. The predictor variables were oculocephalic reflex, motor score of the GCS and midline shift score. They developed a predictor score that could then be transformed in a binary outcome (favourable or unfavourable). In the validation set they reported sensitivity (75%), specificity (67%), predictive value of unfavourable outcome (50%), predictive value of favourable outcome (86%), percentage of false optimistic results (25%), and percentage of false pessimistic results (33%). They did not report the model's performance measured in the derivation set. Confidence intervals of the estimates were not reported. Although the authors reported how to calculate the prediction score, they did not present it in a user-friendly fashion.

Signorini and collaborators developed two prognostic models, for one they used only clinical variables and for the other they added variables on secondary insults. In both models the outcome was survival at 1 year.[91] The first model was validated in 520 patients who attended the same centre. The predictors were age, GCS, Injury Severity Scale (ISS), pupils reactivity and presence of haematoma on the CT scan. They reported measures of discrimination: area under the ROC (0.835), error rate (15.2%) and calibration: Brier score (0.1160), Hosmer-Lemeshow ($p < 0.0001$). They included a graph with the 95% confidence interval of the calibration of the model. The second model was validated in 140 patients who attended the same centre. The predictor variables were the same as the first model plus ICP measures. Although they mentioned that Brier score, error rate, area under the ROC were higher than the original dataset they did not report the actual estimates. They reported a nomogram to predict probability of survival that is user-friendly for physicians.

Hukkelhoven and collaborators reported four different models.[92,93] The predicted outcomes were: raised intracranial pressure (ICP), surgically removable lesions (SRL), unfavourable outcome (death, vegetative state or severe disability) and mortality. For the validation of the first two outcomes they used an historical (previous) sample of 205 patients from the same centre. The predictors for ICP were age, motor score, pupil size, pupillary reactivity, hypotension and ISS. For SRL the predictors were the same except for motor score which was not, and cause of injury that was added. For the validation of the model for unfavourable outcome they used one database and for mortality two databases, none of these databases was related with the population of the derivation set. The predictor variables were age, motor score, pupil reactivity, hypotension, hypoxia, CT classification and traumatic subarachnoid haemorrhage. They

reported the models discrimination: area under the ROC of 0.50 (95% CI 0.41-0.58), 0.67 (95% CI 0.60-0.75), 0.83 (95% CI 0.80-0.86), and 0.87(0.84-0.89) for ICP, SRL, unfavourable outcome, and mortality respectively. They also reported the model calibration: Hosmer-Lemeshow test of 0.18, 0.01, 0.05 and 0.42 (<0.001), for ICP, SRL, unfavourable outcome and mortality respectively (the calibration of the mortality model was validated in two different databases). The mortality and unfavourable outcome models were presented in a user-friendly fashion (a numerical score accompanied by a figure).

In relation to the 35 models that validated pre-existing models, 29 (83%) validated general trauma score. One validation model was reported in a letter and the information provided was limited. Three other models validated prognostic scores that were developed before 1990. Therefore only two studies that reported the validation of (post 1990) TBI prognostic models are described.

Bush and collaborators validated a model previously developed by the same group.[94] Their model was intended to allow better understanding of factors influencing functional outcomes and was not intended to predict individual outcomes. It was not clearly reported whether the patients came from the same original population. They used path analysis to evaluate the predictors (functional status, injury severity and cognitive status) on functional outcomes (disability rating scale, community integration questionnaire and return to employment). The reported different indexes of goodness of fit showed that the original model fitted better than the validation model. They did not report any discrimination measures.

Benzer and collaborators validated a model that used an existing scale, although they did not provide details of when and how it was developed.[95] They did not use any kind of multivariable analysis. The used a score based on the following variables: reaction to acoustic stimuli, reaction to pain, body posture, eye opening, pupil size, pupil response to light, position and movements of eyeballs and oral automatisms to predict mortality at 21 days. They did not report any performance measure, but just the chi square test for survival of those with low versus high score. They presented the score in a user-friendly way.

3.5 Discussion

3.5.1 **Principal findings**

This systematic review shows that although publications of prognostic models for TBI patients are very frequent, their quality is relatively poor. In addition they are rarely validated on external populations or presented to physicians in a user-friendly way. Furthermore, few are developed using populations from low and middle income countries where most trauma occurs.

Patients from all severity spectra were investigated but prognostic models for moderate and severe TBI patients were more frequent. It is noteworthy that only 2% of the models included patients from low income countries taking into account that 90% of trauma occur in these countries. Although biologically prognostic factors should be the same worldwide, it is reasonable to consider that the baseline risk and the strength of the association could differ depending on the medical care received. This difference could affect the accuracy of the prognostic models in different settings.

GCS, age and pupil reactivity were the most common variables analyzed as predictors whereas, GOS and mortality were the most common outcomes investigated. Multiple logistic regression was the multivariable analysis most frequently used.

Several limitations in the quality of the models were found. The majority did not include a thorough discussion of the rationale for including predictor variables. Only a minority had a loss to follow-up of less than 10%. Potentially this is an important limitation as the loss to follow-up could be related to prognosis and this could lead to bias. Furthermore only four models handled the missing data by using statistical imputation. In relation to the multivariable analysis, automatic procedures (stepwise) were quite common in logistic regression. There is no agreement in relation to the appropriateness of this strategy. This is shown, for example, in conflicting recommendations in quality assessment for prognostic studies; while in one study the use of stepwise was considered as good quality in another it was considered as a flaw.[96] One of the limitations was that most of the studies did not explicitly consider clinical criteria to enter the variables in the model beyond the automatic procedures. Interactions were hardly ever explored, although this is strongly recommended in multivariable analysis.[97] Another common weakness was the lack of power of the models, most of them were derived from small samples and it is well established that large samples are required for reliable selection of predictors.[98] Only one third included at least 10 events per variable, and it has been proposed that this is the minimum ratio of events to variables which is large enough to allow an adequate precision of the

estimates.[97] I did not attempt to obtain an overall quality assessment and instead I evaluated its different components. This approach makes a cross comparison between different analytical strategies difficult because, for example, many of the criteria only apply to logistic regression analysis.

It is also important to report how well the model works and to report performance measures. Remarkably only two thirds reported a measure of discrimination and only one fifth reported a measure of calibration. Even when a discrimination measure was reported, less than half presented confidence intervals to provide readers with an estimation of its precision.

For a model to be generalizable to other populations it is important to conduct an external validation.[43] Only seven models (three reports) developed and validated a model, but in only two of them was the validation performed on patients of a different centre.

Finally, to be useful, the method to estimate prognosis should be clearly reported and, to be clinically practical, it should be user-friendly. Only half of the models clearly explained how to obtain the prognostic score, and only one tenth were reported in such a way that could be easily applicable in a clinical setting. None of the models evaluated the clinical credibility of the different presentations.

Two models developed by Hukkelhoven and collaborators, one for mortality and the other for unfavourable outcome, were those which fulfilled most of the methodological and clinical criteria.[93] A thorough discussion of the predictors was included, missing data were handled appropriately, the assumptions of the model were tested, external validation in two different populations was performed, and discrimination and calibration measures were presented. The sample size was 2,269 patients and 1,542 patients for the validation. Furthermore a simple score chart was developed to estimated the outcome probability. The predictors included in the final model were age, GCS motor score, pupil reactivity, hypoxia, hypotension, CT scan abnormalities and presence of traumatic subarachnoid haemorrhage. The discrimination of the model was higher than 0.8 for both outcomes, however, the calibration was poor. As a limitation these models only included patients from developed countries and were restricted to moderate and severe TBI cases.

3.5.2 **Comparison with other studies**

There has only been one previous systematic review of prognostic models in TBI by Hukkelhoven and collaboratos.[83] These authors found 10 reports, all of which were also identified in my systematic review. Unlike my systematic review, Hukkelhoven and

collaborators restricted their systematic review to models that used baseline characteristics to predict mortality or unfavourable outcome (defined by GOS) in moderate and severe TBI patients. Furthermore the search strategy was not specified.

After the publication of my systematic review a new systematic review was by Mushkudiani and collaborators. These authors included prognostic models with variables available at hospital admission and they restricted the population to moderate and severe TBI.[99] They found thirty one studies and their conclusions were very similar to the one reported in this chapter. Most prognostic models included GCS, age and pupil reactivity, they were developed from small sample sizes, and they were rarely externally validated or presented in a practical way.

Systematic reviews of prognostic models for other diseases have found similar results to the one reported in this chapter. For example Counsel and collaborators conducted a systematic review of prognostic models in patients with acute stroke.[89] They found 83 prognostic models but they concluded that none of them was sufficiently well developed and validated.

3.5.3 **Strengths and weaknesses of this systematic review**

This is the first comprehensive systematic review of prognostic models for TBI. I developed a comprehensive search strategy and included TBI of all severities. I also critically analysed the development and validation strategies from an epidemiological perspective to highlight limitations, to inform potential users of prognostic models, and to improve future designs of prognostic models in this area.

Some limitations should be acknowledged in this systematic review. Firstly, only studies that explicitly combined two predictors were included, and by doing so some reports could have been missed. Secondly, studies that assessed clinical decision rules were not included. The methodological framework to assess such studies is fundamentally different from prognostic models. Thirdly, the search was restricted to 1990 onwards so, some relevant prognostic models published prior to that date could have been missed. However, because of changes in management and diagnostic technology in recent years it is doubtful whether prognostic models prior to 1990 could be useful for the current medical care of TBI patients. Finally, there is not yet a clear methodological framework to conduct systematic reviews of prognostic models, therefore, I faced some common challenges that any author conducting systematic reviews of this study design confront. For example, although I used a search strategy recommended for prognostic studies there is not a validated one for prognostic

models, and I developed my own methodological framework to assess the quality of the included studies.

3.5.4 **Implication of the findings and future research**

Reasons for the poor quality found in the studies included in this systematic review are not clear but, as mentioned before, it seems that this is not an exception in systematic reviews of prognostic studies. Among the possible explanations is that most of these studies include a small series of patients from single hospitals and give the impression that they were conducted by physicians who "wanted to know" the predictors among their patients. Most of the studies lack a clear theoretical framework and a clear question to address. One way to move forward and improve the quality of this type of study is to set up clear guidelines for prognostic studies and establish international collaboration to address prognostic questions. There are some signs of improvement in this field, for a example there is an initiative to create a Prognosis Systematic Review Methods Group within the Cochrane Collaboration.[100]

The findings of this systematic review could be used to inform researchers who are involved in the development of prognostic models in TBI. Future studies should consider the following issues to develop valid prognostic models:

- Thorough discussion with physicians of potential predictors that are clinically relevant
- Clear description of the measurement and validity of variables included in the model
- Large sample size to ensure precise estimates
- Adequate handling of continuous variables and missing data
- Assessment of interaction in the multivariable analysis
- Internal validation
- External validation
- Adequate report of model performance measures
- Clear description of the calculation of the prognostic score and
- User-friendly presentation

It should also be encouraged that more studies include populations from low and middle income countries where most of the burden of TBI occurs.

The three main predictors that were consistently associated with poor outcome in TBI patients were age, GCS and pupil reactivity. According to these findings these variables should always be present in a prognostic model for TBI patients.

Chapter 4 Development and validation of prognostic models

4.1 Introduction

In the previous chapters I described the importance of prognostic information for clinical practice, in particular in the context of TBI patients, and have shown that physicians who routinely treat TBI patients would welcome a more accurate way of assessing prognosis. I have also described why prognostic models are likely to be more accurate than simple clinical predictions, and how some studies have shown that the use of a prognostic model can influence TBI patient management.

The systematic review presented in Chapter 3 showed that although hundred of prognostic models for TBI have been reported, the majority have limitations. Most models are developed on small samples, most are methodologically flawed and few are validated in external populations. Only a small number are presented in a clinically practical way, or developed in populations from low and middle income countries where most trauma occurs.

The CRASH (Corticosteroid Randomisation After Significant Head injury) trial is the largest clinical trial conducted in patients with TBI and presents an unique opportunity to develop a prognostic model for TBI.[63,101] The trial had prospective inclusion of patients within eight hours of the injury, used standardised definitions of variables and achieved almost complete follow-up at 14 days and at six months. Furthermore, the large sample size (10,008 patients) allows precise estimates. The high recruitment of patients from low and middle income countries means that models developed using these data are relevant to these settings.

Using data from the CRASH trial, I have developed and validated prognostic models for death at 14 days, and death and severe disability at six months in TBI patients.

4.2 Methods

4.2.1 The sample of patients

The study cohort was all 10,008 patients enrolled in the CRASH trial. Adults with TBI, who had a score on the Glasgow Coma Scale (GCS) of 14 or less, and who were within eight hours of injury, were eligible for inclusion in the trial.

4.2.2 **Outcomes**

The primary outcomes were:

Death at 14 days: Patient death was recorded on an early outcome form which was completed at hospital discharge, death or 14 days after randomisation (whichever occurred first).

Unfavourable outcome at six months: This outcome was defined using the Glasgow Outcome Scale (GOS). The GOS was assessed at six months with a validated questionnaire that was mailed to patients or their carers, administered by telephone interview, or undertaken during a home visit or hospital appointment (Appendix 4.1).[102] The questionnaire included six questions addressing disability: three dealing with functional disability or dependency (extent of help needed in the home, ability to shop and travel), and three addressing psychosocial disability (ability to work, take part in social and leisure activities, and relationship problems). Patients were classified as having a good recovery (able to return to work), moderate disability (able to live independently but unable to return to work or school), severe disability (able to follow commands but unable to live independently) or death at six months.
For the purpose of this analysis, outcome at six months was dichotomised into favourable (moderate disability or good recovery) and unfavourable outcome (dead or severe disability).[11]

4.2.3 **Prognostic variables**

As I intended to develop clinically practical models for the early management of TBI patients, I selected those variables which are generally available at hospital admission. The following variables were considered for the prognostic model: age, gender, cause of injury, time from injury to randomisation, GCS score at randomisation, pupil reactivity, whether the patient sustained a major extracranial injury, computerised tomography (CT) scan results, and country income region (high income countries(HIC) or low & middle income countries(LMIC)).

4.2.4 **Analysis**

4.2.4.1 General strategy

For descriptive purposes, proportions were calculated for each variable for the total population and by country income region. Chi squared tests were performed to evaluate the differences in characteristics between the different regions.

Most of the variables initially considered for the prognostic models have been previously associated with prognosis in TBI, so all of them were included in a first multivariable logistic regression analysis.[65] Analyses were adjusted for trial treatment since this has been reported to be related with mortality.[63,101] Interactions between country income level and all the other predictors were evaluated using a likelihood ratio test.

Different models were developed for each of the two outcomes: a basic model and a CT model. For the basic models, all the clinical and demographic variables were first analysed and only those variables that were statistically significant at the 5% level remained in the final basic models.

For the CT models I included only those clinical and demographic variables that remained in the basic models, and added all the CT scan variables. A multivariable analysis was performed and only those variables that were statistically significant at the 5% level remained in the final CT models.

Data were explored for missing values. There were 30 (0.3%) patients without data on mortality at 14 days, 454 (4.5%) without data on disability at six months, 335 (3.3%) without data on mortality at six months, 143 (1.4%) without data on cause of injury, 238 (2.4%) without data on presence of major extra-cranial injury and 130 (1.3%) without data on the CT scan result. Due to the small number of missing data a complete case analysis was performed.

4.2.4.2 Analysis of individual predictors

4.2.4.2.1 Age

Age has been modelled in many different ways when included in prognostic models of TBI populations, some studies have treated age as a continuous variable whereas others have identified age thresholds.[69] Age was available as a continuous variable. To assess the best way to analyse age, I grouped it into five year intervals and graphically displayed its relationship with 14 day mortality.

4.2.4.2.2 Gender

Gender was analysed as a binary variable (male or female).

4.2.4.2.3 Cause of injury

Cause of injury was included in the analysis as a categorical variable. Three categories were used: road traffic crash (RTC), fall and other.

4.2.4.2.4 Time from injury

Time from injury to randomisation to trial treatment was available as a continuous variable (in hours). Three categories were pre-specified in the CRASH Trial: less than one hour, between one and three hours, and more than three hours. For descriptive purposes it was reported both as continuous and categorical. In the multivariable analysis it was treated as a categorical as defined by the CRASH trial protocol.

4.2.4.2.5 GCS

GCS was measured on admission before randomisation. If the patient was intubated and the GCS could not be assessed, the most recent GCS was reported instead. GCS was coded as "current" or "recent" according to the time of measurement.

The GCS was available as categorical variable for the motor, verbal and eye components and also as categorical for total GCS (3 to 14). For descriptive purposes total GCS was reported as mild (14-13), moderate (9-12), or severe (8 or less).

For deciding how GCS would be included in the model, I performed likelihood ratio tests comparing a model for predicting 14 days mortality with all the components of the GCS (total GCS) against different models with each of the three components (eye, verbal or motor). I also estimated the discrimination ability of each of them through the c statistic.

To assess if GCS could be analyzed as a continuous variable, I graphically displayed the relationship between GCS categories and the log odds ratio for 14 days mortality.

4.2.4.2.6 Pupil reactivity

Pupil response was evaluated for each eye and coded as reactive to light, not reactive, or unable to assess. For the analysis it was coded as both pupils reactive, one reactive, none reactive, or unable to assess.

4.2.4.2.7 Major extracranial injury

The presence of major extracranial injury was reported as present, if according to the physician, the patient presented with an extra-cranial injury requiring hospital admission within its own right. For the analysis it was coded as a binary variable (yes or no).

4.2.4.2.8 CT scan results

The following CT scan results (from the first CT scan available) were reported by the principal investigator as present or absent: normal scan, one or more petechial

haemorrhages within the brain, obliteration of the third ventricle or basal cisterns, subarachnoid bleed, midline shift over 5 mm, non evacuated haematoma, and evacuated haematoma. All the CT scan results were analyzed as binary variables (yes or no). The collaborators completed an outcome form with all the possible CT scan results (Appendix 4.2). A guideline was provided to the collaborators with definitions and images for each of the different possible CT scan results (Appendix 4.3).

4.2.4.2.9 Country income

Data on the country of origin were available for each patient. To explore the influence of different regions I divided the countries according to the World Bank Classification[103]. The following countries were considered high income (HIC): Australia, Austria, Belgium, Eire, Germany, Greece, Italy, New Zealand, Saudi Arabia, Singapore, Spain, Switzerland and United Kingdom. The following countries were considered low and middle income (LMIC): Albania, Argentina, Brazil, Chile, China, Colombia, Costa Rica, Cuba, Czech Republic, Ecuador, Egypt, Georgia, Ghana, India, Indonesia, Iran, Ivory Coast, Kenya, Malaysia, Mexico, Nigeria, Pakistan, Panama, Paraguay, Peru, Romania, Serbia & Montenegro, Slovakia, South Africa, Sri Lanka, Thailand, Tunisia, Turkey, Uganda and Vietnam. This classification was not defined a priori in the CRASH protocol as this was a secondary analysis, but it is a well accepted one and widely used.

4.2.5 **Performance of the models**

The performance of the models was assessed in terms of calibration and discrimination.

Discrimination was assessed using the c statistic (an equivalent concept to area under the receiver operator characteristic curve).[98] I used the terminology suggested by some authors that a c statistic over 0.7 is "acceptable", and over 0.8 as "good".[44]

Calibration was assessed graphically (plotting the observed versus expected probabilities of the outcomes by deciles of risk) and with the Hosmer-Lemeshow test. This statistic evaluates the difference between observed and expected probabilities where a small p value indicates lack of fit.

4.2.5.1 Internal validation

The internal validity of the final model was assessed by the bootstrap re-sampling technique. Regression models were estimated in 100 models. For each of 100

bootstrap samples the model was refitted and tested on the original sample to obtain an estimate of predictive accuracy corrected for overfitting.[44]

4.2.5.2 External validation

It is considered that a good prognostic model should be generalizable to populations different from that in which it was derived.[43] The external validation was conducted in an external cohort of 8509 patients with moderate and severe TBI from 11 studies (eight randomised controlled trials and three observational studies) conducted between 1984 and 1997 in HIC (the IMPACT dataset).[104]

4.2.6 **Web based score development**

A web calculator was planned to allow clinicians to estimate probabilities of the outcomes for individual patients. The estimated probability is obtained by combining the predictor values with the regression coefficients and obtaining the linear predictor for the model, which is then transformed to a predicted probability through the logistic transformation. The web calculator was planned that would be available at the CRASH-2 web page and would be accessible to clinicians internationally.

4.3 Results

4.3.1 **General characteristics**

The characteristics of the patient cohort are shown in table 4.1. The patients were more frequently men (81%) and from LMIC (75%). Over half (58%) of participants were included within three hours of injury. Road traffic crash (RTC) was the most common cause (65%) of injury.

For the majority of patients (80%) the current GCS was reported. Almost 40% presented with a severe TBI, as defined by GCS, while 30% presented with a moderate TBI and 30% with a mild TBI.

In 83% of the patients both pupils were reactive and in only 3% the pupil reactivity could not be assessed. Approximately 23% of patients presented a major extra-cranial injury.

A CT scan was performed in the majority (79%) of participants.

A total of 1,948 patients (19%) died in the first two weeks, 2,323 patients (24%) were dead at six months, and 3,556 patients (37%) were dead or severely dependent at six months.

4.3.1.1 Comparison between patients from LMIC and HIC

In comparison with patients from HIC, those from LMIC were on average 5 years younger and were more frequently male.

On average patients from LMIC were recruited 0.8 hours (48 minutes) later than patients from HIC and RTC was reported more frequently as a cause of TBI in LMIC.

All the patients had a GCS value but for a higher proportion of LMIC patients (86% vs. 61%) GCS was obtained at the moment of recruitment (current GCS). This means that in 14% of the patients from LMIC and in 39% of the patients from HIC the GCS was obtained before recruitment into the CRASH Trial (e.g. pre hospital assessment).

LMIC patents were categorized as less severe in comparison with patients from HIC. While almost 44% presented with a severe TBI in HIC, the proportion in this category was 38% for patients from LMIC.

In LMIC CT scan was performed in 76% of the patients, while for patients from HIC it was performed in 88%. Abnormal CT scan results were more commonly reported in LMIC patients.

Although patients from LMIC experienced higher mortality at 14 days (21% vs. 16%), there was no strong evidence of a difference in unfavourable outcome at six months.

Table 4-1 General characteristics of the study population

Prognostic Variables	Categories	Total (n=10,008) %	Low & Middle income countries (n=7,526) %	High income countries (n=2,482) %	p value*
Age mean in years (SD)		**37± (17.1)**	(35.8 ± 16)	(40.6 ± 19.4)	<0.001
	<20	**12.3**	12.5	11.8	
	20-24	**17.0**	17.8	14.4	
	25-29	**13.0**	13.5	11.2	
	30-34	**10.7**	10.9	10.1	
	35-44	**17.9**	18.5	15.9	
	45-54	**12.5**	12.3	13.3	
	≥ 55	**16.7**	14.5	23.4	
Gender					
	Female	**19.0**	18.3	21.1	0.002
	Male	**81.0**	81.7	78.9	
Hours since injury mean (SD)		**3.4 ±(2.7)**	(3.6 ± 2.8)	(2.8 ± 2.0)	<0.001
	< 1	**26.8**	24.0	35.2	
	1 to 3	**31.0**	30.1	33.7	
	>3	**42.3**	45.9	31.1	
Cause of head injury					
	RTC	**65.1**	69.9	50.2	<0.001
	Fall >2 meters	**13.3**	11.1	20.0	
	Other	**21.7**	19.0	29.8	

continuation of table 4.1

Prognostic Variables	Categories	Total (n=10,008) %	Low & Middle income countries (n=7,526) %	High income countries (n=2,482) %	p value*
Current Glasgow Coma Scale		**79.9**	86.3	60.7	<0.001
Total Glasgow coma score					
	Mild (14-13)	**30.2**	29.4	32.6	<0.001
	Moderate (12-9)	**30.4**	32.6	23.6	
	Severe (3-8)	**39.5**	38.0	43.8	
Pupil reactivity					
	Both reactive	**82.8**	83.5	80.7	
	One reactive	**6.3**	6.2	6.3	<0.001
	None reactive	**8.2**	8.0	9.1	
	Unable to assess	**2.7**	2.3	3.9	
Major extra-cranial injury					
	No	**77.3**	77.3	77.5	0.801
	Yes	**22.7**	22.7	22.5	

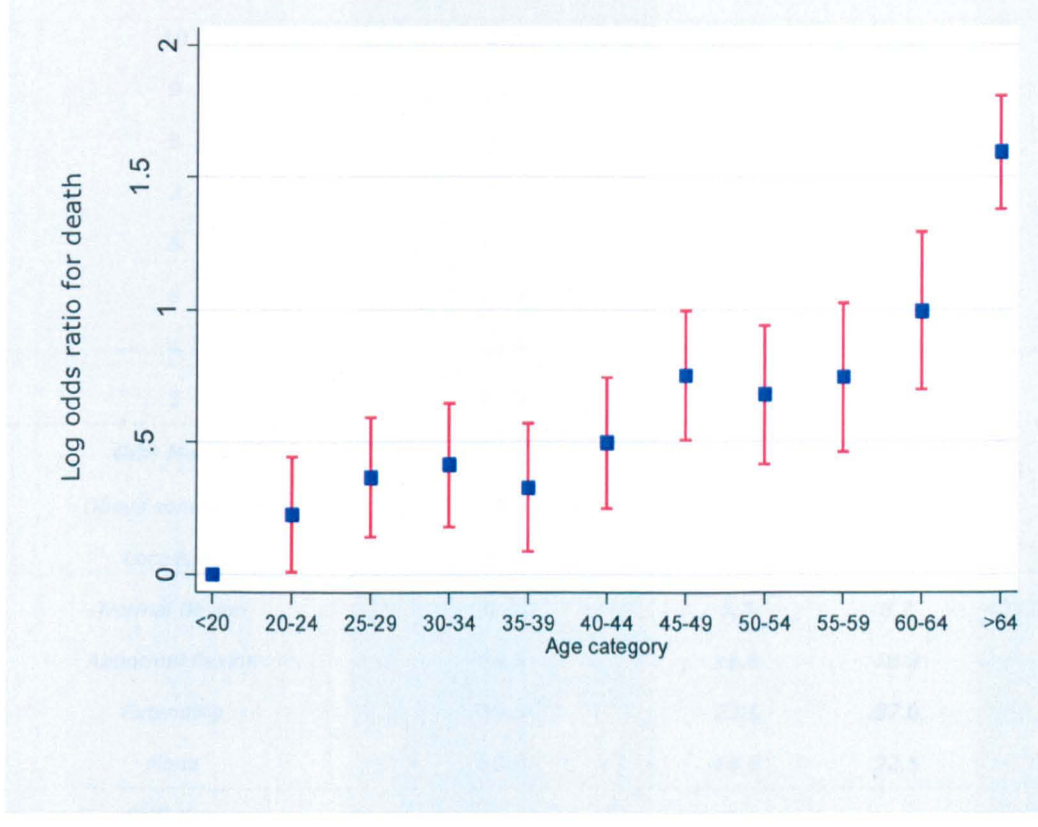
continuation of table 4.1

Prognostic Variables	Categories	Total (n=10,008) %	Low & Middle income countries (n=7,526) %	High income countries (n=2,482) %	p value*
CT scan					
	No scan	**21.1**	24.0	12.0	<0.001
	Normal scan	**22.8**	20.0	30.2	<0.001
	Petechial haemorrhages	**28.7**	28.7	28.7	0.970
	Obliteration of the third ventricle or basal cisterns	**23.4**	28.6	9.6	<0.001
	Subarachnoid bleed	**31.6**	33.5	26.4	<0.001
	Non evacuated haematoma	**27.1**	27.3	26.5	0.475
	Midline shift	**14.6**	15.9	11.1	<0.001
	Evacuated haematoma	**12.7**	14.4	7.9	<0.001
Outcomes					
	14 days mortality	**19.5**	20.7	16.0	<0.001
	6 months death or severe disability	**37.2**	36.8	38.5	0.150

4.3.2 **Relationship between age and 14 days mortality**

The relation between age and the log odds ratio for death at 14 days showed no clear association until the age of 40, after which there was a linear increase. Age was therefore modelled as 0 until 40 years and then equal to age minus 40 after this age (age=0 if age≤40, age-40 if age>40) (Figure 4.1).

Figure 4-1 Relation between age and mortality



4.3.3 **Selection of GCS variable**

In table 4.2 it is shown that there was strong evidence of an association between total GCS (and each of its components) and 14 days mortality.

Table 4-2 Association between GCS and mortality

	Odds ratios	95 % CI		p value for trend
Total Glasgow coma Scale				
14	**1**			
13	**1.5**	*1.05*	*2.17*	
12	**2.0**	*1.40*	*2.97*	
11	**2.5**	*1.73*	*3.71*	
10	**5.2**	*3.70*	*7.40*	
9	**6.1**	*4.4*	*8.7*	*<0.001*
8	**7.7**	*5.5*	*10.8*	
7	**9.7**	*7.1*	*13.3*	
6	**14.7**	*10.6*	*20.2*	
5	**26.7**	*19.2*	*37.3*	
4	**43.4**	*31.0*	*60.8*	
3	**29.8**	*21.7*	*40.9*	
GCS Motor				
Obeys commands	**1**			
Localising	**3.2**	*2.6*	*3.9*	
Normal flexion	**6.6**	*5.3*	*8.2*	*<0.001*
Abnormal flexion	**14.6**	*11.6*	*18.3*	
Extending	**29.5**	*23.1*	*37.6*	
None	**18.0**	*14.5*	*22.5*	
GCS Eye				
Spontaneous	**1**			
To sound	**0.9**	*0.7*	*1.2*	*<0.001*
To pain	**3.0**	*2.4*	*3.7*	
None	**8.3**	*6.9*	*9.9*	
GCS Verbal				
Orientated	**1.0**			
Confused speech	**1.3**	*0.9*	*2.1*	*<0.001*
Words	**2.7**	*1.7*	*4.1*	
Sounds	**8.6**	*5.6*	*13.0*	
None	**14.3**	*9.6*	*21.5*	

When different models to predict mortality at 14 days containing either total GCS or each of the different components were compared, there was strong evidence that the model with total GCS was better than any of the models containing each of the individual components (motor, eye or verbal) (p values for LHR tests were <0.001 for all of comparisons).

In table 4.3 it is shown that in terms of discrimination, total GCS also was superior to each of its components.

Table 4-3 Discrimination of total, motor, verbal and eye GCS and for mortality

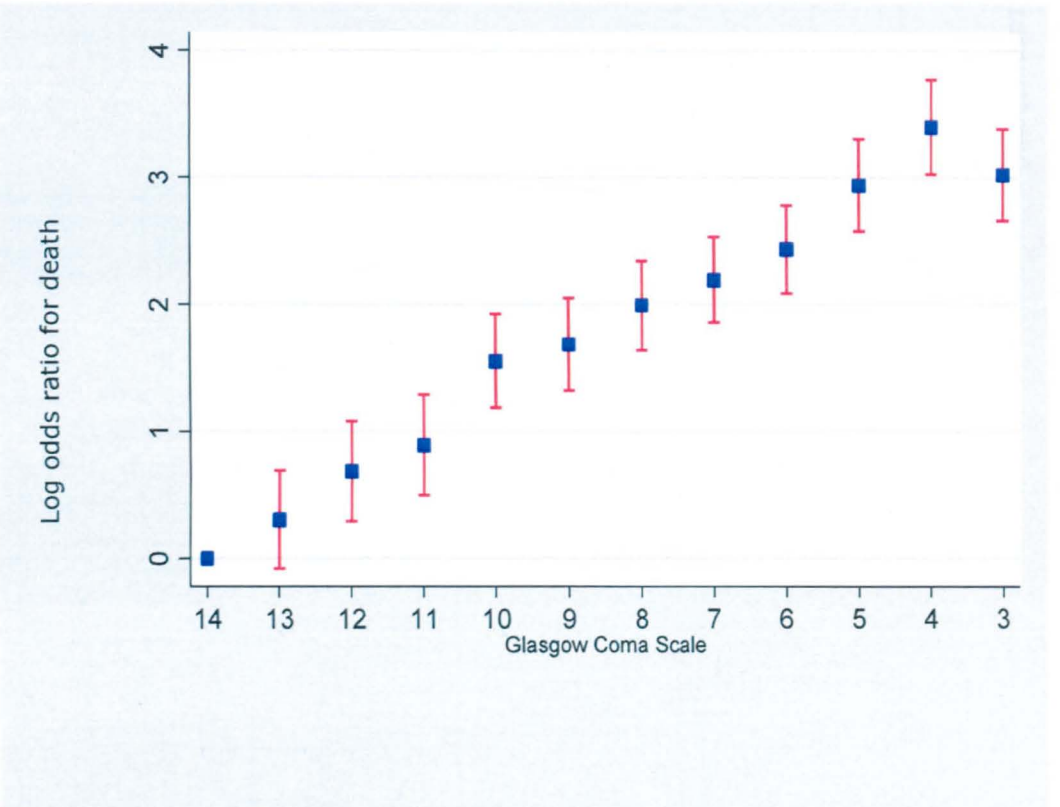
	GCS Total	GCS Motor	GCS Verbal	GCS Eye
C statistic	0.79	0.77	0.75	0.74
(95% CI)	(0.78-0.80)	(0.75-0.78)	(0.74-0.76)	(0.73-0.75)
P value¥	----	<0.001	<0.001	<0.001

¥ All comparison with GCS total model

4.3.4 **Relationship between GCS and 14 days mortality**

The relationship between GCS and 14 day mortality was reasonably linear and therefore GCS was included as a continuous variable (Figure 4.2).

Figure 4-2 Relation between GCS and mortality



4.3.5 **Interactions between income level and predictors**

Statistically significant interactions were found between country income level and several predictors, therefore, two models were developed, one for LMIC and another for HIC. Older age was a stronger predictor of 14 day mortality in HIC (interaction $p<0.001$). On the other hand, lower GCS was a stronger predictor in LMIC (interaction $p=0.003$) (figures 4.3 and 4.4). Obliteration of the third ventricle and a non-evacuated haematoma were both associated with a higher risk in HIC (interaction $p<0.001$ and $p=0.03$ respectively).

Figure 4-3 Relationship between GCS and mortality according to region

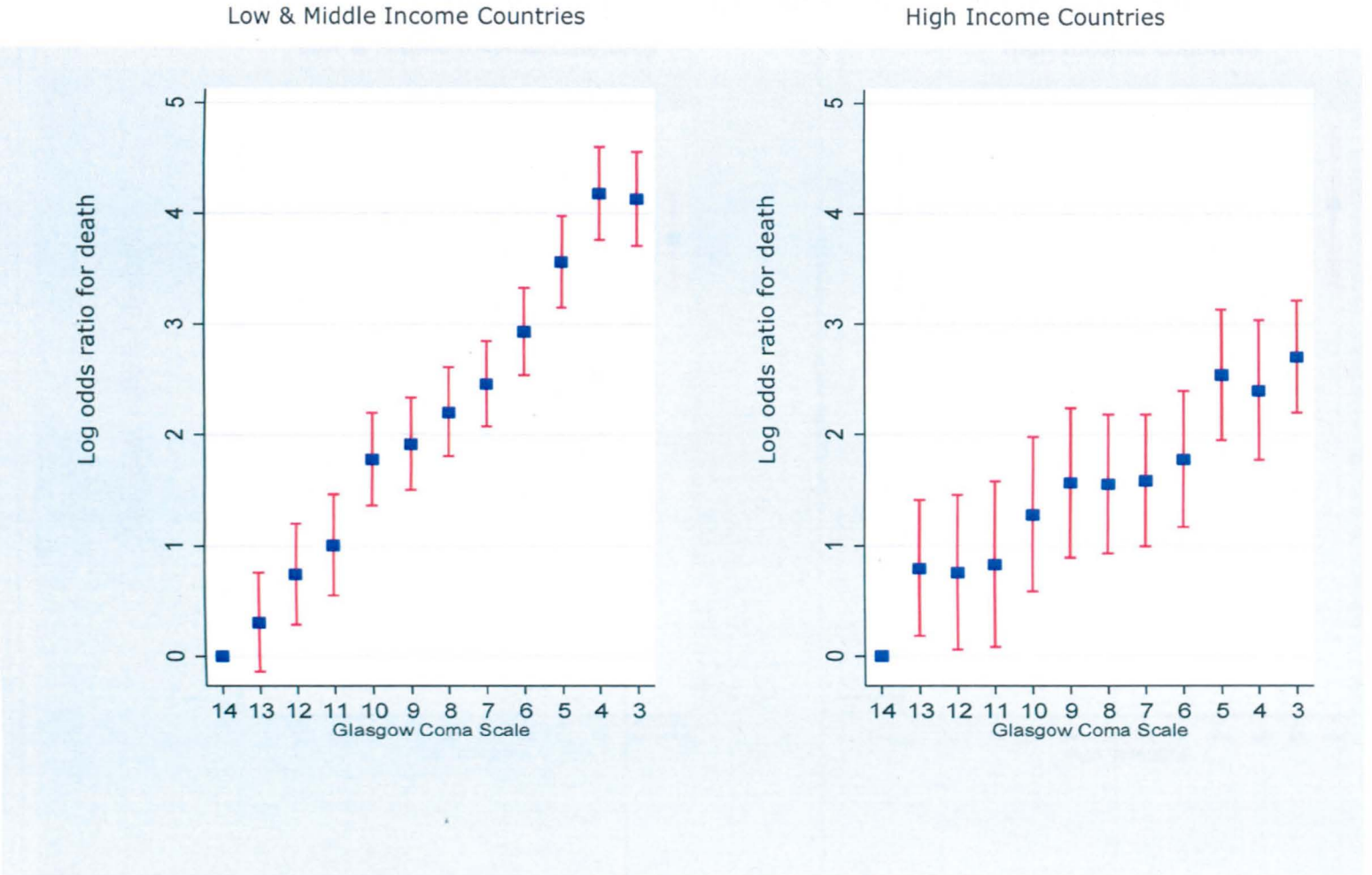
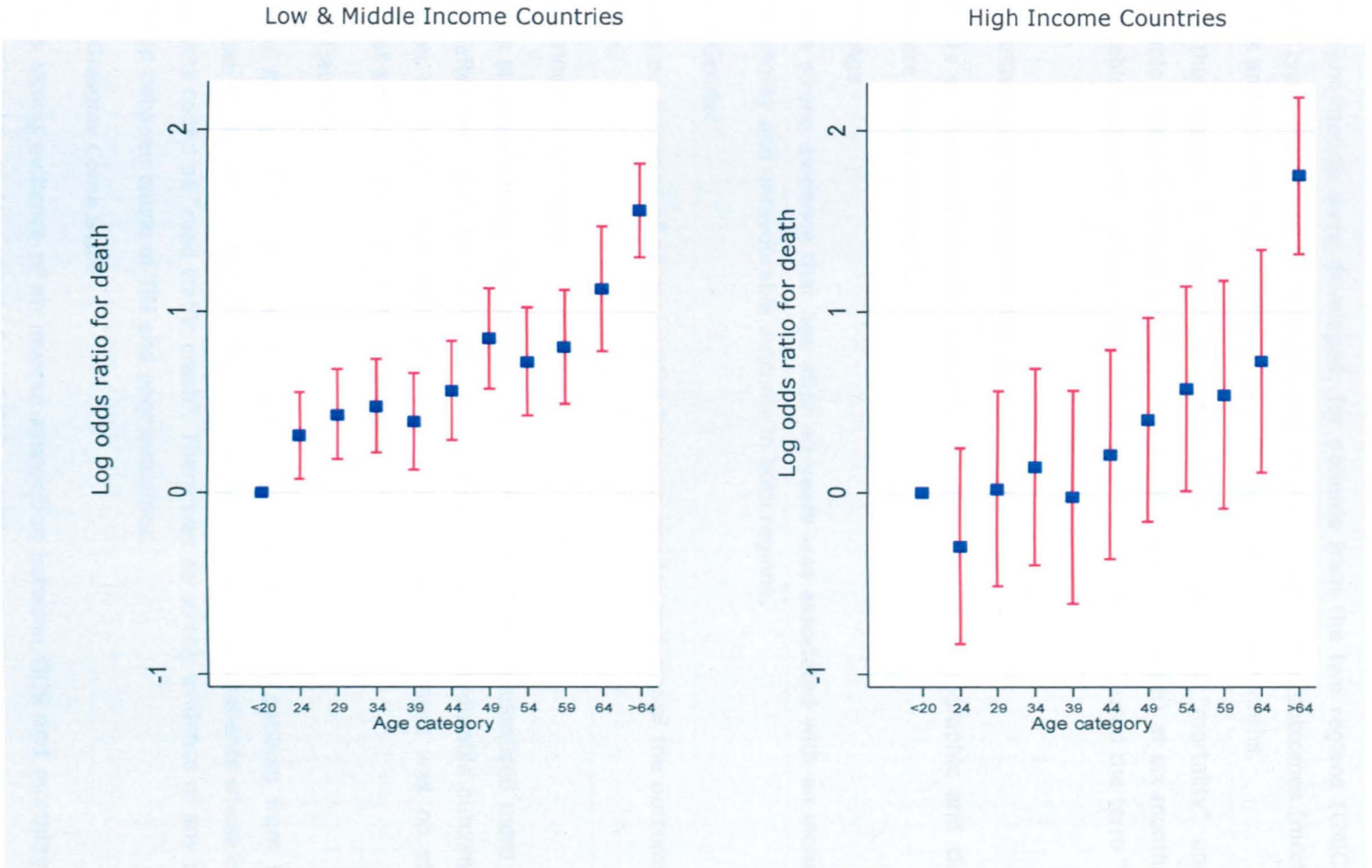


Figure 4-4 Relationship between age and mortality according to region



4.3.6 **Multivariable predictive models**

In total eight models were developed: for patients from the two regions (LMIC and HIC), two types (basic and CT models), and for predicting the two outcomes (mortality at 14 days and unfavourable outcome, as defined by GOS, at six months.

Through this section I referred to mortality at 14 days as "mortality" and to unfavourable outcome defined by GOS (severe disability and death) at six months, as "unfavourable outcome". When I am referring to both outcomes I used the term "poor outcome".

4.3.6.1 Multivariable analysis for the basic models

The results of the multivariable analysis including all the demographic and clinical variables are shown in table 4.4.

4.3.6.1.1 Age

There was strong evidence that age after 40 years was associated with an increased risk of mortality and unfavourable outcome in both regions.

4.3.6.1.2 Gender

There was no evidence of an association between gender and any of the outcomes in any region.

4.3.6.1.3 Hours since injury

There was strong evidence that patients from LMIC who were randomised more than one hour after the injury had a higher risk of mortality and unfavourable outcome, in comparison to those randomised in the first hour. However, there was no strong evidence of such an association in HIC patients.

4.3.6.1.4 Cause of TBI

There was strong evidence of a decreased risk in mortality in patients from LMIC whose cause of TBI was coded as "other", in comparison to those patients whose cause of injury was coded as "road traffic crash". There was no strong evidence of any other relationship between cause of TBI and poor outcomes.

4.3.6.1.5 Glasgow Coma Scale

There was strong evidence of an inverse association between GCS and mortality and unfavourable outcome for patients from LMIC and HIC.

4.3.6.1.6 Pupil reactivity

There was strong evidence of an increased risk of mortality and unfavourable outcome for patients with either one or both pupils not reactive in comparison to those patients with two reactive pupils in both LMIC and HIC.

4.3.6.1.7 Major extra-cranial injury

There was strong evidence of an increased risk of unfavourable outcome for patients with a major extracranial injury in comparison to those without such an injury in both LMIC and HIC. There was also strong evidence of an increased risk in mortality for HIC, and a weaker association was found for LMIC.

Considering the Z score value, GCS was the strongest predictor of outcome in LMIC and age was the strongest predictor in HIC, while the absence of pupil reactivity was the third strongest predictor in both regions.

4.3.6.1.8 Final basic models

For the basic model only the four predictors which were associated with poor outcome in both regions were retained in the model: age, GCS, pupil reactivity and the presence of major extra-cranial injury.

4.3.6.2 Multivariable analysis for the CT models

The results of the multivariable analysis including those variables selected for the basic models (age, GCS, pupil reactivity and presence of major extra-cranial injury) and all the CT characteristics are presented in table 4.5.

4.3.6.2.1 Basic model predictors

There was still strong evidence of association for all the variables from the basic model with poor outcome in both regions. There was a further decrease in the strength of the association between presence of major extra-cranial injury and mortality in LMIC.

4.3.6.2.2 CT scan predictors

The following CT characteristics were strongly associated with the outcomes in addition to the predictors included in the basic models: obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift and non-evacuated haematoma. There was also strong evidence of an association between presence of petechial haemorrhages and poor outcome in LMIC and some weaker evidence of such an association in patients from HIC. There was no evidence of an association between evacuated haematoma and poor outcome in either region.

Considering the Z value, obliteration of the third ventricle and midline shift were the strongest CT scan predictors of mortality, and non-evacuated haematoma was the strongest predictor of unfavourable outcome.

4.3.6.2.3 Final CT models

For the CT model I included all the variables in the basic model and the following CT scan characteristics: presence of petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift and non-evacuated haematoma.

Table 4-4 Multivariable analysis for the basic models‡

	14 days mortality								Six months death or severe disability							
	High Income Countries (n=2,294)				Low & Middle Income Countries (n=7,412)				High Income Countries (n=2,185)				Low & Middle Income Countries (n=7,119)			
Prognostic Variables	O.R. (95% CI)	95% CI Lower	95% CI Upper	Z score	O.R. (95% CI)	95% CI Lower	95% CI Upper	Z score	O.R. (95% CI)	95% CI Lower	95% CI Upper	Z score	O.R. (95% CI)	95% CI Lower	95% CI Upper	Z score
Age*	1.72	1.62	1.83	14.08	1.47	1.40	1.54	14.10	1.73	1.64	1.82	15.99	1.70	1.63	1.77	18.58
Gender																
Female	1				1				1							
Male	0.95	0.69	1.31	-0.32	1.19	0.99	1.43	1.82	0.86	0.67	1.12	-1.12	0.96	0.82	1.13	-0.45
Hours since injury																
< 1 hour	1				1				1				1			
1-3hours	1.04	0.74	1.46	0.21	1.23	1.01	1.50	2.05	0.98	0.76	1.27	-0.13	1.19	1.00	1.42	1.99
> 3 hours	0.85	0.61	1.19	-0.96	1.26	1.05	1.51	2.46	1.03	0.80	1.33	0.24	1.30	1.11	1.53	3.23
Cause																
Road Traffic Crash	1				1				1				1			
Fall	1.21	0.86	1.70	1.08	1.05	0.84	1.31	0.43	1.25	0.95	1.64	1.61	1.07	0.88	1.30	0.65
Other	0.83	0.58	1.18	-1.04	0.75	0.62	0.91	-2.88	1.00	0.78	1.30	0.03	0.96	0.81	1.13	-0.54
Glasgow Coma Scale¥	1.24	1.19	1.29	10.22	1.39	1.35	1.42	25.60	1.22	1.18	1.25	12.84	1.42	1.39	1.45	30.64
Pupil reactivity																
Both reactive	1				1				1				1			
One reactive	2.57	1.65	4.00	4.17	1.91	1.53	2.39	5.69	2.43	1.62	3.66	4.26	2.01	1.59	2.56	5.81
None reactive	5.49	3.70	8.15	8.45	3.92	3.14	4.90	12.07	3.28	2.20	4.89	5.85	4.54	3.38	6.11	10.03
Major extra-cranial injury																
No	1				1				1				1			
Yes	1.53	1.11	2.09	2.62	1.15	0.99	1.34	1.78	1.62	1.26	2.07	3.82	1.73	1.51	1.99	7.76

‡ *Includes age, gender, hours since injury, cause of injury, GCS, pupil reactivity and presence of major extra cranial injury * Per 10 years increase after 40 years* ¥ *Per decrease of each value of the GCS*

Table 4-5 Multivariable analysis for the CT models ‡

	14 days mortality								Six months death or severe disability							
	High Income Countries (n=2,030)				Low & Middle Income Countries (n=5,635)				High Income Countries (n=1,955)				Low & Middle Income Countries (n=5,394)			
Prognostic Variables	**O.R. (95% CI)**	**95% CI**		**Z score**	**O.R. (95% CI)**	**95% CI**		**Z score**	**O.R. (95% CI)**	**95% CI**		**Z score**	**O.R. (95% CI)**	**95% CI**		**Z score**
		Lower	*Upper*			*Lower*	*Upper*			*Lower*	*Upper*			*Lower*	*Upper*	
Age*	1.73	*1.62*	*1.84*	*13.33*	1.46	*1.39*	*1.54*	*12.54*	1.73	*1.63*	*1.83*	*14.94*	1.72	*1.64*	*1.81*	*17.74*
Glasgow Coma Scale ¥	1.18	*1.12*	*1.23*	*6.87*	1.27	*1.24*	*1.31*	*16.68*	1.18	*1.14*	*1.22*	*9.83*	1.34	*1.30*	*1.37*	*22.32*
Pupil reactivity																
Both reactive	1				1				1				1.00			
One reactive	2.00	1.25	3.20	2.88	1.45	1.14	1.86	2.97	2.12	1.39	3.24	3.47	1.54	1.20	1.99	3.35
None reactive	4.00	*2.58*	*6.20*	*6.21*	3.12	*2.46*	*3.97*	*9.31*	2.83	*1.84*	*4.35*	*4.73*	3.56	*2.60*	*4.87*	*7.92*
Major extra-cranial injury																
No	1				1				1				1.00			
Yes	1.53	*1.10*	*2.13*	*2.53*	1.08	*0.91*	*1.28*	*0.89*	1.55	*1.20*	*1.99*	*3.37*	1.61	*1.38*	*1.88*	*6.03*
CT scan																
Petechial haemorrhages	1.15	*0.83*	*1.59*	*0.84*	1.26	*1.07*	*1.47*	*2.82*	1.21	*0.95*	*1.55*	*1.56*	1.49	*1.29*	*1.73*	*5.33*
Obliteration of the third ventricle or basal cisterns	4.46	*2.97*	*6.68*	*7.23*	1.99	*1.69*	*2.35*	*8.25*	2.21	*1.49*	*3.30*	*3.95*	1.53	*1.31*	*1.79*	*5.30*
Subarachnoid bleed	1.48	*1.09*	*2.02*	*2.51*	1.33	*1.14*	*1.55*	*3.60*	1.62	*1.26*	*2.08*	*3.79*	1.20	*1.04*	*1.39*	*2.49*
Midline shift	2.77	*1.82*	*4.21*	*4.77*	1.78	*1.44*	*2.21*	*5.35*	1.93	*1.30*	*2.87*	*3.24*	1.86	*1.48*	*2.32*	*5.42*
Evacuated haematoma	0.78	*0.45*	*1.35*	*-0.89*	1.01	*0.80*	*1.28*	*0.08*	1.35	*0.87*	*2.08*	*1.33*	1.18	*0.94*	*1.48*	*1.46*
Non evacuated haematoma	2.06	*1.49*	*2.84*	*4.40*	1.48	*1.24*	*1.76*	*4.43*	1.72	*1.33*	*2.22*	*4.15*	1.68	*1.43*	*1.97*	*6.34*

‡ Includes age, GCS, pupil reactivity, presence of major extra cranial injury, and all CT characteristics Per 10 years increase after 40 years¥ Per decrease of each value the GCS*

4.3.7 **Performance**

4.3.7.1 Discrimination

4.3.7.1.1 Comparison of discrimination between the models

Table 4.6 displays the discrimination of each of the eight models developed. For estimating the c statistic for the basic and CT models, I used the same samples (those with a CT scan available) so a direct comparison could be made.

All models showed good discrimination, with c statistics over 0.80. CT models showed higher discrimination than basic models.

Table 4-6 Discrimination of the prognostic models

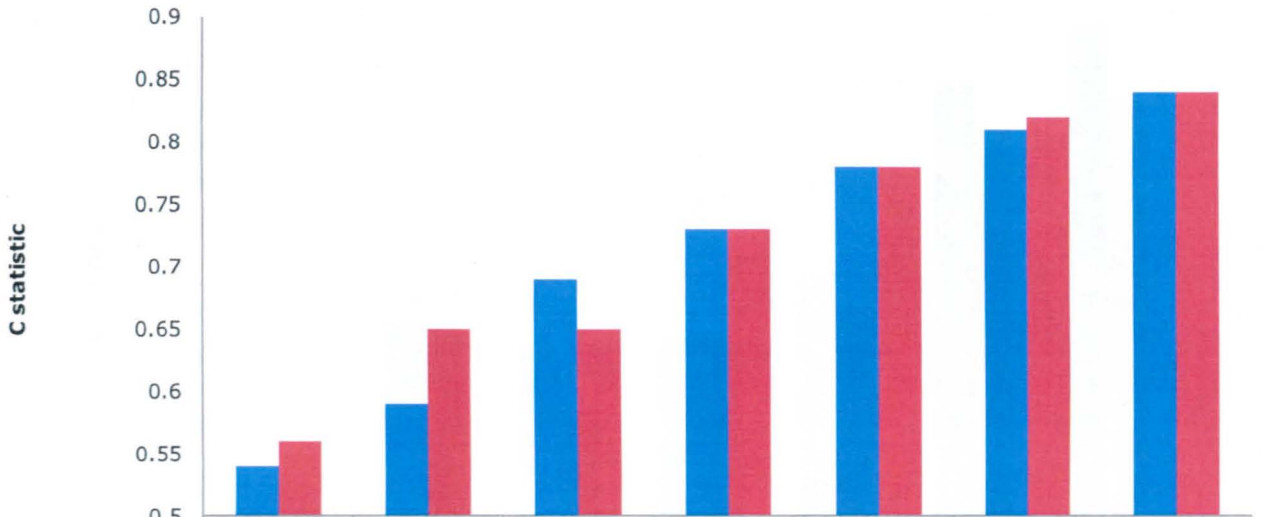
		C statistic	
		14 days Mortality	**Six months unfavourable outcome**
LMIC	Basic Model	0.81 (0.80-0.83)	0.82 (0.81-0.84)
	CT Model	0.84 (0.82-0.85)	0.84 (0.83-0.85)
HIC	Basic Model	0.84(0.82-0.87)	0.81 (0.79-0.83)
	CT Model	0.88 (0.86-0.90)	0.83 (0.81-0.84)

The internal validation through bootstrapping revealed no over optimism in any of the final model's predictive c statistics.

4.3.7.1.2 Comparison of discrimination between the models and individual predictors

Figure 4.5 and figure 4.6 display the c statistics for each of the individual predictors included in the basic model for both outcomes (14 days mortality and six months unfavourable outcome) and compares them with the c statistics of the basic and CT models. For all the estimation I used the same sample (those patients with a CT scan available) so that a direct comparison could be made. GCS was analysed as a continuous variable and as a categorical variable (mild, moderate and severe) because the latter is the way it is commonly used in clinical practice. It can be seen that both models had higher discrimination than any of the single predictors; although GCS in LMIC when analyzed as a continuous variable showed an acceptable discrimination.

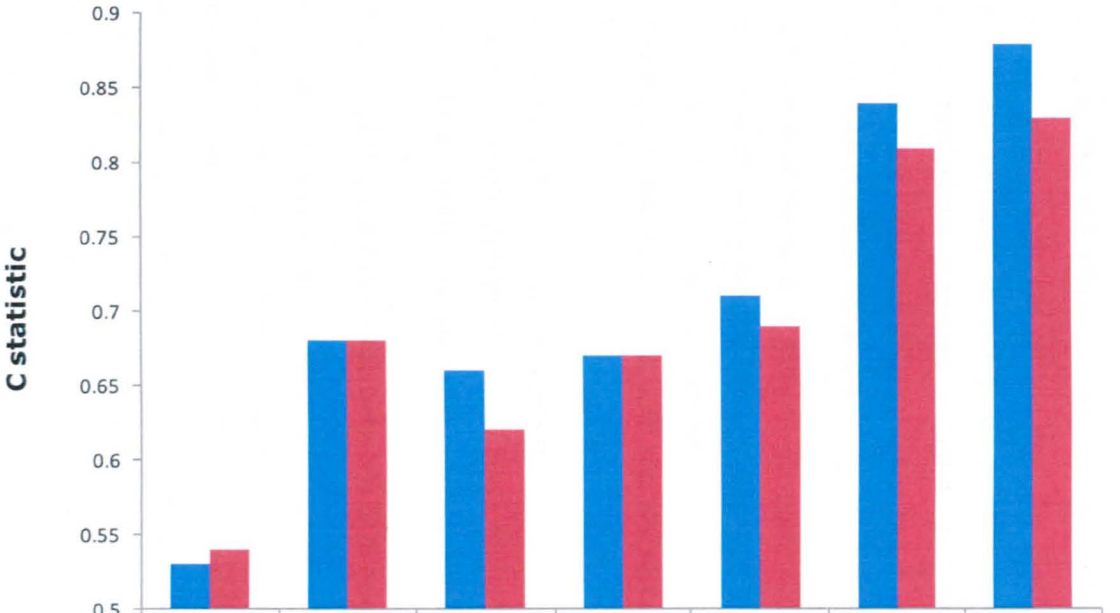
Figure 4-5 Discrimination of predictors, basic and CT models in LMIC



	MEI	Age	Pupils	GCS categorical	GCS continuous	Basic Model	CT Model
14 days mortality	0.54	0.59	0.69	0.73	0.78	0.81	0.84
UO at 6 months	0.56	0.65	0.65	0.73	0.78	0.82	0.84

MEI: Major extracranial injury GCS: Glasgow Coma Scale UO: Unfavourable outcome

Figure 4-6 Discrimination of predictors, basic and CT models in HIC



	MEI	Age	Pupils	GCS categorical	GCS continuous	Basic Model	CT Model
14 days mortality	0.53	0.68	0.66	0.67	0.71	0.84	0.88
UO at 6 months	0.54	0.68	0.62	0.67	0.69	0.81	0.83

MEI: Major extracranial injury GCS: Glasgow Coma Scale UO: Unfavourable outcome

4.3.7.2 Calibration

4.3.7.2.1 Calibration of basic and CT models

Figures 4.7 to 4.14 display the relationship between predicted and observed probability of outcome according to deciles of risk for the CT and basic models in both regions.

The red line on the 45° shows the line of perfect prediction. Each circle represents a decile of risk. If the circle lies on the line, the prediction coincided with the observed frequency of the outcome in that group of patients. If the circle lies below the line it means that the model predicted a higher probability in comparison to what was observed, and if the circle lies above the line the model predicted a lower probability of the outcome than that was observed.

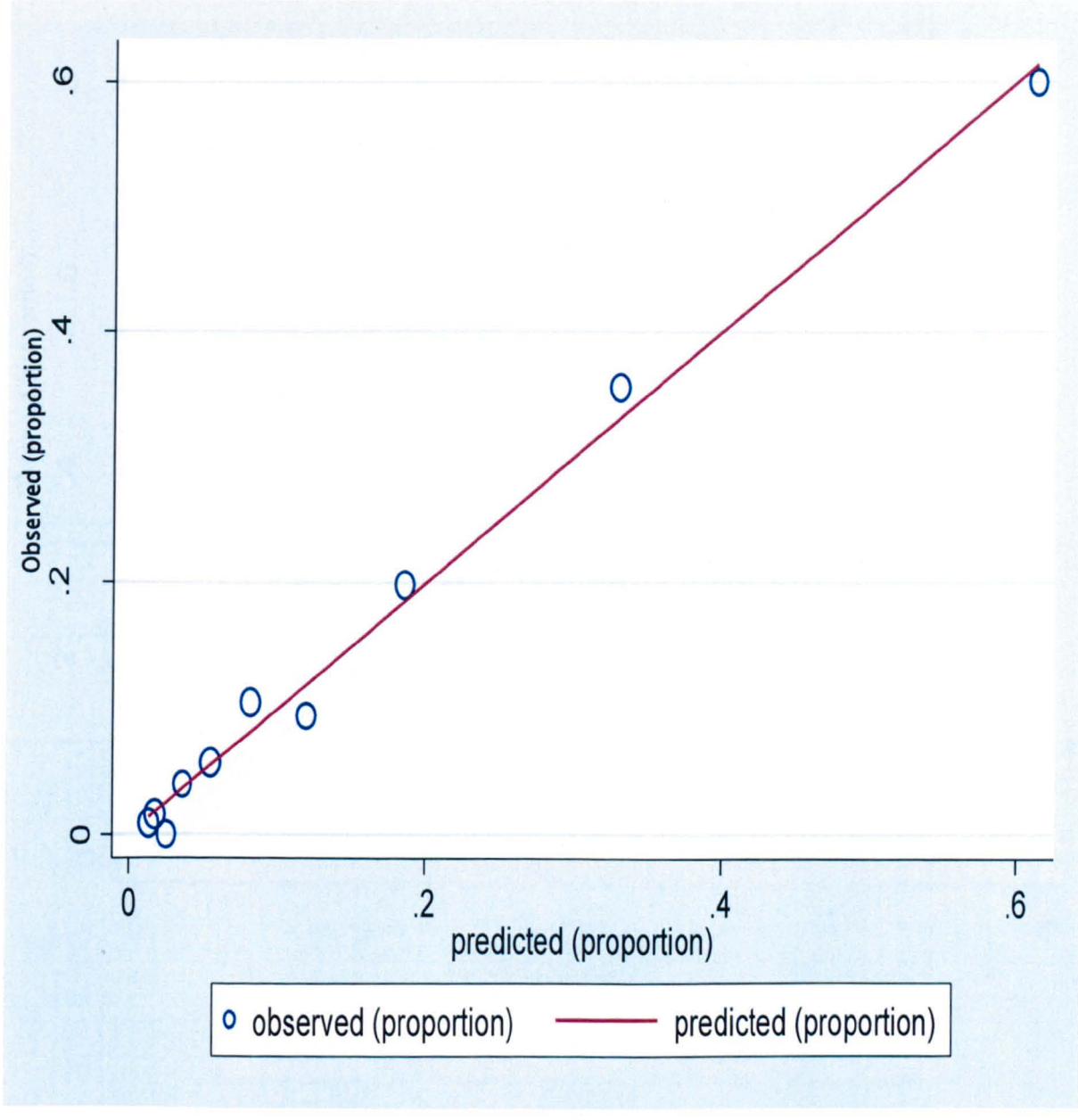
Graphically most of the models showed adequate calibration, but in general the models from LMIC showed an apparently better calibration as the circles were closer to the 45° line. The other common pattern observed in most of the models was that the groups with intermediate risk were those with the worse calibration (circles further from the 45° line).

When analysed with the Hosmer Lemeshow test all the models had good calibration, that is with $p > 0.05$, except the CT models for LMIC patients for 14 days mortality ($p=0.04$) and six months unfavourable outcome ($p=0.03$).

4.3.7.2.2 Comparison of calibration between GCS and basic model in LMIC

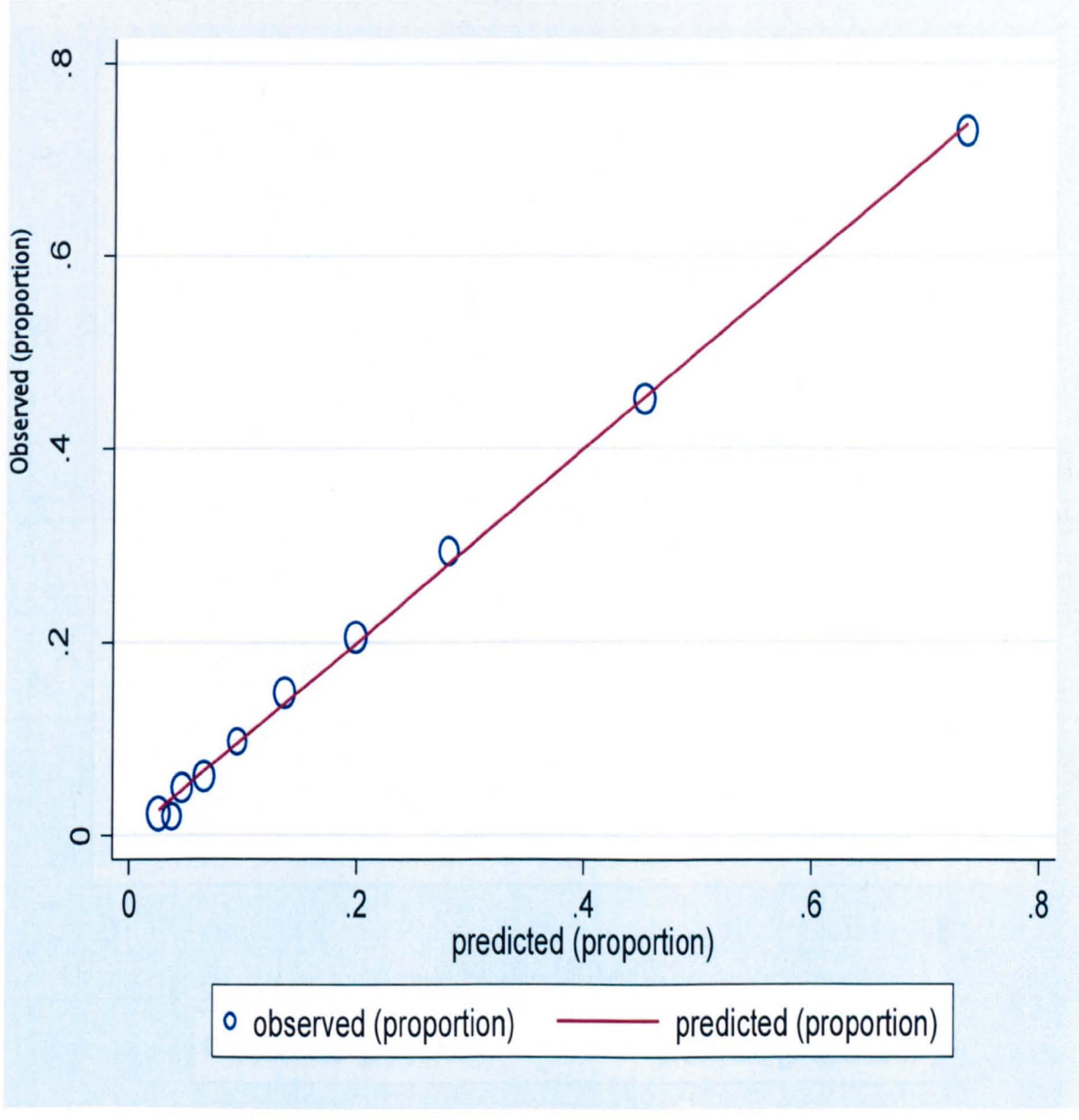
As GCS when analyzed as a continuous variable showed good discrimination for predicting mortality in LMIC, I compared graphically its calibration with that of the basic model. It can be seen in figure 4.15 that calibration, particularly for groups with intermediate risk was better for the basic model. Similarly when tested with the Homer Lemeshow test, there was evidence of lack of calibration for GCS (Hosmer Lemeshow test=0.03) but not for the basic model (Hosmer Lemeshow test=0.39).

Figure 4-7 Calibration of basic model for mortality in HIC



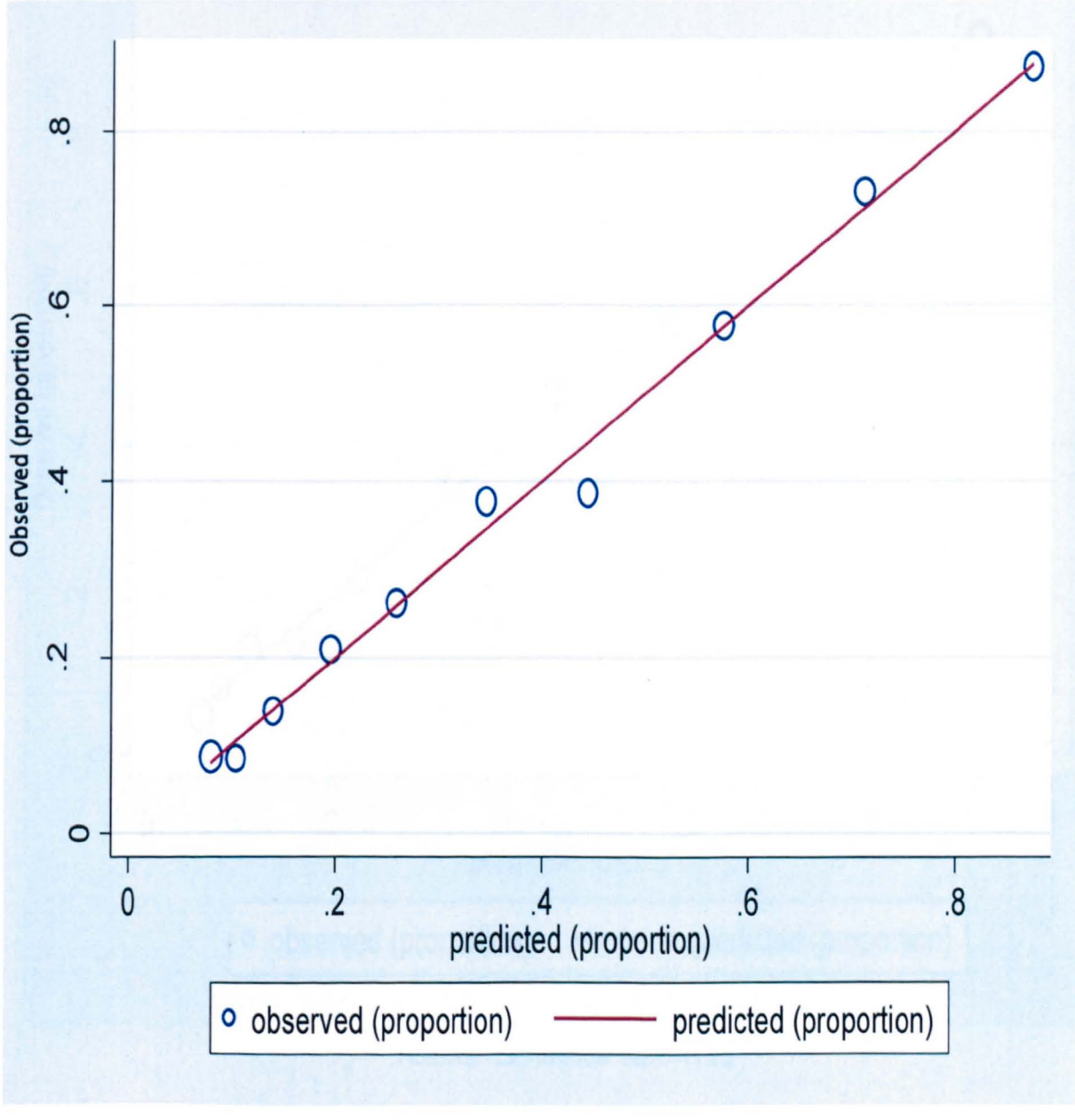
Hosmer-Lemeshow test=0.27

Figure 4-8 Calibration of basic model for mortality in LMIC



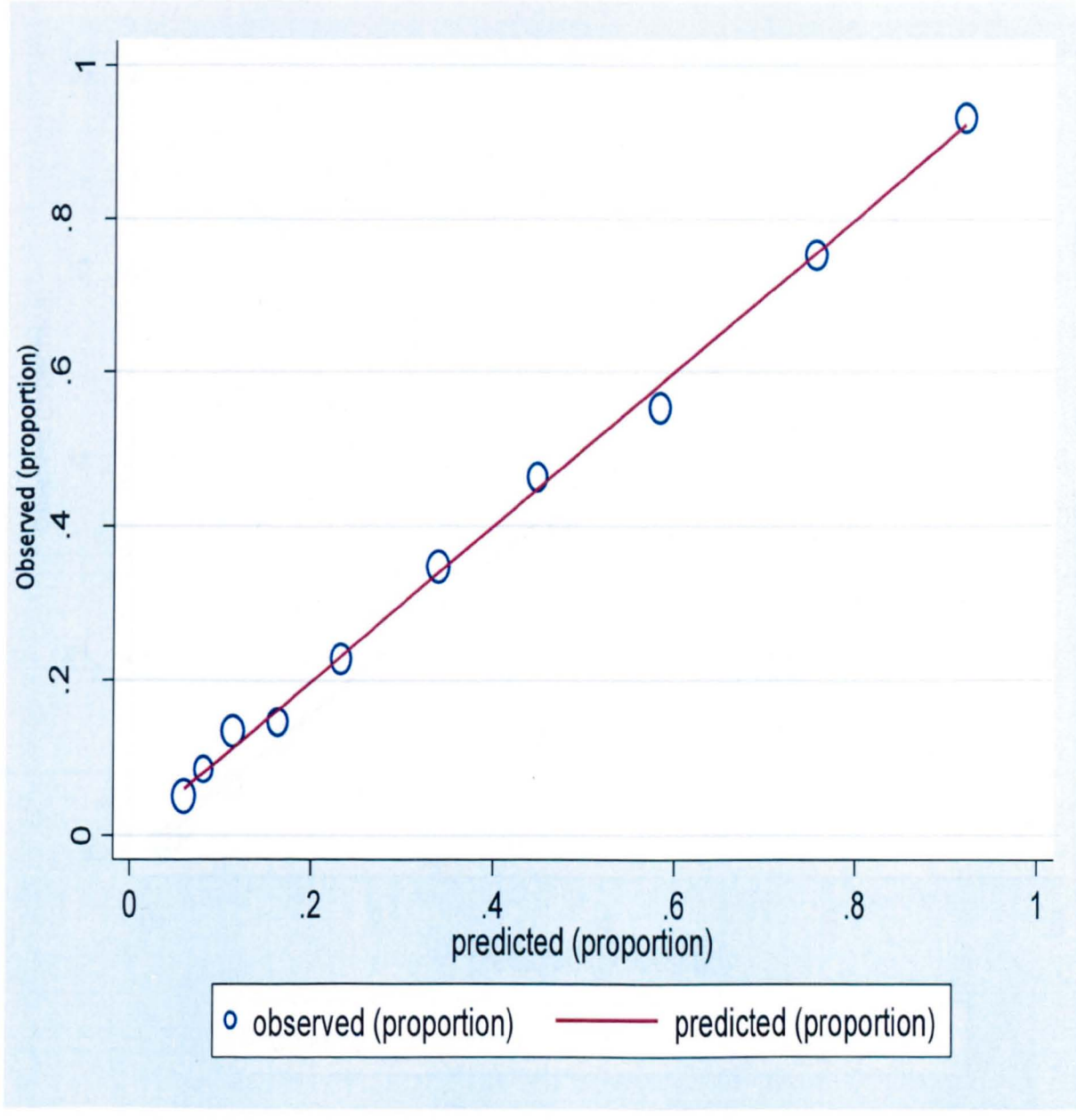
Hosmer-Lemeshow test=0.39

Figure 4-9 Calibration of basic model for unfavourable outcome in HIC



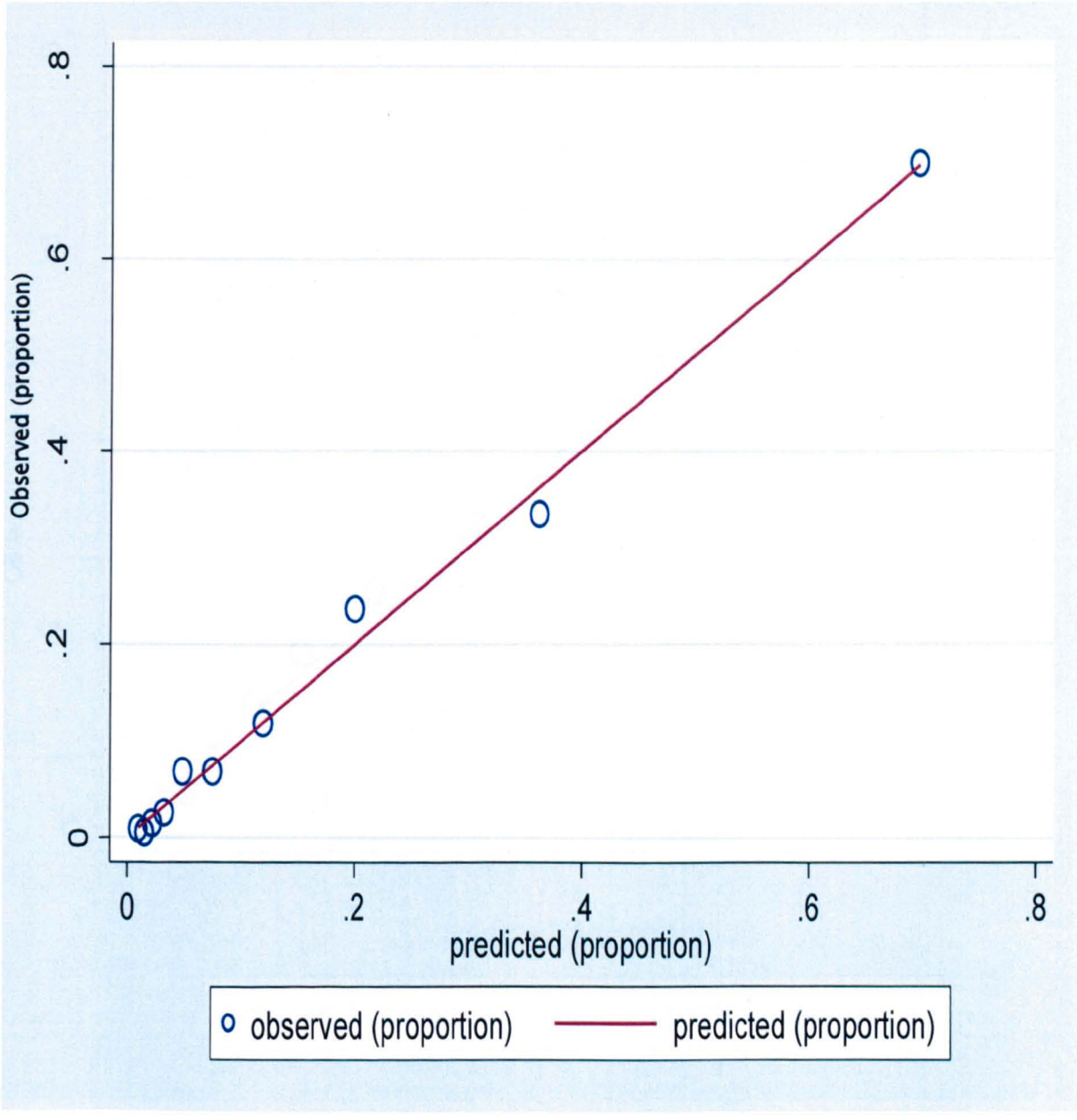
Hosmer-Lemeshow test=0.70

Figure 4-10 Calibration of basic model for unfavourable outcome in LMIC



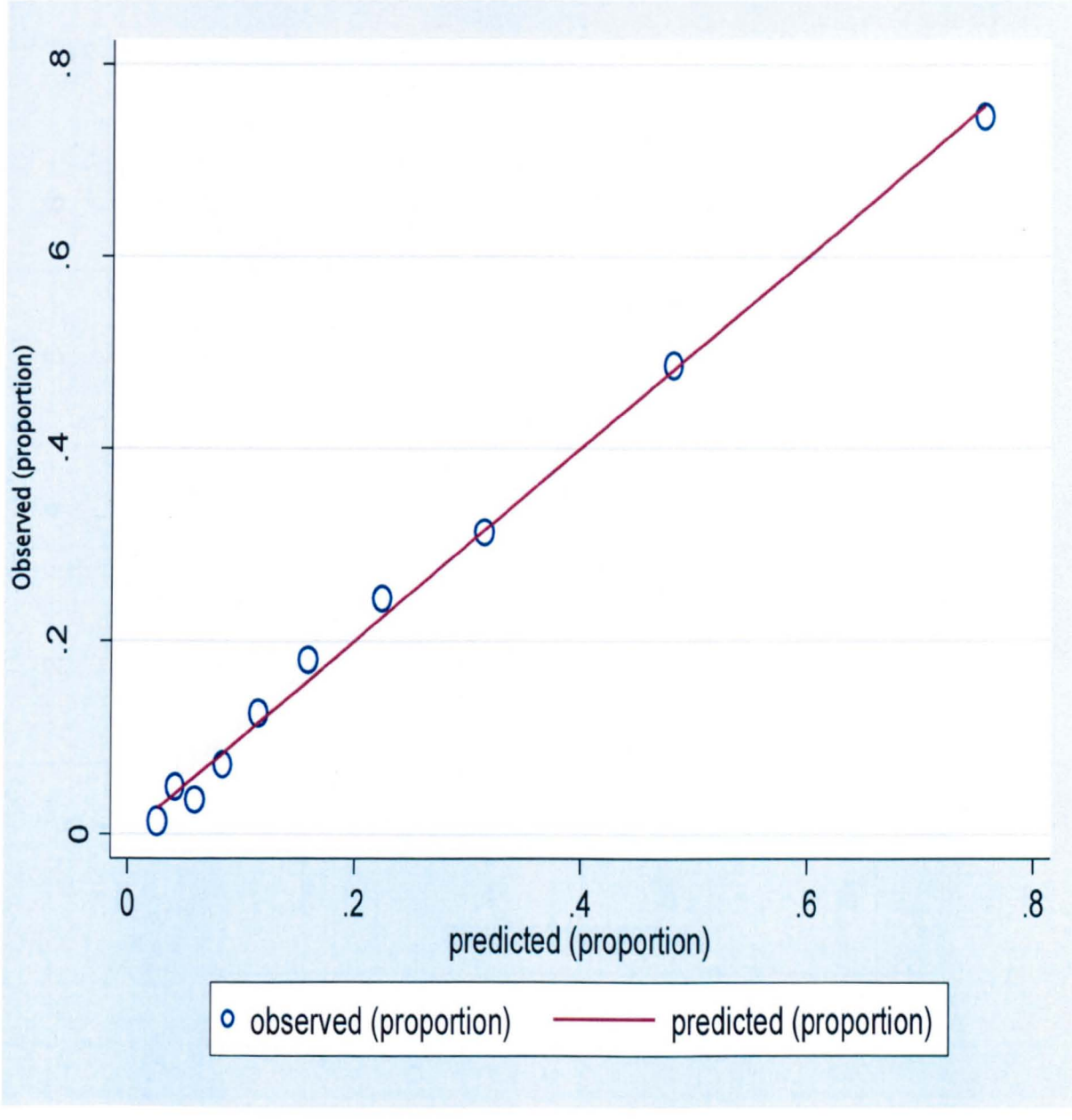
Hosmer-Lemeshow test=0.20

Figure 4-11 Calibration of CT model for mortality in HIC



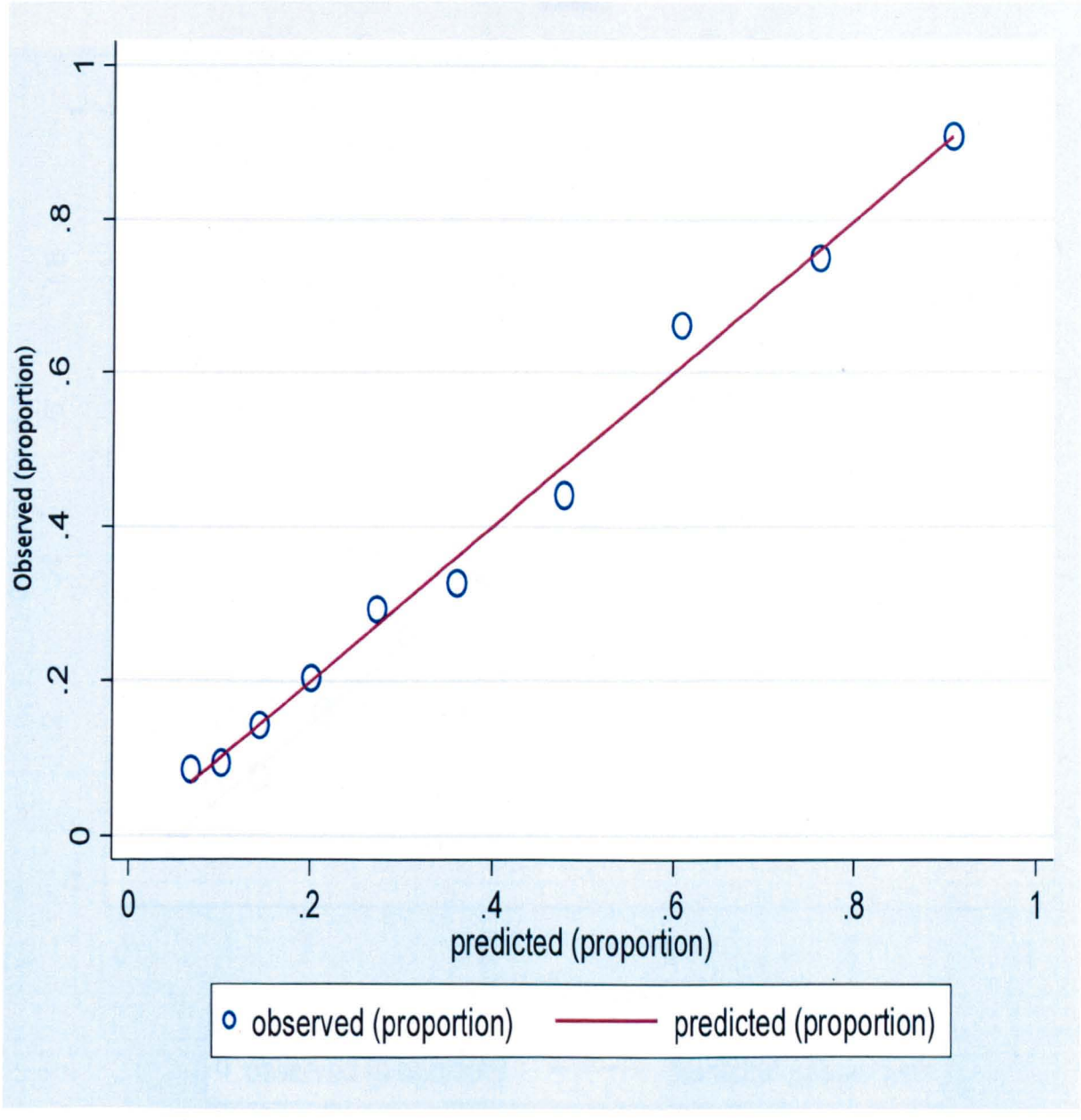
Hosmer-Lemeshow test =0.60

Figure 4-12 Calibration of CT model for mortality in LMIC



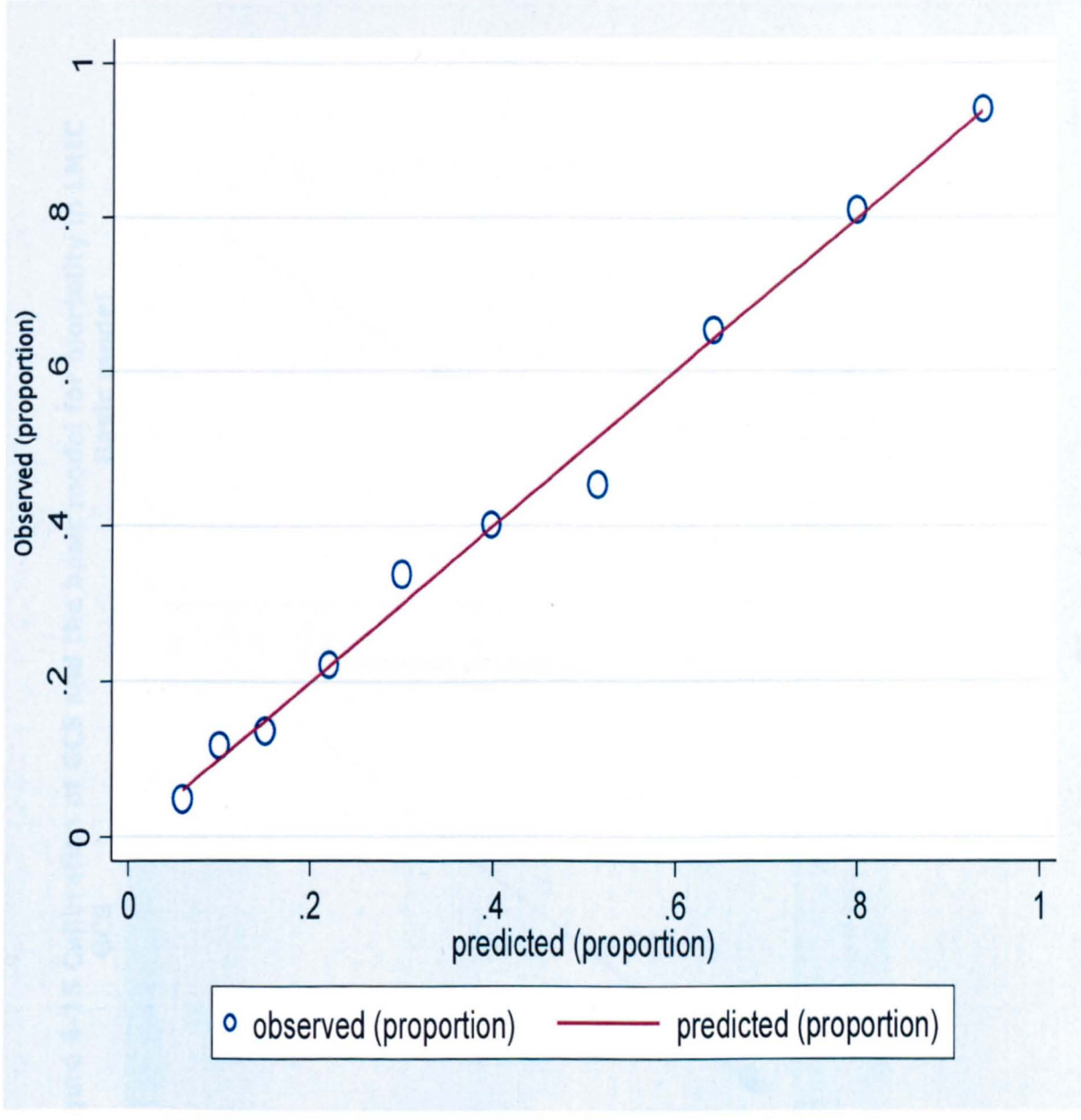
Hosmer-Lemeshow test =0.04

Figure 4-13 Calibration of CT model for unfavourable outcome in HIC



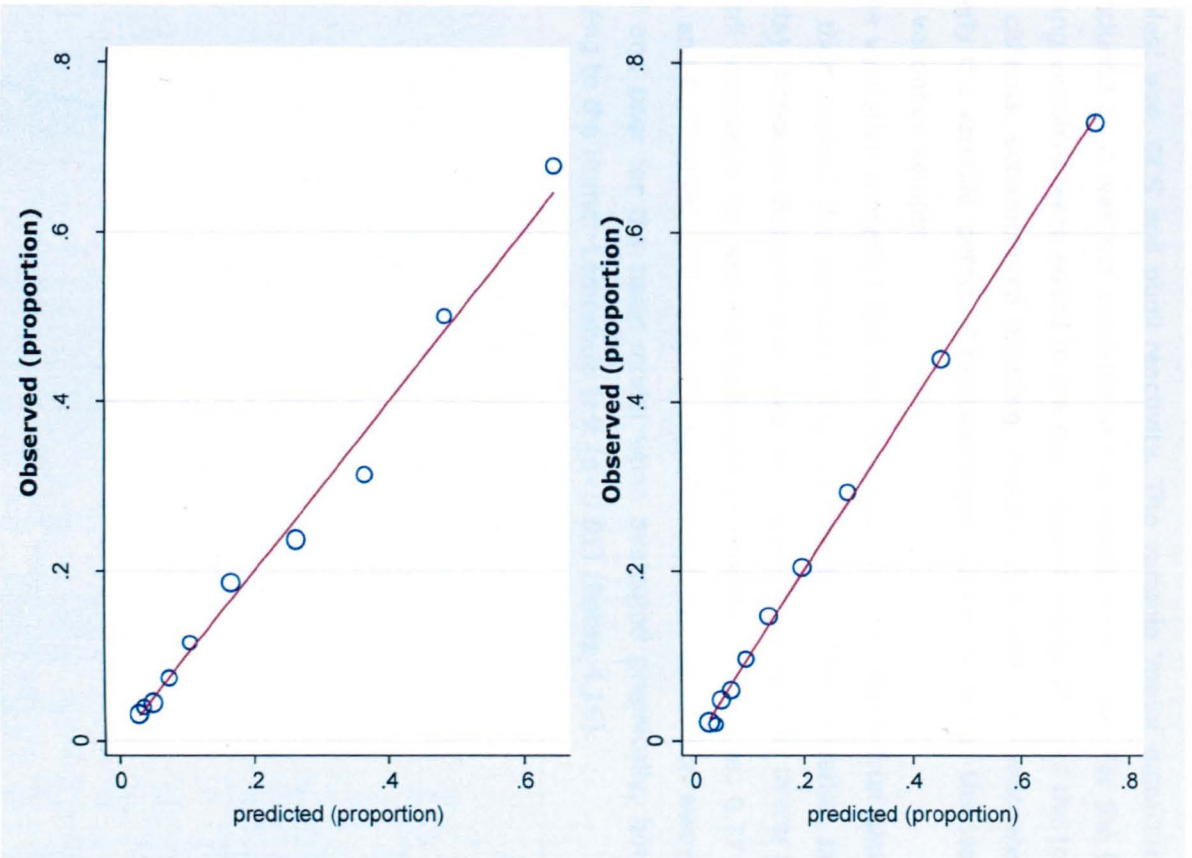
Hosmer-Lemeshow test =0.63

Figure 4-14 Calibration of CT model for unfavourable outcome in LMIC



Hosmer-Lemeshow test =0.03

Figure 4-15 Calibration of GCS and the basic model for mortality in LMIC

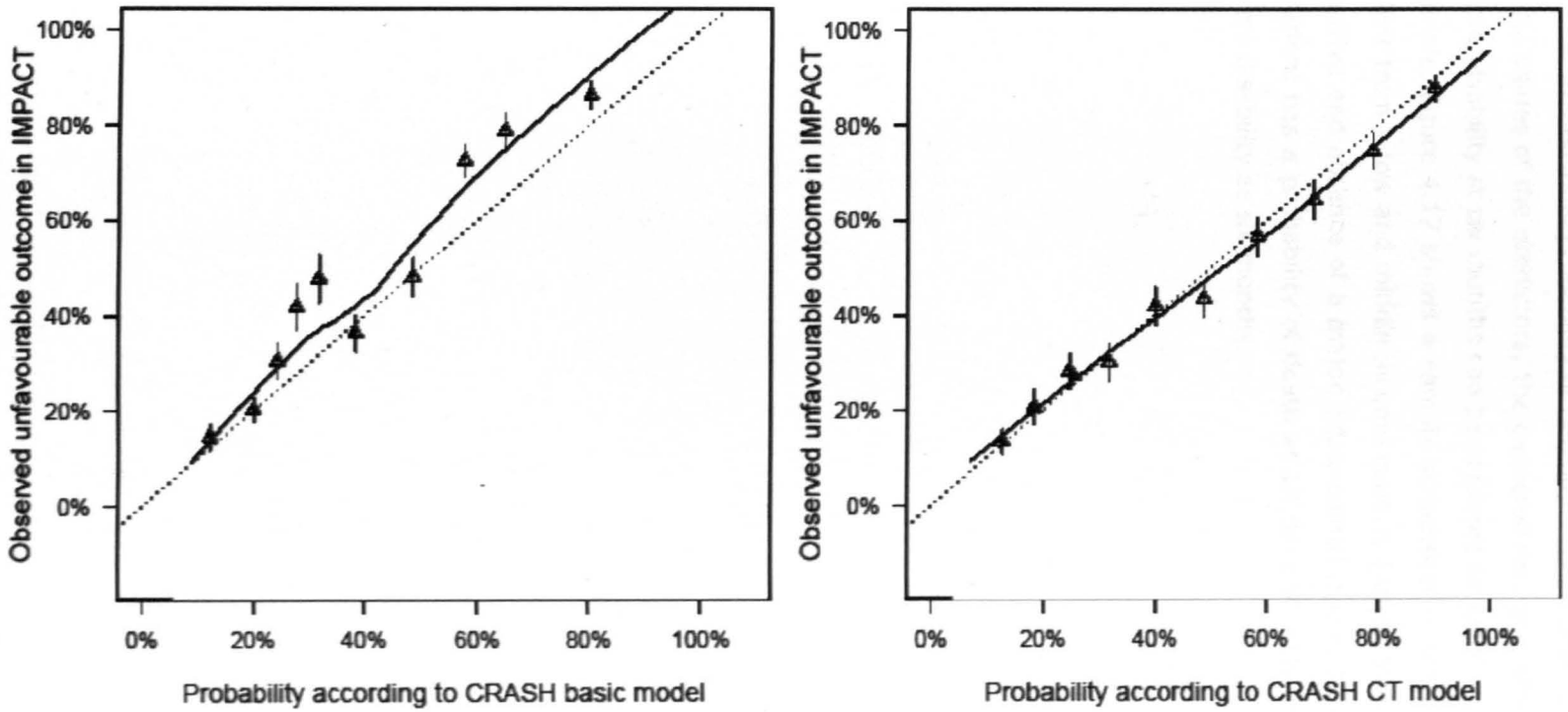


4.3.8 **External validation**

Because an external cohort of patients from LMIC was not available, validation was only performed for the models that included patients from HIC. The IMPACT dataset used for the validation included patients with moderate and severe TBI (GCS <13) and did not include data on mortality at 14 days. Therefore only models for unfavourable outcome at six months could be validated. I validated the basic model with the variables: age, GCS and pupil reactivity. The variable 'major extracranial injury' was not included as it was not available in the validation sample. For the CT models, the following variables were added to the basic model: obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift and non-evacuated haematoma. Similarly the variable 'petechial haemorrhages' was excluded as this was not available in the validation sample.

For the validation process I first ran the models in the CRASH trial patients from HIC, and I then applied the corresponding coefficients in the validation sample (IMPACT patients). Although discrimination was, as expected, lower than in the original data, it was still acceptable for both the basic and CT models (c statistics: 0.77 (95% CI 0.76-0.78), and 0.77 (95% CI 0.76-0.78) respectively). The calibration was good for the CT model and poor for the basic model when evaluated graphically, but poor for both according to the Homer-Lemeshow test ($p<0.01$) (figure 4.16).

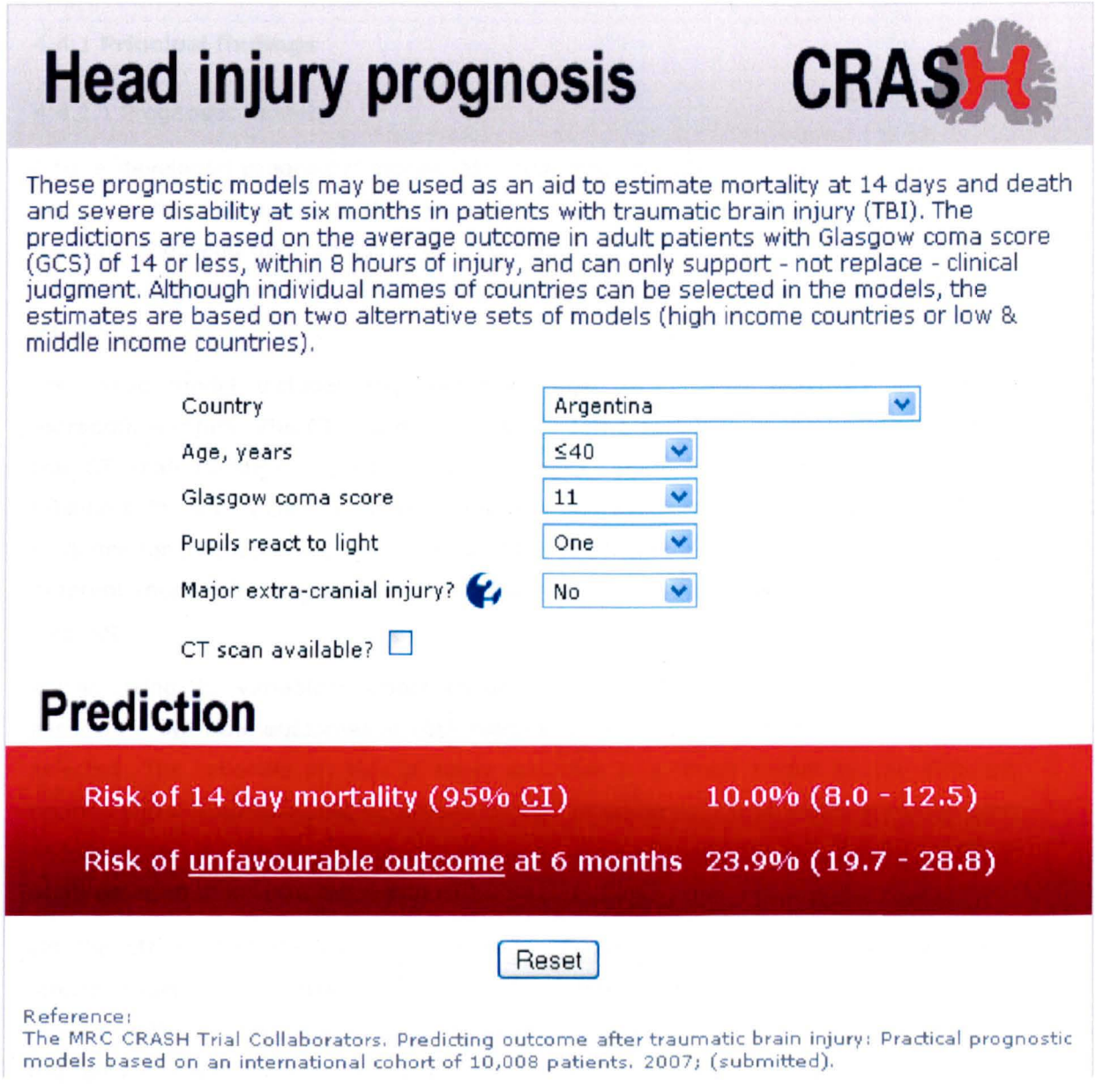
Figure 4-16 Calibration of models for unfavourable outcome in IMPACT



4.3.9 **Clinical score**

A web based calculator, which is available at the CRASH-2 trial web page, was developed to obtain individual probability of outcomes (www.crash2.lshtm.ac.uk).
By entering the values of the predictors, the expected risk of death at 14 days, and of death or severe disability at six months can be obtained for LMIC and HIC patients with TBI. For example, figure 4.17 shows a sample screenshot of the predictions for a 26 year old patient from a low and middle income country (Argentina), with a GCS of 11, one pupil reactive and absence of a major extra-cranial injury. According to the basic model this patient has a probability of death at 14 days of 10% and a 23.9% risk of death or severe disability at six months.

Figure 4-17 Screenshot of the web based calculator



4.4 Discussion

4.4.1 **Principal findings**

4.4.1.1 Prognostic models

I have developed prognostic models for predicting two clinically relevant outcomes in TBI patients using variables that are available at the bedside. The models have good discrimination and good fit using internal validation. The models for HIC patients with moderate or severe TBI also showed acceptable discrimination when externally validated.

The basic model includes the variables: age, GCS, pupil reactivity and major extracranial injury. The CT model includes the same variables as the basic model plus the CT scan findings: petechial haemorrhages, obliteration of the third ventricle, midline shift, subarachnoid haemorrhage and non evacuated haematomas. As I found evidence for interaction between some of the predictors and the outcome I developed different models for HIC and LMIC. All the models have been made available on the internet.

For selecting the variables I chose those for which there was strong evidence of an association with the outcomes in both regions, so a common core of variables could be selected. The rationale for this strategy was that a common model for the different regions and outcomes would be simpler to use. In doing so, I took some decisions such as leaving out the variable "*hours since injury*" because there was only strong evidence of an association in LMIC but not in HIC.

On the other hand the variable "presence of major extracranial injury" there was strong evidence of an association with six month unfavourable outcome in the two regions with both models, but there was weaker evidence for an association with mortality in LMIC. In the basic model the OR was 1.15 (95% CI: 0.99-1.34) and the strength of this association weakened in the CT model (OR: 1.08 95% CI: 0.91-1.28). Nevertheless, I decided to keep this variable in all the models. Similarly, for the presence of petechial haemorrhages in the CT scan there was strong evidence for an association in LMIC but in HIC the association was weaker However, I decided to keep this variable in all the models.

4.4.1.2 Individual predictors

4.4.1.2.1 Demographic and clinical predictors

Age

Increasing age was associated with worse outcomes but this association was most apparent after 40 years of age. A similar threshold has been reported elsewhere.[91] In their systematic review, Hukkelhoven and collaborators, found the best fitting threshold for age was 39 years, however, they reported that the best way to analyse age was a linear and a quadratic term.[69] For simplicity in the model I did not explore a more complex relationship, such as quadratic, but is possible that a better fit could have been found. Furthermore, I acknowledge that it does not seem biologically plausible that the increased risk associated with age only starts after 40 years of age. However, it is unlikely that a different way of analysing the variable age would have changed the main conclusions of this study. Whatever way is chosen to describe this variable, the evidence is quite consistent that a positive association exists between age and poor outcome. Plausible explanations for this relationship include extracranial co-morbidities, changes in brain plasticity, or differences in clinical management associated with increasing age. Further research is needed to explore these mechanisms.

Glasgow Coma Scale

Total GCS showed a clear inverse linear relationship with mortality. The finding that mortality was lower in patients with a GCS of three than in patients with a GCS of four, may be due to GCS scores of sedated patients being reported as GCS three. Previous studies have suggested that total GCS could have become less useful as a predictor in an era of early sedation and pre-hospital intubation. [105] However, in this study I found that GCS has still an acceptable discriminative ability, particularly in LMIC. When GCS was analyzed as a categorical variable (mild, moderate and severe), although it showed an acceptable discrimination, it was lower than the discrimination of GCS as a continuous variable.

In terms of discrimination, total GCS was superior to each of the individual components (motor, eye and verbal). Among the different GCS components the motor item was the most discriminative. In a recent review of the predictive ability of GCS in TBI patients, the authors concluded that the motor has the same discriminative ability as total GCS.[106] However, according to the findings of this study total GCS is superior in terms of discrimination to the motor component.

Pupil reactivity

In concordance with previous studies the absence of pupil reactivity was a strong predictor of poor outcome.[65] The effect estimate (odds ratio) for mortality was among the highest of all predictors. However, when measured with the Z score it was only the third strongest predictor after age and GCS. This finding could be explained by random error as the standard errors were larger due to the relatively low frequency of the pupil abnormalities (6% with one pupil non reactive and 8% with both pupils no reactive).

Major extracranial injury

Patients with a major extracranial injury showed an increased risk of poor outcome. Other studies have shown an independent effect of extracranial injury, for example Signorini and collaborators found that extracranial injuries measured with the Injury Severity Score (ISS) were an independent predictor of mortality. Other studies have failed to report such an association.[107] Nevertheless, it is well accepted that other variables which could be a consequence of presenting extra-cranial injuries, such as hypotension, are associated with poor outcome in TBI patients.[65]

Hours since injury

Only patients from LMIC who were randomised more than one hour after the injury had a higher risk of poor outcome in comparison to those randomised in the first hour. Although there is clinical consensus that TBI patients should receive rapid treatment after the injury, there is no empirical evidence demonstrating the association between time since injury and poor outcome.[108] Furthermore, the CRASH trial cohort has limited information to evaluate this association as only patients who were hospitalized within eight hours of injury were included.

Cause of injury

Cause of injury was not found to be an independent predictor of poor outcome. In a previous meta-analysis of 11 studies, the authors found that the cause "fall" was associated with increased mortality, but after adjustment for age it did not remain as an independent predictor.[109] The only category that I found to be an independent predictor was "other", which was associated with a decreased risk in mortality when compared to the cause "*road traffic crash*" in LMIC. Unfortunately the "other" category includes a wide range of diverse causes, and it was not possible to disentangle the possible explanation for this finding.

Gender

There was no strong evidence for an association between gender and poor outcome. Although some studies have claimed a better outcome in female patients, a recent systematic review concluded that there is no evidence of such a difference in outcome according to gender.[110]

4.4.1.2.2 CT scan predictors

Most of the previous publications have used the CT Marshall classification, referred to in chapter 1, to evaluate the predictive ability of CT findings, however I used instead the individual CT scan variables as these are more useful and practical from a clinical perspective. Recently it has been suggested that analysis of individuals' CT predictors is a better strategy.[111] I found that all abnormal CT scan results, except evacuated haematoma, were associated with poor outcome.

The CT category *"obliteration of third ventricle or basal cisterns"* was most strongly associated with poor outcome. This result is in keeping with the recent findings that absence of basal cisterns is the strongest predictor of six month mortality.[78] As previously reported, traumatic subarachnoid haemorrhage was found to be an independent predictor.[73] The finding that a non-evacuated haematoma was associated with an increased risk of poor outcome, is consistent with studies that have showed that there is an increased risk of poor outcome with different type of intracranial haematomas.[112] Unfortunately, in the CRASH trial there were not enough data to explore this association more in detail. In appendix 7 of this thesis I will further investigate this association using a dataset that includes more details in relation to intracranial bleeding.

4.4.1.3 Differences between patients from LMIC and HIC

4.4.1.3.1 Early mortality

Patients from LMIC had worse early outcome than those from HIC. Regional differences in TBI outcome between Europe and North America have been reported previously, but the difference in mortality between LMIC and HIC has not been explored.[113]
The adjusted odds ratio for mortality at 14 days, for LMIC patients in comparison with HIC patients, was 1.94 (95% CI: 1.64-2.30). It is not possible with the data available in the CRASH trial to reach a conclusion about the causes for this difference. However, some findings could raise hypotheses that might be evaluated in future studies. For example, in comparison with patients from HIC, those from LMIC arrived later to the hospital, and a larger proportion had data on current GCS, a possible indicator that a smaller proportion of LMIC patients was intubated or sedated at hospital admission.

These two findings indicate that patients from LMIC receive more delayed treatment in comparison with those from HIC. Furthermore, a lower proportion of LMIC patients underwent a CT scan. It is therefore possible that for some patients, with potential treatable lesions, a CT scan was not performed.

4.4.1.3.2 Six months unfavourable outcome

Another interesting finding was that although 14 days mortality was higher in patients from LMIC, there was no evidence for a difference for six months unfavourable outcome. This could be related to the fact that those patients from HIC who did not die early had a higher frequency of disability, or alternatively it could be related to the fact that disability at six months was measured with the GOS, the interpretation of which could vary between different settings.

4.4.1.3.3 Differences in the strength of association of predictors

The other important finding was the difference observed in the strength of association between some of the important predictors and outcome. Mine is the first study to report such differences according to whether patients are from HIC or LMIC.

Although GCS was a strong predictor of poor outcome in both regions, I found that it had a higher discriminative ability in LMIC patients than in HIC patients. This might relate to quality of care, or alternatively it could be that low GCS scores in HIC are related to greater use of sedation, rather than to TBI severity.
In LMIC GCS (measured as a continuous variable) had a similar C statistic to both the basic and CT model, while in HIC the difference between the C statistic of the models and the GCS was larger.

The other variable that showed different strength of associations according to the region was age. Increasing age had an even worse prognosis in HIC compared with LMIC. This result probably arises because of even lower risks at younger ages in HIC, while HIC and LMIC have similar risks at older ages. In Figure 4.4 it can be seen that the log odds ratio are similar in both regions for patients older than 64 years but are substantially lower for patients under 40 years in HIC.

Regarding CT scans, some abnormal findings were stronger predictors in HIC compared to LMIC. This could be due to better technology and therefore more accurate CT diagnosis in HIC.

4.4.2 **Comparison with previous studies**

The systematic review reported in Chapter 3 identified over 100 prognostic models for TBI patients, but methodological quality was considered adequate in only a few. [114] Two of the more methodologically robust models showed similar findings to my models, with good discrimination but worse calibration.[91,93] They too included GCS, age, pupil reactivity and CT scan results as predictors but, unlike my models, they did not include the presence of major extracranial injury, and none of them included patients from LMIC.

Subsequent to the publication of the prognostic models reported in this thesis, the investigators from the IMPACT study, the database in which the CRASH prognostic models were validated, developed a series of prognostic models for TBI patients.[115] The IMPACT dataset included 8,509 patients with moderate and severe TBI from 11 studies conducted between 1984 and 1997 in HIC. They developed three types of model to predict mortality at six months, and unfavourable outcome at six months, as defined by the GCS. The three models developed were:

1) A core model that included age, motor score component from GCS, and pupillary reactivity.

2) An extended model that included the core model variables plus secondary insults (hypoxia and hypotension) and CT characteristics (Marshall CT classification, subarachnoid haemorrhage and epidural haematoma).

3) A laboratory model that included the variables from the core and extended models, plus glucose and haemoglobin.

The core model was developed using the whole dataset (8,509 patients), while the other two models used smaller samples as they were forced to restrict the sample to studies with the relevant variables. The extended model was derived from 6,999 patients and the laboratory model from 3,554 patients. Because of missing values on some of the variables in the different studies, they used the method of chained equations to impute missing data. A total of 5%, 13% and 8% of the values were imputed on the core model, the extended model and the laboratory model respectively. For the internal validation, the area under the ROC was calculated with a cross-validation procedure, were each study omitted in turn. The discriminatory ability of models for predicting six months unfavourable outcome in the internal validation ranged from 0.66 to 0.87, with the highest discrimination (0.87) reported for the extended model to predict mortality when evaluated in one of the observational studies. The core and extended models were externally validated in 6,681 patients

with moderate or severe TBI from the CRASH trial. The area under the ROC for unfavourable outcome was 0.78 and 0.80 for the core and extended model respectively. The calibration was poor for all the models when assessed with the Hosmer Lemeshow test ($p<0.001$) and graphically. When they restricted their validation to patients from HIC the discrimination did not change but the graphical display of the calibration for the extended model to predict six months mortality improved. The authors presented a simple score with an accompanying figure and also made the models available as a web-based calculator.

The models developed by the IMPACT study have some aspects in common with those that I derived from the CRASH trial. In relation to the variables included the IMPACT core model and my basic model, both included age, pupil reactivity and the GCS (although IMPACT models included the motor component, while mine included total GCS). Both CT models included subarachnoid haemorrhage as a predictor. A direct comparison of the strength of the association was not possible as the IMPACT study reported proportional odds logistical regression and the effect measures are not comparable with the one that I reported in this thesis. Importantly for both studies, internal and external validation was performed. Finally both studies attempted to make the models easily available to doctors worldwide with a web-based calculator and a simple paper based format. However, they presented some differences. My models were derived from patients from HIC and LMIC while the IMPACT models were derived only from patients from high income countries. Another difference is that the CRASH models were derived from a more recent period (1999-2004) in comparison with the IMPACT models (1984-1997). Also, the CRASH dataset had very few missing variables, while for the IMPACT models imputation methods were necessary to handle the extent of missing data. On the other hand, some of the IMPACT models included more variables which also have been shown to be strong predictors of poor outcome (i.e. hypotension, hypoxia, haemoglobin and glycaemia). Finally, a common feature of the external validation for both studies was that the discrimination was acceptable but the calibration was poor when assessed with the Hosmer-Lemeshow test.

4.4.3 **Strengths and weaknesses of the study**

Among the study's strengths are the use of a well-described inception cohort of patients, prospective and standardised collection of data on prognostic factors, few missing data, very low loss to follow-up, and the use of a validated outcome measure at a fixed time following the injury. All of these factors provide reassurance about the internal validity of the models. The large sample size in relation to the number of prognostic variables examined is also another particular strength.

In relation to its external validity, there are only a few prognostic models developed from LMIC patients, and to the best of my knowledge the models I have developed are the first with a large sample size and adequate methodology.[114] The external validation confirmed the discriminatory ability of the models in patients from HIC and showed, graphically, good calibration for the CT model. Unlike most published prognostic models, my models included the complete spectrum of TBI patients ranging from mild to severe. Finally, the data required to make predictions with the model are easily available to clinicians, and a web based risk calculator was developed.

There are some limitations: the data from which the models were developed originate from a clinical trial and this could therefore limit its external validity. For example, the patients were recruited within eight hours of injury and the accuracy of the models for patients evaluated beyond this time window cannot be estimated. Nevertheless, the CRASH trial was a pragmatic trial that did not require any additional tests and therefore included a diversity of 'real life' patients. I did not include other known or potential predictors of poor outcome in TBI patients. Following the conceptual framework referred to in chapter 1 (figure 1.1) some of these "missing" variables are related to a) the environment, for example variables related with quality of care, b) the host, for example social class, and c) the condition, for example type of bleeding, such as hypotension or hypoxia. Unfortunately none of these variables were available in the CRASH dataset. However, although these variables have been reported to be independent predictors, it has also been shown that they do not add much to the performance of prognostic models when most important predictors (age, GCS and pupil reactivity) are already in the model.[116] A limitation, common in this type of study is related to "self-fulfilling prophecy" bias. All of the variables included in the model are known predictors of poor outcome and it is possible that physicians changed their medical behaviour according to the presence of these variables. For example, an order not to resuscitate could be recommended for elderly patients with low GCS therefore influencing the association of these variables with poor outcome.

Another limitation was that I was only able to validate the models from HIC. Furthermore, the variables "*major extracranial injury*" and "*petechial haemorrhages*" were excluded, because they were not available in the IMPACT sample. However, neither of these variables was among the strongest predictors. The external validation showed good discriminatory ability, but this was somewhat lower than in the original data.

4.4.4 **Implications of the study**

I have developed methodologically valid, simple and accurate models that may help health care decisions for individual patients and counselling for relatives and patients. It is important to emphasise, however, that whilst prognostic models may complement clinical decision making they cannot replace clinical judgement. This is particularly important in the context of judgements regarding the withdrawal of care or clinical triage.

4.4.4.1 Implications for patients in LMIC

Most of the burden of TBI is in LMIC countries where case fatality is high and resources are limited. I found that several predictors differed in their strength of association with outcome according to country income level, suggesting that it may be inappropriate to extrapolate from models based on HIC populations to poorer settings. My models showed good discrimination and my basic models also show good calibration in this setting.

However, GCS used as a continuous variable demonstrated an acceptable ability to discriminate poor outcome in TBI patients from LMIC. When the discrimination was evaluated with categorical GCS, as it is used in clinical practice (i.e. mild, moderate and severe), its discrimination was lower.

In terms of calibration the models showed a better agreement than total GCS between predicted probability and observed outcomes when analysed graphically, in particular for those patients with intermediate risk.

4.4.4.2 Implications for patients in HIC

Basic and CT models showed good discrimination and the latter showed the highest discrimination of all the models developed. The calibration of the models was good when evaluated with the Hosmer-Lemeshow test. Most importantly, when externally evaluated, the models kept an acceptable discriminative ability although the calibration, in particular for the basic model, was poorer.

In HIC both models substantially increased the discriminatory ability of GCS. The CT model in particular showed a very good calibration in the external validation in this setting.

4.4.4.3 Other implications

These prognostic models can also help for research purposes such as in the design and analysis of clinical trials, through prognostic stratification, or can be used in clinical

audit by allowing adjustment for case-mix.[33] In this thesis I focused in the implications of prognostic models for clinical practice, but in the final chapter I will discuss briefly some of the potential implications of prognostic models for research.

4.4.5 **Future research**

The differences found between the prognostic models for LMIC and HIC patients are important. Although most of the burden of trauma occurs in LMIC, most research takes place in HIC.[19] My systematic review reported in Chapter 3 found that very few prognostic models for TBI were developed in LMIC.[114] More research is therefore needed in LMIC in order to obtain reliable data from these settings. An improved understanding of the differences between these regions might also clarify the mechanisms of predictors that are not immediately obvious when analysing a homogeneous population.

The models were developed and validated throughout a rigorous methodological process to ensure their internal validity. However, for prognostic models to be used they need to be user friendly. Research in relation to the different ways for presenting the models to physicians is needed to ensure that the models are practical to use in the clinical setting.

As the models were developed using data from a clinical trial, further prospective validation in independent cohorts is needed to strengthen the generalizability of the models. This particularly applies to the models from LMIC which I have been unable to validate in an external cohort.

I acknowledge that the development, validation and presentation of models are some of the necessary stages of prognostic models research, but the final challenge is the evaluation of their impact on patient outcomes.[40] One of the problems in relation to the impact of risk scores for TBI patients is related to the lack of treatment recommendations according to the baseline risk. However, even in the absence of evidence of effective treatments according to baseline risk, the use of risk scores could unveil how medical care is already strongly influenced by prognosis, although in an implicit way. In the final chapter I will discuss in more detail potential uses of prognostic models for TBI patients.

Chapter 5 Development of the CRASH score card

Focus group and survey

5.1 Introduction

In Chapter 4 I developed prognostic models based on the CRASH trial cohort of traumatic brain injury (TBI) patients to predict mortality at 14 days and unfavourable outcome at six months.[117] The models showed good discrimination and good calibration, when measured graphically, for both internal and external validation. One of the major strengths of these models is that, unlike previous studies, the CRASH models included patients from low and middle income countries.

But even a valid prognostic model will not be used if its presentation is inadequate or complicated. Methodological guidelines stress the importance that prognostic models should be easy and simple to use and well accepted by physicians.[36,38] Simplicity of presentation is even more relevant in the context of the emergency situation when treating patients with TBI.[3]

In the systematic review reported in Chapter 3 I found 102 prognostic models for TBI patients but very few of them presented the prognostic model in a simple way.[114] Two different formats were identified: Hukkelhoven and collaborators presented a numerical score accompanied by a figure.[93] and Signorini and collaborators presented a nomogram.[91] However, none of them investigated if they were considered practical or used appropriately by doctors.

There are many issues that are important to consider when developing a simple prognostic model, for example: the platform by which they are presented (internet or paper based); the way the risk is estimated (regression formula or numerical scores); the precision with which the estimated probability is reported (i.e. when using scores the exact estimated probability can be obtained using a table, or alternatively an approximate estimate can be obtained using different graphs such as score charts or nomograms) or whether or not there is a need to display the confidence interval of the estimated probability.

The CRASH models are available online. However, to be used in the emergency setting, particularly in low and middle income countries, a practical and simple paper based prognostic model is needed.

The aim of this study was to develop a practical and easy to use format for the CRASH prognostic model that predicts mortality at 14 days for patients from low and middle income countries.

I conducted this study in two phases. In the first phase, focus groups were undertaken with the purpose of obtaining information to develop a CRASH score card. In the second phase, a survey was conducted to evaluate if the CRASH score card developed in the first phase was used appropriately and whether it was considered "practical" by doctors.

5.2 First phase: *Focus groups*

5.2.1 **Methods**

Two focus groups were conducted one with Peruvian doctors and a second with Indian doctors.

5.2.1.1 Research Team

I conducted the focus group in Peru. I am a physician and epidemiologist with experience in emergency medicine. I coordinate the CRASH-2 Trial in Latin America. No previous relationship existed with the participants. Ian Roberts conducted the focus group in India; he is a physician and epidemiologist with experience in emergency medicine. He is the principal investigator of the CRASH-2 Trial. No previous relationship existed between the interviewer and the participants.

5.2.1.2 Study design

5.2.1.2.1 Participant selection

The participants were selected by convenience. Doctors participating in CRASH-2 national meetings and with experience in treating TBI patients were asked if they would participate in the focus group. There were four participants in each of the focus groups. None of the doctors approached refused to participate

5.2.1.2.2 Setting

The focus group in Peru took place in the emergency department of Hospital Unanue in Lima Peru, and the focus group in India met in a hotel room where the CRASH-2 national meeting was being held. Only interviewers and participants were present during the focus groups.

The participants in Peru were three males and one female, three of them were emergency physicians and one a neurosurgeon. The four Indian participants were all male and neurosurgeons.

5.2.1.2.3 Data Collection

Participants were first asked to sign an informed consent. Then they were introduced to the concept of prognostic models, in particular for TBI patients, and were informed about the CRASH prognostic model. As previously mentioned, based on the systematic review of prognostic models for TBI patients, I identified two different formats of presentation (score accompanied by a figure, and a nomogram).[114] Furthermore, I also selected two other formats of presentation used in cardiovascular risk scores that I considered could be adapted for TBI patients (a numerical score without figure, and a coloured chart) [118,119] We presented physicians with the four different formats and explained how they work. (Appendix 5.1)
We then asked them: In the context of the clinical management of TBI patients, what would you use a prognostic score for?
For each of the formats the following questions were asked: Would you use this format? Mention some strengths and/or weaknesses for this format and how would you improve it? These questions were only the initial focus from which to develop a non structured discussion. Each focus group discussion was recorded. The duration of the focus groups was between 30 and 60 minutes.

5.2.1.2.4 Analysis

Transcripts were assessed and data derived from the themes identified in advance. Participant quotations were presented to illustrate the different themes.

5.2.2 **Result**

5.2.3 **The potential role of a prognostic model for the management of TBI patients**

Almost all the participants emphasized the importance of prognostic information for patients with TBI. Two main potential uses were identified for a prognostic model available at the bedside: making treatment decisions and communicating with patients or relatives.

Some of the participants highlighted the importance of prognosis for making management decisions, for example one physician said: "*Prognosis is important to assess if the patient is recoverable or not to decide what treatment they should*

receive" this was, according to another physician, particularly important in the context of limited resources where they work, he said: *"We have scarce resources so it would be good for triage patients to decide for example which patients go to intensive care."*

The other main use mentioned was for communicating with relatives, one physician said: *"As emergency doctors it is very useful to provide information to the relatives, it will give us more support and we would be better covered for potential legal problems"* Although another physician mentioned that he would prefer to provide a very pessimistic prognosis so that if the patient recovered, the relatives would be happy and if the patient dies, they have been forewarned.

5.2.4 **Format of the prognostic model**

There was consensus, in both focus groups, that the first presentation (a numerical score with an accompanying figure) was preferred. One doctor, summarizing his preference said:

"This is the best option; it is clear to calculate at the bedside, graphs are easier"

Nomograms on the other hand were not well accepted, one physician observed:

"It is not very exact and adding up is complicated" other comments were *"Not very practical, I don't go with a ruler!" "This requires more time than the previous one""I find it difficult" and "it does not look precise"*

The third option, a numerical score without a graph, was in general well accepted but a doctor argued that it is not very practical if the score range is large:

"It is practical and simple because it is more straightforward than the graph but only if the score value is less than 50 otherwise a graph would be better."

Another doctor added:

"Better with a graph because we are familiar with them"

Finally the coloured chart was not considered a good option, some of the comments regarding this format were:

"It is nice but less precise" "It will need more training" "the first impression makes me dizzy!"

Another doctor argued that it will not be very useful:

"if the number of variables included in the model is very large there would be too many boxes"

There were also some general recommendations to improve the preferred format (numerical score plus figure) some of the suggestions were:

"The line of the figure should be larger and the confidence interval line it is not very helpful!" "It would be easier to interpret with a squared paper" "definitions of the variables used should be included"

Among other suggestions some of the respondents recommended that we include other variables such as hypotension or hypoxia which they considered to be strong predictors of mortality. One doctor expressed the preference to report outcome as mortality instead of survival:

"..if we use the word survival relatives will stick to that no matter how low the probability"

5.3 Second phase: *Survey*

5.3.1 **Methods**

5.3.1.1 Development of the CRASH score card

With the information obtained in the first phase and the participant's suggestions I developed a CRASH score card which included a numerical score and accompanying figure on a squared paper.

To develop the CRASH Score, I first needed to adapt the way of including some variables so they can be used in a paper based score.

Age was included as a categorical variable (<40, 40-49, 50-59, 60-69 and >69). Total GCS was also included as a categorical variable (3 to 14). In addition, the same variables included in the CT model for predicting mortality at 14 days were included (pupil reactivity, presence of major extra-cranial injury and the CT scan results: petechial haemorrhages, obliteration of the third ventricle, midline shift, subarachnoid haemorrhage and non evacuated haematomas).

The discrimination of the CRASH Score model (using age and GCS as categorical variables) was similar to the one reported for the CT model for LMIC reported in Chapter 4 (C statistic 0.84 in both).

For estimating the values corresponding to each category of predictor I ran the CRASH Score model and multiplied the predictors' β coefficients by 10 and rounded them. For example the β coefficient for the age categories (40-49, 50-59, 60-69 and >69) were 0.40, 0.59, 1.10 and 1.77 so after multiplying by 10 their value for the CRASH Score were 4, 6, 11 and 18 respectively.

In designing the figure, I obtained the predicted probability associated with each CRASH score value. For example, a score of 31 corresponds to a probability of 25% of morality at 14 days, and a score of 41 to a probability of 50%.

The CRASH score card presents the value for each category of the predictors and includes the figure displaying the predicted probability according to the total score obtained by adding the different predictors' values (Figure 5.1).

For example, if a patient is from Colombia , is 77 years old, has a total GCS of 6, has one pupil reactive, no major extra-cranial injury and a non-evacuated haematoma, the CRASH score value would be 47 (18+21+4+4) and would correspond to a probability of death of about 62%.

Figure 5-1 CRASH score card for predicting mortality in LMIC

CRASH SCORE CARD

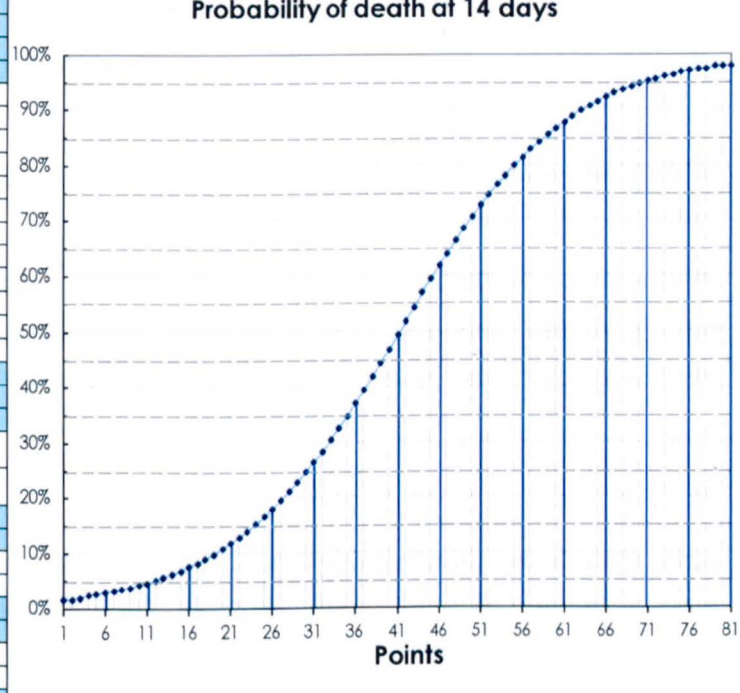
Glasgow Coma Scale					
EYE OPENING		MOTOR RESPONSE		VERBAL RESPONSE	
4	Spontaneous	6	Obeys commands	5	Orientated
3	To sound	5	Localising	4	Confused speech
2	To pain	4	Normal flexion	3	Words
1	None	3	Abnormal flexion	2	Sounds
		2	Extending	1	None
		1	None		

DEFINITION OF MAJOR EXTRA-CRANIAL INJURY:
Any injury that requires hospital admission within its own right

INSTRUCTIONS

1. Select one value for each **PREDICTOR**
2. Add up the corresponding points
3. Find the total value on the **POINTS** scale in the graph and follow the upright line to where it crosses the blue curved line
4. Find the **PROBABILITY** of death at 14 days on the vertical scale

PREDICTOR		VALUE	POINTS
AGE		<40	0
		40-49	4
		50-59	6
		60-69	11
		>69	18
GCS		14	0
		13	5
		12	9
		11	9
		10	15
		9	15
		8	17
		7	18
		6	21
		5	25
		4	30
		3	30
PUPIL REACTIVITY		BOTH	0
		ONE	4
		NONE	11
MAJOR EXTRA CRANIAL INJURY		NO	0
		YES	1
CT SCAN	Petechial haemorrhages	NO	0
		YES	2
	Obliteration of the 3rd ventricle	NO	0
		YES	7
	Subarachnoid bleed	NO	0
		YES	3
	Midline shift >5mm	NO	0
		YES	5
	Non evacuated haematoma	NO	0
		YES	4



5.3.1.2 Sample

I conducted the survey in a convenience sample of doctors participating in the CRASH-2 Trial who routinely treat TBI patients. The CRASH network consists of approximately 200 doctors from 36 countries. Doctors who were good recruiters and considered communicative by regional coordinators were identified and selected. A total of 40 doctors from low and middle income countries were asked to participate.

5.3.1.3 Data collection

Physicians received the material (CRASH score card) by post and completed a questionnaire (which was available on paper and electronically) (appendix 5.2)

I first asked the respondents about their demographics characteristics (age, sex, country of residence) medical related characteristics (speciality, average number of TBI patients treated by month). I then presented the physicians with the following vignette: *"Male aged 52 years who had a road traffic crash; on physical examination the total Glasgow Coma Score is 10 and both pupils are reactive. The CT scan shows midline shift of 8 mm. The patient does not have a major extra cranial injury"* and asked them to calculate the CRASH score value and the corresponding probability of death according to the CRASH score card. I also presented them the following statement *"The format of the CRASH Score Card is practical for use in the clinical setting"* and asked them to answer a Likert scale

a) Strongly agree b) Agree c) Neither agree nor disagree d) Disagree e) Strongly disagree

I finally left a space for them to write any further suggestion to improve the CRASH score card.

The questionnaire was tested on a convenience sample (4 respondents) and written comments were obtained regarding the instructions and the face validity of the questionnaire. The questionnaire and the CRASH score card were translated into Spanish as some of the respondents were Spanish speakers.

5.3.1.4 Data analysis

I reported the frequencies of the characteristics of the respondents. The CRASH score card uses a figure to display the probability of death, therefore, there would be some expected variations on the probability estimated. The correct score for the vignette presented was 26 and the corresponding probability of death was 17% but I defined an interval of ± 2% as acceptable (15-19%). I reported the frequencies of responses for

the score, death probability and the different categories of the Likert scale. I did a content analysis of the last open question about how to improve the CRASH score card.

5.3.2 **Results**

37 out of 40 doctors responded to the survey (92% response rate).Table 5.1 shows their general characteristics.

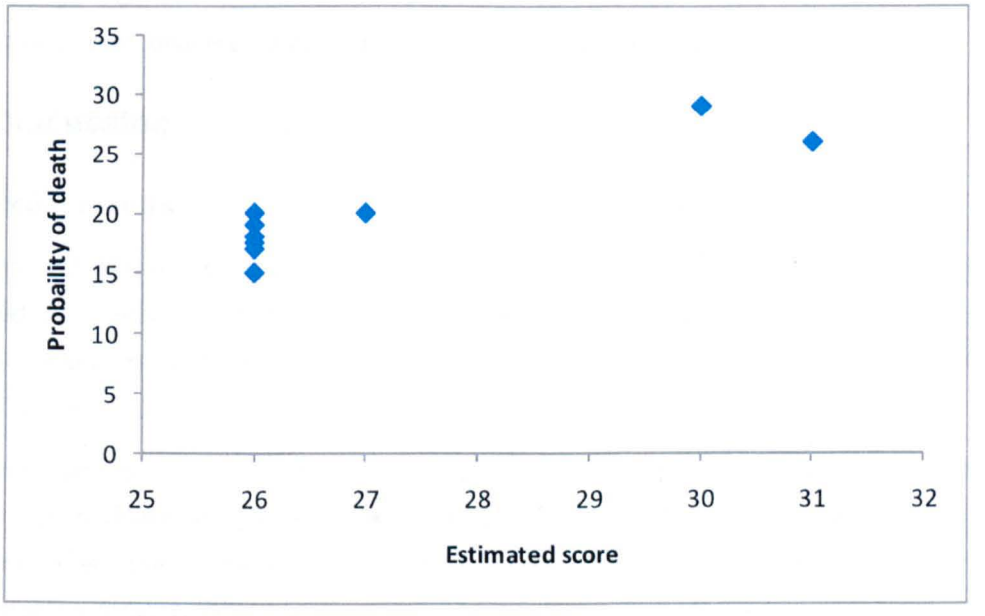
The average age was 41.8 ± 8.2 years. More than half of the respondents were either surgeons or neurosurgeons. The doctors were from seventeen different countries from Latin America, Africa and Asia. The countries that contributed with the larger number of doctors were India, Thailand and Colombia. The median number of TBI patients that doctors reported to treat each month was 30.

Table 5-1 Characteristics of participants

Speciality	n (%)	Countries	n (%)	Patients per month	n (%)
Surgery	13 (35)	*India*	6 (16)	*Less than 10*	10 (27)
Neurosurgery	10 (27)	*Thailand*	6 (16)	*10 to 100*	19 (51)
Orthopaedics	4 (11)	*Colombia*	4 (11)	*>100*	8 (23)
Emergency	3 (8)	*Mexico*	3 (8)		
Anaesthesiology	3 (8)	*Georgia*	2 (5)		
internal medicine	2 (5)	*Malaysia*	2 (5)		
Community Medicine	1 (3)	*Nigeria*	2 (5)		
Paediatrics	1 (3)	*Peru*	2 (5)		
		Zambia	2 (5)		
		Argentina	1 (3)		
		Cameroon	1 (3)		
		Ecuador	1 (3)		
		Egypt	1 (3)		
		Ghana	1 (3)		
		Indonesia	1 (3)		
		Iran	1 (3)		
		Tanzania	1 (3)		

A total of 34 respondents (92%) correctly calculated a score of 26 and of these 33 (97%) estimated a death probability that was in the range defined as acceptable (15-19%) (Figure 5.2).

Figure 5-2 Estimated score and probability of death by the respondents



The three doctors who answered incorrectly obtained the following scores: 27, 30 and 31. The first two estimated a death probability that were within the range of acceptable for the score that they calculated (20% vs. 21%, 29% vs. 27%) but the doctor who calculated a score of 31 estimated a death probability of 26% while the correct death probability for that score value was 29%.

A total of 30 respondents (81%) agreed or totally agreed that the CRASH score card was practical 6 were unsure and only one doctor disagreed.

The last open section allowed the respondent to include comments and suggestions.

These were some of the positive comments included: *"Very practical!" "..looks very useful and from my point of view might have practical usage in a clinical setting" "I think this score card is simple and practical" "It will be useful tool" "... liked the way I could show relatives a more objective measure of outcome instead of just giving personal opinion" "Completely practical"*

Some of the recommendations suggested to improve the CRASH score card were: *"should instruct when the timeframe of the observed parameters be recorded" "Add footnote referring to original article" "card should be plastic or laminated for durability"*

"CT Scan is not routinely available for many patients. Exclusion of imaging investigation like CT Scan will make the scoring system universally applicable" "CT Scans are not available in Resource Limited settings" "CT Scan not readily available in all settings" Other respondents suggested to include more variables such as: aspiration, heart disease, liver disease, renal disease or diabetes.

5.4 Discussion

5.4.1 **Main results**

I developed a score card to predict in-hospital mortality for TBI patients which was deemed as practical and was appropriately used by the majority of respondents in a survey conducted among doctors from low and middle income countries who routinely treat TBI patients.

The focus groups showed that the main potential two uses of a prognostic model for TBI at the bedside are for decision making and for communicating with relatives. It also identified the numerical score with an accompanying figure as the preferred presentation format. Nomogram and coloured chart formats were considered difficult to interpret by the participants.

The survey allowed improving the original version of the CRASH score card. Some of the participants suggested developing a similar CRASH score card but without CT scan as this is not always available in poor resource settings. Others suggested including more variables (i.e. hypoxia and hypotension) which they considered to be important predictors. For "clinical acceptability" it is always desirable to include variables which are considered relevant by doctors. However, in this case these data were not available in the CRASH trial so could not be incorporated. Furthermore, it has been shown that hypotension and hypoxia are strongly associated with poor outcome in crude analysis but their relative prognostic value decreased markedly after adjustment by age, GCS and pupil reactivity, predictors which are included in the CRASH model.[116]

5.4.2 **Comparison with other studies**

In the previous survey described in Chapter 2 respondents reported that prognostic information was important to make treatment decisions and counselling TBI patients (or relatives).[120] The studies reported in this chapter, which included a different sample

of doctors, agreed with the previous finding and highlighted the relevance of prognostic information particularly when dealing with limited resources.

To the best of my knowledge this is the first study to explore different ways of presenting prognostic models to physicians treating TBI. However, research in related areas reported similar findings. A qualitative study with general practitioners showed that graphical presentation of information was favoured over numerical information.[121]

Some of the participants referred to the importance of a negative framing (predict mortality instead of survival) to avoid unrealistic expectations from the relatives. This attitude of physicians to give a pessimistic prediction (which means a worse prediction than they really believe) when treating critically ill patients has been previously reported.[29]

5.4.3 **Strengths and weaknesses**

The two phases of this research allowed to complement qualitative and quantitative information to better understand the role of prognostic information in the context of emergency treatment of TBI. It also helped to develop a score tailored to the needs of doctors, from low and middle income countries, who routinely treat TBI patients. The focus group methodology capitalized the interaction with a peer group to obtain data. The survey provided quantitative information about the accuracy of the implementation of this model and the doctor's preferences. One of the strengths is that I included physicians from low and middle income countries from all the regions of the world, who would be potential users of this score. Finally it should also be highlighted that response rate for the survey was high (92%).

The study is not free of limitations, the sample was a convenience sample and therefore not necessarily representative of doctors from low middle income countries who treat TBI patients. The other limitation is related with the use of a vignette, which does not represent real life, so the real accuracy in the emergency situation could be lower than the one we reported. However, vignettes have been extensively used in different settings as a way of standardizing patient information.[122-124] Because the respondents knew that this prognostic models was developed by our group it is possible that a type of "desirability" bias could have been introduced when assessing the practicality of the CRASH score card. Finally, the discrimination reported for the CRASH score card applies for the exact estimates of the model but when used by doctors, as reported in this chapter, only an approximation of the exact probability provided by the model is obtained. Total GCS has been shown to be highly discriminative and it is possible that the CRASH score card, when used in normal

practice, does not add substantial discrimination in comparison to the routine use of the GCS by physicians.

5.4.4 **Implications and future research**

I have previously developed prognostic models which have shown good performance and are available online. However, to be practical in low and middle income countries where internet availability in the emergency room is very unlikely, they should be paper based so they can be used at the bedside. I developed a paper based score card and showed that doctors from low and middle income countries can use it appropriately and that they consider it practical. Nevertheless, a few doctors made mistakes which could be avoided when using the web based model. The practicality of using the CRASH score card should be further evaluated in a larger sample in a "real life" situation and, ideally, future studies should compare the performance of the CRASH score card against routine clinical prediction by doctors without the CRASH score card.

Chapter 6 Association between the Modified Oxford Handicap Scale and GOS

6.1 Introduction

In evaluating the effectiveness of clinical interventions for TBI two main outcomes are commonly evaluated, mortality and disability.

The Glasgow Outcome Scale (GOS) is the most widely used disability outcome measure in randomised controlled trials (RCT) in TBI patients, and is usually completed at six months after the injury.[12] [14] The GOS classifies patients into five categories, dead, persistent vegetative state, severe disabled, moderate disabled and good recovery. There are structured interviews to assess the GOS that have shown to increase its reliability.[11,13] Among the strengths of the GOS, is that it covers all possible outcomes, forms clinically meaningful categories and can be completed by patients or proxies.[16] However, the use of the GOS for research and for clinical practice is not without problems.

A practical problem with the use of six months GOS as an outcome is that it lags behind in time, so that it takes time to be available. In the context of RCT in TBI patients, Data Monitoring Committees (DMC) do not have an early disability outcome and can only use in-hospital mortality and six month GOS, which is not available for all the patients included. If an early measure of disability was on hand that could predict long term disability, it might also be potentially useful to inform interim analysis.

Another problem with the measurement of the GOS at six months is that loss to follow-up is frequent in clinical trials of TBI patients. In a review of 208 clinical trials conducted on TBI patients the average loss to follow-up was 19%.[14] Large losses to follow-up reduce the precision of estimates and may introduce bias. If an early outcome measure was available that could predict long term disability, it might be useful for dealing with loss to follow-up, which can be considered as a type of missing data. Missing data has been defined as; missing completely at random (MCAR), when loss to follow-up is not related to any patient characteristics, missing at random (MAR), when loss to follow-up is dependent on patient characteristics for which information is available, and missing not at random (MNAR), where loss to follow-up depends on information that is not observed even after conditioning on the observed data. Most of the techniques for dealing with missing data assume MAR, and use available information to impute missing data. In other words prognostic models are used to

predict, and eventually impute, the missing value.[125] In the context of TBI if an early disability outcome adds predictive information, to the already known predictors of unfavourable outcome at six months, it would be useful for using in imputation methods when dealing with loss to follow-up.

Finally, most of TBI related death occurs during hospitalisation. So from the patient and relative perspectives, the most important prognostic information, after hospital discharge, is related to long term disability.[126] An early and valid predictor of long term disability, available at hospital discharge, could be a useful tool for the communication among doctors and patients about long term prognosis.

The CRASH trial presents an opportunity to evaluate the predictive validity of an early disability outcome measure, the Modified Oxford Handicap Scale (mOHS), which was completed at hospital discharge, and the GOS which was completed at six months after hospital discharge.

The aim of this chapter was to analyze the association between the mOHS and GOS at six months. The three specific objectives were

1) Evaluate the potential uses of the mOHS for informing DMC
2) Evaluate the potential use of the mOHS for dealing with loss to follow-up
3) Evaluate the potential use of the mOHS for communicating with patients and relatives at hospital discharge.

6.2 Methods

6.2.1 **The sample of patients**

Of 10,008 study participants enrolled in the CRASH trial, 99 (1%) had missing data on the mOHS, 418 (4.2%) had missing data on the GOS at six months, and 36 (0.3%) had missing data for both mOHS and GOS. A further 8 patients were excluded from analysis as they had a Glasgow Coma Scale (GCS) score of 15 at randomisation. In total 9,447 patients (94, 4%) were available for this analysis. For objective one I used all the patients from the sample, while for objectives two and three I based my analysis only on survivors at hospital discharge.

6.2.2 **Exposure**

Modified Oxford Handicap Scale (mOHS)

The mOHS is the result of various modifications of previous disability scales. In table 6.1 it is shown that the original source was the Rankin Scale (RS). The RS was

developed in Scotland in 1957 to describe recovery in stroke patients at hospital discharge.[127] In 1988, as part of a study of aspirin in stroke, it was modified and renamed to Modified Rankin Scale (mRS). An additional grade (0 *no symptoms*) was added.[128] In 1989 the mRS was modified for a study of stroke patients in the community, and renamed Oxford Handicap Scale (OHS).[129] All three scales have been used in different settings, at different times, and administered in a variety of ways.[130] The last column of table 6.1 displays the six categories of the mOHS used in the CRASH trial. In relation to the OHS, moderate handicap and moderately severe handicap were combined in one, and a further category, death, was added.

Table 6-1 Comparison of disability scales

Rankin Scale 1957	Modified Rankin Scale 1988	Oxford Handicap Scale 1989	Modified Oxford Handicap Scale as used in the CRASH trial
No significant disability: able to carry out all usual duties	No symptoms at all	No symptoms	No symptoms
	No significant disability despite symptoms: able to carry out all usual duties and activities	Minor symptoms that do not interfere with lifestyle	Minor symptoms
Slight disability: unable to carry out some previous activities but able to look after own affairs without assistance	Slight disability: unable to carry out all previous activities but able to look after own affairs without assistance	Minor handicap, symptoms that lead to some restriction in lifestyle but do not interfere with the patient's capacity to look after himself	Some restriction in lifestyle but independent
Moderate disability: requiring some help, but able to walk without assistance	Moderate disability: requiring some help, but able to walk without assistance	Moderate handicap, symptoms that significantly restrict lifestyle and prevent totally independent existence	Dependent but not requiring constant attention
Moderately severe disability: unable to walk without assistance, and unable to attend to own bodily needs without assistance	Moderately severe disability: unable to walk without assistance, and unable to attend to own bodily needs without assistance	Moderately severe handicap, symptoms that clearly prevent independent existence though not needing constant attention	
Severe disability: bedridden, incontinent and requiring constant nursing care and attention	Severe disability: bedridden, incontinent and requiring constant nursing care and attention	Severe handicap, totally dependent patient requiring constant attention night and day	Fully dependant requiring attention day and night
			Death

6.2.3 **Other variables considered in the analysis**

The variables that have been reported to be associated with six month unfavourable outcome (as measured with the GOS) in Chapter 4 were also included in some of the analysis reported in this chapter. These variables were: age, Glasgow Coma Scale (GCS) at randomisation, pupil reactivity, whether the patient sustained a major extra cranial injury and computerised tomography (CT) scan results (petechial haemorrhages, obliteration of the third ventricle, midline shift, subarachnoid haemorrhage and non evacuated haematomas).[131] I also analysed income region, high income countries (HIC) or low & middle income countries (LMIC), as a potential effect modifier of the association between mOHS and GOS. All these variables were defined as referred in Chapter 4.

6.2.4 **Outcome**

Unfavourable outcome at six months

This outcome was defined using the GOS, which was assessed at six months with a validated questionnaire that was mailed to patients or their carers, administered by telephone interview, or undertaken during a home visit or hospital appointment (Appendix 4.1).[102] GOS was dichotomised as for the analysis in the CRASH trial into favourable outcome (good recovery or moderate disability) and unfavourable outcome (severe disability or death). I created two further dichotomies: good recovery versus other outcomes, and survival versus death.

6.2.5 **Analysis**

I estimated the association between mOHS and 6 month unfavourable outcome according to income regions, and conducted a likelihood ratio test to evaluate the presence of interaction between income and mOHS. For descriptive purposes I performed a cross-tabulation between the OHS and GOS categories.

6.2.5.1 Use of the mOHS to inform DMC

I evaluated the relationship between mOHS and GOS. Their relation was assessed with the Spearman rank correlation coefficient. The validity of all the possible disability dichotomies of the mOHS for predicting unfavourable outcome measured with GOS at six months was assessed by calculating their sensitivity and specificity.

6.2.5.2 Use of the mOHS for dealing with loss to follow-up

Because the problem of loss to follow-up occurs after hospitalisation, I conducted this analysis only among survivors at hospital discharge. I first fitted a logistic regression model to predict unfavourable outcome as defined by GOS (severe disability or death) including all the variables which have been already reported to be predictors in Chapter 4. This model included age, GCS, pupil reactivity, whether the patient sustained a major extra cranial injury and CT scan variables (petechial haemorrhages, obliteration of the third ventricle, midline shift, subarachnoid haemorrhage and non evacuated haematomas). I then fitted a second model which also included mOHS. For each model I analysed its discrimination using the c statistic. For descriptive purposes I also reported the crude and adjusted odds ratio for GOS at six months for each of the categories of the mOHS.

6.2.5.3 Use of the mOHS for communicating with relatives and patients

Among survivors at hospital discharge I estimated the positive predictive value (with 95% confidence intervals) of each mOHS category for three different outcome defined by GOS at six months (good recovery, good recovery or moderate disability and survival).

6.3 Results

There was not strong evidence of an interaction between mOHS and income region to predict six months unfavourable outcome (p=0.12), so the analysis was performed for patients from both HIC and LMIC.

6.3.1 **General characteristics of the population**

Table 6.2 shows a cross tabulation between mOHS at 14 days and GOS at six months for the sample included in the analysis. At 14 days 1,092 (11%) were fully dependent, and 1,948 (21%) patients had died. A total of 4,869 patients (50%) were discharged with no or minor symptoms, and 1,538 (16%) were with some restriction or dependent but without need of constant care.

At six months, 1,208 (13%) patients were severely disabled, and 2,317 (24%) patients had died, while 5,922 (63%) reported to have a good recovery or to be moderately disabled.

Table 6-2 Cross tabulation between mOHS and GOS

	Glasgow Outcome Scale at six months								
Modified Oxford Handicap Scale at 14 days	*Good recovery*		*Moderate disability*		*Severe disability*		*Death*		***Total***
	n	%	n	%	n	%	n	%	**N**
No symptoms	1,910	79	334	14	150	6	17	1	**2,411**
Minor symptoms	1,646	67	537	22	233	9	42	2	**2,458**
Some restriction in lifestyle but independent	354	46	246	32	147	19	20	3	**767**
Dependent but not requiring constant attention	232	30	273	35	221	29	45	6	**771**
Fully dependent requiring attention day & night	148	14	242	22	457	42	245	22	**1,092**
Dead	*0*	*0*	*0*	*0*	*0*	*0*	*1948*	*100*	***1948***
Total	**4,290**	**46**	**1,632**	**17**	**1,208**	**13**	**2317**	**24**	**9,477**

6.3.2 **Use of the mOHS Scale to inform DMC**

For evaluating the potential role of the mOHS to inform DMC I analysed the correlation among mOHS and GOS and the validity of the different possible dichotomies. The mOHS at 14 days and GOS at six months were highly correlated (Spearman rank correlation coefficient 0.75). Four dichotomies of disability for the mOHS were considered (Table 6.3). The first separated patients with no symptoms from the rest, the second considered patients with no or minor symptoms in the baseline group, the third included patients with no symptoms, minor symptoms or with some restriction (but independent) as the baseline group. Finally the forth dichotomy separated patients who were fully dependent or dead, from the rest.

Table 6-3 Dichotomies of the mOHS

	A	***B***	***C***	***D***
No Symptoms	**No**	**No**	**No**	**No**
Minor Symptoms	Yes	**No**	**No**	**No**
Some restriction in lifestyle but independent	Yes	Yes	**No**	**No**
Dependent but not requiring constant attention	Yes	Yes	Yes	**No**
Fully dependent requiring attention day and night	Yes	Yes	Yes	Yes
Dead	Yes	Yes	Yes	Yes

Table 6.4 shows the validity measures for each dichotomy in relation to unfavourable outcome as defined by the GOS (severe disability or death). As expected there was a trade-off between sensitivity and specificity. The increase in specificity was obtained at expense of a decrease in sensitivity

Table 6-4 Validity of the mOHS dichotomies for unfavourable outcome

mOHS dichotomy	Sensitivity	Specificity
A	95.3	37.9
B	87.5	74.8
C	82.7	84.9
D	75.2	93.4

6.3.3 Use of the mOHS for dealing with loss to follow-up

For evaluating the potential role of the mOHS for dealing with loss to follow-up I analysed the incremental predictive ability of the mOHS to the other known predictors of unfavourable outcome at six months.

The c statistic, for predicting unfavourable outcome at six months, for the model including age, GCS, pupil reactivity, major extra cranial injury and CT scan variables was 0.78 (95% CI 0.77-0.80). The c statistic for the model including the previous variables plus the mOHS was 0.83 (95% CI 0.82-0.84). There was strong evidence that the discrimination of the model including mOHS was superior to the one without it ($p < 0.001$). The c statistic for mOHS alone was 0.77 (0.78-0.79).

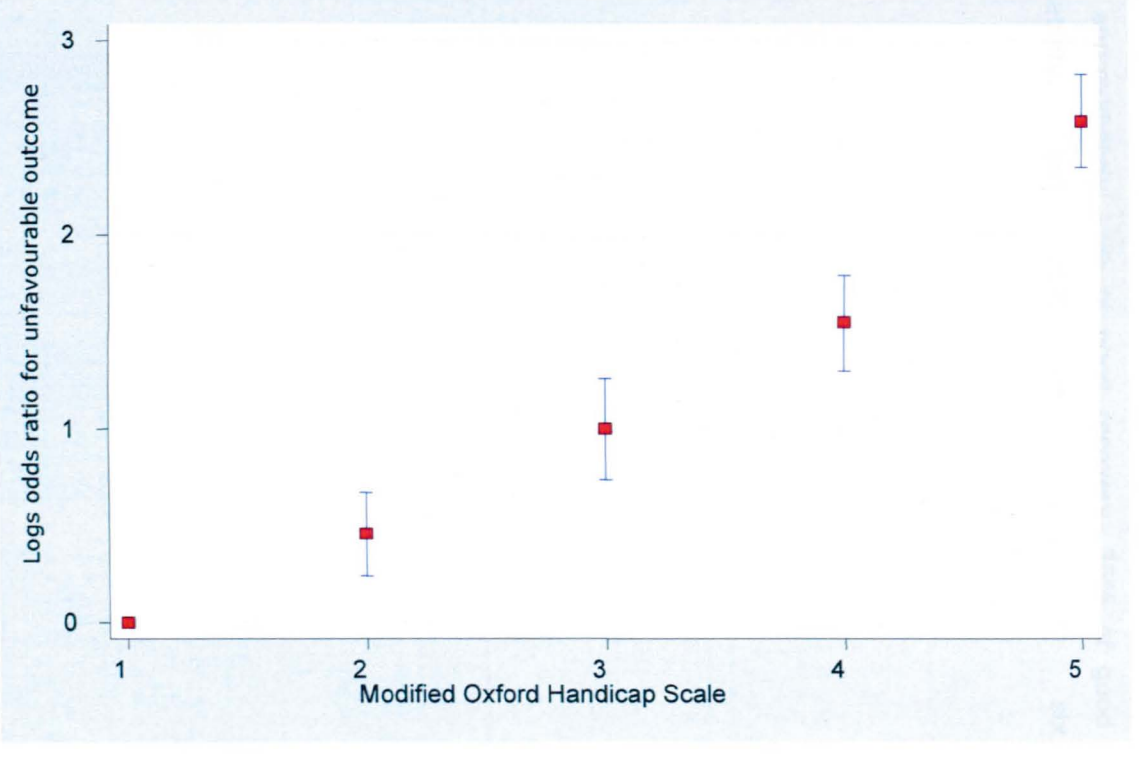
In table 6.5 it can be seen that in the crude analysis there was a strong association between the mOHS and six months unfavourable outcome, as measured with GOS. After adjusting for the other predictors, although there was an attenuation in the effect measures, there was still strong evidence of an association between mOHS and GOS. Patients who were fully dependent at hospital discharge were thirteen times more likely to have an unfavourable outcome at six months, in comparison with those patients who were discharged without symptoms. There were no changes in the estimates when adjusted for treatment. Figure 6.1 shows that after adjustment there was a linear relationship among mOHS and unfavourable outcome.

Table 6-5 Crude and adjusted association between mOHS and unfavourable outcome

mOHS	Crude			Adjusted¥		
	OR	95% CI		OR	95% CI	
		Lower	Upper		Lower	Upper
No symptoms	1.0			1.0		
Minor symptoms	1.7	1.4	2.1	1.6	1.3	2.0
Some restriction in lifestyle but independent	3.7	3.0	4.7	2.7	2.1	3.5
Dependent but not requiring constant attention	7.1	5.7	8.8	4.7	3.7	6.0
Fully dependent requiring attention day & night	24.1	19.8	29.4	13.3	10.4	16.9

¥Adjusted by GCS, pupil reactivity, major extra-cranial injury ,age and plus CT findings (petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleed, midline shift, non evacuated haematoma)

Figure 6-1 Relationship between mOHS and unfavourable outcome



6.3.4 **Use of the mOHS for communicating with relatives and patients**

For evaluating the potential use of the mOHS for communicating with relatives and patients I analysed the predictive ability of each category of the mOHS for different relevant outcomes as defined by GOS at six months. Table 6.6 shows the prediction of different disability status, measured with the GOS at six months, according to mOHS categories. For example, a patient who was fully dependent at hospital discharge, had a probability of approximately 13% of good recovery, 36% of good recovery or moderate disability, and 78% of survival at six months.

Table 6-6 Prediction of GOS according to mOHS categories

OHS	Good recovery			Good recovery or Moderate disability			Survival		
		95% CI			95% CI			95% CI	
	%	Lower	Upper	%	Lower	Upper	%	Lower	Upper
No symptoms	79.2	*77.6*	*80.8*	93.1	*92.1*	*94.1*	99.3	*99.0*	*99.6*
Minor symptoms	67.0	*65.1*	*68.8*	88.8	*87.6*	*90.1*	98.3	*97.8*	*98.8*
Some restriction in lifestyle but independent	46.1	*42.6*	*49.7*	78.2	*75.3*	*81.2*	97.4	*96.3*	*98.6*
Dependent but not requiring constant attention	30.1	*26.8*	*33.3*	65.5	*62.1*	*68.9*	94.2	*92.5*	*95.8*
Fully dependent requiring attention day & night	**13.6**	*11.5*	*15.6*	**35.8**	*32.9*	*38.6*	**77.6**	*75.1*	*80.0*

6.4 Discussion

6.4.1 **Principal findings**

I found that the mOHS was highly correlated with GOS at six months.

Among the different dichotomies of the mOHS explored, the dichotomy that considered patients as dead or fully dependent (dichotomy D) was the one with the highest specificity in comparison with unfavourable outcome defined by the GOS.

The mOHS showed to add predictive information to the prognostic models for unfavourable outcome at six months which included age GCS, pupil reactivity, major extra cranial injury and CT scan variables.

I reported the predictive ability of the mOHS for different GOS categories at six months in a way that could be easily communicate to relatives and patients.

6.4.2 **Comparison with other studies**

The incidence of unfavourable GOS outcome at six months in the CRASH trial cohort was lower (37%) than the one reported in the IMPACT study (48%).[104] However, unlike the CRASH trial, the IMPACT study included only moderate and severe cases.

The mOHS was first used in the CRASH trial, so this is the first study reporting on the predictive validity of this scale in relation to GOS at six months. However, a previous study reported a good agreement between the Modified Rankin Scale (the scale from which the OHS was derived) and the GOS in a trial of stroke patients.[132]

6.4.3 **Strengths and weaknesses of the study**

To my knowledge this is the first study that evaluated the predictive validity of a simple scale for disability at hospital discharge in TBI patients. The main strengths of our study include the large sample size, which ensures precision in our estimates and provides some information in an area which is virtually unexplored.

However, I acknowledge that there are some limitations. Because the main objective of the correlation analysis was to evaluate the usefulness of mOHS when GOS is not available, I used all the mOHS categories for this analysis, including dead patients. But, it is self evident that as the GOS also included dead patients this means that a large part of the high correlation observed might be explained by the correlation between this category. I repeated the correlation analysis excluding dead patients and the Spearman rank correlation coefficient was lower (0.43). This result confirms that in terms of disability both scales are not highly correlated. Nevertheless, it is important to

stress that the purpose of this analysis was to evaluate if the early use of the mOHS was a practical way to tackle the problem of missing GOS data in the context of RCT of TBI patients, and not to evaluate if the scales were correlated in term of disability. In fact this finding is not surprising as the scales measure different aspects of disability. The mOHS is a very rough scale of disability at the moment of hospital discharge, while the GOS measures more subtle levels of functional and psychosocial disability when the patients is already back in his/her usual social life, six months after hospital discharge. This difference explains why for example 20% of patients discharged with no symptoms according to the mOHS where considered with disability when evaluated with the GOS six months later.

The measurement of the mOHS was not conducted with a structured interview and the inter rater reliability was not assessed. Nevertheless, a good inter-observer reliability has been reported for the original OHS (weighted kappa of 0.72).[129] Furthermore, it is possible that some of the investigators who completed the early disability scale also completed the GOS questionnaire at 6 months. However, it is unlikely they would remember the mOHS result at hospital discharge when completing the GOS questionnaire after six months, and furthermore, there are no clear parallel categories in the two forms completed.

In addition, the results of the validity of the mOHS to predict GOS at six months reported in this chapter were not externally validated in a different sample of patients. Unfortunately, at the time of writing this thesis there is not a TBI patient sample with availability of mOHS data at hospital discharge which could be used for validation purposes.

6.4.4 **Implications**

6.4.4.1 Use of the mOHS Scale to inform DMC

DMC dealing with trials of TBI patients usually have data on in-hospital mortality and six months GOS for interim analysis. The data for the latter lags behind and is not available for all the patients included. The findings from this chapter in relation to the high correlation and validity of the mOHS in comparison with GOS, suggest that the mOHS have a potential role providing some additional useful data to the DMC. Among the different dichotomies, the one that categorises patients as fully dependent or dead against the other categories (dichotomy D) had the higher specificity (93.4%), and would be the most useful for interim analysis. In my opinion, the additional information provided by the mOHS should only be used for safety rather than for efficacy. I would not recommend to claim efficacy for an intervention based only on the early disability

outcome. According to the results from this chapter there are some changes in the disability status after hospital discharge and it would not be appropriate to assume that the effect on disability at hospital discharge reflects exactly the long term effect. However, it is possible that DMC could use mOHS for suspending a trial for safety reasons while awaiting the six months disability outcome for all the patients.

6.4.4.2 Use of the mOHS for dealing with loss to follow-up

Loss to follow-up is a common problem in RCT involving TBI patients. Ideally trials should be designed and carried out in such a way to minimise loss to follow-up, but sometimes missing data on outcomes is unavoidable. Different strategies are used to deal with missing data. Some of these strategies, like the parametric imputation method, use predictions based on statistical models fitted with variables associated with the missing variable.[133] The evidence that the mOHS adds predictive information to the other predictors already identified in this thesis suggests that mOHS is a good candidate to be used for such techniques when GOS is missing. Furthermore, as the mOHS is simple to obtain, its application is very practical in the context of clinical trials.

6.4.4.3 Use of the mOHS for communicating with relatives and patients

Finally, the positive predictive value of the different categories of mOHS to predict GOS at six months could be potentially useful for informing patients and relatives and give doctors a simple way to estimate prognosis at hospital discharge. From the patient and relative perspective it might be important to know that approximately 1 out of 10 patients who were discharged as fully dependent had a good recovery by six months. It could be argued that a prognostic model including all the predictors has a better performance, as measured with the c statistic, however a simple scale of disability has the advantage that it might be easier to use. First, because it uses data available at that moment and doctors don't need to rely on admission variables, such as GCS, pupil reactivity or CT scans results that could not be easily available at hospital discharge. Also it might be easier to remember because of its face value, as it predicts disability at six months based on another disability scale.

6.4.5 **Future research**

The association between mOHS and GOS, and the predictive validity of the mOHS should be examined in new cohorts of patients in order to confirm these findings. Ideally, new studies should include measurement of the mOHS at hospital discharge and also at 6 months. The analysis of the correlation of the mOHS at these two

different points in time could further inform about the usefulness of the mOHS when there is loss to follow-up of patients in RCT of TBI patients.

Simulation studies could evaluate the use of the mOHS for dealing with loss to follow up and its potential use for interim analysis.

Further studies should also explore the needs of prognosis in relation to disability from the patients' and relatives' perspectives, and should also evaluate the feasibility and practicality of the use of the mOHS for communication with them at hospital discharge.

Finally, new scales capturing more complex domains of disability, such as cognitive function should be devised and evaluated in future studies.

Chapter 7 Conclusions

7.1 Principal Findings

This thesis has shown that doctors consider prognostic information to be very important in the clinical management of patients with TBI, such as when deciding on medical treatment, surgery or withdrawing medical care. However, only a minority of clinicians think that they currently predict accurately, and most of them believe that a more accurate way of making prognosis, such as a prognostic model, would change the way they treat patients and communicate with patients and relatives.[120]

A systematic review, reported in this thesis, found that although prognostic models for TBI were frequently published, these have many limitations. Most models were developed on small samples, many were methodologically flawed, and few were validated in external populations. Few were presented in a clinically practical way, nor were they developed in populations from low and middle income countries, where most trauma occurs.[114]

In this thesis I also reported the development and validation of prognostic models using data from the CRASH trial.[117] These models, derived from a cohort of 10,008 patients with TBI , overcame many of the limitations from previous prognostic models. Two types of models were derived: one using only clinical and demographic variables (age, GCS, pupil reactivity, and extracranial injury), and the other using the previous variables plus CT scan results (petechial haemorrhages, obliteration of the third ventricle, midline shift, subarachnoid haemorrhage and non evacuated haematomas). The outcomes predicted in the models were mortality at 14 days and unfavourable outcome at six months, as defined by the GOS. Because some interactions were identified among predictors and outcomes by income region, different models were developed for HIC and LMIC. All of the models showed good internal validity. The models for HIC were validated in an external sample and showed good discrimination and calibration. An important finding was that GCS, measured as a continuous variable, had a good discriminative ability, similar to the one reported for the models in LMIC. In HIC, GCS showed lower discrimination. One hypothesis for this difference is that in HIC pre-hospital care is more common, so when patients arrive they are already sedated and so GCS does not only reflect TBI severity but also medical sedation. Intracranial bleeding (IB) was shown to be a strong predictor. Because of lack of data collected in the CRASH Trial I could not explore further the relationship between size if IB and poor outcome. But using a different dataset I explored this relationship, see appendix 6.

All the models developed are available in a web based calculator. Because internet availability is unlikely in the emergency department in LMIC, a paper based score was also developed for this setting. In this thesis I showed that this paper based score, named *"CRASH score card"*, was considered practical by doctors, and was used correctly by most of them.

I also showed that the mOHS, which is a simple disability scale obtained at hospital discharge, was strongly correlated with GOS at six months.[134] The mOHS dichotomy that showed the highest specificity in relation to the GOS was the one that separates patients dead or fully dependent from the rest (dichotomy D). It was shown that the mOHS adds predictive information to the prognostic models reported in Chapter 4. The predictive value of each Modified Oxford Handicap Scale (mOHS) category for three different outcomes at six months, according to GOS, was also reported.

7.2 Comparison with other studies

7.2.1 Doctors' perception about the importance of prognosis for TBI patients

There have been very few studies assessing the importance of prognosis from the perspective of physicians in the context of the care of patients. To the best of my knowledge the only previous study in the context of TBI patients was a survey conducted more than 20 years ago in a sample of a similar size as the one reported in this thesis.[85] This survey found similar results, most of the respondents thought that prognosis was a frequent practice, and that statistical prediction would improve the way they make predictions.

7.2.2 Systematic review of prognostic models for TBI

Previous to the publication of the systematic review reported in Chapter 3, only one systematic review had been published addressing the same question. This systematic review was more limited as it only included models with early indicators that predict mortality or unfavourable outcome defined by GOS in moderate or severe TBI.[83] The authors found 10 prognostic models and did not attempt to carry out a critical appraisal of the models. They evaluated their performance in an external cohort and reported that calibration and discrimination was worse than the original measures reported. The authors' conclusion emphasized the need for external validation. Subsequent to the publication of the systematic review reported in this thesis, a new systematic review was published which was also focused on the methodological appraisal of existing models.[99] They found 31 models and their conclusions were

similar to those reported in this thesis, the most common predictors included were GCS, age and pupil reactivity, models were derived from small samples, they were rarely validated, and they do not report adequate measures of performance.

7.2.3 **Prognostic models for TBI patients**

The prognostic models reported in Chapter 4 of this thesis differ from the prognostic models included in the systematic review in different ways: Firstly, specific models were derived for HIC and LMIC patients. Secondly, they were derived from the largest dataset of TBI patients, so the predictions are more precise. Thirdly, they were validated in a large dataset, and finally the models are available both as a web based calculator, and as a paper based score, for which practicality and acceptability was formally evaluated. To the best of my knowledge, this the first time that the presentation of a prognostic model for TBI patients has been evaluated by physicians.

Subsequent to the publication of the prognostic models reported in this thesis, the investigators from the IMPACT study developed and published similar prognostic models for six months mortality and unfavourable outcome.[115] Both, the IMPACT study and the models reported in this thesis, used large datasets, adequate methodology (e.g. discussion of the rationale of the predictors, clear definition of variables, correct strategy to build multivariable analysis, adequate handling of missing data). Furthermore, both studies reported important aspects such as characteristics of the sample, performance of the models so that physicians can make an informed judgment about the applicability in their own settings. The IMPACT "core" model and the CRASH "basic" model were very similar. Both CT models included subarachnoid haemorrhage as a predictor. Importantly, for both studies, an internal and external validation was performed. Finally, both studies made the models easily available with a web based calculator and in a simple paper based format. However, one of the strengths of the models I developed in this thesis is that the CRASH models were derived, not only from patients from HIC, but also from patients from LMIC. Another important difference is that the CRASH models were derived from a more recent period (1999-2004) in comparison with the IMPACT models (1984-1997), and that the CRASH dataset had very few missing variables, while for the IMPACT models imputation methods were necessary to handle missing data. Another difference was that, unlike for the IMPACT chart scores, the practicality of the CRASH Score Card was formally evaluated. On the other hand some of the IMPACT models included more variables that have been shown to be strong predictors (i.e. hypotension, hypoxia, haemoglobin and glycaemia). Although their added advantage to the more simple models is limited, it is possible that they add clinical credibility for physicians. Finally, for both studies the

discrimination was acceptable when evaluated in the external database but calibration was poor when tested with the Hosmer-Lemeshow test.

7.2.4 **Use of the Modified Oxford Handicap Scale**

Because the CRASH trial was the first to used the mOHS at hospital discharge, no previous study has assessed its relationship with disability at six months. The only study that compared one of the scales from which the mOHS was derived (the Modified Rankin Scale) also showed good agreement with the Glasgow Outcome Scale.[132]

7.3 Strengths and weaknesses

7.3.1 **Doctors' perception of the importance of prognosis for TBI patients**

Among the strengths of the survey reported in Chapter 2 of this thesis, it can be mentioned that it included mainly doctors from LMIC where most of the trauma occurs. However, the survey has some limitations. For example, it was a small convenience sample from a selected group of doctors who participated in the CRASH trial. It is also possible that because respondents were aware of the intention of developing prognostic models with the CRASH data, they responded more positively about the need for prognostic models, than what they really believe (desirability bias).

7.3.2 **Systematic review**

The systematic review reported in Chapter 3 had several strengths: the inclusion criteria were broad, so studies including patients with all the spectram of severity and predicting all types of outcomes were selected, and a thorough methodological description was conducted. However, it was not free of limitations, only studies from 1990 were included and some prognostic models could have been missed. Finally, systematic reviews of prognostic models are still in their infancy, and therefore there are still many methodological challenges, such as, valid search strategies, accepted risk of bias framework, or statistical methods to synthesize the results from the included studies.

7.3.3 **Prognostic models**

The prognostic models reported in Chapter 4 of this thesis, have many strengths. A large inception cohort of patients was used to develop the models; there was a prospective and standardized way of collecting exposures and outcomes, and very low loss to follow-up and missing data. One of its major strengths is that, as far as I am aware, it is the first time prognostic models for patients with TBI from LMIC have been developed from a large sample and using adequate methodology. Prognostic models

for HIC were externally validated and showed an acceptable discrimination. Another strength is that the models are available to clinicians in a web based calculator and as a practical paper based score to predict 14 day mortality in low & middle income countries.

Among the limitations of the prognostic models reported in this thesis is that they were derived from a clinical trial and this could influence the judgment about its external validity for some of the users. Some strong predictors were not available, such as hypotension and hypoxia. It has been shown that these variables do not add much discriminative ability once age, GCS, and pupil reactivity are already in the models. However, the clinical acceptability of the models could be influenced by this omission, as doctors have expressed their belief that hypotension and hypoxia are important predictors. Finally the prognostic models for LMIC were not validated in an external cohort of patients.

7.3.4 **Modified Oxford Handicap Scale**

The main strength of the analysis reported in this thesis in relation to the predictive ability of the mOHS is that it provides a simple and useful tool in an area largely unexplored. The substantial number of patients included in the analysis is a further strength. Nevertheless, there are limitations, the mOHS was not used in a structured way. In addition, it is possible that some of the investigators who completed the early disability scale also completed the GOS questionnaire at 6 months. However, it is unlikely they would remember the mOHS result at hospital discharge when completing the GOS questionnaire after six months, and furthermore, there are no clear parallel categories in the two forms completed.

7.4 Implications

7.4.1 **Prognostic models**

I developed user-friendly and practical prognostic models which showed good performance in internal and external validation and that can be used in the clinical setting.

Although external validity is always a matter of judgement, the fact that these were the first models to be developed from a large sample of patients from LMIC, makes them more likely to be relevant for this setting.

According to the results of this thesis, when externally evaluated, the models showed better discrimination than calibration. The relevance of the different components of accuracy will vary according to the setting where the prognostic model is applied. For example, in a setting where physicians need to allocate limited resources they might

be more interested in discrimination. While in other settings, where physicians use the model to inform patients and relatives, calibration might be more relevant.

However, two important questions should be answered to establish the clinical usefulness of these models. The first question is: *Do these prognostic models add predictive capacity to clinical prediction?*

In the survey reported in Chapter 2, the doctors identified as important predictors the variables that were subsequently included into the models. In addition, they reported that they frequently use the GCS to assess prognosis. In LMIC I found that the GCS had a very similar discriminative ability in comparison to the models I developed. So, it is possible that doctors, who are already using the main predictors included in the model, can discriminate as accurately as the models reported in this thesis. However, it is not clear how doctors combine the difference predictors and transform them into probabilities. There is no recommendation on how to do this, and it is likely that there is a great variability in this process.

The second question is related to the previous one: If the use of the prognostic models adds to clinical prediction, *does this have an impact on TBI patients' outcomes?*

This question remains unanswered. The main challenge to answer this question is that, currently, there is no evidence from randomised trials of effective interventions for the management of TBI patients.[61] Hopefully, in the future, trials will provide evidence of effective interventions for these patients. Randomised clinical trials will provide the relative effect of treatment, and with the baseline risk estimated with the prognostic models, specific individual recommendations will be able to be derived. This approach is useful for clinical practice, so that for any specific patient, one can judge the absolute benefit of a treatment (set against the absolute risk of side effects) and decide whether the treatment is indicated for that particular patient.[135]

The hypothetical scenario presented below exemplifies the potential usefulness of this approach for the management of patients with TBI. One of the potential targets of treatment for TBI patients is IB. As shown in appendix 6 there is evidence that the larger the IB the worse the outcome. There is also some evidence that IB progress within the first 24-48 hours, so haemostatic drugs administered in the first hours after a TBI could potentially be effective in reducing the progression of IB and therefore could be clinically useful. However, as there is yet no empirical evidence of this potential effect, I will assume the following facts for this example:

Tranexamic acid (TXA), a haemostatic, reduces mortality in TBI patients with a relative risk of 0.9 (in comparison to placebo), which is constant for patients with different

baseline risk. Deep venous thrombosis (DVT) is an adverse event associated with TXA which is observed in 0.8% of patients, and which is constant for all patients. Table 8.1 shows the baseline absolute risk of death, the relative risk of death associated with the use of TXA, the absolute risk reduction of death for patients treated with TXA, the absolute risk of DVT for patients treated with TXA. For clarity, I also present the absolute risk of death and DVT per 1,000 patients treated with TXA. Although the judgment about whether or not to use the intervention will depend on different circumstances, such as costs, resource constraints and patient values, it is likely that for patients with a higher baseline risk the intervention will be recommended, while in the low risk group, adverse events might offset the potential benefit. It is important to remember that there are two general assumptions that might not hold for all the situations and should be confirmed for this approach. The first assumption is that the relative reduction of the outcome we want to reduce with the treatment (in this case death) is constant for patients with different baseline risk (no evidence of interaction).The second assumption is that the adverse event rate (in this case DVT) is also constant for all the groups.

Table 7-1 Risks associated with TXA

Baseline risk of death	Relative risk of death	Absolute risk reduction	Absolute risk of DVT	Lives saved per 1,000 patients	DVT events per 1,000 patients
1%	0.9	0.1%	0.8%	1	8
30%	0.9	3%	0.8%	30	8
70%	0.9	7%	0.8%	70	8

Unfortunately, the current state of knowledge in TBI does not allow to use this approach. In a hypothetical future with evidence of effectiveness (and harms) of interventions, such an approach could be taken. Meanwhile, even in the absence of evidence of effective interventions, physicians still make routine clinical management decisions based on their own idiosyncratic way of estimating prognosis. But it is likely that the use of prognostic information differs by setting. For example, in the focus groups reported in Chapter 5, one of the respondents from Peru mentioned that prognostic models would be useful to define the need for intensive care because they had very limited resources. The photograph shown in figure 8.1, taken from an emergency department in Peru, speaks for itself about the need to make triage

decisions, and the potential use of an accurate prognostic model, in the context of limited resources.

Figure 7-1 Emergency department in a hospital in Peru

Facilities in most high income countries are much more widely available, and the use of prognostic information is likely to have different implications. So, as the use of prognostic information has different uses in different settings, it is quite difficult to assess the overall impact of prognostic models, even if they add to clinical prediction. However, it could be argued that if prognostic models are more predictive than clinical prediction, they would allow physicians to make better informed decisions. It is also possible that the existence of formal tools to assess prognosis will encourage physicians to discuss more frequently with patients and relatives about prognosis, and perhaps will enhance shared decision making. Ultimately, in my opinion, even discussing about prognostic models could make doctors more aware of how much prognosis is neglected in clinical practice. Prognosis is generally implicit, non systematic and not evidence based, something that nowadays would be unacceptable in the context of diagnosis and therapy practice.[28]

The prognostic models developed and validated in this thesis represent a step forward, as they provide a valid and practical tool for those doctors who are willing to explicitly use prognostic information in the context of the management of patients with TBI. This implication is backed up by the fact that a recent review article about the management of TBI patients published in *Lancet Neurology* referred to the models included in this

thesis as overcoming the limitations, and with greater validity and generalizability in comparison with previous models.[4]

Another potential use of prognostic models is in the design and analysis of clinical trials. In this thesis I concentrate on prognostic models in the context of clinical practice, but there is a large bibliography in relation to the use of prognostic models for research. For example, prognostic models could be used to select patients to be included in a trial, excluding those patients with very low or very high risk and less likely to benefit from any treatment. In addition, some authors suggest that adjustment by prognostic models could enhance the precision of the estimates increasing the power of clinical trials.[136,137] If these findings are confirmed, it is possible that simple pre-specified prognostic models, including the widely accepted three predictors (age, GCS and pupil reactivity), could be used when analysing clinical trials in TBI. However, if predictors are well balanced among treatment arms, it is unlikely that adjusted estimates will differ from the crude estimates. Furthermore, any increase in precision should be judged against the drawback of reporting a measure (adjusted odds ratio) which, possibly, would be more difficult to interpret and might have less acceptance by physicians.

7.4.2 **Modified Oxford Handicap Scale**

If the findings of the association between the mOHS and the GOS are replicated in a different population, then the mOHS could be used to inform interim analysis in randomised clinical trials in TBI patients. It could be a useful tool with which provide early data on disability while awaiting the six months data on GOS. The mOHS could also help tackling the problem of loss to follow-up by being used in imputation techniques for dealing with missing data. The fact that this scale is very simple is a further advantage as it can be easily collected at hospital discharge. Finally, it could also be used as a simple tool to inform patients and relatives about their prognosis at hospital discharge.

7.5 Future research

7.5.1 **Prognostic models**

There are two main aspects that should be further explored in relation to the prognostic models reported in this thesis, their validity and their clinical usefulness.

Validity should be evaluated in other populations, particularly the models from LMIC which were not externally validated. The models could subsequently be updated with the data of the new patients in the validation study. Two simple methods can be used to update the calibration of the models. The first is to adjust the intercept of the model

according to the incidence of the outcome in the new population. The second named "logistic recalibration" adjusts the regression coefficients if there is evidence of overfitting in the development study.[138] To update the discrimination, other methods referred as "model revision methods" could be used.[138] These include re-estimating the regression coefficient for some predictors for which the strength of the association is different from the original setting, or adding new predictors not included in the original models.

As discussed above, to determine whether the models are clinically useful there are two further questions that should be answered. First, do prognostic model add to clinical prediction? The predictive ability of physicians with and without the model should be compared in future studies. One difficulty in designing these studies will be related to the metrics used to compare their performance. The setting and purpose of the prediction will define the relative importance of the accuracy component to be evaluated (discrimination or calibration). An additional difficulty is related to the fact that although statistical "significance" can be defined for both components, clinical "significance" is not easy to define. There are some standard definitions for the c statistic as "very good" "good" "acceptable" "poor" but these definitions are arbitrary. The calibration measurement is not without difficulties either. In general, it is estimated graphically, and statements as "there is evidence of good calibration" are always a matter for judgment. The statistical test to evaluate calibration, Hosmer Lemeshow test, has also some drawbacks. If the sample is too large the test could be "statistically significant" but the difference detected among "observed" and "predicted" might not be clinically meaningful. Furthermore, it compares deciles of risk, a concept that is not very meaningful to doctors.

New evidence about effective treatments may in the future provide meaningful threshold of probabilities for which specific interventions are recommended, and more useful metrics, such as net reclassification improvement, could be reported.

In this scenario, in addition, studies to answer the second important question "*Do prognostic models improve patients' outcomes?*" would be feasible to conduct. Once effective interventions become available according to the estimated risk, then impact analysis through cluster randomised clinical trials could be conducted. Some could argue that randomised clinical trials should be conducted even in the absence of effective interventions according to baseline risk. But, in my opinion, the problem with this approach is that the "intervention" (i.e. the use of prognostic models) has a complex causal pathway between its implementation and patients' outcomes. It is likely that the "uptake" of prognostic models, its interpretation, and the decisions made according to the risk estimated will differ from one setting to another. Because

there are no clear recommendations according to baseline risk, the use of prognostic information implies different actions in different settings, so the effect of the "prognostic model strategy" is likely to differ according to the setting. If a randomised clinical trial is conducted to evaluate the impact of prognostic models in TBI it should report, not only the main clinical outcomes, mortality and disability, but it should also report the effect on intermediate steps, and different outcomes according to the setting's priorities (e.g. physician's or patient's satisfaction, cost, etc). This approach would allow a better understanding of the impact of the use of prognostic model in TBI in different settings.

7.5.2 **Modified Oxford Handicap Scale**

The simple disability scale at hospital discharge showed good correlation and predictive ability in relation to unfavourable outcome, as defined by the GOS. Future research could validate these findings and formally assess the potential use of this scale to tackle the problem of loss to follow-up in trials including TBI patients. Simulation studies using different techniques such as multiple imputation could evaluate this approach. In addition the potential use of this scale as an early outcome for interim analysis for these trials could also be assessed through simulation studies. Finally, future studies could also explore, the practicality and acceptability of using the mOHS when communicating with patients at hospital discharge.

REFERENCES

1. CDC. What is Traumatic Brain Injury? http://www.cdc.gov/ncipc/tbi/TBI.htm accessed March 2009.

2. Ghajar J. Traumatic brain injury. *Lancet* 2000;**356**(9233):923-9.

3. Vincent JL, Berre J. Primer on medical management of severe brain injury. *Crit Care Med* 2005;**33**(6):1392-9.

4. Maas AI, Stocchetti N, Bullock R. Moderate and severe traumatic brain injury in adults. *Lancet Neurol* 2008;**7**(8):728-41.

5. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974;**2**(7872):81-4.

6. Anonymous. Part 2: Prognosis in penetrating brain injury. *J Trauma* 2001;**51**(2 Suppl):S44-86.

7. Peek-Asa C, McArthur D, Hovda D, et al. Early predictors of mortality in penetrating compared with closed brain injury. *Brain Inj* 2001;**15**(9):801-10.

8. Bilbao A, Kennedy C, Chatterji S, et al. The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *NeuroRehabilitation* 2003;**18**(3):239-50.

9. Van Baalen B, Odding E, Maas AI, et al. Traumatic brain injury: classification of initial severity and determination of functional outcome. *Disabil Rehabil* 2003;**25**(1):9-18.

10. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet* 1975;**1**(7905):480-4.

11. Bullock MR, Merchant RE, Choi SC, et al. Outcome measures for clinical trials in neurotrauma. *Neurosurg Focus* 2002;**13**(1):ECP1.

12. Wilson JT, Pettigrew LE, Teasdale GM. Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. *J Neurotrauma* 1998;**15**(8):573-85.

13. Dickinson K, Bunn F, Wentz R, et al. Size and quality of randomised controlled trials in head injury: review of published studies. *Bmj* 2000;**320**(7245):1308-11.

14. Edwards P, Fernandes J, Roberts I, et al. Young men were at risk of becoming lost to follow-up in a cohort of head-injured adults. *J Clin Epidemiol* 2007;**60**(4):417-24.

15. Lopez AD, Mathers CD, Ezzati M, et al. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006;**367**(9524):1747-57.

16. Hofman K, Primack A, Keusch G, et al. Addressing the growing burden of trauma and injury in low- and middle-income countries. *Am J Public Health* 2005;**95**(1):13-7.

17. Langlois JA, Rutland-Brown W, Wald MM. The epidemiology and impact of traumatic brain injury: a brief overview. *J Head Trauma Rehabil* 2006;**21**(5):375-8.

18. Thurman DJ, Alverson C, Dunn KA, et al. Traumatic brain injury in the United States: A public health perspective. *J Head Trauma Rehabil* 1999;**14**(6):602-15.

19. Hyder AA, Wunderlich CA, Puvanachandra P, et al. The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation* 2007;**22**(5):341-53.

20. Puvanachandra P, Hyder AA. Traumatic brain injury in Latin America and the Caribbean: a call for research. *Salud Publica Mex* 2008;**50 Suppl 1**(1):S3-5.

21. Bruns J, Jr., Hauser WA. The epidemiology of traumatic brain injury: a review. *Epilepsia* 2003;**44 Suppl 10**:2-10.

22. Narayan RK, Michel ME, Ansell B, et al. Clinical trials in head injury. *J Neurotrauma* 2002;**19**(5):503-57.

23. Windeler J. Prognosis - what does the clinician associate with this notion? *Stat Med* 2000;**19**(4):425-30.

24. Hayden JA, Cote P, Steenstra IA, et al. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *J Clin Epidemiol* 2008;**61**(6):552-60.

25. Hemingway H. Prognosis research: why is Dr. Lydgate still waiting? *J Clin Epidemiol* 2006;**59**(12):1229-38.

26. Christakis NA. Death foretold: prophecy and prognosis in medical care. Chicago IL: University of Chicago Press, 1999.

27. Christakis NA. The ellipsis of prognosis in modern medical thought. *Soc Sci Med* 1997;**44**(3):301-15.

28. Kellett J. Prognostication--the lost skill of medicine. *Eur J Intern Med* 2008;**19**(3):155-64.

29. Christakis NA, Iwashyna TJ. Attitude and self-reported practice regarding prognostication in a national sample of internists. *Arch Intern Med* 1998;**158**(21):2389-95.

30. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;**19**(4):453-73.

31. Altman DG. Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;**323**(7306):224-228.

32. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules. Applications and methodological standards. *N Engl J Med* 1985;**313**(13):793-9.

33. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med* 1999;**33**(4):437-47.

34. Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? *BMJ* 1995;**311**(7019):1539-1541.

35. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 1997;**277**(6):488-94.

36. Redelmeier DA, Lustig AJ. Prognostic indices in clinical practice. *Jama* 2001;**285**(23):3024-5.

37. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006;**144**(3):201-9.

38. Rothwell PM. Prognostic models. *Pract Neurol* 2008;**8**(4):242-53.

39. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Third ed. Philadelphia: Lippincott Williams & Wilkins, 2008.

40. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**(6):515-24.

41. Steyerberg EW. Clinical Prediction Models. New York: Springer, 2009.

42. Bartfay E, Bartfay WJ. Accuracy assessment of prediction in patient outcomes. *J Eval Clin Pract* 2008;**14**(1):1-10.

43. McGeechan K, Macaskill P, Irwig L, et al. Assessing new biomarkers and predictive models for use in clinical practice: a clinician's guide. *Arch Intern Med* 2008;**168**(21):2304-10.

44. Poses RM, Bekes C, Copare FJ, et al. The answer to "What are my chances, doctor?" depends on whom is asked: prognostic disagreement and inaccuracy for critically ill patients. *Crit Care Med* 1989;**17**(8):827-33.

45. Glare P, Virik K, Jones M, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *Bmj* 2003;**327**(7408):195-8.

46. Brandt HE, Ooms ME, Ribbe MW, et al. Predicted survival vs. actual survival in terminally ill noncancer patients in Dutch nursing homes. *J Pain Symptom Manage* 2006;**32**(6):560-6.

47. Pignone M, Phillips CJ, Elasy TA, et al. Physicians' ability to predict the risk of coronary heart disease. *BMC Health Serv Res* 2003;**3**(1):13.

48. Knaus WA, Wagner DP, Lynn J. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. *Science* 1991;**254**(5030):389-94.

49. Liao L, Mark DB. Clinical prediction models: are we building better mousetraps? *J Am Coll Cardiol* 2003;**42**(5):851-3.

50. Grove WM. Clinical versus statistical prediction: the contribution of Paul E. Meehl. *J Clin Psychol* 2005;**61**(10):1233-43.

51. Grove WM, Zald DH, Lebow BS, et al. Clinical versus mechanical prediction: a meta-analysis. *Psychol Assess* 2000;**12**(1):19-30.

52. Anonymous. A controlled trial to improve care for seriously ill hospitalized patients. The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). The SUPPORT Principal Investigators. *Jama* 1995;**274**(20):1591-8.

53. Murray GD, Murray LS, Barlow P, et al. Assessing the performance and clinical impact of a computerized prognostic system in severe head injury. *Stat Med* 1986;**5**(5):403-10.

54. Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *Jama* 2005;**293**(10):1223-38.

55. Jackson J. Primary prevention of cardiovascular disease: the absolute-risk-based approach. In: Lancet T, ed. Treating individuals: from randomised trials to personalised medicine, 2007.

56. Schull MJ, Ferris LE, Tu JV, et al. Problems for clinical judgement: 3. Thinking clearly in an emergency. *Cmaj* 2001;**164**(8):1170-5.

57. Alexander M. Bias and asymmetric loss in expert forecasts: a study of physician prognostic behavior with respect to patient survival. *J Health Econ* 2008;**27**(4):1095-108.

58. Ker K, Perel P, Blackhall K, et al. How effective are some common treatments for traumatic brain injury? *Bmj* 2008;**337**(337):a865.

59. Schouten JW. Neuroprotection in traumatic brain injury: a complex struggle against the biology of nature. *Curr Opin Crit Care* 2007;**13**(2):134-42.

60. CRASH trial Collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004;**364**(9442):1321-8.

61. Narayan RK, Maas AI, Servadei F, et al. Progression of traumatic intracerebral hemorrhage: a prospective observational study. *J Neurotrauma* 2008;**25**(6):629-39.

62. Brain Trauma Foundation (U.S.) AAoNS. Management and prognosis of severe traumatic brain injury. New York, 2000.

63. Udekwu P, Kromhout-Schiro S, Vaslef S, et al. Glasgow Coma Scale score, mortality, and functional outcome in head-injured patients. *J Trauma* 2004;**56**(5):1084-9.

64. Miller KJ, Schwab KA, Warden DL. Predictive value of an early Glasgow Outcome Scale score: 15-month score changes. *J Neurosurg* 2005;**103**(2):239-45.

65. Eftekhar B, Zarei MR, Ghodsi M, et al. Comparing logistic models based on modified GCS motor component with other prognostic tools in prediction of mortality: results of study in 7226 trauma patients. *Injury* 2005;**36**(8):900-4.

66. Hukkelhoven CW, Steyerberg EW, Rampen AJ, et al. Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. *J Neurosurg* 2003;**99**(4):666-73.

67. Boto GR, Gomez PA, De la Cruz J, et al. [Prognostic factors in severe head injury]. *Neurocirugia (Astur)* 2004;**15**(3):233-47.

68. Lobato RD, Cordobes F, Rivas JJ, et al. Outcome from severe head injury related to the type of intracranial lesion. A computerized tomography study. *J Neurosurg* 1983;**59**(5):762-74.

69. Marshall L F, Marshal SB, Klauber MR, et al. A new classification of head injury based on computerized tomography. *J Neurosurg (Suppl) 75:14-20* 1991.

70. Wardlaw JM, Easton VJ, Statham P. Which CT features help predict outcome after head injury? *J Neurol Neurosurg Psychiatry* 2002;**72**(2):188-92; discussion 151.

71. Knuckey NW, Gelbard S, Epstein MH. The management of "asymptomatic" epidural hematomas. A prospective study. *J Neurosurg* 1989;**70**(3):392-6.

72. Chen TY, Wong CW, Chang CN, et al. The expectant treatment of "asymptomatic" supratentorial epidural hematomas. *Neurosurgery* 1993;**32**(2):176-9; discussion 179.

73. Bullock R, Golek J, Blake G. Traumatic intracerebral hematoma--which patients should undergo surgical evacuation? CT scan features and ICP monitoring as a basis for decision making. *Surg Neurol* 1989;**32**(3):181-7.

74. Yanaka K, Kamezaki T, Yamada T, et al. Acute subdural hematoma--prediction of outcome with a linear discriminant function. *Neurol Med Chir (Tokyo)* 1993;**33**(8):552-8.

75. Maas AI, Hukkelhoven CW, Marshall LF, et al. 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* 2005;**57**(6):1173-82; discussion 1173-82.

76. Houlden H, Greenwood R. Apolipoprotein E4 and traumatic brain injury. *J Neurol Neurosurg Psychiatry* 2006;**77**(10):1106-7.

77. Zhou W, Xu D, Peng X, et al. Meta-analysis of APOE4 allele and outcome after traumatic brain injury. *J Neurotrauma* 2008;**25**(4):279-90.

78. Townend WJ, Guy MJ, Pani MA, et al. Head injury outcome prediction in the emergency department: a role for protein S-100B? *J Neurol Neurosurg Psychiatry* 2002;**73**(5):542-6.

79. Jennett B, Teasdale G, Braakman R, et al. Predicting outcome in individual patients after severe head injury. *Lancet* 1976;**1**(7968):1031-4.

80. Hukkelhoven CW, Rampen AJ, Maas AI, et al. Some prognostic models for traumatic brain injury were not valid. *J Clin Epidemiol* 2006;**59**(2):132-43.

81. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med* 1996;**125**(5):406-12.

82. Barlow P, Teasdale G. Prediction of outcome and the management of severe head injuries: the attitudes of neurosurgeons. *Neurosurgery* 1986;**19**(6):989-91.

83. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001;**323**(7303):42-6.

84. Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol* 2005;**5**:19.

85. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaboration, 2008.

86. Counsell C, Dennis M. Systematic review of prognostic models in patients with acute stroke. *Cerebrovasc Dis* 2001;**12**(3):159-70.

87. Pillai SV, Kolluri VR, Praharaj SS. Outcome prediction model for severe diffuse brain injuries: development and evaluation. *Neurol India* 2003;**51**(3):345-9.

88. Signorini DF, Andrews PJD, Jones PA, et al. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *J Neurol Neurosurg Psychiatry* 1999;**66**(1):20-25.

89. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Admission of patients with severe and moderate traumatic brain injury to specialized ICU facilities: a search for triage criteria. *Intensive Care Med* 2005;**31**(6):799-806.

90. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. *J Neurotrauma* 2005;**22**(10):1025-39.

91. Bush BA, Novack TA, Malec JF, et al. Validation of a model for evaluating outcome after traumatic brain injury. *Arch Phys Med Rehabil* 2003;**84**(12):1803-7.

92. Benzer A, Mitterschiffthaler G, Marosi M, et al. Prediction of non-survival after trauma: Innsbruck Coma Scale. *Lancet* 1991;**338**(8773):977-8.

93. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006;**144**(6):427-37.

94. Concato J, Feinstein AR, Holford TR. The Risk of Determining Risk with Multivariable Models. *Ann Intern Med* 1993;**118**(3):201-210.

95. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**(4):361-87.

96. Mushkudiani NA, Hukkelhoven CW, Hernandez AV, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. *J Clin Epidemiol* 2008;**61**(4):331-43.

97. CRASH trial Collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months. *Lancet* 2005;**365**(9475):1957-9.

98. Wilson JT, Edwards P, Fiddes H, et al. Reliability of postal questionnaires for the Glasgow Outcome Scale. *J Neurotrauma* 2002;**19**(9):999-1005.

99. World Bank. World Development Indicators 2006.

100. Maas AI, Marmarou A, Murray GD, et al. Prognosis and clinical trial design in traumatic brain injury: the IMPACT study. *J Neurotrauma* 2007;**24**(2):232-8.

101. Marion DW, Carlier PM. Problems with initial Glasgow Coma Scale assessment caused by prehospital treatment of patients with head injuries: results of a national survey. *J Trauma* 1994;**36**(1):89-95.

102. McNett M. A review of the predictive ability of Glasgow Coma Scale scores in head-injured patients. *J Neurosci Nurs* 2007;**39**(2):68-75.

103. Heinzelmann M, Platz A, Imhof HG. Outcome after acute extradural haematoma, influence of additional injuries and neurological complications in the ICU. *Injury* 1996;**27**(5):345-9.

104. Stiver SI, Manley GT. Prehospital management of traumatic brain injury. *Neurosurg Focus* 2008;**25**(4):E5.

105. Butcher I, McHugh GS, Lu J, et al. Prognostic value of cause of injury in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):281-6.

106. Slewa-Younan S, van den Berg S, Baguley IJ, et al. Towards an understanding of sex differences in functional outcome following moderate to severe traumatic brain injury: a systematic review. *J Neurol Neurosurg Psychiatry* 2008;**79**(11):1197-201.

107. Maas AI, Steyerberg EW, Butcher I, et al. Prognostic value of computerized tomography scan characteristics in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):303-14.

108. Bullock R, Chesnut R, Ghajar J, et al. Guidelines for the Surgical Management of Traumatic Brain Injury. *Neurosurgery* 2006;**58**(3):S1-1 S2-62.

109. Hukkelhoven CW, Steyerberg EW, Farace E, et al. Regional differences in patient characteristics, case management, and outcomes in traumatic brain injury: experience from the tirilazad trials. *J Neurosurg* 2002;**97**(3):549-57.

110. Perel P, Edwards P, Wentz R, et al. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* 2006;**6**(38):38.

111. Steyerberg EW, Mushkudiani N, Perel P, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med* 2008;**5**(8):e165; discussion e165.

112. Murray GD, Butcher I, McHugh GS, et al. Multivariable prognostic analysis in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):329-37.

113. CRASH trial Collaborators. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;**12:**12.

114. Anonymous. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *Jama* 2001;**285**(19):2486-97.

115. Anonymous. Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association. *Bmj* 2000;**320**(7236):705-8.

116. Perel P, Wasserberg J, Ravi RR, et al. Prognosis following head injury: a survey of doctors from developing and developed countries. *J Eval Clin Pract* 2007;**13**(3):464-5.

117. Edwards A, Elwyn G, Gwyn R. General practice registrar responses to the use of different risk communication tools in simulated consultations: a focus group study. *Bmj* 1999;**319**(7212):749-52.

118. Junghans C, Feder G, Timmis AD, et al. Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. *Arch Intern Med* 2007;**167**(2):195-202.

119. Sirovich BE, Gottlieb DJ, Welch HG, et al. Variation in the tendency of primary care physicians to intervene. *Arch Intern Med* 2005;**165**(19):2252-6.

120. Bachmann LM, Muhleisen A, Bock A, et al. Vignette studies of medical choice and judgement to study caregivers' medical decision behaviour: systematic review. *BMC Med Res Methodol* 2008;**8**(50):50.

121. Wilson JT. Assessing outcome in head injury trials. *Curr Pharm Des* 2001;**7**(15):1537-52.

122. Donders AR, van der Heijden GJ, Stijnen T, et al. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;**59**(10):1087-91.

123. Junque C, Bruna O, Mataro M. Information needs of the traumatic brain injury patient's family members regarding the consequences of the injury and associated perception of physical, cognitive, emotional and quality of life changes. *Brain Inj* 1997;**11**(4):251-8.

124. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott Med J* 1957;**2**(5):200-15.

125. Van Swieten JC, Koudstaal PJ, Visser MC, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;**19**(5):604-7.

126. Bamford JM, Sandercock PA, Warlow CP, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1989;**20**(6):828.

127. New PW, Buchbinder R. Critical appraisal and review of the Rankin scale and its derivatives. *Neuroepidemiology* 2006;**26**(1):4-15.

128. MRC CRASH Trial Collaborators. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;**12**:12.

129. Tilley BC, Marler J, Geller NL, et al. Use of a global test for multiple outcomes in stroke trials with application to the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial. *Stroke* 1996;**27**(11):2136-42.

130. Wang D, Bakhai A. Clinical Trials A Practical Guide to Design, Analysis, and Reporting. London: Remedica, 2006.

131. Bullock R, Chesnut R, Clifton GL, et al. Management and prognosis of severe traumatic brain injury. *J Neurotrauma* 2000;**17**:451-627.

132. Oertel M, Kelly DF, McArthur D, et al. Progressive hemorrhage after head trauma: predictors and consequences of the evolving injury. *J Neurosurg* 2002;**96**(1):109-16.

133. AAM. The Abbreviated injury scale 1990 revision. Illinois, 1990.

134. Patel HC, Bouamra O, Woodford M, et al. Trends in head injury outcome from 1989 to 2003 and the effect of neurosurgical care: an observational study. *Lancet* 2005;**366**(9496):1538-44.

135. Baker SP, O'Neill B. The injury severity score: an update. *J Trauma* 1976;**16**(11):882-5.

136. Perel P, Edwards P, Shakur H, et al. Use of the Oxford Handicap Scale at hospital discharge to predict Glasgow Outcome Scale at 6 months in patients with traumatic brain injury. *BMC Med Res Methodol* 2008;**8**(72):72.

137. Pocock SJ, Lubsen J. More on subgroup analyses in clinical trials. *N Engl J Med* 2008;**358**(19):2076; author reply 2076-7.

138. Hernandez AV, Eijkemans MJ, Steyerberg EW. Randomized controlled trials with time-to-event outcomes: how much does prespecified covariate adjustment increase power? *Ann Epidemiol* 2006;**16**(1):41-8.

139. Hernandez AV, Steyerberg EW, Habbema JD. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol* 2004;**57**(5):454-60.

140. Toll DB, Janssen KJ, Vergouwe Y, et al. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008;**61**(11):1085-94.

1. CDC. What is Traumatic Brain Injury? http://www.cdc.gov/ncipc/tbi/TBI.htm accessed March 2009.

2. Ghajar J. Traumatic brain injury. *Lancet* 2000;**356**(9233):923-9.

3. Vincent JL, Berre J. Primer on medical management of severe brain injury. *Crit Care Med* 2005;**33**(6):1392-9.

4. Maas AI, Stocchetti N, Bullock R. Moderate and severe traumatic brain injury in adults. *Lancet Neurol* 2008;**7**(8):728-41.

5. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. *Lancet* 1974;**2**(7872):81-4.

6. Fischer J, Mathieson C. The history of the Glasgow Coma Scale: implications for practice. *Crit Care Nurs Q* 2001;**23**(4):52-8.

7. Anonymous. Part 2: Prognosis in penetrating brain injury. *J Trauma* 2001;**51**(2 Suppl):S44-86.

8. Peek-Asa C, McArthur D, Hovda D, et al. Early predictors of mortality in penetrating compared with closed brain injury. *Brain Inj* 2001;**15**(9):801-10.

9. Bilbao A, Kennedy C, Chatterji S, et al. The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. *NeuroRehabilitation* 2003;**18**(3):239-50.

10. Van Baalen B, Odding E, Maas AI, et al. Traumatic brain injury: classification of initial severity and determination of functional outcome. *Disabil Rehabil* 2003;**25**(1):9-18.

11. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet* 1975;**1**(7905):480-4.

12. Bullock MR, Merchant RE, Choi SC, et al. Outcome measures for clinical trials in neurotrauma. *Neurosurg Focus* 2002;**13**(1):ECP1.

13. Wilson JT, Pettigrew LE, Teasdale GM. Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: guidelines for their use. *J Neurotrauma* 1998;**15**(8):573-85.

14. Dickinson K, Bunn F, Wentz R, et al. Size and quality of randomised controlled trials in head injury: review of published studies. *Bmj* 2000;**320**(7245):1308-11.

15. Edwards P, Fernandes J, Roberts I, et al. Young men were at risk of becoming lost to follow-up in a cohort of head-injured adults. *J Clin Epidemiol* 2007;**60**(4):417-24.

16. Wilson JT. Assessing outcome in head injury trials. *Curr Pharm Des* 2001;**7**(15):1537-52.

17. Bullinger M, Azouvi P, Brooks N, et al. Quality of life in patients with traumatic brain injury-basic issues, assessment and recommendations. *Restor Neurol Neurosci* 2002;**20**(3-4):111-24.

18. Lopez AD, Mathers CD, Ezzati M, et al. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006;**367**(9524):1747-57.

19. Hofman K, Primack A, Keusch G, et al. Addressing the growing burden of trauma and injury in low- and middle-income countries. *Am J Public Health* 2005;**95**(1):13-7.

20. Langlois JA, Rutland-Brown W, Wald MM. The epidemiology and impact of traumatic brain injury: a brief overview. *J Head Trauma Rehabil* 2006;**21**(5):375-8.

21. Thurman DJ, Alverson C, Dunn KA, et al. Traumatic brain injury in the United States: A public health perspective. *J Head Trauma Rehabil* 1999;**14**(6):602-15.

22. Hyder AA, Wunderlich CA, Puvanachandra P, et al. The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation* 2007;**22**(5):341-53.

23. Puvanachandra P, Hyder AA. Traumatic brain injury in Latin America and the Caribbean: a call for research. *Salud Publica Mex* 2008;**50 Suppl 1**(1):S3-5.

24. Bruns J, Jr., Hauser WA. The epidemiology of traumatic brain injury: a review. *Epilepsia* 2003;**44 Suppl 10**:2-10.

25. Narayan RK, Michel ME, Ansell B, et al. Clinical trials in head injury. *J Neurotrauma* 2002;**19**(5):503-57.

26. Windeler J. Prognosis - what does the clinician associate with this notion? *Stat Med* 2000;**19**(4):425-30.

27. Hayden JA, Cote P, Steenstra IA, et al. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *J Clin Epidemiol* 2008;**61**(6):552-60.

28. Hemingway H. Prognosis research: why is Dr. Lydgate still waiting? *J Clin Epidemiol* 2006;**59**(12):1229-38.

29. Christakis NA. Death foretold: prophecy and prognosis in medical care. Chicago IL: University of Chicago Press, 1999.

30. Christakis NA. The ellipsis of prognosis in modern medical thought. *Soc Sci Med* 1997;**44**(3):301-15.

31. Kellett J. Prognostication--the lost skill of medicine. *Eur J Intern Med* 2008;**19**(3):155-64.

32. Christakis NA, Iwashyna TJ. Attitude and self-reported practice regarding prognostication in a national sample of internists. *Arch Intern Med* 1998;**158**(21):2389-95.

33. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;**19**(4):453-73.

34. Altman DG. Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;**323**(7306):224-228.

35. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules. Applications and methodological standards. *N Engl J Med* 1985;**313**(13):793-9.

36. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med* 1999;**33**(4):437-47.

37. Wyatt JC, Altman DG. Commentary: Prognostic models: clinically useful or quickly forgotten? *BMJ* 1995;**311**(7019):1539-1541.

38. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA* 1997;**277**(6):488-94.

39. Redelmeier DA, Lustig AJ. Prognostic indices in clinical practice. *Jama* 2001;**285**(23):3024-5.

40. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006;**144**(3):201-9.

41. Rothwell PM. Prognostic models. *Pract Neurol* 2008;**8**(4):242-53.

42. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Third ed. Philadelphia: Lippincott Williams & Wilkins, 2008.

43. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**(6):515-24.

44. Steyerberg EW. Clinical Prediction Models. New York: Springer, 2009.

45. Bartfay E, Bartfay WJ. Accuracy assessment of prediction in patient outcomes. *J Eval Clin Pract* 2008;**14**(1):1-10.

46. McGeechan K, Macaskill P, Irwig L, et al. Assessing new biomarkers and predictive models for use in clinical practice: a clinician's guide. *Arch Intern Med* 2008;**168**(21):2304-10.

47. Poses RM, Bekes C, Copare FJ, et al. The answer to "What are my chances, doctor?" depends on whom is asked: prognostic disagreement and inaccuracy for critically ill patients. *Crit Care Med* 1989;**17**(8):827-33.

48. Glare P, Virik K, Jones M, et al. A systematic review of physicians' survival predictions in terminally ill cancer patients. *Bmj* 2003;**327**(7408):195-8.

49. Brandt HE, Ooms ME, Ribbe MW, et al. Predicted survival vs. actual survival in terminally ill noncancer patients in Dutch nursing homes. *J Pain Symptom Manage* 2006;**32**(6):560-6.

50. Pignone M, Phillips CJ, Elasy TA, et al. Physicians' ability to predict the risk of coronary heart disease. *BMC Health Serv Res* 2003;**3**(1):13.

51. Knaus WA, Wagner DP, Lynn J. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. *Science* 1991;**254**(5030):389-94.

52. Liao L, Mark DB. Clinical prediction models: are we building better mousetraps? *J Am Coll Cardiol* 2003;**42**(5):851-3.

53. Grove WM. Clinical versus statistical prediction: the contribution of Paul E. Meehl. *J Clin Psychol* 2005;**61**(10):1233-43.

54. Grove WM, Zald DH, Lebow BS, et al. Clinical versus mechanical prediction: a meta-analysis. *Psychol Assess* 2000;**12**(1):19-30.

55. Anonymous. A controlled trial to improve care for seriously ill hospitalized patients. The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). The SUPPORT Principal Investigators. *Jama* 1995;**274**(20):1591-8.

56. Murray GD, Murray LS, Barlow P, et al. Assessing the performance and clinical impact of a computerized prognostic system in severe head injury. *Stat Med* 1986;**5**(5):403-10.

57. Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *Jama* 2005;**293**(10):1223-38.

58. Jackson J. Primary prevention of cardiovascular disease: the absolute-risk-based approach. In: Lancet T, ed. Treating individuals: from randomised trials to personalised medicine, 2007.

59. Schull MJ, Ferris LE, Tu JV, et al. Problems for clinical judgement: 3. Thinking clearly in an emergency. *Cmaj* 2001;**164**(8):1170-5.

60. Alexander M. Bias and asymmetric loss in expert forecasts: a study of physician prognostic behavior with respect to patient survival. *J Health Econ* 2008;**27**(4):1095-108.

61. Ker K, Perel P, Blackhall K, et al. How effective are some common treatments for traumatic brain injury? *Bmj* 2008;**337**(337):a865.

62. Schouten JW. Neuroprotection in traumatic brain injury: a complex struggle against the biology of nature. *Curr Opin Crit Care* 2007;**13**(2):134-42.

63. CRASH trial Collaborators. Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004;**364**(9442):1321-8.

64. Narayan RK, Maas AI, Servadei F, et al. Progression of traumatic intracerebral hemorrhage: a prospective observational study. *J Neurotrauma* 2008;**25**(6):629-39.

65. Brain Trauma Foundation (U.S.) AAoNS. Management and prognosis of severe traumatic brain injury. New York, 2000.

66. Udekwu P, Kromhout-Schiro S, Vaslef S, et al. Glasgow Coma Scale score, mortality, and functional outcome in head-injured patients. *J Trauma* 2004;**56**(5):1084-9.

67. Miller KJ, Schwab KA, Warden DL. Predictive value of an early Glasgow Outcome Scale score: 15-month score changes. *J Neurosurg* 2005;**103**(2):239-45.

68. Eftekhar B, Zarei MR, Ghodsi M, et al. Comparing logistic models based on modified GCS motor component with other prognostic tools in prediction of mortality: results of study in 7226 trauma patients. *Injury* 2005;**36**(8):900-4.

69. Hukkelhoven CW, Steyerberg EW, Rampen AJ, et al. Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. *J Neurosurg* 2003;**99**(4):666-73.

70. Boto GR, Gomez PA, De la Cruz J, et al. [Prognostic factors in severe head injury]. *Neurocirugia (Astur)* 2004;**15**(3):233-47.

71. Lobato RD, Cordobes F, Rivas JJ, et al. Outcome from severe head injury related to the type of intracranial lesion. A computerized tomography study. *J Neurosurg* 1983;**59**(5):762-74.

72. Marshall L F, Marshal SB, Klauber MR, et al. A new classification of head injury based on computerized tomography. *J Neurosurg (Suppl) 75:14-20* 1991.

73. Wardlaw JM, Easton VJ, Statham P. Which CT features help predict outcome after head injury? *J Neurol Neurosurg Psychiatry* 2002;**72**(2):188-92; discussion 151.

74. Knuckey NW, Gelbard S, Epstein MH. The management of "asymptomatic" epidural hematomas. A prospective study. *J Neurosurg* 1989;**70**(3):392-6.

75. Chen TY, Wong CW, Chang CN, et al. The expectant treatment of "asymptomatic" supratentorial epidural hematomas. *Neurosurgery* 1993;**32**(2):176-9; discussion 179.

76. Bullock R, Golek J, Blake G. Traumatic intracerebral hematoma--which patients should undergo surgical evacuation? CT scan features and ICP monitoring as a basis for decision making. *Surg Neurol* 1989;**32**(3):181-7.

77. Yanaka K, Kamezaki T, Yamada T, et al. Acute subdural hematoma--prediction of outcome with a linear discriminant function. *Neurol Med Chir (Tokyo)* 1993;**33**(8):552-8.

78. Maas AI, Hukkelhoven CW, Marshall LF, et al. 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* 2005;**57**(6):1173-82; discussion 1173-82.

79. Houlden H, Greenwood R. Apolipoprotein E4 and traumatic brain injury. *J Neurol Neurosurg Psychiatry* 2006;**77**(10):1106-7.

80. Zhou W, Xu D, Peng X, et al. Meta-analysis of APOE4 allele and outcome after traumatic brain injury. *J Neurotrauma* 2008;**25**(4):279-90.

81. Townend WJ, Guy MJ, Pani MA, et al. Head injury outcome prediction in the emergency department: a role for protein S-100B? *J Neurol Neurosurg Psychiatry* 2002;**73**(5):542-6.

82. Jennett B, Teasdale G, Braakman R, et al. Predicting outcome in individual patients after severe head injury. *Lancet* 1976;**1**(7968):1031-4.

83. Hukkelhoven CW, Rampen AJ, Maas AI, et al. Some prognostic models for traumatic brain injury were not valid. *J Clin Epidemiol* 2006;**59**(2):132-43.

84. Braitman LE, Davidoff F. Predicting clinical states in individual patients. *Ann Intern Med* 1996;**125**(5):406-12.

85. Barlow P, Teasdale G. Prediction of outcome and the management of severe head injuries: the attitudes of neurosurgeons. *Neurosurgery* 1986;**19**(6):989-91.

86. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001;**323**(7303):42-6.

87. Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol* 2005;**5**:19.

88. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaboration, 2008.

89. Counsell C, Dennis M. Systematic review of prognostic models in patients with acute stroke. *Cerebrovasc Dis* 2001;**12**(3):159-70.

90. Pillai SV, Kolluri VR, Praharaj SS. Outcome prediction model for severe diffuse brain injuries: development and evaluation. *Neurol India* 2003;**51**(3):345-9.

91. Signorini DF, Andrews PJD, Jones PA, et al. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *J Neurol Neurosurg Psychiatry* 1999;**66**(1):20-25.

92. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Admission of patients with severe and moderate traumatic brain injury to specialized ICU facilities: a search for triage criteria. *Intensive Care Med* 2005;**31**(6):799-806.

93. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. *J Neurotrauma* 2005;**22**(10):1025-39.

94. Bush BA, Novack TA, Malec JF, et al. Validation of a model for evaluating outcome after traumatic brain injury. *Arch Phys Med Rehabil* 2003;**84**(12):1803-7.

95. Benzer A, Mitterschiffthaler G, Marosi M, et al. Prediction of non-survival after trauma: Innsbruck Coma Scale. *Lancet* 1991;**338**(8773):977-8.

96. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med* 2006;**144**(6):427-37.

97. Concato J, Feinstein AR, Holford TR. The Risk of Determining Risk with Multivariable Models. *Ann Intern Med* 1993;**118**(3):201-210.

98. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**(4):361-87.

99. Mushkudiani NA, Hukkelhoven CW, Hernandez AV, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. *J Clin Epidemiol* 2008;**61**(4):331-43.

100. Riley RD, Ridley G, Williams K, et al. Prognosis research: toward evidence-based results and a Cochrane methods group. *J Clin Epidemiol* 2007;**60**(8):863-5; author reply 865-6.

101. CRASH trial Collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months. *Lancet* 2005;**365**(9475):1957-9.

102. Wilson JT, Edwards P, Fiddes H, et al. Reliability of postal questionnaires for the Glasgow Outcome Scale. *J Neurotrauma* 2002;**19**(9):999-1005.

103. World Bank. World Development Indicators 2006.

104. Maas AI, Marmarou A, Murray GD, et al. Prognosis and clinical trial design in traumatic brain injury: the IMPACT study. *J Neurotrauma* 2007;**24**(2):232-8.

105. Marion DW, Carlier PM. Problems with initial Glasgow Coma Scale assessment caused by prehospital treatment of patients with head injuries: results of a national survey. *J Trauma* 1994;**36**(1):89-95.

106. McNett M. A review of the predictive ability of Glasgow Coma Scale scores in head-injured patients. *J Neurosci Nurs* 2007;**39**(2):68-75.

107. Heinzelmann M, Platz A, Imhof HG. Outcome after acute extradural haematoma, influence of additional injuries and neurological complications in the ICU. *Injury* 1996;**27**(5):345-9.

108. Stiver SI, Manley GT. Prehospital management of traumatic brain injury. *Neurosurg Focus* 2008;**25**(4):E5.

109. Butcher I, McHugh GS, Lu J, et al. Prognostic value of cause of injury in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):281-6.

110. Slewa-Younan S, van den Berg S, Baguley IJ, et al. Towards an understanding of sex differences in functional outcome following moderate to severe traumatic brain injury: a systematic review. *J Neurol Neurosurg Psychiatry* 2008;**79**(11):1197-201.

111. Maas AI, Steyerberg EW, Butcher I, et al. Prognostic value of computerized tomography scan characteristics in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):303-14.

112. Bullock R, Chesnut R, Ghajar J, et al. Guidelines for the Surgical Management of Traumatic Brain Injury. *Neurosurgery* 2006;**58**(3):S1-1 S2-62.

113. Hukkelhoven CW, Steyerberg EW, Farace E, et al. Regional differences in patient characteristics, case management, and outcomes in traumatic brain injury: experience from the tirilazad trials. *J Neurosurg* 2002;**97**(3):549-57.

114. Perel P, Edwards P, Wentz R, et al. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* 2006;**6**(38):38.

115. Steyerberg EW, Mushkudiani N, Perel P, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med* 2008;**5**(8):e165; discussion e165.

116. Murray GD, Butcher I, McHugh GS, et al. Multivariable prognostic analysis in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;**24**(2):329-37.

117. CRASH trial Collaborators. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;**12**:12.

118. Anonymous. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *Jama* 2001;**285**(19):2486-97.

119. Anonymous. Joint British recommendations on prevention of coronary heart disease in clinical practice: summary. British Cardiac Society, British Hyperlipidaemia Association, British Hypertension Society, British Diabetic Association. *Bmj* 2000;**320**(7236):705-8.

120. Perel P, Wasserberg J, Ravi RR, et al. Prognosis following head injury: a survey of doctors from developing and developed countries. *J Eval Clin Pract* 2007;**13**(3):464-5.

121. Edwards A, Elwyn G, Gwyn R. General practice registrar responses to the use of different risk communication tools in simulated consultations: a focus group study. *Bmj* 1999;**319**(7212):749-52.

122. Junghans C, Feder G, Timmis AD, et al. Effect of patient-specific ratings vs conventional guidelines on investigation decisions in angina: Appropriateness of Referral and Investigation in Angina (ARIA) Trial. *Arch Intern Med* 2007;**167**(2):195-202.

123. Sirovich BE, Gottlieb DJ, Welch HG, et al. Variation in the tendency of primary care physicians to intervene. *Arch Intern Med* 2005;**165**(19):2252-6.

124. Bachmann LM, Muhleisen A, Bock A, et al. Vignette studies of medical choice and judgement to study caregivers' medical decision behaviour: systematic review. *BMC Med Res Methodol* 2008;**8**(50):50.

125. Donders AR, van der Heijden GJ, Stijnen T, et al. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;**59**(10):1087-91.

126. Junque C, Bruna O, Mataro M. Information needs of the traumatic brain injury patient's family members regarding the consequences of the injury and associated perception of physical, cognitive, emotional and quality of life changes. *Brain Inj* 1997;**11**(4):251-8.

127. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott Med J* 1957;**2**(5):200-15.

128. Van Swieten JC, Koudstaal PJ, Visser MC, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;**19**(5):604-7.

129. Bamford JM, Sandercock PA, Warlow CP, et al. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1989;**20**(6):828.

130. New PW, Buchbinder R. Critical appraisal and review of the Rankin scale and its derivatives. *Neuroepidemiology* 2006;**26**(1):4-15.

131. MRC CRASH Trial Collaborators. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;**12**:12.

132. Tilley BC, Marler J, Geller NL, et al. Use of a global test for multiple outcomes in stroke trials with application to the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial. *Stroke* 1996;**27**(11):2136-42.

133. Wang D, Bakhai A. Clinical Trials A Practical Guide to Design, Analysis, and Reporting. London: Remedica, 2006.

134. Perel P, Edwards P, Shakur H, et al. Use of the Oxford Handicap Scale at hospital discharge to predict Glasgow Outcome Scale at 6 months in patients with traumatic brain injury. *BMC Med Res Methodol* 2008;**8**(72):72.

135. Pocock SJ, Lubsen J. More on subgroup analyses in clinical trials. *N Engl J Med* 2008;**358**(19):2076; author reply 2076-7.

136. Hernandez AV, Eijkemans MJ, Steyerberg EW. Randomized controlled trials with time-to-event outcomes: how much does prespecified covariate adjustment increase power? *Ann Epidemiol* 2006;**16**(1):41-8.

137. Hernandez AV, Steyerberg EW, Habbema JD. Covariate adjustment in randomized controlled trials with dichotomous outcomes increases statistical power and reduces sample size requirements. *J Clin Epidemiol* 2004;**57**(5):454-60.

138. Toll DB, Janssen KJ, Vergouwe Y, et al. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008;**61**(11):1085-94.

139. Bullock R, Chesnut R, Clifton GL, et al. Management and prognosis of severe traumatic brain injury. *J Neurotrauma* 2000;**17**:451-627.

140. Oertel M, Kelly DF, McArthur D, et al. Progressive hemorrhage after head trauma: predictors and consequences of the evolving injury. *J Neurosurg* 2002;**96**(1):109-16.

141. AAM. The Abbreviated injury scale 1990 revision. Illinois, 1990.

142. Patel HC, Bouamra O, Woodford M, et al. Trends in head injury outcome from 1989 to 2003 and the effect of neurosurgical care: an observational study. *Lancet* 2005;**366**(9496):1538-44.

143. Baker SP, O'Neill B. The injury severity score: an update. *J Trauma* 1976;**16**(11):882-5.

APPENDICES

Appendix 2.1

Survey of CRASH trial collaborators' beliefs, behaviours and attitudes in relation to prognosis for head injury patients

INSTRUCTIONS

We intend to develop a useful clinical prognostic model for head injury patients. Your answers will be very helpful for this. There are no right or wrong answers.

Please add any other comments on the blank page at the end

A. What you think about prognosis in your practice

Please say how much you agree or disagree with the following statements

For each question, tick one box ☑

	Totally disagree	Disagree	Unsure	Agree	Totally agree
1) I currently assess prognosis for head injury patients accurately	☐	☐	☐	☐	☐
2) Assessing prognosis with an accurate prognostic model would change the way I manage head injury patients	☐	☐	☐	☐	☐
3) Assessing prognosis with an accurate prognostic model would change the way I tell the prognosis to a patient or relative	☐	☐	☐	☐	☐

B. What you do in relation to prognosis in your practice

For each question, tick one box ☑

	Always	Usually	Sometimes	Almost Never	Never
4) Do you use a specific score to assess prognosis for head injury patients?	☐	☐	☐	☐	☐

Please specify the names of the scores you use:

.....

.....

5) Do you record the prognosis for head injury patients in their clinical notes?	☐	☐	☐	☐	☐

C. What you would like from a prognosis model

6) Which outcomes do you consider most important to predict accurately in head injury patients?

For each outcome, tick one box ☑

	Very important	Important	Not important
In-hospital death	☐	☐	☐
6 month death	☐	☐	☐
Need for surgery	☐	☐	☐
Need for Intensive Care Unit	☐	☐	☐
Days of stay in hospital	☐	☐	☐
Major disability (e.g. Persistent Vegetative State or severely disabled - conscious but dependent)	☐	☐	☐
Minor disability (i.e. independent but disabled)	☐	☐	☐
Need for rehabilitation	☐	☐	☐
Other (please specify).....	☐	☐	☐
Other (please specify).....	☐	☐	☐

7) Which of the following ways of expressing prognosis are useful to you?

For each option, tick one box ☑

	Very useful	Useful	Not useful
As a percentage (e.g. 90% of survivors at 1 week)	☐	☐	☐
Qualitatively (e.g. excellent, good, etc)	☐	☐	☐
As survival time (e.g. days, weeks, months, years)	☐	☐	☐
Other (please, specify).....	☐	☐	☐

8) For which situations do you think an accurate prognosis is important?

For each situation, tick one box ☑

	Very important	Important	Not important
To decide which patients should receive treatment (e.g. hyperventilation, barbiturates, mannitol)	☐	☐	☐
To decide in which patients treatment should be withdrawn	☐	☐	☐
To decide in which patients CT scan should be done	☐	☐	☐
To decide in which patients intracranial pressure should be monitored	☐	☐	☐
To decide which patients need Intensive Care Unit	☐	☐	☐
To decide which patients need rehabilitation	☐	☐	☐
To decide which patients need surgery	☐	☐	☐
To decide which patients need decompressive craniotomy	☐	☐	☐
To give counselling to patients and/or relatives	☐	☐	☐

Please specify any other important uses that you would have for accurate prognostic information in the management of head injury patients:

.....

.....

.....

9) Please specify the most important uses of accurate prognostic information:

1 .....

2 .....

3 .....

10) Which 3 variables do you consider to be most important for making prognosis for head injury patients?

1 .....

2 .....

3 .....

11) Which variables do you consider to be important for making prognosis for head injury patients?

For each variable, tick one box ☑

	Very important	Important	Not important
Age	☐	☐	☐
Gender	☐	☐	☐
Cause of injury	☐	☐	☐
Presence of major extracranial injury	☐	☐	☐
Time since injury to hospital arrival	☐	☐	☐
Total GCS	☐	☐	☐
Eye component of the GCS	☐	☐	☐
Motor component of the GCS	☐	☐	☐
Verbal component of the GCS	☐	☐	☐
Pupil reactions	☐	☐	☐
Presence of complications	☐	☐	☐
Wound infection	☐	☐	☐
Gastrointestinal bleeding	☐	☐	☐
Seizures	☐	☐	☐

Pneumonia	☐	☐	☐
Abnormal CT scan	☐	☐	☐
One or more petechial haemorrhages within the brain	☐	☐	☐
Obliteration of the third ventricle or basal cisterns	☐	☐	☐
Subarachnoid bleed	☐	☐	☐
Midline shift over 5mm	☐	☐	☐
Non evacuated haematoma	☐	☐	☐
Evacuated haematoma	☐	☐	☐
Cortical Contusion	☐	☐	☐
Other, please specify:	☐	☐	☐

To help us interpret the results, please answer these questions.

12) What is your age? _____ **13)** What is your gender? ☐ Female ☐ Male

14) What is your primary specialty in medicine?

☐ Orthopaedic & Trauma

☐ Emergency Medicine

☐ Anaesthesiology

☐ Intensive Care

☐ Neurological Surgery

☐ General Surgery

☐ Other (*please specify*)..... .....

15) In which region of the world is your hospital located?

☐ Australia & New Zealand

☐ Caribbean

☐ East Asia & Pacific

☐ Eastern Europe & Central Asia

☐ Latin America

☐ Middle East & North Africa

☐ North America

☐ South & Southeast Asia

☐ Sub-Saharan Africa

☐ Western Europe

16) Does your hospital have?

*(Tick **all** that apply)*

☐ Intensive Care Unit

☐ Computed Tomography

☐ Neurosurgery Service

Any other comment:

_____

_____

_____

Thank you very much for your collaboration!

Appendix 3.1

Electronic bibliographical databases and search strategies for the systematic review

Database (time period or version)	Search Strategy
Medline(PUBMED version) limit to 1990 - 2005	[brain injuries OR traumatic brain injury OR craniocerebral trauma] OR ["brain injuries" OR "Traumatic brain injury " OR "brain Trauma" OR "brain Trauma" Field: Title] AND [brain[ti] OR brain*[ti] OR coma[ti] OR conscious*[ti] OR cranio*[ti] OR skull[ti]] AND ["Case-Control Studies"[MeSH] OR "Cohort Studies"[MeSH] OR "Follow-Up Studies"[MeSH] OR prognos* OR predict* Field: Title]
Embase(OVID version): yr=1990-2005	1. traumatic brain injury .mp. or exp traumatic brain injury / or exp *traumatic brain injury / or brain injur$.ti. 2. (brain$ or brain$ or coma$ or conscious$ or cranio$ or skull$).ti. 3. 1 and 2 4. (prognos$ or predict$).mp. 5. 3 and 4 6. case control study.mp. or (cohort study or cohort analysis).mp. or exp follow up/ or exp case control study/ or follow up.mp. or systematic review.mp. or trial.mp. or randomi$.mp. 7. 5 and 6 8. limit 7 to

Appendix 3.2

Studies included in the systematic review

1. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Admission of patients with severe and moderate traumatic brain injury to specialized ICU facilities: a search for triage criteria. *Intensive Care Med* 2005;**31**(6):799-806.
2. Hukkelhoven CW, Steyerberg EW, Habbema JD, et al. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. *J Neurotrauma* 2005;**22**(10):1025-39.
3. Hsu MH, Li YC, Chiu WT, et al. Outcome prediction after moderate and severe head injury using an artificial neural network. *Stud Health Technol Inform* 2005;**116**:241-5.
4. Poon WS, Zhu XL, Ng SC, et al. Predicting one year clinical outcome in traumatic brain injury (TBI) at the beginning of rehabilitation. *Acta Neurochir Suppl* 2005;**93**:207-8.
5. Wechsler B, Kim H, Gallagher PR, et al. Functional status after childhood traumatic brain injury. *J Trauma* 2005;**58**(5):940-9; discussion 950.
6. Levin HS, McCauley SR, Josic CP, et al. Predicting depression following mild traumatic brain injury. *Arch Gen Psychiatry* 2005;**62**(5):523-8.
7. Carter BG, Butt W. A prospective study of outcome predictors after severe brain injury in children. *Intensive Care Med* 2005;**31**(6):840-5.
8. Eftekhar B, Mohammad K, Ardebili HE, et al.. Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data. *BMC Med Inform Decis Mak* 2005;**5**(1):3.
9. Rovlias A, Kotsou S. Classification and regression tree for prediction of outcome after severe head injury using simple clinical and laboratory variables. *J Neurotrauma* 2004;**21**(7):886-93.
10. Demetriades D, Kuncir E, Murray J, et al. Mortality prediction of head Abbreviated Injury Score and Glasgow Coma Scale: analysis of 7,764 head injuries. *J Am Coll Surg* 2004;**199**(2):216-22.
11. Ibanez J, Arikan F, Pedraza S, et al. Reliability of clinical guidelines in the detection of patients at risk following mild head injury: results of a prospective study. *J Neurosurg* 2004;**100**(5):825-34.
12. Fabbri A, Servadei F, Marchesini G, et al. Prospective validation of a proposal for diagnosis and management of patients attending the emergency department for mild head injury. *J Neurol Neurosurg Psychiatry* 2004;**75**(3):410-6.

13. Tender GC, Awasthi D. Risk stratification in mild head injury patients: the head injury predictive index. *J La State Med Soc* 2003;**155**(6):338-42.
14. Bush BA, Novack TA, Malec JF, et al. Validation of a model for evaluating outcome after traumatic brain injury. *Arch Phys Med Rehabil* 2003;**84**(12):1803-7.
15. Pillai SV, Kolluri VR, Praharaj SS. Outcome prediction model for severe diffuse brain injuries: development and evaluation. *Neurol India* 2003;**51**(3):345-9.
16. Brenner T, Freier MC, Holshouser BA, et al. Predicting neuropsychologic outcome after traumatic brain injury in children. *Pediatr Neurol* 2003;**28**(2):104-14.
17. Cassidy LD, Potoka DA, Adelson PD, et al. Development of a novel method to predict disability after head trauma in children. *J Pediatr Surg* 2003;**38**(3):482-5.
18. Ratanalert S, Chompikul J, Hirunpat S, et al. Prognosis of severe head injury: an experience in Thailand. *Br J Neurosurg* 2002;**16**(5):487-93.
19. Andrews PJ, Sleeman DH, Statham PF, et al. Predicting recovery in patients suffering from traumatic brain injury by using admission variables and physiological data: a comparison between decision tree analysis and logistic regression. *J Neurosurg* 2002;**97**(2):326-36.
20. Heard C, Li V, Heard A. A useful tool for predicting outcome for the pediatric head trauma patient. *Crit Care Med* 2002;**30**(6):1403-4.
21. Schaan M, Jaksche H, Boszczyk B. Predictors of outcome in head injury: proposal of a new scaling system. *J Trauma* 2002;**52**(4):667-74.
22. Schreiber MA, Aoki N, Scott BG, et al. Determinants of mortality in patients with severe blunt head injury. *Arch Surg* 2002;**137**(3):285-90.
23. Sustic A, Turina D, Ticac Z, et al. War head injury score: an outcome prediction model in War casualties with acute penetrating head injury. *Mil Med* 2001;**166**(4):331-4.
24. Sinha M, Kennedy CS, Ramundo ML. Artificial neural network predicts CT scan abnormalities in pediatric patients with closed head injury. *J Trauma* 2001;**50**(2):308-12.
25. Mukherjee KK, Sharma BS, Ramanathan SM, et al.. A mathematical outcome prediction model in severe head injury: a pilot study. *Neurol India* 2000;**48**(1):43-8.
26. Ashwal S, Holshouser BA, Shu SK, et al. Predictive value of proton magnetic resonance spectroscopy in pediatric closed head injury. *Pediatr Neurol* 2000;**23**(2):114-25.

27. Wagner AK, Hammond FM, Sasser HC, et al. Use of injury severity variables in determining disability and community integration after traumatic brain injury. *J Trauma* 2000;**49**(3):411-9.
28. Vath A, Meixensberger J, Dings J, et al. Prognostic significance of advanced neuromonitoring after traumatic brain injury using neural networks. *Zentralbl Neurochir* 2000;**61**(1):2-6.
29. Adachi S, Hirano N, Tanabe M, et al. Multivariate analysis of patients with head injury using quantification theory type II--with special reference to prediction of patient outcome. *Neurol Med Chir (Tokyo)* 2000;**40**(4):200-4; discussion 204-5.
30. Wagner AK, Hammond FM, Grigsby JH, et al. The value of trauma scores: predicting discharge after traumatic brain injury. *Am J Phys Med Rehabil* 2000;**79**(3):235-42.
31. Stuss DT, Binns MA, Carruth FG, et al. Prediction of recovery of continuous memory after traumatic brain injury. *Neurology* 2000;**54**(6):1337-44.
32. Lannoo E, Van Rietvelde F, Colardyn F, et al. Early predictors of mortality and morbidity after severe closed head injury. *J Neurotrauma*. 2000;**17**(5):403-14.
33. Nissen JJ, Jones PA, Signorini DF, et al. Glasgow head injury outcome prediction program: an independent assessment. *J Neurol Neurosurg Psychiatry* 1999;**67**(6):796-9.
34. Sakellaropoulos GC, Nikiforidis GC. Development of a Bayesian Network for the prognosis of head injuries using graphical model selection techniques. *Methods Inf Med* 1999;**38**(1):37-42.
35. Signorini DF, Andrews PJ, Jones PA, et al.. Adding insult to injury: the prognostic value of early secondary insults for survival after traumatic brain injury. *J Neurol Neurosurg Psychiatry* 1999;**66**(1):26-31.
36. Signorini DF, Andrews PJ, Jones PA, et al. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *J Neurol Neurosurg Psychiatry* 1999;**66**(1):20-5.
37. Lai YC, Chen FG, Goh MH, et al.. Predictors of long-term outcome in severe head injury. *Ann Acad Med Singapore* 1998;**27**(3):326-31.
38. Alvarez M, Nava JM, Rue M, Quintana S. Mortality prediction in head trauma patients: performance of Glasgow Coma Score and general severity systems. *Crit Care Med* 1998;**26**(1):142-8.
39. Lang EW, Pitts LH, Damron SL, et al. Outcome after severe head injury: an analysis of prediction based upon comparison of neural network versus logistic regression analysis. *Neurol Res* 1997;**19**(3):274-80.

40. Cho DY, Wang YC. Comparison of the APACHE III, APACHE II and Glasgow Coma Scale in acute head injury for prediction of mortality and functional outcome. *Intensive Care Med* 1997;**23**(1):77-84.

41. Combes P, Fauvage B, Colonna M, et al. Severe head injuries: an outcome prediction and survival analysis. *Intensive Care Med* 1996;**22**(12):1391-5.

42. Zafonte RD, Hammond FM, Mann NR, et al. Revised trauma score: an additive predictor of disability following traumatic brain injury? *Am J Phys Med Rehabil* 1996;**75**(6):456-61.

43. Mamelak AN, Pitts LH, Damron S. Predicting survival from head trauma 24 hours after injury: a practical method with therapeutic implications. *J Trauma* 1996;**41**(1):91-9.

44. Walder AD, Yeoman PM, Turnbull A. The abbreviated injury scale as a predictor of outcome of severe head injury. *Intensive Care Med* 1995;**21**(7):606-9.

45. Cooke RS, McNicholl BP, Byrnes DP. Use of the Injury Severity Score in head injury. *Injury* 1995;**26**(6):399-400.

46. Temkin NR, Holubkov R, Machamer JE, et al. Classification and regression trees (CART) for prediction of function at 1 year following head trauma. *J Neurosurg* 1995;**82**(5):764-71.

47. Fearnside MR, Cook RJ, McDougall P, et al. The Westmead Head Injury Project outcome in severe head injury. A comparative analysis of pre-hospital, clinical and CT variables. *Br J Neurosurg* 1993;**7**(3):267-79.

48. Vilalta J, Vaque J, Olona M, et al. [Predictive factors of mortality in severe craniocerebral trauma]. *Med Clin (Barc)* 1992;**99**(12):441-3.

49. Ross SE, O'Malley KF, Stein S, et al. Abbreviated injury scaling of head injury as a prognostic tool for functional outcome. *Accid Anal Prev* 1992;**24**(2):181-5.

50. Benzer A, Mitterschiffthaler G, Marosi M, et al. Prediction of non-survival after trauma: Innsbruck Coma Scale. *Lancet* 1991;**338**(8773):977-8.

51. Choi SC, Muizelaar JP, Barnes TY, et al. Prediction tree for severely head-injured patients. *J Neurosurg* 1991;**75**(2):251-5.

52. Feldman Z, Contant CF, Robertson CS, et al. Evaluation of the Leeds prognostic score for severe head injury. *Lancet* 1991;**337**(8755):1451-3.

53. Zagara G, Scaravilli P, Mastorgio P, et al. Validation of a prognostic system in severe brain-injured patients. *J Neurosurg Sci* 1991;**35**(2):77-81.

Appendix 3.3

Characteristics of the models included in the systematic review

Study number	Author	Year of publication	Age group	Severity	Objective	Outcome	Multivariable analysis	N° of patients included
1	choi	1991	nr	severe	develop	GOS	CART	555
2	feldman	1991	nr	all	validate	mortality	na	479
2	feldman	1991	nr	all	validate	mortality	na	131
3	benzer	1991	all	severe	validate	mortality	na	421
4	zagara	1991	adults	severe	validate	mortality	na	76
5	vilalta	1992	all	severe	develop	mortality	logistic	173
6	ross	1992	all	severe	validate	GOS	na	503
7	fearnside	1993	all	severe	develop	mortality	logistic	315
7	fearnside	1993	all	severe	develop	functional	logistic	315
8	walder	1995	adults	severe	validate	GOS	na	109
9	temkin	1995	nr	all	develop	functional	CART	448
9	temkin	1995	nr	all	develop	functional	CART	448
9	temkin	1995	nr	all	develop	functional	CART	448
9	temkin	1995	nr	all	develop	functional	CART	448
9	temkin	1995	nr	all	develop	functional	CART	448
9	temkin	1995	nr	all	develop	functional	CART	448
10	cooke	1995	nr	severe	validate	GOS	na	131
11	mamelak	1996	all	severe	develop	GOS	logistic	672
12	combes	1996	nr	severe	develop	GOS	logistic	198
13	zafonte	1996	adults	all	validate	functional	na	501
14	lang	1997	all	severe	develop	mortality	logistic	799
14	lang	1997	all	severe	develop	mortality	logistic	799
14	lang	1997	all	severe	develop	mortality	neural network	799
14	lang	1997	all	severe	develop	mortality	neural network	799
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	functional	na	200
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	mortality	na	200
15	cho	1997	adults	all	validate	functional	na	200
16	alvarez	1998	adults	all	validate	mortality	na	247
16	alvarez	1998	adults	all	validate	mortality	na	247
16	alvarez	1998	adults	all	validate	mortality	na	247
16	alvarez	1998	adults	all	validate	mortality	na	247
17	lai	1998	adults	severe	validate	GOS	na	70
18	signorini	1999	adults	severe and moderate	develop	mortality	logistic	110
19	signorini	1999	adults	severe and moderate	develop	mortality	logistic	372
20	nissen	1999	all	all	validate	GOS	na	324
21	sakellaropoulos	1999	nr	nr	develop	GOS	bayesian	525
22	stuss	2000	adults	all	develop	functional	CART	187
22	stuss	2000	adults	all	develop	functional	weibull regression	187
23	sustic	2000	adults	all	develop	mortality	no	41
23	sustic	2000	adults	all	develop	mortality	no	43
24	mukherjee	2000	all	severe	develop	GOS	logistic	103
25	sinha	2000	children	nr	develop	CT scan lesion	neural network	351
25	sinha	2000	children	nr	develop	CT scan lesion	logistic	351

Study number	Author	Year of publication	Age group	Severity	Objective	Outcome	Multivariable analysis	N° of patients included
26	ashwal	2000	children	nr	develop	functional	linear discriminant analysis	27
26	ashwal	2000	children	nr	develop	functional	linear discriminant analysis	26
27	lannoo	2000	adults	severe	develop	GOS	logistic	78
27	lannoo	2000	adults	severe	develop	mortality	logistic	158
28	vath	2000	nr	severe	develop	GOS	neural network	95
28	vath	2000	nr	severe	develop	GOS	neural network	95
28	vath	2000	nr	severe	develop	GOS	neural network	95
28	vath	2000	nr	severe	develop	GOS	neural network	95
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
29	wagner	2000	adults	all	validate	functional	na	378
30	wagner	2000	adults	all	validate	functional	na	378
30	wagner	2000	adults	all	validate	functional	na	378
31	adachi	2000	all	nr	nr	GOS	linear discriminant analysis	63
32	schaan	2001	all	nr	develop	GOS	not clear	554
33	schreiber	2002	adults	severe	develop	mortality	logistic	368
34	ratanalert	2002	all	severe	develop	GOS	logistic	337
35	andrews	2002	all	all	develop	GOS	CART	124
35	andrews	2002	all	all	develop	GOS	CART	124
35	andrews	2002	all	all	develop	GOS	CART	124
36	heard	2002	children	nr	validate	mortality	na	119
37	pillai	2003	adults	severe	develop	GOS	logistic	289
38	tender	2003	all	mild	develop	GOS	no	255
39	brenner	2003	children	severe	develop	functional	discriminant analysis	22
39	brenner	2003	children	severe	develop	functional	linear discriminant analysis	22
39	brenner	2003	children	severe	develop	functional	discriminant analysis	22
40	cassidi	2003	children	all	develop	functional	logistic	3491
41	bush	2003	adults	all	validate	functional	na	294
42	rovlias	2004	adults	severe	develop	GOS	CART	345
43	ibañez	2004	adults	mild	develop	CT scan lesion	logistic	1101
43	ibañez	2004	adults	mild	develop	CT scan lesion	logistic	1101
43	ibañez	2004	adults	mild	develop	CT scan lesion	CART	1101
43	ibañez	2004	adults	mild	develop	CT scan lesion	CART	1101
44	demetriades	2004	all	severe and moderate	validate	mortality	na	7764
45	fabrri	2004	adults	mild	develop	CT scan lesion	logistic	5578
45	fabrri	2004	adults	mild	develop	neurosurgical intervention	logistic	5578
45	fabrri	2004	adults	mild	develop	GOS	logistic	5578
46	carter	2005	children	severe	develop	GOS	na	102
47	levin	2005	adults	mild	develop	functional	logistic	129
48	eftekhar	2005	adults	all	develop	mortality	logistic	1271
48	eftekhar	2005	adults	all	develop	mortality	neural network	1271
49	hsu	2005	all	severe and moderate	develop	GOS	neural network	3345
50	wechsler	2005	children	all	develop	functional	logistic	4439
50	wechsler	2005	children	all	develop	functional	logistic	4439
51	poon	2005	nr	severe and moderate	develop	GOS	logistic	68
52	hukkelhoven	2005	nr	severe and moderate	develop	neurosurgical intervention	logistic	275
52	hukkelhoven	2005	nr	severe and moderate	develop	raised IPC	logistic	275
53	hukkelhoven	2005	nr	severe and moderate	develop	GOS	logistic	2269
53	hukkelhoven	2005	nr	severe and moderate	develop	mortality	logistic	2269

Appendix 4.1

Questionnaire for six months follow-up to assess Glasgow Outcome Scale

INTERNATIONAL STUDY OF RECOVERY AFTER HEAD INJURY

These questions are about changes in your lifestyle since your injury. They can be answered by you, a relative or friend, or by you both together. If you have any questions about this form, please contact Phil Edwards on 020 7958 8112. Please answer each question below by ticking one box ☑ which is true for you.

Your answers will help us improve the care of people following a head injury.

Please say who filled out this form:

☐ Patient alone ☐ Relative, friend or carer alone ☐ Patient and relative, friend or carer together

1. At present, where do you live most of the time?

☐ In own home ☐ In hospital ☐ In residential care

2. As a result of your injury, do you now need help in the home?

☐ No ☐ Yes. I need some help in the home, but not every day. ☐ Yes. I need help in the home every day. ☐ I need help in the home, but not because of the injury.

3. As a result of your injury, do you now need help to shop?

☐ No ☐ Yes. I need some help, but can go to the local shops on my own. ☐ Yes. I need help to shop even locally, or cannot shop at all. ☐ I need help to shop, but not because of the injury.

4. As a result of your injury, do you now need help to travel?

☐ No ☐ Yes. I need some help, but can travel locally on my own (e.g. by arranging a taxi). ☐ Yes. I need help to travel even locally, or I cannot travel at all. ☐ I need help to travel, but not because of the injury.

5. As a result of your injury, has there been a change in your ability to work? (or to study if you were a student; or to look after your family)

☐ No ☐ Yes. I still work, but at a reduced level (e.g. a change from full-time to part-time, or a change in level of responsibility). ☐ Yes. I am unable to work at present. ☐ My ability to work is restricted, but not because of the injury, or I have retired.

6. As a result of your injury, has there been a change in your ability to take part in social and leisure activities outside home?

☐ No ☐ Yes. I take part a bit less, but at least half as often. ☐ Yes. I take part much less, or do not take part at all. ☐ My ability to take part is restricted for some other reason, not because of the injury.

7. As a result of your injury, are there now problems in how you get on with friends or relatives?

☐ No ☐ Yes. There are occasional problems (less than once a week). ☐ Yes. There are frequent or constant problems. ☐ There are problems for some other reason, not because of the injury.

Thank you for your help. Please return this form in the envelope provided to: Dr Ian Roberts, International Study of Recovery after Head Injury, LSHTM, University of London, Keppel Street, London WC1E 7HT

Appendix 4.2

MOISTEN EDGES AND SEAL TOGETHER

MOISTEN EDGES AND SEAL TOGETHER

CRASH

EARLY OUTCOME FORM

Complete at **discharge, death in hospital, or 14 days after injury** whichever occurs first

Please **PRINT** clearly and answer **EVERY** question

Attach treatment pack label here

1. Hospital name

2. Patient details or attach a label with these details (for 6-month follow-up)

Family name: | Patient identification no. (if appropriate)

Given name(s): | NHS number (if appropriate)

Sex: ☐ **M** ☐ **F** Date of Birth: / / (day/month/year)

Address:

Postcode: | Telephone:

3. Cause of injury: ☐ Road traffic accident ☐ Fall > 2 metres Other:

4. Outcome (please complete questions **a,b,c** and **d**)

a. ☐ Death in hospital ☐ Transferred to other acute care hospital ☐ Discharged to rehabilitation centre or nursing home ☐ Discharged home ☐ Still in this hospital now

b. Date of death, transfer or discharge: / /

c. If transferred give consultant name/department, and name of hospital

d. Tick ✓ **one box** that best describes the patient's **head injury-related symptoms now** (i.e. at 14 days or prior discharge):

☐ No symptoms ☐ Minor symptoms ☐ Some restriction in lifestyle but independent ☐ Dependent, but not requiring constant attention ☐ Fully dependent, requiring attention day and night ☐ Dead

OLD FOLD

5. Management and complications
(please tick ✓ **ONE** box on **EACH** line)

Yes	No	
☐	☐	Admitted to Intensive Care Unit If **Yes**, please write number of days in ICU ☐
☐	☐	Seizure
☐	☐	Haematemesis or melaena requiring transfusion
☐	☐	Wound infection with pus
☐	☐	Pneumonia treated with antibiotics
☐	☐	Other treatment with antibiotics
☐	☐	Neurosurgical operation
☐	☐	Major extracranial injury

6. Head CT scan ☐ **Yes** ☐ **No** If **No** go to section 7

Date of first head CT scan: / / Time (24 hr clock) :

Result: (please ✓ tick all that apply)

☐ Obliteration of the 3rd ventricle or basal cisterns
☐ Midline shift >5mm
☐ Intracranial haematoma – *evacuated*
☐ Intracranial haematoma – *non-evacuated*
☐ One or more petechial haemorrhages within the brain
☐ Cortical contusion > 1cm in diameter
☐ Subarachnoid bleed
☐ Normal scan

7. Trial treatment **a)** Loading dose: ☐ Yes ☐ No **b)** Hours of maintenance dose: ☐ hours (1-48)

Sections 8 and 9 are only required if the patient is alive

8. Reliable contact (Next of kin or friend)

Name:
Address:
Post code:
Tel:

9. Family doctor

Name:
Address:
Post code:
Tel:

10. Person completing form (please PRINT):

Name: Position: Date: / /

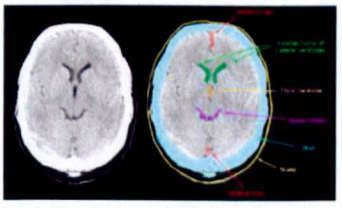
When complete fold form as indicated, stick together and post to:
CRASH Co-ordinating Centre, FREEPOST, LON14211, London WC1N 1BR OR FAX +44 (0)20 7299 4663

CRASH/OC/3/101

Appendix 4.3

CT Scan Guidance

Case 1 Normal Scan



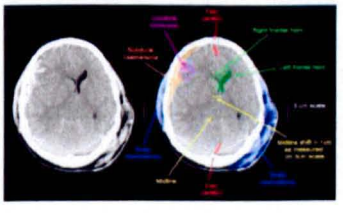
Case 2 Acute subdural haematoma demonstrating midline shift

Midline shift > 5mm

Intracranial haematoma - non-evacuated

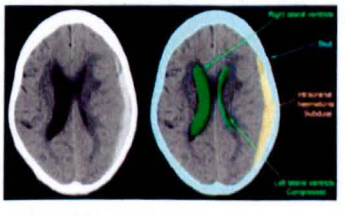
Cortical contusion > 1cm in diameter

Obliteration of the 3rd ventricle



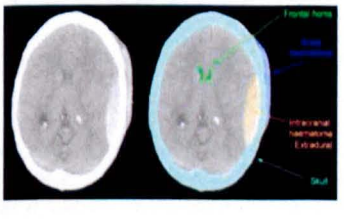
Case 3 Acute subdural haematoma

Intracranial haematoma - non-evacuated



Case 4 Acute extradural haematoma

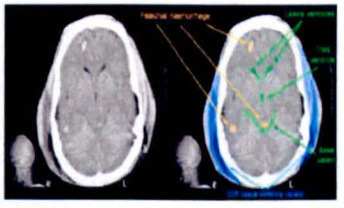
Intracranial haematoma - non-evacuated



appendix 4.3 continuation

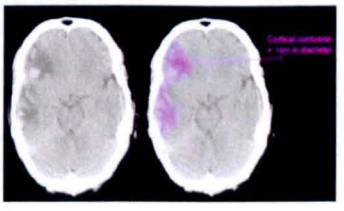
Case 5 Diffuse axonal injury

One or more petechial haemorrhages within the brain



Case 6 Cerebral contusion

Cortical contusion > 1cm in diameter



Obliteration of the third ventricle : The third ventricle is demonstrated well in Case One - Normal Scan. Here you can see that it is effectively a small cleft within the brain. If there is any pressure, or swelling of the brain, this is one of the first structures to disappear on the scan, as the walls of the cleft are pushed against each other. It is a sign of increased pressure within the head and can either result from a blood clot pressing on the brain, or of swelling of the brain itself.

Midline shift - Greater than 5mm : If you look at Case Two, you will see that we have drawn the midline using a series of yellow dots. You can do the same thing by using a ruler and joining the falx cerebri anteriorly and posteriorly, as labelled in Case One. The third ventricle and the septum between the frontal horns of the lateral ventricle should not deviate from the midline at all, although there may be slight variations less than 5mm. You can use the scale to the right hand side of each scan to work out if the shift is greater than 5mm.

Intracranial haematoma - Evacuated Often, when you look at a post operative scan, there is still residual blood clot. Even though an extradural, or subdural haematoma has been removed, one will often see evidence of a blood clot post operatively, which is hopefully much smaller and not causing pressure on the brain. It does, however, indicate that there has been evacuation of a significant intracranial haematoma. These cases should be classified as "Intracranial haematoma - Evacuated"

Intracranial haematoma -Non-evacuated: This refers to blood clot lying either on the surface of the brain in the extradural, or subdural space, or, indeed, to haematomas within the brain substance itself. If these have not been removed surgically, then they should be classified as non-evacuated.

One, or more petechial haemorrhages within the brain: This refers to very small haemorrhages seen as small, white dots on the scan. They usually occur at the interface between grey and white matter (See Case Five). Other classical sites are the dorsolateral quadrant of the midbrain and also the corpus callosum. They are an indicator of diffuse axonal injury, which is a form of severe primary injury to the brain. This usually carries quite a poor prognosis.

Appendix 5.1

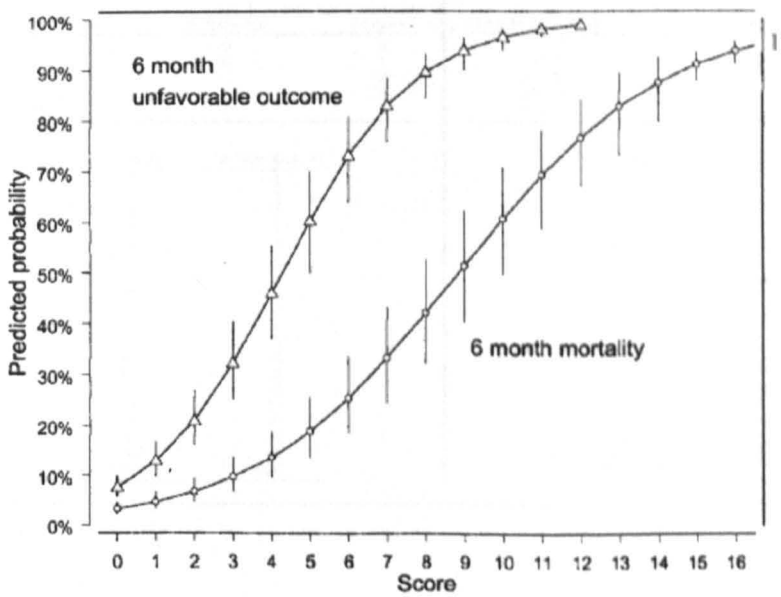
Score and figure

PROGNOSTIC SCORE CHART FOR THE PROBABILITY OF MORTALITY AND UNFAVORABLE OUTCOME IN PATIENTS WITH SEVERE OR MODERATE TRAUMATIC BRAIN INJURY ACCORDING TO THE PROGNOSTIC MODELS

Predictor	*Value*	*Mortality*	*Unfavorable outcome*
Age (years)	15–39	0	0
	40–54	1	1
	55–64	2	2
	≥65	3	3
Motor score	None/extensor	3	3
	Abnormal flexion	2	2
	Withdraws	1	1
	Localizes/obeys	0	0
Pupillary reactivity	Both react	0	0
	One reacts	1	1
	None reacts	2	2
Hypoxia	No	0	0
	Yes	1	1
Hypotension	No	0	0
	Yes	2	1
CT classification[a]	I or II	0	0
	III	2	1
	IV	4	1
	V or VI	2	1
Traumatic subarachnoid hemorrhage	No	0	0
	Yes	2	1
Sum score[b]: add relevant scores		—	—

[a]CT classification I = no visible intracranial pathlogy on CT scan; II = midline shift 0–5 mm; III = cisterns compressed or absent with midline shift 0–5 mm; IV = midline shift > 5 mm; V = any lesion surgically evacuated; VI = high- or mixed-density lesion >25 mm, not surgically evacuated.

[b]The sum score can be used to obtain the predicted probability of mortality or unfavorable outcome from Figure 2.



Continuation appendix 5.1
Numerical score

Estimate of 10-Year Risk for **Men**
(Framingham Point Scores)

Age, y	Points
20-34	-9
35-39	-4
40-44	0
45-49	3
50-54	6
55-59	8
60-64	10
65-69	11
70-74	12
75-79	13

Total Cholesterol, mg/dL	Points				
	Age 20-39 y	Age 40-49 y	Age 50-59 y	Age 60-69 y	Age 70-79 y
<160	0	0	0	0	0
160-199	4	3	2	1	0
200-239	7	5	3	1	0
240-279	9	6	4	2	1
≥280	11	8	5	3	1

	Points				
	Age 20-39 y	Age 40-49 y	Age 50-59 y	Age 60-69 y	Age 70-79 y
Nonsmoker	0	0	0	0	0
Smoker	8	5	3	1	1

HDL, mg/dL	Points
≥60	-1
50-59	0
40-49	1
<40	2

Systolic BP, mm Hg	If Untreated	If Treated
<120	0	0
120-129	0	1
130-139	1	2
140-159	1	2
≥160	2	3

Point Total	10-Year Risk, %
<0	<1
0	1
1	1
2	1
3	1
4	1
5	2
6	2
7	3
8	4
9	5
10	6
11	8
12	10
13	12
14	16
15	20
16	25
≥17	≥30

Estimate of 10-Year Risk for **Women**
(Framingham Point Scores)

Age, y	Points
20-34	-7
35-39	-3
40-44	0
45-49	3
50-54	6
55-59	8
60-64	10
65-69	12
70-74	14
75-79	16

Total Cholesterol, mg/dL	Points				
	Age 20-39 y	Age 40-49 y	Age 50-59 y	Age 60-69 y	Age 70-79 y
<160	0	0	0	0	0
160-199	4	3	2	1	1
200-239	8	6	4	2	1
240-279	11	8	5	3	2
≥280	13	10	7	4	2

	Points				
	Age 20-39 y	Age 40-49 y	Age 50-59 y	Age 60-69 y	Age 70-79 y
Nonsmoker	0	0	0	0	0
Smoker	9	7	4	2	1

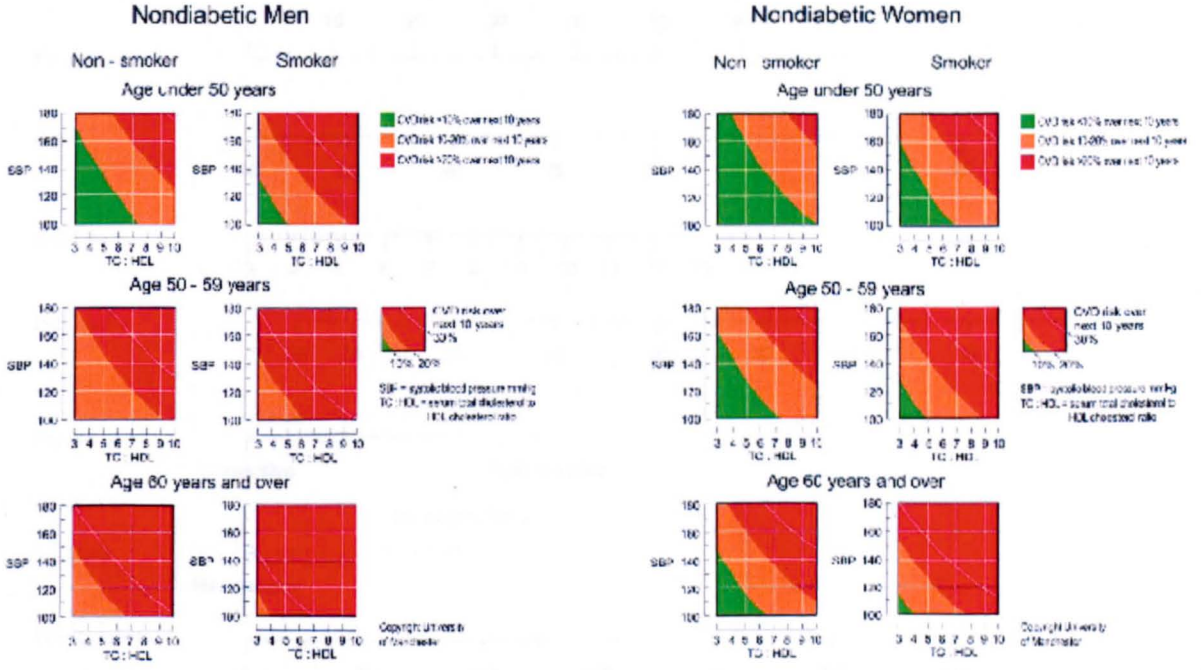
HDL, mg/dL	Points
≥60	-1
50-59	0
40-49	1
<40	2

Systolic BP, mm Hg	If Untreated	If Treated
<120	0	0
120-129	1	3
130-139	2	4
140-159	3	5
≥160	4	6

Point Total	10-Year Risk, %
<9	<1
9	1
10	1
11	1
12	1
13	2
14	2
15	3
16	4
17	5
18	6
19	8
20	11
21	14
22	17
23	22
24	27
≥25	≥30

Continuation appendix 5.1
Coloured chart

Cardiovascular Disease Risk Prediction Chart for Primary Prevention



These charts are for estimating cardiovascular disease (CVD) risk (non-fatal MI and stroke, coronary and stroke death and new angina pectoris) for individuals without major atherosclerotic disease. They are intended to aid clinical decisions about how intensively to intervene on lifestyle and whether to use antihypertensive, lipid lowering medication and aspirin. The use of these charts is not appropriate for those with CHD or other major atherosclerotic disease, inherited dyslipidaemias, chronic renal dysfunction, and type 1 and 2 diabetes mellitus.

To estimate an individual's absolute 10 year risk of developing CVD choose the table for his or her gender, smoking status (smoker/non smoker) and age. Within this square define the level of risk according to the point where the coordinates for systolic blood pressure (SBP) and the ratio of total cholesterol to HDL cholesterol meet. If no HDL cholesterol result is available, then assume this is 1.00mmol/l and the lipid scale can be used for total serum cholesterol alone. Higher risk individuals (red areas) are defined as those whose 10 year CVD risk exceeds 20%, which is approximately equivalent to the CHD risk of >15% over the same period indicated by the previous version of these charts. As a minimum those at highest CVD risk (greater than 30% shown by the line within the red area) should be targeted and treated now. When resources allow, others with a CVD risk of >20% should be progressively targeted.

Smoking status should reflect lifetime exposure to tobacco and not simply tobacco use at the time of assessment. For example, those who have given up smoking within 5 years should be regarded as current smokers for the purposes of the charts. The initial BP and the first random (non-fasting) total cholesterol and HDL cholesterol can be used to estimate an individual's risk. However, the decision on using drug therapy should generally be based on repeat risk factor measurements over a period of time.

Men and women do not reach the level of risk predicted by the charts for the three age bands until they reach the ages 40, 50, and 60 years respectively. Everyone aged 70 years and over should be considered at higher risk. The charts will overestimate current risk most in the under forties.

These charts (and all other currently available methods of CVD risk prediction) are based on groups of people with untreated levels of BP, total cholesterol and HDL cholesterol. In patients already receiving antihypertensive therapy in whom the decision is to be made about whether to introduce lipid-lowering medication or vice versa the charts can act as a guide, but unless recent pre-treatment risk factor values are available it is generally safest to assume that CVD risk is higher than that predicted by current levels of BP or lipids on treatment.

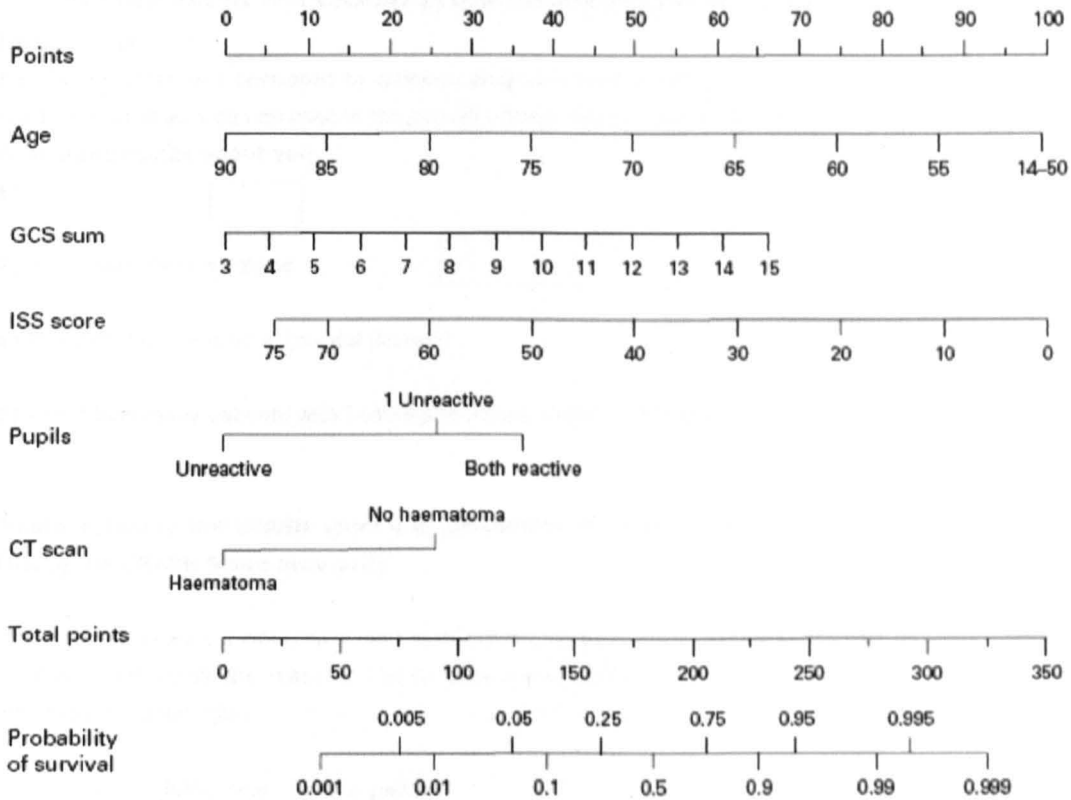
CVD risk is also higher than indicated in the charts for:-

- Those with a family history of premature CVD or stroke (male first degree relatives aged <55 years and female first degree relatives aged <65 years) which increases the risk by a factor of approximately 1.5
- Those with raised triglyceride levels
- Women with premature menopause
- Those who have not been diagnosed with diabetes, but have impaired fasting glucose (6.1-6.9 mmol/l)

In some ethnic minorities the risk charts underestimate CVD risk, because they have not been validated in these populations. For example, in people originating from the Indian subcontinent it is safest to assume that the CVD risk is higher than predicted from the charts (1.5 times).

The charts may be used to illustrate the direction of impact of risk factor intervention on estimated level of CVD risk. However, such estimates are crude and are not based on randomised trial evidence. Nevertheless, this approach maybe helpful in motivating appropriate intervention. The charts are primarily to assist in directing intervention to those who typically stand to benefit most.

Continuation appendix 5.1
Nomogram



-Nomogram for predicted probability of survival at 1 year. For each of the five variables, points are calculated by reading from the top scale. The total point score is then translated into a probability of survival by using the bottom two scales. For example, a patient of age 70 (50 points), with a GCS of 12 (50 points), an ISS of 20 (75 points), reactive pupils (37 points), and no haematoma on CT (25 points) has a total score of 237 points, and a corresponding probability of survival at 1 year of about 0.93.

Appendix 5.2

Survey about the CRASH score for prognosis in traumatic head injury

Instructions

The CRASH score was developed to estimate prognosis for traumatic head injury patients. We want it in a practical format so it can be used in the clinical setting. Your answers will help this.

First some things about you

1) Your age ☐

2) Your specialty in medicine ☐

4) In which country is your hospital located? ☐

5) About how many patients with head injury do you treat per month? ☐

Please estimate the CRASH score and probability of death at 14 days for the following patient. (Using the CRASH Score provided)

Male aged 52 years old who had a road traffic crash; on physical examination the total Glasgow Coma Scale is 10 and both pupils are reactive. The CT scan shows midline shift of 8 mm. The patient does not have a major extra cranial injury.

7) What is the CRASH score for this patient? ☐

6) What is the probability of death at 14 days? ☐

SO that we know your opinion about the practicality of the CRASH score Card, please say how much you agree or disagree with the following statement (tick one box)

"The format of the CRASH score is practical for use in the clinical setting"

Totally agree	Agree	Unsure	Disagree	Totally disagree
☐	☐	☐	☐	☐

Finally write any suggestion you have to improve the CRASH score.

Thank you very much for your collaboration!

Appendix 6

Prognosis of size of intracranial bleeding in patients with TBI

Introduction

Intracranial bleeding (IB) is a common and serious consequence of traumatic brain injury (TBI). In the CRASH trial 56% of the patients had at least one IB.[63]

IB can be classified according to the location into epidural haemorrhage (EDH), subdural haemorrhage (SDH), intraparenchymal haemorrhage (IPH), and subarachnoid haemorrhage (SAH).

A review by the Brain Trauma Foundation found that all types of IB are associated with a worse prognosis, with increased in-hospital mortality and disability at six months.[139]

In the multivariable analysis reported in Chapter 4 I found that in the CRASH trial SAH and non evacuated haematoma were independently associated with a worse outcome at 2 weeks and six months.[117]

Among all the variables identified in the prognostic models IB is one of the few that could be "etiologically" related with the outcome. Furthermore, the increasingly recognized fact that IB "evolves" in the first hours after a TBI makes it a potential target for interventions. Studies involving repeated CT scanning of patients with TBI have found that intracranial bleeding can develop or expand in the first 24-72 hours after injury. Oertel and colleagues found that among patients who had their first scan within 2 hours of injury, 49% had radiological evidence of progressive bleeding within 24 hours.[140] More recently Narayan and collaborators showed that approximately 50% of IPH expand in the first 72 hours after a TBI.[64] Evidence that IB can enlarge following TBI has generated interest in potential therapeutic approaches, such as haemostatic drugs, that could prevent or decrease the growth of IB.[64] If IB enlarges after hospital admission, and larger bleeds have a worse prognosis, this would strengthen the therapeutic rationale for agents to prevent an increase in the extent of bleeding. Although there have been some studies on the association between size of IB and prognosis, the empirical evidence is limited, most studies having small sample sizes and restricted populations.[74-77] Many of these studies were conducted in one centre and a limited set of data were collected; this explains why, although biologically plausible, there is still uncertainty regarding the strength and characteristics of the association between size of IB and mortality in TBI patients.

The CRASH trial did not collect data about the different types and size of IB to study

this relationship. Therefore, to further explore this topic I contacted the Trauma Audit & Research Network (TARN), a large European trauma registry, which has detailed data on EDH, SDH, and IPH, and I set up a collaboration. I designed and wrote the full extent of the study included in this appendix. Because of confidentiality aspects associated with the database a statistician from TARN conducted the analysis.

The aim of this chapter was to evaluate the association between the size of IB (EDH, SDH and IPH) and mortality and haematoma evacuation in patients with TBI.

Methods

Sample

TARN was established in 1989 to benchmark and improve hospital trauma care (using case fatality measures). Membership is voluntary and includes 60% of hospitals receiving trauma patients in England and Wales and some hospitals in European centres. Data are collected on patients who arrive at hospital alive and meet any of the subsequent criteria:

- Death from injury at any point during admission
- Stay in hospital for longer than 3 days
- Require intensive or high dependency care
- Require inter – hospital transfer for specialist care.

Patients with isolated closed limb injuries are excluded, as are patients over 65 years with isolated fracture neck of femur or pubic ramus fracture. Data are collated by trained staff in participating hospitals and submitted via the TARN Electronic Data Collection and Reporting (EDCR) system (ref www.tarn.ac.uk). Each submission is checked for consistency and accuracy by trained coders at the University of Manchester. All injuries are coded using the Abbreviated Injury Scale (AIS) 1998 Dictionary which allocates each injury a severity code between 1 (minimal) and 6 (maximal).[141] AIS severity coding is derived from the precise injury descriptions given by imaging, operative and post mortem reports.

For this study adult patients, hospitalized between 2001 and 2008, with a Glasgow Coma Score (GCS) less than 15 at presentation or any head injury with AIS severity code 3 and above were selected.

Variables

Main Exposures

The extent of intracranial bleeding was determined from the AIS code.[141] IB was coded as epidural (EDH), subdural (SDH), and intraparenchymal (IPH). Each type was coded as absent, present small, present large or present size unspecified, referred as "no

further specification" (NFS) in this paper. There are differences in the volume of blood that attract "small/large" codes depending on the site of bleeding. (Appendix 6.1)

There was no data about the size of subarachnoid haemorrhage (SAH).

Potential confounders

Potential confounders of the relationship between size of bleeding and patient outcome were selected for the multivariable analysis. These variables were: age, GCS, SAH, brain contusions, brain swelling, petechial haemorrhages, presence of other brain injuries (skull fractures and any brain lesion no further specified), presence of extracranial injuries (AIS with severity score >2), and whether or not the patient has been treated at a neurosurgical unit (NSU). These variables have previously been reported to be associated with poor outcome. [65,115,117,142]

Outcome

The main outcome was in hospital mortality. I also explored the association between size of bleeding and evacuation of haematoma.

Other variables

Other variables reported for descriptive purposes of the sample were: gender, cause of injury (road traffic crash, fall <2 m, fall > 2 m) and Injury Severity Score (ISS). ISS is a summary of the overall severity of anatomical injury for each patient.[143] It has an ordinal scale from 1-75 and is derived from the AIS severity scores for each injury.

Analysis

Analysis of Age and GCS

To determine the functional form of the predictors age and GCS in the model, fractional polynomials, quadratic and cubic spline and Lowess smoothing were explored.

Conceptual framework of the multivariable analysis

Deciding which variables should be considered confounders and which should be considered mediators that are on the causal pathway between bleeding and outcome requires a conceptual framework. I could consider as confounders all variables shown to be associated with poor prognosis in TBI such as age, severity of the TBI (as defined by GCS), and other CT scan abnormalities. However, some of these variables (i.e. brain swelling and GCS) might be on the causal pathway between bleeding and patient outcome. Adjusting for these variables would attenuate a true association between bleeding and outcome. Because of the uncertainty in determining which factors are confounders and which are on the causal pathway, I analysed the data from two

conceptual frameworks in the hope that the two different analyses would provide a better understanding of the association between IB and outcome. The first includes all potential confounders, the second excludes brain swelling and GCS as these variables could be on the causal pathway between IB and patient outcome.

Statistical approach for the multivariable analysis

Firstly I reported the crude association for each of the exposures of interest. Secondly for the adjusted models all the potential confounding variables were entered into a multivariable logistic regression to analyse their relationship with the outcome. All the exposures and the potential confounders for which there was strong evidence for a relationship with mortality ($p < 0.05$), were retained in the adjusted models.
An initial analysis considered no bleeding as the baseline category. Because I was interested in quantifying the mortality risk associated with large, as opposed to small IB, I also conducted a second analysis evaluating the effect of IB size on mortality using small IB as the baseline.

Results

General characteristics

Between 2001 and 2008 18,055 adult patients meeting study inclusion criteria presented to TARN hospitals. In 2,507 (14%) patients the outcome was unknown, in 1,586 (9%) patients, the GCS was missing. The remaining 13,962 (77%) were used for this study.

Table A6.1 describes the characteristics of the study population. Almost three quarters of the patients were male. The median age was 41 years old, the median GCS was 13 and the median ISS was 18. The most common mechanism of injury was road traffic crashes and in-hospital mortality was 22%. About 46% of patients had some type of IB. SDH was the most common type, present in 30% of the patients. EDH, IPH and SAH were present in 22% each. The size of IB (either large or small) was reported in 30% of patients with EDH, in 53% with SDH, and in 27% of patients with IPH. Patients with IB were generally older, had more severe TBI (as defined by GCS), and had higher in-hospital mortality. Among the different types of IB, patients with EDH were the youngest, and those with SAH had the highest in-hospital mortality. Patients with IPH were less frequently hospitalized in services with neurosurgery units (NSU).

Table A6-1 Characteristics of the population

	All patients	EDH	SDH	IPH	SAH	No bleeding
N(%)	13962 (100%)	3140 (22.5%)	4204 (30.1%)	2990 (21.8%)	3025 (21.7%)	7517 (53.8%)
NFS		2,185 (70%)	1985 (47%)	2193 (73%)		
Small	NA	536 (17%)	1168 (28%)	321(11%)	NA	NA
Large		419 (13%	1051 (25%)	476 (16%)		
Median Age	40.7	43.4	48.9	47.1	46.6	37.8
Male	10229 (73.3%)	2352 (74.9%)	3050 (72.5%)	2187 (73.1%)	2257 (74.6%)	5456 (72.6%)
Female	3733 (26.7%)	788 (25.1%)	1154 (27.5%)	803 (26.9%)	768 (25.4%)	2061 (27.4%)
Median GCS	13	11	10	11	8	14
Median ISS	18	25	25	25	25	13
RTC	6125 (43.9%)	1053 (33.5%)	1337 (31.8%)	1025 (34.3%)	1299 (42.9%)	3756 (50.0%)
Fall more > 2m	2312 (16.6%)	753 (24.0%)	1026 (24.4%)	690 (23.1%)	715 (23.6%)	892 (11.9%)
Fall <2m	2706 (19.4%)	720 (22.9%)	1119 (26.6%)	763 (25.5%)	561 (18.5%)	1238 (16.5%)
Other	2819 (20.2%)	614 (19.6%)	722 (17.2%)	512 (17.1%)	450 (14.9%)	1631 (21.7%)
NSU	6055 (43.4%)	1617 (51.5%)	2160 (51.4%)	1356 (45.4%)	1599 (52.9%)	2704 (36.0%)
Mortality	3065 (22.0%)	869 (27.7%)	1380 (32.8%)	950 (31.8%)	1222 (40.4%)	1098 (14.6%)

NA: Not applicable NFS: No further specified GCS: Glasgow coma scale ISS: Injury severity score RTC: Road traffic crash NSU: Neurosurgical unit EDH: Epidural haemorrhage SDH: Subdural haemorrhage IPH: Intracerebral haemorrhage SAH Subarachnoid haemorrhage

Relationship between age and GCS with mortality

Figures A6.1 and A6.2 show the fit of the three functional forms to the observed data. It can be seen that fractional polynomials (FP) fit the data well for both age and GCS, therefore they were included in this way in the analysis. For age the optimal functional form was the sum of square root age and age, for GCS it was the sum of inverse GCS cubed and GCS.

Figure A6-1 Functional form for age in TARN

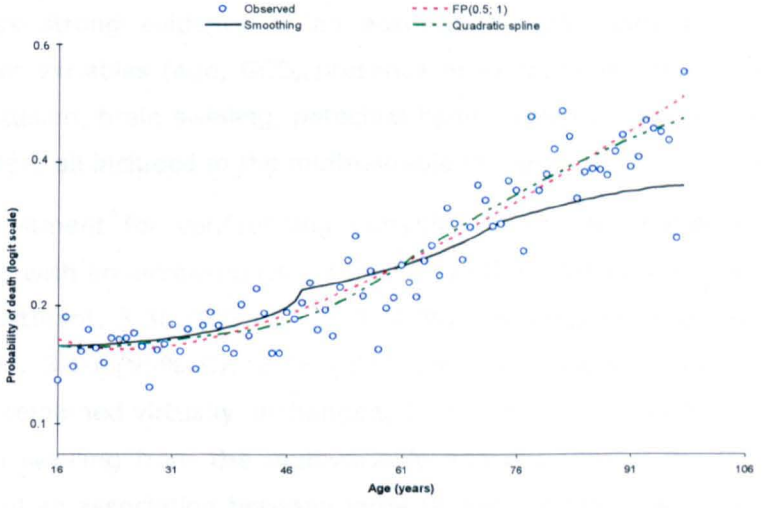
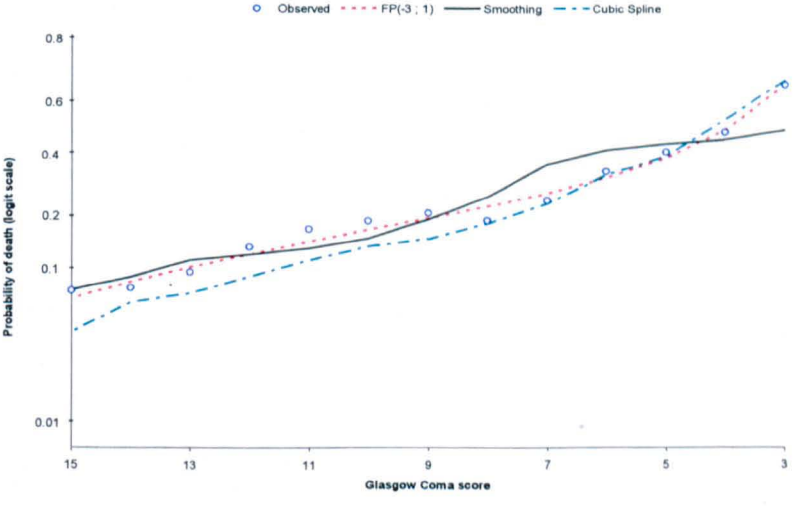


Figure A6-2 Functional form for GCS in TARN



In-hospital Mortality

Table A6.2 shows the crude and adjusted effect (odds ratio) for mortality of the different types and sizes of IB.

Crude analysis

IB either coded as large or NFS in all locations were associated with an increased risk of mortality in comparison with no bleeding. Large SDH and large IPH were associated with a worse prognosis, with an odds ratio (OR) for mortality of 6.30 (*95% CI 5.50-7.21*) and OR 4.19 (*95% CI 3.46-5.06*) respectively. Small SDH were the only small lesions associated with an increase in mortality.

Adjusted analysis

There was strong evidence of an association with mortality for all the potential confounder variables (age, GCS, presence of extracranial injury, treatment at a NSU, brain contusion, brain swelling, petechial haemorrhages, *SAH and other brain injuries*) so they were all included in the multivariable model.

After adjustment for confounding variables, large IB irrespective of location was associated with an increased risk of mortality. The odds ratio for large SDH was halved after adjustment, 3.36 (*95% CI: 2.76-4.08*), the odds ratio for large IPH was slightly attenuated, 3.10 *(95% CI: 2.38-4.03*) and the association between large EDH and mortality remained virtually unchanged, 1.85(*95% CI: 1.36-2.51*). After excluding GCS and brain swelling from the multivariable analysis (model 2), there was still strong evidence of an association between large IB and mortality, with values for the OR that were more extreme than those reported in model 1.

Table A6-2 Association between haemorrhage size and mortality

Type of IB	Size	Crude odds ratio	Adjusted odds ratio *(model 1†)*	Adjusted odds ratio *(model 2‡)*
Epidural haemorrhage (EDH)	**No EDH**	1	1	1
	NFS	1.77 (1.60-1.96)	1.28 (0.84-1.93)	1.27 (0.89-1.83)
	Small	0.57 (0.43-0.74)	0.67 (0.47-0.95)	0.61 (0.45-0.83)
	Large	1.61(1.29-2.01)	1.85 (1.36-2.51)	2.11 (1.62-2.75)
Subdural haemorrhage (SDH)	**No SDH**	1	1	1
	NFS	1.75 (1.57-1.96)	0.98 (0.71-1.35)	1.05 (0.80-1.40)
	Small	1.31 (1.13-1.53)	0.99 (0.81-1.22)	1.21 (1.02-1.44)
	Large	6.30 (5.50-7.21)	3.36 (2.76-4.08)	7.09 (6.01-8.37)
Intraparenchymal haemorrhage (IPH)	**No IPH**	1	1	1
	NFS	1.79 (1.61-1.98)	1.13 (0.75-1.69)	1.40 (0.98-1.99)
	Small	0.88 (0.65-1.19)	0.89 (0.61-1.30)	0.83 (0.60-1.15)
	Large	4.19 (3.46-5.06)	3.10 (2.38-4.04)	3.45 (2.74-4.33)

Variables included in the model:

†Model 1: Age, GCS, NSU, EDH, SDH, IPH, Brain contusion, Swelling, Petechial, SAH, Other brain injuries, Extracranial injuries.

‡Model 2: Age, NSU, EDH, SDH, IPH, Brain contusion, Petechial, Penetrating, SAH, Other brain injuries, Extracranial injuries.

Evacuation of haematoma

Table A6.3 shows the crude and adjusted effect (odds ratio) for haematoma evacuation of the different types, and size, of IB.

Crude analysis

IB from all the locations and from all the categories (large, small and NFS) were associated with an increased risk of evacuation, except for small IPH. EDH and SDH showed the largest odds ratio (22.6 and 13.7 respectively).

Adjusted analysis

After adjusting for all the potential confounding variables, there was an increased risk of haematoma evacuation for both SDH and EDH. The magnitude of the association was larger for large haematomas, intermediate for those coded as NFS and smallest for the small ones. The odds ratio for large EDH and SDH were, respectively, 25.58 (*95% CI:18.80-34.81*) and 15.47 (*95% CI: 11.88-20.13*). After multivariate analysis none of the categories of IPH remained positively associated with evacuation. Similar results were obtained when excluding GCS and brain swelling from the multivariable adjustment.

Table A6-3 Association between haemorrhage size and haematoma evacuation

Type of IB	Size	Crude odds ratio	Adjusted odds ratio *(model 1†)*	Adjusted odds ratio *(model 2‡)*
Epidural haemorrhage (EDH)	No EDH	1	1	1
	NFS	4.69 (4.02-5.49)	2.78 (1.75-4.44)	2.95 (1.85-4.70)
	Small	3.96 (2.98-5.20)	2.99 (2.15-4.20)	3.08 (2.22-4.27)
	Large	22.56 (18.05-28.16)	25.58 (18.80-34.81)	28.87 (21.27-39.20)
Subdural haemorrhage (SDH)	No SDH	1	1	1
	NFS	6.43 (5.40 -7.65)	5.58 (3.78-8.25)	5.77 (3.91-8.50)
	Small	3.96 (3.14-4.98)	3.29 (2.50-4.33)	3.59 (2.73-4.72)
	Large	13.70 (11.40-16.47)	15.47 (11.88-20.13)	19.40 (15.07-24.97)
Intraparenchymal haemorrhage (IPH)	No IPH	1	1	1
	NFS	3.54 (3.07-4.08)	0.58 0.36-0.95)	0.58 (0.36-0.95)
	Small	0.82 (0.44-1.40)	0.617 (0.31-1.22)	0.66 (0.34-1.29)
	Large	1.91 (1.36-2.63)	0.91 (0.57-1.44)	1.04 (0.66-1.63)

Variables included in the model:

†Model 1: Age, GCS, NSU, EDH,SDH,IPH, Brain contusion, Swelling, Petechial, SAH, Other brain injuries, Extracranial injuries.

‡Model 2: Age, NSU, EDH, SDH, IPH, Brain contusion, Petechial, SAH, Other brain injuries, Extracranial injuries.

Comparison between large and small haemorrhages

In table A6.4 it is shown that large IB, wherever the location, were associated with an increased risk of mortality in comparison with small IB lesions. After adjusting for potential confounders (model 1) the odds ratio for mortality was 2.86 (*95% CI: 1.86-4.38*) for large EDH, 3.41 (*95% CI: 2.68-4.33*) for large SDH and 3.47 (*95% CI: 2.26-5.33*) for large IPH. Patients with EDH coded as NFS had an odds ratio for mortality of 1.89 (95% CI: 1.20-2.99) in comparison with those with small EDH. Patients with no EDH showed an increase risk in comparison with those patients with a small EDH. There was no strong evidence of increased risk of mortality for those patients with SDH or IPH coded as NFS when compared with patients with corresponding lesions coded as small.

Table A6-4 Comparison between large and small haemorrhages

Type of IB	Size	Adjusted odds ratio †
Epidural haemorrhage (EDH)	Small	1
	No EDH	1.49 (1.05 – 2.12)
	NFS	1.89 (1.20 -2.99)
	Large	2.86 (1.86 – 4.38)
Subdural haemorrhage (SDH)	Small	1
	No SDH	1.07 (0.85 - 1.35)
	NFS	0.99 (0.72 - 1.37)
	Large	3.41 (2.68 - 4.33)
Intraparenchymal haemorrhage (IPH)	Small	1
	No IPH	1.23 (0.84 - 1.80)
	NFS	1.39 (0.84 – 2.28)
	Large	3.47 (2.26 – 5.33)

Variables included in the model: † Age, GCS, NSU, EDH, SDH, IPH, Brain contusion, Swelling, Petechial, SAH, Other brain, Extracranial injuries.

Discussion

General findings

This analysis of over 13,000 patients with TBI showed that patients with a large EDH, SDH or IPH have a substantially higher mortality than patients with either no bleeding or a small bleed in the same location. Even after adjusting for other potential confounding variables, such as age, GCS, extracranial injuries and CT findings, large bleeds substantially increased the probability of death. Patients with large IPH or large SDH had more than a threefold increase in mortality odds in comparison with patients with small IB in the same location, while large EDH showed more than a doubling in mortality odds in comparison with patients with small EDH. Small IBs were not associated with an increase in mortality after adjustment for other potentially confounding variables. Patients with IB coded as NFS had generally a risk which was intermediate between that reported for patients with large IB and the one reported for patients with small IB. Patients with no EDH had a higher risk of mortality in comparison with patients small EDH. This finding although counterintuitive has been previously reported.[104] In the analysis reported in this appendix I adjusted for other type of IB as I considered them potential confounders. However, because of the large amount of missing data in relation to size of IB it is very likely that there is residual confounding. This means that those patients coded as no EDH could still have some type of IB not recorded in the TARN database. Other explanations for this paradoxical finding could be related to other potential confounders not recorded in the TARN database. For example the presence of small EPH could be related to a certain mechanism of TBI which has a better prognosis than the mechanism seen in patients who do not have an EPH, such as those with diffuse axonal injury, a CT scan result not recorded in this dataset.

Comparison with previous studies

These results are consistent, but more precise, than those of previous studies showing that IB is associated with increased mortality. There has not been any systematic review describing the association between size of IB and prognosis in TBI but a comprehensive review has been reported in the *Guideline for the Surgical Management of Traumatic Brain Injury*".[112] In this guideline bleeding size is taken into account to recommend surgical evacuation. However, the evidence presented in the guideline is very limited. For EDH they reported 18 studies with a median of 67 patients included in each (range:11 to 200), seven of the studies evaluated the effect of size, and only three reported a positive association between size and poor outcome. In relation to SDH there were 21 studies and the median number of patients was also 67 (range 15-

211), five reported on the effect of size, and only two reported a positive association between size and poor outcome. For IPH 51 studies were reported, with a median of 35 patients included (range 8-1,107), and only five reported a positive association between size and poor outcome.

Strengths and weaknesses

The strength of this analysis is that it included more than 13,000 patients with traumatic brain injury, and so the precision of the estimates of the risk associated with IB is high. I also adjusted for most of the relevant potential confounding variables.

One limitation of this analysis is that for a large proportion of patients it was not reported whether their IB was small or large. These patients, with IB size coded as NFS, presented intermediate risk between patients with small and large lesions. The large proportion of patients with missing data on size of bleeding makes the results presented in this appendix of limited value and further studies with more complete data should be carried out to confirm the reported associations.

Another limitation is the potential of residual confounding. Although I adjusted for known confounders, it is possible that some of them were measured in a limited way or other confounding variables may have remained unaccounted for. For example, I did not have information on pupil reactivity which has been shown to be an important prognostic factor. A further limitation is that I had no continuous measurements of the size of the bleeding, nor did I analyse the timing of the CT scan.

Finally, 23% of the patients who fulfilled the inclusion criteria were excluded because either their outcome was unknown or there was lack of data on GCS. This is a potential cause of selection bias and could have influenced the effect estimates obtained in this study. Selection bias in this context would have occurred if the relationship between size of IB and mortality in the excluded patients was different from the one observed in the sample analysed. Although there are no evident reasons for such a difference this cannot be ruled out and should be considered as a limitation of this study.

Implications of this study

Patients with TBI and large intracranial bleeds have a substantially worse outcome than patients with either no bleeding or a small bleed. There is evidence from other studies that bleeds enlarge in the first 24-72 hours after injury. This raises the possibility that interventions to prevent the enlargement of intracranial bleeds could improve patient outcome.

Future research

However, it is possible that some of the increased risk of death associated with having large rather than small IB found in this study is due to bias or confounding. Therefore, the effect on mortality of interventions that reduce the extent of intracranial bleeding would need to be established in randomised controlled trials.

Appendix 6.1

AIS 1990 revision, update 1998

Small EDH

Cerebellum ≤ 30cc

Cerebrum ≤ 50mls

Small SDH

Cerebellum ≤ 30cc

Cerebrum ≤ 50cc

Small IPH

Cerebellum (≤15cc; or 3cm diameter)

Cerebral ≤ 30cc; ≤ 4cm diameter

Large EDH

Cerebellum >30cc, > 2cm diameter/ thick, massive / extensive, bilateral

Bilateral cerebral

Cerebral > 50 cc, >1cm thick, massive / extensive

Large SDH

Cerebellar > 30cc, > 2cm diameter/ thick, massive /extensive, bilateral

Bilateral cerebral

Cerebral > 50 cc, >1cm thick, massive / extensive

Large IPH

Cerebellar > 15cc; >3cm diameter, bilateral

Bilateral cerebral

Cerebral > 30cc / 4cm diameter

EDH NFS

Epidural haematoma (haemorrhage) to cerebellum not further specified

Cerebrum epidural l haematoma/haemorrhage not further specified,

SDH NFS

Cerebellar subdural haematoma not further specified

Cerebrum subdural haematoma not further specified

IPH NFS

Intracerebellar including petechial and subcortical haematoma (haemorrhage) not further specified

Cerebrum intracerebral haematoma not further specified

Appendix with published papers

Perel P, Wasserberg J, Ravi RR, Shakur H, Edwards P, Roberts I. Prognosis following head injury: a survey of doctors from developing and developed countries. J Eval Clin Pract. 2007 Jun;13(3):464-5 (Chapter 2).

Perel P, Edwards P, Wentz R, Roberts I. Systematic review of prognostic models in traumatic brain injury BMC Med Inform Decis Mak. 2006 Nov 14;6:38 (Chapter 3).

MRC CRASH Trial Collaborators, Perel P, Arango M, Clayton T, Edwards P, Komolafe E, Pocock S, Roberts I, Shakur H, Steyerberg E, Yutthakasemsunt S. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. BMJ. 2008 Feb 23;336(7641):425-9 (Chapter 4).

Perel P, Edwards P, Shakur H, Roberts I. Use of the Oxford Handicap Scale at hospital discharge to predict Glasgow Outcome Scale at 6 months in patients with traumatic brain injury BMC Med Res Methodol. 2008 Nov 6;8:72 (Chapter 6).

Journal of Evaluation in Clinical Practice ISSN 1356-1294

RESEARCH LETTER

Prognosis following head injury: a survey of doctors from developing and developed countries

Pablo Perel MSc MD,[1] Jonathan Wasserberg BSc MBBChir FRCS,[2] Ramalingam R. Ravi MD,[3] Haleema Shakur BSc MSc RGN,[4] Phil Edwards BSc PhD CStat[5] and Ian Roberts MB BCh MRCP PhD[6]

[1]Research Fellow, [4]Trial Manager, [5]Lecturer, [6]Professor of Epidemiology, London School of Hygiene and Tropical Medicine, London, UK
[2]Senior Lecturer, Department of Neurosurgery, Queen Elizabeth Hospital, Birmingham, UK
[3]Chief of Neurosciences, Medical Trust Hospital – Kochi, India

Correspondence
Pablo Perel
CRASH Co-ordinating Centre
London School of Hygiene and Tropical Medicine
Keppel Street
London WC1E 7HT
UK
E-mail: pablo.perel@lshtm.ac.uk

Accepted for publication: 30 November 2005

doi:10.1111/j.1365-2753.2006.00713.x

Introduction

Head injury is an important cause of death and disability worldwide with most of the burden in low- and middle-income countries. Accurate information on prognosis has the potential to improve clinical decisions in patients with head injury [1]. Prognostic models are statistical models that combine two or more items of patient data to predict clinical outcome. Some studies have shown that correctly interpreted prognostic models can be more reliable than clinical judgement [2]. We conducted a survey among doctors who routinely treat patients with head injury to assess their needs in relation to prognostic information.

Participants, methods and results

The sampling frame for the survey was all doctors participating in the final results meeting of a large-scale international multi-centre clinical trial in head injury [3]. Of the 67 doctors attending, 60 completed the study questionnaire. Prior to the meeting, the criterion and face validity of the questionnaire was assessed in a convenience sample of doctors treating head injury (not included in the final sample). The questionnaire was translated into Spanish for the one quarter of respondents who were Spanish speakers. The main specialities of the respondents were intensive care (34% of respondents) and neurosurgery (32%). Most doctors worked in low- and middle-income countries: 25% from Latin America and the Caribbean, 20% from Africa, 19% from Southeast Asia and 9% from Eastern Europe and Central Asia; while 27% were from Western Europe. Most worked in hospitals with intensive care, computed tomography and neurosurgical facilities.

The outcome considered most important to predict was in-hospital mortality (73%); the most favoured way of expressing the prognosis was probability of the outcome (97%) followed by a qualitative scale (82%) or as survival time (77%).

The majority (85%) reported routinely using a score to assess prognosis, of which the summated Glasgow Coma Scale was the most popular used by three quarters of the doctors. In relation to individual prognostic variables, 80% of doctors felt that the motor component of the Glasgow Coma Scale was very important. Other variables identified as very important were pupillary reaction (80%) and a midline shift greater than 5 mm on the CT scan (88%).

Although most doctors routinely used a score to assess prognosis, only 37% agreed that they currently assess prognosis accurately. The majority (67%) reported that a more accurate prognostic model would change the way that they manage patients and the way that they currently tell the prognosis to a patient's relative (88%). Accurate prognostic information was considered to be very important for a number of clinical decisions, including the need to undertake a decompressive craniotomy, who should receive intensive care, and in which patients treatment should be withdrawn (see Table 1).

Table 1 Doctors' opinions in relation to the situations for which accurate prognostic information is important ($n = 60$)

Situations	Very important (%)	Important (%)	Not important (%)
To decide which patients need decompressive craniotomy	61	27	12
To decide which patients need Intensive Care Unit	60	26	14
To decide which patients should receive treatment (e.g. hyperventilation, barbiturates, mannitol)	55	26	19
To give counselling to patients and/or relatives	54	37	9
To decide in which patients treatment should be withdrawn	52	34	14
To decide in which patients intracranial pressure should be monitored	50	40	10
To decide which patients need surgery	49	32	19
To decide in which patients CT scan should be done	19	49	32
To decide which patients need rehabilitation	14	53	33

Comments

Doctors around the world make important decisions about the care of patients with head injury, including the decision to withdraw care, based on judgments about prognosis. More widespread use of accurate methods of assessing prognosis such as statistical prognosis models may improve clinical management in head injury.

A key strength of our survey is that it includes doctors treating head-injured patients from diverse regions of the world, mainly from developing countries, where the major burden of head injury occurs. The response rate was high at 90%.

A previous survey of 59 neurosurgeons [4] showed that two-thirds of them thought that computer predictions would be helpful in the clinical management of patients with head injury. Twenty years later, it appears that there is still demand for accurate prognostic information. Although there has been a systematic review of individual prognostic factors in head injury [5], such an approach is lacking for prognostic models. A systematic and critical appraisal of existing prognostic models would enable doctors to know which of the available models are accurate and clinically useful.

References

1. Altman, D. G. (2001) Systematic reviews in health care: systematic reviews of evaluations of prognostic variables. *British Medical Journal*, 323, 224–228.
2. Lee, K. L., Pryor, D. B., Harrell, F. E. Jr, Califf, R. M., Behar, V. S., Floyd, W. L., Morris, J. J., Waugh, R. A., Whalen, R. E. & Rosati, R. A. (1986) Predicting outcome in coronary disease. Statistical models versus expert clinicians. *The American Journal of Medicine*, 80, 553–560.
3. CRASH Trial Collaborators. (2004) Effect of intravenous corticosteroids on death within 14 days in 10 008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet*, 364, 1321–1328.
4. Barlow, P. & Teasdale, G. (1986) Prediction of outcome and the management of severe head injuries: the attitudes of neurosurgeons. *Neurosurgery*, 19, 989–991.
5. Brain Trauma Foundation (U.S.) (2000) Management and Prognosis of Severe Traumatic Brain Injury. The Brain Trauma Foundation Website. Available at: http://www2.braintrauma.org/guidelines

BMC Medical Informatics and Decision Making

BioMed Central

Research article

Open Access

Systematic review of prognostic models in traumatic brain injury

Pablo Perel*, Phil Edwards, Reinhard Wentz and Ian Roberts

Address: Nutrition and Public Health Intervention Research Unit, Epidemiology and Population Health Department, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Email: Pablo Perel* - pablo.perel@lshtm.ac.uk; Phil Edwards - phil.edwards@lshtm.ac.uk; Reinhard Wentz - reinhard.wentz@lshtm.ac.uk; Ian Roberts - ian.roberts@lshtm.ac.uk

* Corresponding author

Published: 14 November 2006

BMC Medical Informatics and Decision Making 2006, 6:38 doi:10.1186/1472-6947-6-38

Received: 03 August 2006
Accepted: 14 November 2006

This article is available from: http://www.biomedcentral.com/1472-6947/6/38

Abstract

Background: Traumatic brain injury (TBI) is a leading cause of death and disability world-wide. The ability to accurately predict patient outcome after TBI has an important role in clinical practice and research. Prognostic models are statistical models that combine two or more items of patient data to predict clinical outcome. They may improve predictions in TBI patients. Multiple prognostic models for TBI have accumulated for decades but none of them is widely used in clinical practice. The objective of this systematic review is to critically assess existing prognostic models for TBI

Methods: Studies that combine at least two variables to predict any outcome in patients with TBI were searched in PUBMED and EMBASE. Two reviewers independently examined titles, abstracts and assessed whether each met the pre-defined inclusion criteria.

Results: A total of 53 reports including 102 models were identified. Almost half (47%) were derived from adult patients. Three quarters of the models included less than 500 patients. Most of the models (93%) were from high income countries populations. Logistic regression was the most common analytical strategy to derived models (47%). In relation to the quality of the derivation models (n:66), only 15% reported less than 10% pf loss to follow-up, 68% did not justify the rationale to include the predictors, 11% conducted an external validation and only 19% of the logistic models presented the results in a clinically user-friendly way

Conclusion: Prognostic models are frequently published but they are developed from small samples of patients, their methodological quality is poor and they are rarely validated on external populations. Furthermore, they are not clinically practical as they are not presented to physicians in a user-friendly way. Finally because only a few are developed using populations from low and middle income countries, where most of trauma occurs, the generalizability to these setting is limited.

Background

Traumatic brain injury (TBI) is a leading cause of death and disability worldwide. Every year, an estimated 1.5 million people die and hundreds of millions require emergency treatment after a TBI. Fatality rates and disability rates vary depending on the severity and mechanisms of the TBI but unfavourable outcomes (death, vegetative state and severe disability) following TBI can be higher than 20%[1,2].

Physicians routinely make diagnostic and therapeutic decisions based on the patient's prognosis. Furthermore, prognostic information is also important in the counselling of patients and relatives in this critical scenario. Nevertheless in general, physicians believe that their predictions are inaccurate. A survey of doctors about prognosis in TBI found that only 37% thought that they currently assess prognosis accurately[3].

Prognostic models are statistical models that combine two or more items of patient data to predict clinical outcome. They may improve predictions in TBI patients. Some studies have shown that they are more reliable than what doctors can foretell [4]. A study conducted with TBI patients demonstrated that the introduction of a computer-based outcome prediction influenced patient management, with a higher use of resources in those patients with better prognosis [5].

Prognostic models could also be used in the design and analysis of Randomized Controlled Trials (RCTs). RCTs in TBI patients face many difficulties. Trauma is one of the most neglected research topics worldwide with a paucity of resources invested in RCTs [6]. Furthermore, unfamiliarity with issues of informed consent in unconscious patients pose further obstacles in this clinical setting [7]. Because of these barriers RCT in TBI are generally underpowered. A review of published RCTs in this area found that the average size was 82 participants per trial and no trial was large enough to detect reliably a 5% absolute reduction in risk [8]. Prognostic models have been proposed as a way to improve the power in TBI and stroke clinical trials [9,10]. With one such approach TBI patients' outcomes are defined taking account their baseline prognosis, instead of using the usual Glasgow Outcome Scale dichotomized in favourable or unfavourable.

Prognostic models can also assist in clinical audit by allowing adjustment for case-mix [11,12].

Multiple prognostic models for TBI have accumulated for decades but none of them is widely used in clinical practice. For a prognostic model to be clinically useful it should fulfil two requirements: it must be clinically valid and methodologically valid [13]. Systematic reviews of prognostic models in different areas of medical care have shown that models often fail in these two aspects [14,15]. Previous reviews of prognostic studies in TBI have only focused on individual predictors or have been restricted to prognostic models of some type of traumatic brain injury or outcome. So far, there has not been any comprehensive systematic review of prognostic models in traumatic brain injury [16,17]. It has then become increasingly important to identify and evaluate prognostic models in TBI patients.

Objective

Our objectives were

(a) identify prognostic models in traumatic brain injury

(b) describe their characteristics

(c) investigate their quality and

(d) described the models that were validated in an external population.

Methods

Type of studies

We included studies that gave an overall prognostic estimation combining the predictive information from at least two variables. Studies could develop new prognostic models (derivation studies) or evaluate previous ones (validation studies). Studies conducted prior to 1990 were excluded because patient management and diagnostic techniques may have changed since this time. Studies that investigate more than one variable but do not combine them for obtaining a prediction were excluded.

Type of exposures

Only variables that were collected before hospital discharge were considered as predictors. Glasgow Coma Scale (GCS) was considered as one predictor variable.

Type of participants

Patients of any age with any type or severity of traumatic brain injury.

Type of outcome measures

Studies that predict any outcome in traumatic brain injury patient (i.e. neurological impairment, disability, survival, etc.). There was no time restriction for the evaluation of the outcomes.

Search strategy for identification of studies [see Additional file 1]

The reference lists of included studies were inspected for further possible studies meeting the inclusion criteria. A forward search (citing references in the Web of Knowledge) was conducted with selected seminal papers and some of the citing papers, not found by the database search, were inspected for relevance and possible inclusion. All records were converted into an Endnote database.

Trial identification and selection

Two reviewers (PP & PE) independently examined titles, abstracts and keywords of records from electronic databases, for eligibility. The full text of all potentially relevant records was obtained and two reviewers (PP & PE) inde-

pendently assessed whether each met the pre-defined inclusion criteria. Disagreement was resolved by a third reviewer (IR).

Quality assessment

Quality assessment scores for controlled clinical trials and diagnostic studies have been criticized [18,19]. The main problem with quality scores is to determine the weight that each item should provide to the overall score. The abundance of quality scores shows that there is no consensus on this issue. Instead, a component approach appraisal allows one to evaluate each methodological aspect. Depending on the question and the study design some components may be more relevant than others (e.g. with a surgical intervention blinding of the patient and caregiver would be unachievable)

In studies of prognostic models in particular, although diverse quality assessment criteria have been proposed, there is none widely accepted [14,20-22]. We analyzed the quality of the prognostic models included in this systematic review considering two main domains:

a) Internal validity. This refers to the systematic error of the study and is related to study design, variables and analysis strategy.

b) External validity or generalizability. This refers to the extrapolation of the study to other settings. For making judgments about generalisability it is important to consider the characteristics of the sample from which the model was derived, the clear presentation of the results and finally the model should, ideally, be evaluated (validated) in a different sample from the original.

Taking into account these two domains, 18 questions were considered for each of the models included [see Additional file 2].

We restricted the quality assessment to the derivation studies.

Performance of models externally validated

We reported the performance of models that were validated in an external sample. We considered as externally validated those models that were reported by the authors as evaluated in a different cohort of TBI patients from the derivation set.

Data extraction

One reviewer (PP) extracted the information from each study for assessing the quality of reporting in each of the questions.

Results

A total of 3354 records were identified. After reading all the records 92 reports were identified and read in full. Thirty nine were excluded for the following reasons: 18 analyzed individual predictors but did not combine them in a single score, eight did not include in-hospital predictors, six included patients without traumatic brain injury, five were not original research (e.g. discussion, letter) and in two the objective was not to evaluate prognosis in TBI patients. (Figure 1)

The remaining 53 reports described 102 prognostic models [see Additional file 3].

General characteristics of the prognostic models [see Additional file 4]

Population included

Almost half (47%) of the models were derived from an adult population, 12% were derived from a child population while 21% were derived from a population that included both adults and children. In 21% of the models it was not clearly reported from which population they were derived.

In relation to the severity of the TBI studied, forty five models (44%) included all grades of severity, thirty one (30%) included severe TBI, nine (9%) moderate or severe TBI, nine (9%) mild TBI and in eight (8%) the severity of TBI was not clearly reported.

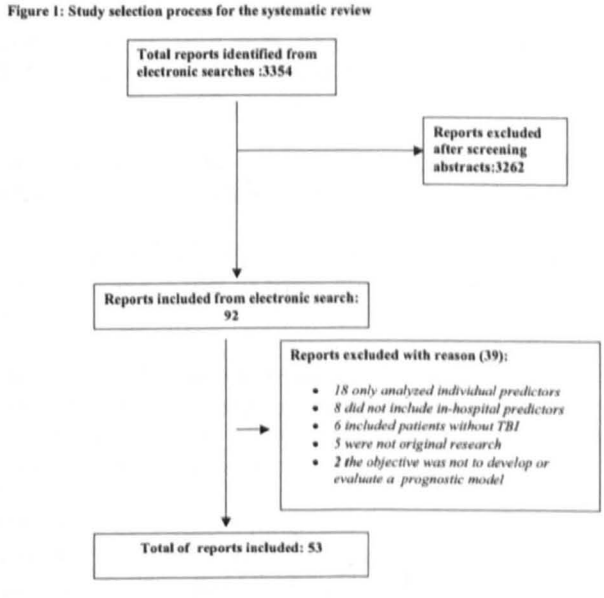


Figure 1
Study selection process for the systematic review of prognostic models in head injury.

A median of 319 patients (range 22–7764 patients) were included per model. Three quarters included less than 500 patients.

A total of ninety five models (93%) included populations from high income countries, five (5%) included populations from middle income countries and in two (2%) the population was from a low income country.

Objectives

Most of the models reported (65%), were derived for the first time (derivation models) while in 35% the models reported were validating pre-existing models (validation models). The majority of the validation models (29 out of 35) validated general trauma score. The remaining 6 models validated specific TBI prognostic score. One validation model was reported as a letter and information was limited. Three models validated prognostic scores that were developed before 1990.

Variables included as predictors

A total of 89 variables were included in the prognostic models. A mean of 5 variables were included in each model (range 2 to 13). GCS was the most common predictor included in the models,(50%) followed by age (46%) and pupil reactivity (26%). Overall clinical variables were included in 66% of the models, demographic variables were included as predictors in 50% of the models, CT scan predictors were used in 19% of the models and 7% included variables related to characteristics of the injury. In 7% of the models other predictors were included (e.g. other complementary tests or existing scores).

Outcomes

Mortality was the main outcome in 30% of the models and GOS in 28%. Other functional outcomes were reported in 31% of the models. The presence of a CT scan lesion was the main outcome in 7%, the need of neurosurgical intervention in 2% and raised intracranial pressure in 1%.

Analysis

In the multivariate analysis for the development of prognostic models (n:66) logistic regression was used in 31 (47%) models. Regression tree analysis was reported in 14 (21%) and neural networks in nine (13%). Other methods of analysis were performed in nine (14%) models while in one (2%) it was not clear and in three (5%) no multivariable analysis was performed.

Quality assessment (table 1)

We restricted the quality assessment to the 66 derivation models. Some of the quality assessment items could only be applied to logistic regression models.

Internal validity

In over half of the models loss to follow-up was not reported;15% reported an adequate loss to follow-up (less than 10%).

Most of the models (68%) did not include a discussion about the rationale to include the predictors in the model. A detailed description of the measurement of the predictors was absent in 82% of the models. In one third of the models the validity of the outcome measures was not reported.

In relation to the analysis of those that used multivariate logistic regression, stepwise was the most common approach (81%). Interactions were examined in 13% of the models. Predictor variables were analyzed as continuous in 19% of the models. A third (29%) of the models included at least 10 events per variable analyzed as predictor. The most common strategy to handle missing data was exclusion of observations (55%).

External Validity

The sample was described in almost all the models (83%). The procedure to obtain the score was explained in approximately half of the models (56%), however in those that used logistic regression only 19% included a user-friendly presentation.

In relation to the performance of the models, discrimination was reported in 58% of the models through the area under the receive operator curve (A.U.R.O.C.), 44% of which included the respective confidence interval. Calibration was reported with the Homer-Lemeshow test in 27% of the models. Almost half the models(56%) reported their overall accuracy.

Less than half of the models (38%) were validated, of which 11% were validated in an external population.

None of the models was evaluated prospectively in a randomized clinical trial to assess the effect in clinical practice.

Description of externally validated models

Seven models were developed and also reported an external validation (table 2). Two other models were validation of pre-existing models.

Pillai *et al.* developed a prognostic model to predict unfavourable outcome (death or vegetative state) at one month [23]. They developed the model in a cohort of 289 patients and validated the model in 26 patients from the same centre. The predictor variables were oculocephalic reflex, motor score of the GCS and midline shift score. In the validation set they reported sensitivity (75%), specifi-

Table 1: Quality assessment of prognostic models

INTERNAL VALIDITY	*All models N:66*	*Logistic regression N:31*	*Other analysis N:35*
STUDY			
Loss to follow-up			
< 10%	10 (15%)	5 (16%)	5 (14%)
>10%	19 (29%)	7 (23%)	12 (34%)
Not reported	37 (56%)	19(61%)	18 (52%)
VARIABLES			
Discussion about predictors			
Yes	21(32%)	11 (35%)	10(29%)
No	45(68%)	20 (65%)	25 (71%)
Description of measurement of predictors			
Yes	12 (18%)	8 (26%)	3 (9%)
No	54 (82%)	23(74%)	32 (91%)
Validity of outcome reported			
Yes	31 (47%)	14 (45%)	17(49%)
No	20 (30%)	7 (23%)	13 (37%)
Not applicable	15(23%)	10 (32%)	5 (14%)
Handling of missing data			
Estimated statistically	4 (6%)	4 (13%)	0
Excluded	36(55%)	16(52%)	20 (57%)
Not reported	26(39%)	11(35%)	15 (43%)
ANALYSIS			
Multivariable analysis Stepwise			
Backwards	-	12 (39%)	N/A
Forwards	-	3 (10%)	
Not specified	-	10 (32%)	
Not reported	-	5 (16%)	
Other	-	1 (3%)	
Interactions examined			
Yes	-	4 (13%)	
Not reported	-	27 (87%)	
Handling of predictors variables			
Continuous	-	6 (19%)	
Categorical	-	16 (52%)	
Not clear	-	9 (29%)	
More than 10 events per variable			
Yes	-	9 (29%)	
No	-	16 (52%)	
Not reported	-	6 (19%)	
EXTERNAL VALIDITY	All models	Logistic regression	Other analysis
Description of the sample			
Yes	55 (83%)	28 (90%)	27 (77%)
No	11(17%)	3 (10%)	8 (23%)
Presentation of the prognostic model			
Normogram	1 (1%)	1 (3%)	0
Simplified score	8 (12%)	4 (13%)	4 (11%)
Figure	13 (20%)	1 (3%)	12 (34%)
Regression formula	15 (23%)	12(39%)	3 (9%)
Not explained	29 (44%)	13(42%)	16 (46%)

EXTERNAL VALIDITY	*All models N:66*	*Logistic regression N:31*	*Other analysis N:35*
Performance reported A.U.C (Discrimination)			
Yes	-	18 (58%)	NA
No	-	13(42%)	
C.I. presented	-	8 out of 18 (44%)	
H-L (Calibration)			
Yes	-	7 (23%)	NA
No	-	23(74%)	
Other	-	1 (3%)	
Overall accuracy			
Yes	37 (56%)	15 (48%)	22 (63%)
No	29 (44%)	16 (52%)	13 (17%)
Validation			
Yes	25 (38%)	17 (55%)	8 (23%)
External	*7 (11%)*	*7 (23%)*	*0*
No	41 (62%)	14 (45%)	27 (77%)

Table 2: Characteristics of models externally validated

Author	Derivation sample	Validation sample	Predictors	Outcomes	Performance in the validation sample	Presentation of a simplified score
Pillai et al.	208 patients from India with severe TBI	26 patients from the same centre	1-oculocephalic reflex 2-motor GCS 3-midline shift	Death or vegetative state	Sensitivity (75%) Specificity (67%) PPV 50%	No
Signorini et al	372 patients from Scotland with moderate and severe TBI	520 patients from the same centre	1-GCS 2-ISS 3-pupils reactivity and 4-haematoma (CT scan)	Survival at 1 year	A.U.R.O.C (0.835) Error rate (15.2%) Brier score (0.1160) Hosmer-Lemeshow ($p < 0.001$)	Nomogram
Signorini et al	110 patients from Scotland with moderate and severe TBI	140 patients from the same centre	1-GCS 2-ISS 3-pupils reactivity and 4-haematoma (CT scan) 5-ICP measures	Survival at 1 year	Not reported	No
Hukkelhoven et al.	134 patients from Netherlands with moderate and severe TBI	180 patients from the same centre	1-age 2-motor GCS 3-pupils reactivity 4-pupillary size 5-hypotension 6-ISS	Raised ICP	A.U.R.O.C. (0.50) Hosmer-Lemeshow ($p = 0.18$)	No
Hukkelhoven et al.	275 patients from Netherlands with moderate and severe TBI	250 patients from the same centre	1-age 2-cause of injury 3-pupils reactivity 4-pupillary size 5-hypotension 6-ISS	Surgical removable lesions	A.U.R.O.C. (0.67) Hosmer-Lemeshow ($p = 0.01$)	No
Hukkelhoven et al.	2269 patients from 2 trials in high income countries with moderate and severe TBI	796 patients from Europe	1-age 2-motorGCS 3-pupils reactivity 4-hypoxia 5-hypotension 6-CT classification 7-subarachnoid haemorrhage	Death or disability at 6 months	A.U.R.O.C. (0.83) Hosmer-Lemeshow ($p = 0.05$)	Score chart
Hukkelhoven et al.	2269 patients from 2 trials in high income countries with moderate and severe TBI	796 patients from Europe and 746 from the United States	1-age 2-motor GCS 3-pupils reactivity 4-hypoxia 5-hypotension 6-CT classification 7-subarachnoid haemorrhage	Death at 6 months	A.U.R.O.C. (0.87/0.89) Hosmer-Lemeshow ($p = 0.42/<0.001$)	Score chart

city (67%), predictive value of unfavourable outcome (50%), predictive value of favourable outcome (86%), percentage of false optimistic results (25%), and percentage of false pessimistic results (33%). They did not report the model's performance measured in the derivation set. Confidence intervals of the estimates were not reported. Although the authors reported how to calculate the prediction score, they did not present it in a user-friendly fashion.

Signorini *et al.* developed two prognostic models, for one they used only clinical variables and for the other they added variables on secondary insults. [24,25] In both models the outcome was survival at 1 year. The first model was validated in 520 patients who attended the same centre. The predictors were age, GCS, ISS, pupils reactivity and presence of haematoma on the CT scan. They reported measures of discrimination: A.U.R.O.C. (0.835), error rate (15.2%) and calibration: brier score (0.1160), Hosmer-Lemeshow ($p < 0.0001$). They included a graph with the 95% confidence interval of the calibration of the model. The second model was validated in 140 patients who attended the same centre. The predictor variables were the same as the first model plus ICP measures. Although they mentioned that brier score, error rate, A.U.R.O.C were higher than the original dataset they did not report the actual estimates. They reported a normogram to predict probability of survival that is user-friendly for physicians.

Hukkelhoven *et al.* reported four different models [26,27]. The outcomes were: raised intracranial pressure (ICP), surgically removable lesions (SRL), unfavourable outcome (death, vegetative state or severe disability) and mortality at six months. For the validation of the first two outcomes they use an historical (previous) sample of 205 patients from the same centre. The predictors for ICP were age, motor score, pupil size, pupillary reactivity, hypotension and ISS. For SRL the predictors were the same except for motor score which was not, and cause of injury that was added. For unfavourable outcome they used one database and for mortality two databases, none of these databases were related with the population of the derivation set. The predictor variables were age, gender, cause of injury, pupil reactivity, hypotension, hypoxia, CT classification and traumatic subarchnoid haemorrhage They reported the models discrimination: A.U.R.O.C. of 0.50 (95% CI 0.41–0.58), 0.67 (95% CI 0.60–0.75) and 0.83 (95% CI 0.80–0.86) for ICP, SRL, unfavourable outcome and mortality respectively. They also reported the model calibration: Hosmer-Lemeshow goodness of fit test of 0.18, 0.01,0.05 and 0.42 (<0.001), for ICP, SRL, unfavourable outcome and mortality respectively (the calibration of the mortality model was validated in two different databases). They presented the model as a score chart to facilitate its use in clinical practice.

Bush *et al.* validated a model previously developed by the same group [28]. Their model was intended to allow better understanding of factors influencing functional outcomes and was not intended to predict individual outcomes. It was not clearly reported whether the patients came from the same original population. They used path analysis to evaluate the predictors (functional status, injury severity and cognitive status) on functional outcomes (disability rating scale, community integration questionnaire and return to employment). The reported difference indexes of goodness of fit that showed that the originally model fitted better than the validation model. They did not report any discrimination measures.

Benzer *et al.* validated a model that used an existing scale, although they did not provide details of when and how it was developed [29]. They did not use any kind of multivariable analysis. The used a score based in the following variables : reaction to acoustic stimuli, reaction to pain, body posture, eye opening, pupil size, pupil response to light, position and movements of eyeballs and oral automatisms to predict mortality at 21 days. They did not report any performance measure, but just the chi square test for survival of those with low versus high score. They presented the score in a user-friendly way.

Discussion

This systematic review shows that although publications of prognostic models for TBI patients are very frequent their quality is relatively poor. In addition they are rarely validated on external populations or presented to physicians in a friendly way. Furthermore, only a few are developed using populations from low and middle income countries where most trauma occurs.

Patients from all severity spectra were investigated but prognostic models for moderate and severe TBI patients were more frequent. It is noteworthy that only 2% of the models included patients from low income countries taking into account that 90% of trauma occur in these countries. Although biologically prognostic factors should be the same worldwide, is reasonable to consider that the strength of the association could differ depending on the medical care received. This difference could affect the accuracy of the prognostic models in different settings. Although there is no data about this, an ongoing project analysing the MRC CRASH Trial Cohort is exploring this issue.

GCS, age and pupil reactivity were the most common variables analyzed as predictors whereas, G.O.S. and mortality were the most common outcomes investigated.

Multiple logistic regression was the multivariable analysis most frequently used.

We found several limitations in the quality of the models. The majority did not include a thorough discussion of the rationale for including the predictor variables. Only a minority had a loss to follow-up of less than 10%. This is potentially an important limitation as the loss to follow-up could be related to prognosis and this could lead to biased results. Furthermore only four models handle appropriately the missing data with statistical imputation. In relation to the multivariable analysis, automatic procedures (stepwise) were quite common in logistic regression. There is no agreement in relation to the appropriateness of this strategy. This is shown, for example, in conflict recommendations in quality assessment for prognostic studies; while in one study the use of stepwise was considered as good quality in other it was considered as an indicator of a fatal flaw [30]. One of the limitations we found was that most of the studies did not explicitly consider clinical criteria to enter the variables in the model beyond the automatic procedures. Interactions were hardly ever explored although this is strongly recommended in multivariable analysis [31]. Another common weakness in the logistic regression models was the lack of power of the models, as only one third included at least 10 events per variable. It has been proposed that this is the minimum ratio of events to variables which is large enough to allow an adequate precision of the estimates [31].

We did not attempt to obtain an overall quality assessment and instead we evaluated its different components, this approach makes a cross comparison between different analytical strategies difficult because, for example many of the criteria only apply to logistic regression analysis.

It is also important to report how well the model works and for this performance measures should be reported. Remarkably only two thirds reported a measure of discrimination and only one fifth a measure of calibration. This is of particular concern considering that calibration is the most important performance measure for the application of the models in clinical practice [32]. Even when a discrimination measure was reported, less than half presented confidence intervals to provide readers an estimation of the precision.

For a model to be generalizable to other populations it is very important to conduct an external validation [32]. Only seven models (three reports) developed and validated a model but in only one of them the validation was performed on patients of a different centre. Those models that considered mortality as an outcome found A.U.R.O.C that were higher than 0.70 which is considered as excellent discrimination. However the discrimination for the other outcomes were not as good. Furthermore, the calibration measures were low in all the outcomes considered.

Finally, to be useful, the method to estimate prognosis should be clearly reported and, to be clinically practical, they should be user-friendly. In only half of the models was it clearly explained how to obtain the prognostic score and in only one tenth was it reported in such a way that could be easily applicable in a clinical setting.

From all of the models found in our systematic review we consider that those developed by Hukkelhoven et al. and Signorini et al. are the most clinical useful for patients from high income countries with moderate and severe TBI, as they fulfilled the majority of the methodological requirements and showed an acceptable performance in the external validation, furthermore they are available in a user-friendly way [27,25].

We acknowledge some limitations in our study. Firstly, we only included studies that explicitly combined at least two predictors, in doing so we could have missed some reports that used multivariable to analyze individual predictors and did not report in the abstract the overall estimation although they included the estimate in the full report. Secondly, we did not include studies that assessed clinical predictor rules for which although they considered more than one variable they did not combine them. We considered that the methodological framework to assess such studies is fundamentally different from prognostic models. Thirdly, we restricted our search to 1990 onwards so, we could have missed some relevant prognostic models published prior to that date. However because of changes in management and diagnostic technology in recent years we doubt that prognostic models previous to 1990 could be useful for the current medical care of TBI patients. Finally, another limitation of this paper is that we did not describe the time of prediction assessment of the prognostic models. Although we acknowledge that this information can be clinically very useful unfortunately this data was seldom available in the reports.

To our knowledge there has been only one previous systematic review of prognostic models in TBI [17]. They found 10 reports, all of which were identified in our systematic review. They validated four of these reports (6 models) in four series of patients. Discrimination(A.U.R.O.C.) in the validation series ranged from 0.70 to 0.80. On the other hand calibration was poor. They concluded that large sample sizes and refitting of the original model coefficients are related with a better performance of the models. Unlike ours this systematic review was

restricted to models that use baseline characteristics to predict mortality or unfavourable outcome (defined by G.O.S.) in moderate and severe TBI patients. Furthermore the search strategy was not specified.

Systematic reviews of prognostic models for other diseases have found similar results to ours. For example Counsel *et al.* conducted a systematic review of prognostic models in patients with acute stroke [14]. They found 83 prognostic models but they concluded that none of them has been sufficiently well developed and validated.

Conclusion

This systematic review describes the limitations of published prognostic models in TBI and most importantly inform researchers who are involved in the development of prognostic models in TBI. Future studies should consider the following issues to develop valid prognostic models: thorough discussion with physicians of potential predictors that are "clinically relevant", clear description of the measurement and validity of variables included in the model, large sample size to ensure precise estimates, adequate handling of continuous variables and missing data, assessment of interaction in the multivariable analysis, clear description of the calculation of the prognostic score, external validation and adequate report of model performance measures, such that physicians can interpret their accuracy. It should also be encouraged that more studies include population from low and middle income countries where most of the burden of TBI occurs. Finally, for prognostic models to be clinical useful they should be presented in user-friendly way to be easily applied in the clinical scenario.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

PP, IR and PE developed the protocol and conducted the SR. RW conducted the searching. PP conducted the analysis. All the authors read and approved the final manuscript.

Additional material

Additional File 1

Electronic bibliographical databases and search strategies. This table describe the databases and search strategies used in the systematic review.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1472-6947-6-38-S1.doc]

Additional File 2

Quality assessment of prognostic models. This document describe the different items analysed for the quality assessment.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1472-6947-6-38-S2.doc]

Additional File 3

Studies included in the systematic review. List of references of included studies in the systematic review
Click here for file
[http://www.biomedcentral.com/content/supplementary/1472-6947-6-38-S3.doc]

Additional File 4

General Characteristics of the models. This table describe the characteristics of the included models in the systematic review
Click here for file
[http://www.biomedcentral.com/content/supplementary/1472-6947-6-38-S4.pdf]

Acknowledgements

PP and RW were funded by the Cochrane Injuries Group. IR and PE were funded by the London School of Hygiene & Tropical Medicine.

References

1. Bruns J Jr, Hauser WA: **The epidemiology of traumatic brain injury: a review.** *Epilepsia* 2003, 44(Suppl 10):2-10.
2. Fleminger S, Ponsford J: **Long term outcome after traumatic brain injury.** *BMJ* 2005, 331(7530):1419-20.
3. Perel PWJ, Ravi R, Shakur H, *et al.*: **Prognosis following head injury: a survey of doctors from developing and developed countries.** *JECP* in press.
4. Lee KL, Pryor DB, Harrell FE Jr, *et al.*: **Predicting outcome in coronary disease. Statistical models versus expert clinicians.** *Am J Med* 1986, 80(4):553-60.
5. Murray LS, Teasdale GM, Murray GD, *et al.*: **Does prediction of outcome alter patient management?** *Lancet* 1993, 341(8859):1487-91.
6. Roberts I, Shakur H, Edwards P, *et al.*: **Trauma care research and the war on uncertainty.** *BMJ* 2005, 331(7525):1094-6.
7. Coats TJ, Shakur H: **Consent in emergency research: new regulations.** *Emerg Med J* 2005, 22(10):683-5.
8. Dickinson K, Bunn F, Wentz R, *et al.*: **Size and quality of randomised controlled trials in head injury: review of published studies.** *BMJ* 2000, 320(7245):1308-11.
9. Murray GD, Barer D, Choi S, *et al.*: **Design and analysis of phase III trials with ordered outcome scales: the concept of the sliding dichotomy.** *J Neurotrauma* 2005, 22(5):511-7.
10. Young FB, Lees KR, Weir CJ: **Improving trial power through use of prognosis-adjusted end points.** *Stroke* 2005, 36(3):597-601.
11. Altman DG: **Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables.** *BMJ* 2001, 323(7306):224-228.
12. Lecky F, Woodford M, Yates DW: **Trends in trauma care in England and Wales 1989–97. UK Trauma Audit and Research Network.** *Lancet* 2000, 355(9217):1771-5.
13. Altman DG, Royston P: **What do we mean by validating a prognostic model?** *Stat Med* 2000, 19(4):453-73.
14. Counsell C, Dennis M: **Systematic review of prognostic models in patients with acute stroke.** *Cerebrovasc Dis* 2001, 12(3):159-70.
15. Jacob M, Lewsey JD, Sharpin C, *et al.*: **Systematic review and validation of prognostic models in liver transplantation.** *Liver Transpl* 2005, 11(7):814-25.

16. Brain Trauma Foundation (U.S.) AAoNS: *Management and prognosis of severe traumatic brain injury. New York* 2000.
17. Hukkelhoven CW, Rampen AJ, Maas AI, *et al.*: **Some prognostic models for traumatic brain injury were not valid.** *J Clin Epidemiol* 2006, **59(2)**:132-43.
18. Juni P, Altman DG, Egger M: **Systematic reviews in health care: Assessing the quality of controlled clinical trials.** *BMJ* 2001, **323(7303)**:42-6.
19. Whiting P, Harbord R, Kleijnen J: **No role for quality scores in systematic reviews of diagnostic accuracy studies.** *BMC Med Res Methodol* 2005, **5**:19.
20. Laupacis A, Sekar N, Stiell IG: **Clinical prediction rules. A review and suggested modifications of methodological standards.** *JAMA* 1997, **277(6)**:488-94.
21. Stiell IG, Wells GA: **Methodologic standards for the development of clinical decision rules in emergency medicine.** *Ann Emerg Med* 1999, **33(4)**:437-47.
22. Wasson JH, Sox HC, Neff RK, *et al.*: **Clinical prediction rules. Applications and methodological standards.** *N Engl J Med* 1985, **313(13)**:793-9.
23. Pillai SV, Kolluri VR, Praharaj SS: **Outcome prediction model for severe diffuse brain injuries: development and evaluation.** *Neurol India* 2003, **51(3)**:345-9.
24. Signorini DF, Andrews PJ, Jones PA, *et al.*: **Predicting survival using simple clinical variables: a case study in traumatic brain injury.** *J Neurol Neurosurg Psychiatry* 1999, **66(1)**:20-5.
25. Signorini DF, Andrews PJ, Jones PA, *et al.*: **Adding insult to injury: the prognostic value of early secondary insults for survival after traumatic brain injury.** *J Neurol Neurosurg Psychiatry* 1999, **66(1)**:26-31.
26. Hukkelhoven CW, Steyerberg EW, Habbema JD, *et al.*: **Admission of patients with severe and moderate traumatic brain injury to specialized ICU facilities: a search for triage criteria.** *Intensive Care Med* 2005, **31(6)**:799-806.
27. Hukkelhoven CW, Steyerberg EW, Habbema JD, *et al.*: **Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics.** *J Neurotrauma* 2005, **22(10)**:1025-39.
28. Bush BA, Novack TA, Malec JF, *et al.*: **Validation of a model for evaluating outcome after traumatic brain injury.** *Arch Phys Med Rehabil* 2003, **84(12)**:1803-7.
29. Benzer A, Mitterschiffthaler G, Marosi M, *et al.*: **Prediction of non-survival after trauma: Innsbruck Coma Scale.** *Lancet* 1991, **338(8773)**:977-8.
30. Hayden JA, Cote P, Bombardier C: **Evaluation of the quality of prognosis studies in systematic reviews.** *Ann Intern Med* 2006, **144(6)**:427-37.
31. Concato J, Feinstein AR, Holford TR: **The Risk of Determining Risk with Multivariable Models.** *Ann Intern Med* 1993, **118(3)**:201-210.
32. Justice AC, Covinsky KE, Berlin JA: **Assessing the generalizability of prognostic information.** *Ann Intern Med* 1999, **130(6)**:515-24.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1472-6947/6/38/prepub

BMJ

Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients

MRC CRASH Trial Collaborators

BMJ published online 12 Feb 2008;
doi:10.1136/bmj.39461.643438.25

BMJ

RESEARCH

Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients

MRC CRASH Trial Collaborators

London School of Hygiene and Tropical Medicine, London WC1B 3DP

Correspondence to: P A Perel
Pablo.perel@lshtm.ac.uk

doi:10.1136/bmj.39461.643438.25

ABSTRACT

Objective To develop and validate practical prognostic models for death at 14 days and for death or severe disability six months after traumatic brain injury.

Design Multivariable logistic regression to select variables that were independently associated with two patient outcomes. Two models designed: "basic" model (demographic and clinical variables only) and "CT" model (basic model plus results of computed tomography). The models were subsequently developed for high and low-middle income countries separately.

Setting Medical Research Council (MRC) CRASH Trial.

Subjects 10 008 patients with traumatic brain injury. Models externally validated in a cohort of 8509.

Results The basic model included four predictors: age, Glasgow coma scale, pupil reactivity, and the presence of major extracranial injury. The CT model also included the presence of petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift, and non-evacuated haematoma. In the derivation sample the models showed excellent discrimination (C statistic above 0.80). The models showed good calibration graphically. The Hosmer-Lemeshow test also indicated good calibration, except for the CT model in low-middle income countries. External validation for unfavourable outcome at six months in high income countries showed that basic and CT models had good discrimination (C statistic 0.77 for both models) but poorer calibration.

Conclusion Simple prognostic models can be used to obtain valid predictions of relevant outcomes in patients with traumatic brain injury.

INTRODUCTION

Traumatic brain injury is a leading cause of death and disability worldwide. Every year, about 1.5 million affected people die and several millions receive emergency treatment.[1 2] Most of the burden (90%) is in low and middle income countries.[3]

Clinicians treating patients often make therapeutic decisions based on their assessment of prognosis. According to a 2005 survey, 80% of doctors believed that an accurate assessment of prognosis was important when they made decisions about the use of specific methods of treatment such as hyperventilation, barbiturates, or mannitol.[4] A similar proportion considered that this was important in deciding whether or not to withdraw treatment. Assessment of prognosis was also deemed important for counselling patients and relatives. Only a third of doctors, however, thought that they accurately assessed prognosis.[4]

Prognostic models are statistical models that combine data from patients to predict outcome and are likely to be more accurate than simple clinical predictions.[5] The use of computer based prediction of outcome in patients with traumatic brain injury increases the use of certain therapeutic interventions in those predicted to have a good outcome and reduces their use in those predicted to have a poor outcome.[6]

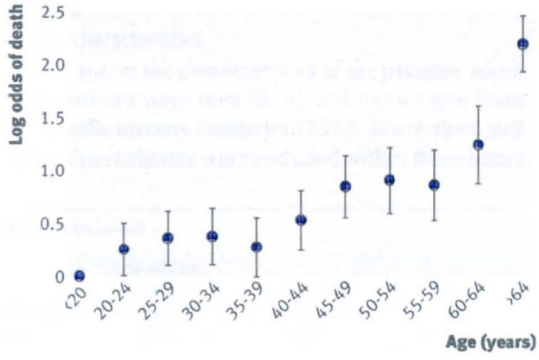
Many prognostic models have been reported but none are widely used. A recent systematic review offers possible explanations.[7] Most models were developed on small samples, most were methodologically flawed, and few were validated in external populations. Few were presented in a clinically practical way, nor were they developed in populations from low and middle income countries, where most trauma occurs.

The Medical Research Council (MRC) CRASH (corticosteroid randomisation after significant head injury) trial is the largest clinical trial conducted in patients with traumatic brain injury and presents a unique opportunity to develop a prognostic model.[8 9] The trial prospectively included patients within eight hours of the injury, used standardised definitions of variables, and achieved almost complete follow-up at six months. Furthermore, the large sample size guarantees precise and valid predictions. The high recruitment of patients from low and middle income countries means that models developed with these data are relevant to these settings.

We have developed and validated prognostic models for death at 14 days and death and disability at six months in patients with traumatic brain injury.

METHODS

Patients—The study cohort was all 10 008 patients enrolled in the trial. Adults with traumatic brain injury, who had a score on the Glasgow coma scale of 14 or



	<20	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	>64
No of patients	1232	1690	1294	1067	960	799	692	561	436	325	902
No of deaths	151	252	217	186	159	149	158	121	99	89	367

Fig 1 | Relation between age and mortality at 14 days

less, and who were within eight hours of injury, were eligible for inclusion in the trial.

Outcomes—Death of a patient was recorded on an early outcome form that was completed at hospital discharge, death, or 14 days after randomisation (whichever occurred first). Unfavourable outcome (death or severe disability) at six months was defined with the Glasgow outcome scale (see box). The scale comprises five categories: death, vegetative state, severe disability, moderate disability, and good recovery. For the purpose of this analysis, we dichotomised outcomes into favourable (moderate disability or good recovery) and unfavourable (dead, vegetative state, or severe disability).[10]

Prognostic variables—For the prognostic model we considered age, sex, cause of injury, time from injury to randomisation, Glasgow coma score at randomisation, pupil reactivity, results of computed tomography, whether the patient had sustained a major extracranial injury, and level of income in country (high or low-middle income countries, as defined by the World Bank) (see table A on bmj.com).[11] We adjusted analyses for treatment within the trial as this was related to outcome, and we did not find interaction between treatment and the potential predictors.[8,9]

Analysis—Most of the variables collected in the CRASH trial have been previously associated with prognosis in traumatic brain injury, so we included all of them in a first multivariable logistic regression analysis.[12] We excluded variables that were not significant at 5% level. We quantified each variable's predictive contribution by its z score (the model coefficient divided by its standard error). We explored linearity between age and mortality at 14 days and Glasgow coma score and mortality at 14 days. Interactions between country income level and all the other predictors were evaluated with a likelihood ratio test. Because there were few data missing, we performed a complete case analysis.

Prognostic models—We developed different models for each of the two outcomes: a basic model, which included only clinical and demographic variables, and a CT model, which also included results of computed tomography.

Performance of the model—We assessed performance of the models in terms of calibration and discrimination. Calibration was assessed graphically and with the Hosmer-Lemeshow test. Discrimination was assessed with the C statistic (an equivalent concept to area under the receiver operator characteristic curve).[13]

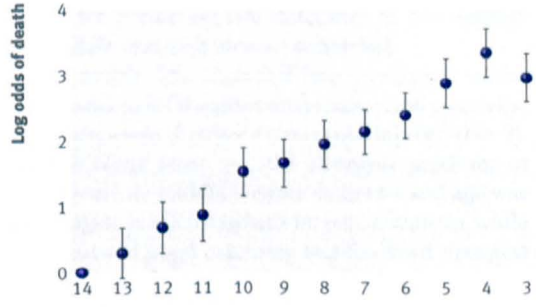
Internal validation—The internal validity of the final model was assessed by the bootstrap re-sampling technique. Regression models were estimated in 100 models. For each of 100 bootstrap samples we refitted and tested the model on the original sample to obtain an estimate of predictive accuracy corrected for bias. This showed no overoptimism in any of the final model's predictive C statistics.

External validation—A good prognostic model should be generalisable to populations different to those in which it was derived.[14] We externally validated the models in an external cohort of 8509 patients with moderate and severe traumatic brain injury from 11 studies conducted in high income countries (the IMPACT (international mission for prognosis and clinical trial) dataset).[15]

Score development—We developed a clinical score based on regression coefficients. A web based version of the model was developed to be accessible to clinicians internationally.

Category and definition on Glasgow outcome scale

- Good recovery: able to return to work or school
- Moderate disability: able to live independently; unable to return to work or school
- Severe disability: able to follow commands/unable to live independently
- Persistent vegetative state: unable to interact with environment; unresponsive
- Dead



	14	13	12	11	10	9	8	7	6	5	4	3
No of patients	1499	1498	962	748	690	633	709	946	677	478	455	667
No of deaths	51	76	65	62	108	113	152	241	231	232	275	342

Fig 2 | Relation between Glasgow coma scale and mortality at 14 days

RESULTS

General characteristics

Table 1 shows the characteristics of the patients. More of the patients were men (81%) and more came from low-middle income countries (75%). More than half (58%) of participants were included within three hours of injury. Road traffic crashes were the most common cause of injury (65%) and 79% of the participants underwent computed tomography. A total of 1948 patients (19%) died in the first two weeks, 2323 patients (24%) were dead at six months, and 3556 patients (37%) were dead or severely dependent at six months.

The relation between age and the log odds of death within 14 days showed no association until the age of 40 and a linear increase afterwards. The relation between Glasgow coma score and mortality at 14 days was reasonably linear and we therefore included the coma score as a continuous variable (figs 1 and 2). The relation with unfavourable outcome at six months showed similar patterns.

Table 1 | General characteristics of study population

	Total (n=10 008)	Low-middle income countries (n=7 526)	High income countries (n=2 482)	P value*
Age (years):				
<20	12.3	12.5	11.8	
20-24	17.0	17.8	14.4	
25-29	13.0	13.5	11.2	
30-34	10.7	10.9	10.1	
35-44	17.9	18.5	15.9	
45-54	12.5	12.3	13.3	
≥55	16.7	14.5	23.4	
Mean (SD)	37 (17.1)	35.8 (16)	40.6 (19.4)	<0.001
Sex:				
Female	19.0	18.3	21.1	0.002
Male	81.0	81.7	78.9	
Hours since injury:				
<1	26.8	24.0	35.2	
1-3	31.0	30.1	33.7	
>3	42.3	45.9	31.1	
Mean (SD)	3.4 (2.7)	3.6 (2.8)	2.8 (2.0)	<0.001
Cause of head injury:				
Road traffic crash	65.1	69.9	50.2	<0.001
Fall >2 meters	13.3	11.1	20.0	
Other	21.7	19.0	29.8	
Total Glasgow coma score:				
Mild (13-14)	30.2	29.4	32.6	<0.001
Moderate (9-12)	30.4	32.6	23.6	
Severe (3-8)	39.5	38.0	43.8	
Pupil reactivity:				
Both reactive	82.8	83.5	80.7	<0.001
One reactive	6.3	6.2	6.3	
None reactive	8.2	8.0	9.1	
Unable to assess	2.7	2.3	3.9	
Major extracranial injury:				
No	77.3	77.3	77.5	0.801
Yes	22.7	22.7	22.5	
Computed tomography:				
No scan	21.1	24.0	12.0	<0.001
Normal scan	22.8	20.0	30.2	<0.001
Petechial haemorrhages	28.7	28.7	28.7	0.970
Obliteration of 3rd ventricle or basal cisterns	23.4	28.6	9.6	<0.001
Subarachnoid bleed	31.6	33.5	26.4	<0.001
Midline shift	14.6	15.9	11.1	<0.001
Non-evacuated haematoma	27.1	27.3	26.5	0.475
Evacuated haematoma	12.7	14.4	7.9	<0.001
Outcomes:				
Mortality at 14 days	19.5	20.7	16.0	<0.001
Death or severe disability at 6 months	37.2	36.8	38.5	0.150

*P value for comparison between low-middle income countries and high income countries.

Low-middle *v* high income countries

In comparison with patients from high income countries, those from low-middle income countries were younger, more likely to be male, were recruited later, had less severe traumatic brain injury (as defined by Glasgow coma score and pupil reactivity), and more often had abnormal results on computed tomography. Road traffic crashes were a more common cause of traumatic brain injury. Although patients from low-middle income countries experienced higher mortality at 14 days (odds ratio 1.94, 95% confidence interval 1.64 to 2.30), there was no significant difference in unfavourable outcome at six months.

There were significant interactions between the country's income level and several predictors and so we developed two models, one for low-middle income countries and another for high income countries. Older age was a stronger predictor of 14 day mortality in high income countries (interaction $P<0.001$), and lower Glasgow coma score was a stronger predictor in low-middle income countries (interaction $P=0.003$). Obliteration of the third ventricle and a non-evacuated haematoma were both associated with a higher risk in high income countries (interaction $P<0.001$ and $P=0.03$, respectively).

Multivariable predictive models

We developed eight models altogether: basic and CT models for predicting two outcomes in two settings (low-middle and high income countries).

Basic models—We included four predictors in the basic model: age, Glasgow coma score, pupil reactivity, and the presence of major extracranial injury (table 2). Glasgow coma score was the strongest predictor of outcome in low-middle income countries and age was the strongest predictor in high income countries, while the absence of pupil reactivity was the third strongest predictor in both regions.

CT models—The following characteristics on computed tomography were strongly associated with the outcomes in addition to the predictors included in the basic models: presence of petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift, and non-evacuated haematoma (table 3). Obliteration of the third ventricle and midline shift were the strongest

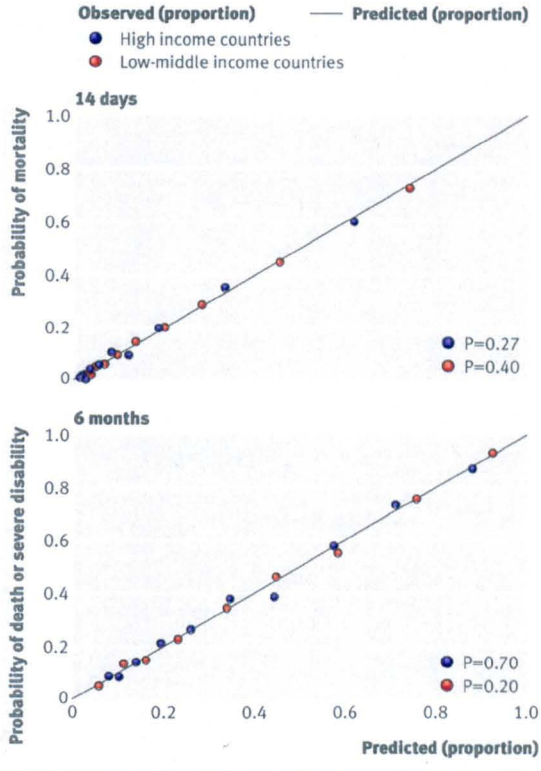


Fig 3 | Calibration of basic models using expected and observed probabilities of mortality at 14 days (top) and death or severe disability at six months (bottom) in patient with traumatic brain injury according to income level of country. P value is for Hosmer-Lemeshow test

predictors of mortality at 14 days, and non-evacuated haematoma was the strongest predictor of unfavourable outcome at six months.

Performance of models—All models showed excellent discrimination, with C statistics over 0.80 (tables 2 and 3). Calibration in all models was adequate and six out of the eight models had good calibration when evaluated with the Hosmer-Lemeshow test (figs 3 and 4).

Clinical score—Individual scores and their respective probability of outcome can be obtained from our web based calculator (www.crash2.lshtm.ac.uk/). By entering the values of the predictors, we can obtain the expected risk of death at 14 days and of death or severe disability at six months. Figure 5 shows a sample screenshot of the predictions for a 26 year old patient from a low and middle income country (Argentina), with a Glasgow coma score of 11, one pupil reactive, and absence of a major extra cranial injury. According to the basic model this patient has a probability of death at 14 days of 10% and a 23.9% risk of death or severe disability at six months. A good agreement is evident between observed and predicted outcome by the web calculator (figs 3 and 4).

External validation—Because an external cohort of patients from low-middle income countries was not available, we validated the models in patients from high income countries only. The IMPACT dataset used for the validation did not include mortality at 14 days and so we could validate only models for unfavourable outcome at six months. We validated the basic model with the variables age, Glasgow coma score, and pupil reactivity. We did not include the variable "major extracranial injury" as it was not available in the validation sample. For the CT models, we added obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, midline shift, and non-evacuated haematoma to the basic model. Similarly, we excluded the variable "petechial haemorrhages" as this was not available in the validation

Table 2 | Multivariable basic predictive models (excluding data from computed tomography*). Figures are odds ratios (95% confidence intervals) with z scores

	Mortality at 14 days		Death or severe disability at 6 months	
Prognostic variables	High income countries (n=2294)	Low-middle income countries (n=7412)	High income countries (n=2185)	Low-middle income countries (n=7119)
Age†	1.72 (1.62 to 1.83), 14.08	1.47 (1.40 to 1.54), 14.10	1.73 (1.64 to 1.82), 15.99	1.70 (1.63 to 1.77), 18.58
GCS‡	1.24 (1.19 to 1.29), 10.22	1.39 (1.35 to 1.42), 25.60	1.22 (1.18 to 1.25), 12.84	1.42 (1.39 to 1.45), 30.64
Pupil reactivity:				
Both	1	1	1	1.00
One	2.57 (1.65 to 4.00), 4.17	1.91 (1.53 to 2.39), 5.69	2.43 (1.62 to 3.66), 4.26	2.01 (1.59 to 2.56), 5.81
None	5.49 (3.70 to 8.15), 8.45	3.92 (3.14 to 4.90), 12.07	3.28 (2.20 to 4.89), 5.85	4.54 (3.38 to 6.11), 10.03
Major extracranial injury:				
No	1	1	1	1.00
Yes	1.53 (1.11 to 2.09), 2.62	1.15 (0.99 to 1.34), 1.78	1.62 (1.26 to 2.07), 3.82	1.73 (1.51 to 1.99), 7.76
C statistic	0.86	0.84	0.81	0.84

GCS=Glasgow coma scale.
*Includes age, GCS, sex, hours since injury, cause of injury, pupil reactivity, and presence of major extracranial injury.
†Per 10 year increase after 40 years.
‡Per decrease of each value of GCS.

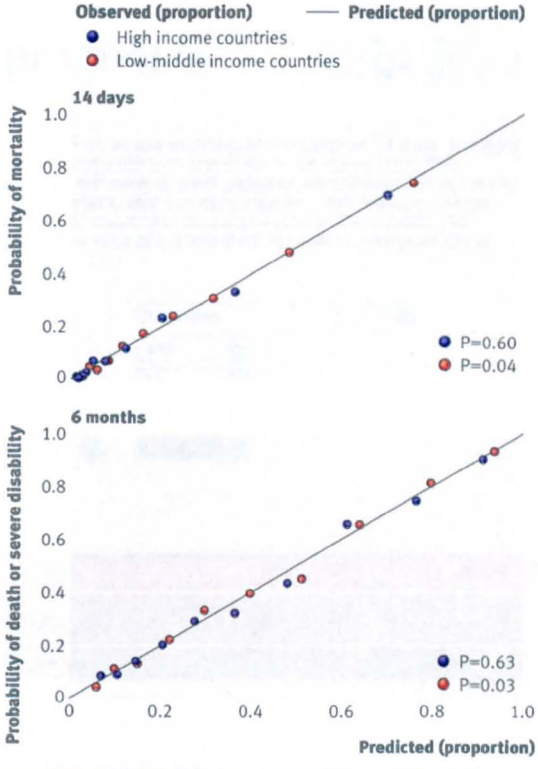


Fig 4 | Calibration of computed tomography models using expected and observed probabilities of mortality at 14 days (top) and death or severe disability at six months (bottom) in patient with traumatic brain injury according to income level of country. P value is for Hosmer-Lemeshow test

sample. For the validation process we first ran these models in the CRASH trial cohort and we then applied the corresponding coefficients in the validation sample. Although discrimination was, as expected, lower than in the original data, it was still quite good for both the basic and CT models (C statistic 0.77 for both models). The calibration was excellent for the CT model but poorer for the basic model (figs 6 and 7).

DISCUSSION

We have developed web based prognostic models for predicting two clinically relevant outcomes in patients with traumatic brain injury using variables that are available at the bedside. The models have excellent discrimination and good fit with both internal and external validation. We have reported on differences in outcomes and on the strength of predictors of outcomes, according to whether patients are from high or low-middle income countries.

Older age, low Glasgow coma score, absent pupil reactivity, and the presence of major extracranial injury predict poor prognosis. All of these variables have been previously identified as prognostic factors for poor outcome in traumatic brain injury.[12] Glasgow coma score showed a clear linear relation with mortality. Our finding that mortality in patients with Glasgow coma score of 3 was lower than in patients with a score of 4 may be because scores of sedated patients are reported as 3. Increasing age was associated with worse outcomes but this association was apparent only after age 40. A similar threshold has been reported elsewhere.[16,17] Plausible explanations for this include extracranial comorbidities, changes in brain plasticity, or differences in clinical management associated with increasing age. The presence of "obliteration of third ventricle or basal cisterns" on

Table 3 | Multivariable predictive models with computed tomography*. Figures are odds ratios (95% confidence intervals) with z scores

	Mortality at 14 days		Death or severe disability at 6 months	
Prognostic variables	High income countries (n=2030)	Low-middle income countries (n=5635)	High income countries (n=1955)	Low-middle income countries (n=5 394)
Age†	1.73 (1.62 to 1.84), 13.33	1.46 (1.39 to 1.54), 12.54	1.73 (1.63 to 1.83), 14.94	1.72 (1.64 to 1.81), 17.74
GCS‡	1.18 (1.12 to 1.23), 6.87	1.27 (1.24 to 1.31), 16.68	1.18 (1.14 to 1.22), 9.83	1.34 (1.30 to 1.37), 22.32
Pupil reactivity:				
Both	1	1	1	1.00
One	2.00 (1.25 to 3.20), 2.88	1.45 (1.14 to 1.86), 2.97	2.12 (1.39 to 3.24), 3.47	1.54 (1.20 to 1.99), 3.35
None	4.00 (2.58 to 6.20), 6.21	3.12 (2.46 to 3.97), 9.31	2.83 (1.84 to 4.35), 4.73	3.56 (2.60 to 4.87), 7.92
Major extracranial injury:				
No	1	1	1	1.00
Yes	1.53 (1.10 to 2.13), 2.53	1.08 (0.91 to 1.28), 0.89	1.55 (1.20 to 1.99), 3.37	1.61 (1.38 to 1.88), 6.03
Findings on computed tomography:				
Petechial haemorrhages	1.15 (0.83 to 1.59), 0.84	1.26 (1.07 to 1.47), 2.82	1.21 (0.95 to 1.55), 1.56	1.49 (1.29 to 1.73), 5.33
Obliteration of 3rd ventricle or basal cisterns	4.46 (2.97 to 6.68), 7.23	1.99 (1.69 to 2.35), 8.25	2.21 (1.49 to 3.30), 3.95	1.53 (1.31 to 1.79), 5.30
Subarachnoid bleed	1.48 (1.09 to 2.02), 2.51	1.33 (1.14 to 1.55), 3.60	1.62 (1.26 to 2.08), 3.79	1.20 (1.04 to 1.39), 2.49
Midline shift	2.77 (1.82 to 4.21), 4.77	1.78 (1.44 to 2.21), 5.35	1.93 (1.30 to 2.87), 3.24	1.86 (1.48 to 2.32), 5.42
Non-evacuated haematoma	2.06 (1.49 to 2.84), 4.40	1.48 (1.24 to 1.76), 4.43	1.72 (1.33 to 2.22), 4.15	1.68 (1.43 to 1.97), 6.34
C statistic	0.88	0.84	0.83	0.84

GCS=Glasgow coma scale.
*Includes age, GCS, pupil reactivity, presence of major extra cranial injury, and all findings on computed tomography.
†Per 10 year increase after 40 years.
‡Per decrease of each value of GCS.

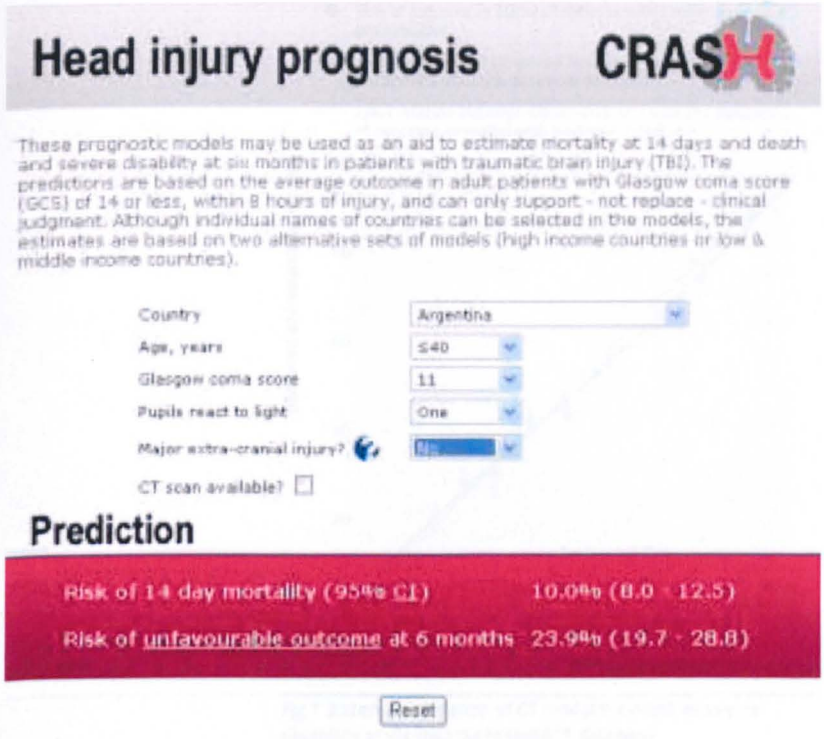


Fig 5 | Screenshot of web based calculator available at www.crash2.lshtm.ac.uk/. If CT scan available box is ticked, calculator displays additional CT variables

computed tomography was associated with the worst prognosis at 14 days. This is supported by recent findings that absence of basal cisterns is the strongest predictor of six month mortality.[18] We also found—as previously reported—the independent prognostic value of traumatic subarachnoid haemorrhage.[19]

Patients from low-middle income countries had worse early prognosis than those from high income countries. Regional differences in outcome after traumatic brain injury have previously been reported between Europe and North America, but the difference in mortality between low-middle and high income countries has not been explored.[20]

The strength of association between some predictors and outcomes differed by region. A low Glasgow coma score had an even worse prognosis in patients from low-middle income countries compared with patients from high income countries. This might relate to quality of care or it could be that low Glasgow coma score in high income countries is associated with greater use of sedation, rather than to severity of traumatic brain injury. Increasing age had a worse prognosis in high income countries compared with low-middle income countries. This is because of even lower risks at younger ages in high income countries, while both have similar risks at older ages. Regarding computed tomography, some abnormal findings were stronger predictors in high income countries. This could be because of better technology and therefore more accurate diagnosis with computed tomography.

A systematic review identified over 100 prognostic models for patients with traumatic brain injury, but methodological quality was adequate in only a few.[7] As with our models, two of the more methodologically robust models showed similarly good discrimination but worse calibration.[17,21] They too included Glasgow coma score, age, pupil reactivity, and results of computed tomography as predictors, but, unlike our models, they did not include the presence of major extracranial injury, and none of them included patients from low-middle income countries.

Strengths and weakness of the study

Our study's strengths are the use of a well described cohort of patients, prospective and standardised collection of data on prognostic factors, low loss to follow-up, and the use of a validated outcome measure at a fixed time after the injury. All of these factors provide reassurance about the internal validity of our models. The large sample size in relation to the number of prognostic variables examined is another particular strength. In relation to its external validity, only a few prognostic models have been developed from patients in low-middle income countries, and to the best of our knowledge the models we developed are the first with a large sample size and adequate methods.[7] The external validation confirmed the discriminatory ability of the models in patients from high income countries and showed good calibration for the computed tomography model. Unlike most published prognostic models,

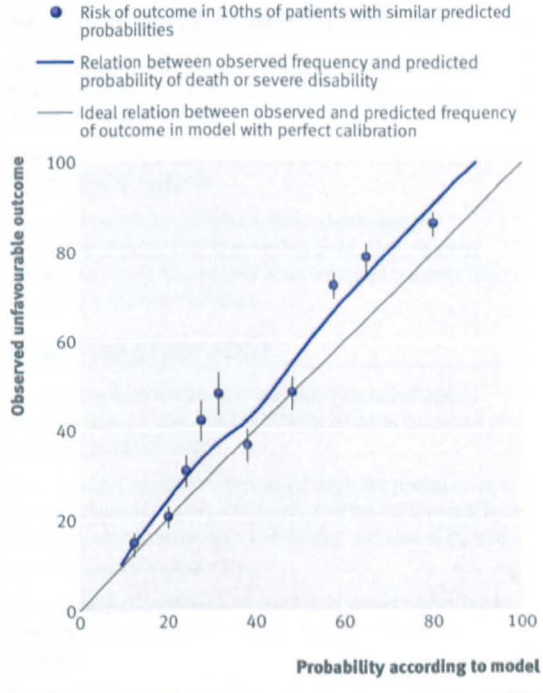


Fig 6 | External validation of basic model for death or severe disability at six months in IMPACT database

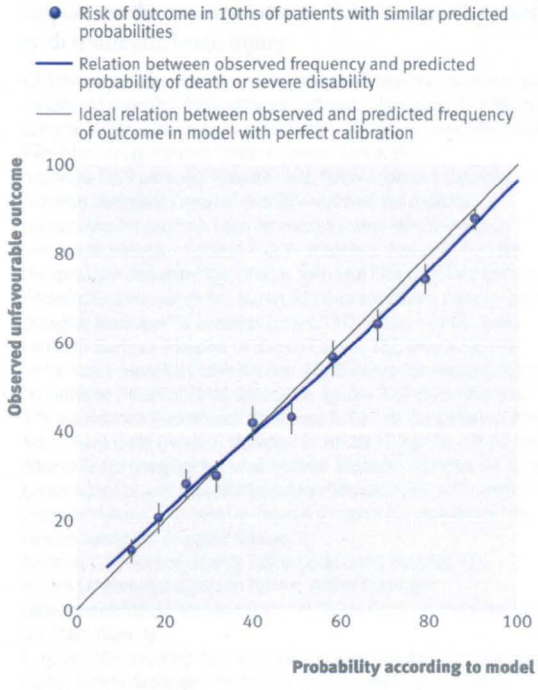


Fig 7 |External validation of CT model for death or severe disability at six months in IMPACT database

we included the complete spectrum of patients with traumatic brain injury, ranging from mild to severe. Finally, the data required to make predictions with the model are easily available to clinicians, and we have developed a web based risk calculator.

There are some limitations. The data from which the models were developed come from a clinical trial and this could therefore limit external validity. For example, patients were recruited within eight hours of injury and we cannot estimate the accuracy of the models for patients evaluated beyond this time. Nevertheless, the CRASH trial was a pragmatic trial that did not require any additional tests and therefore included a diversity of "real life" patients. Another limitation was that for the validation we were forced to exclude the variables major extracranial injury and petechial haemorrhages because they were not available in the IMPACT sample. Neither of these variables, however, was among the stronger predictors. The external validation showed good discriminatory ability, but this was somewhat lower than in the original data. This may be explained by a more homogeneous selected case mix in these other trials, which included only patients with moderate and severe Glasgow coma score.

Implications of the study

Most of the burden of traumatic brain injury is in low-middle income countries, where case fatality is higher and resources are limited. We found that several predictors differed in their strength of association with outcome according to income level of country, suggesting that it may be inappropriate to extrapolate from models for high income countries to poorer settings. We have developed a methodologically valid, simple, and accurate model that may help decisions about health care for individual patients. It is important to emphasise, however, that while prognostic models may complement clinical decision making they cannot and should not replace clinical judgment. This is particularly important in the context of judgments about the withdrawal of care or clinical triage. These prognostic models can also help in the design and analysis of clinical trials, through prognostic stratification, and can be used in clinical audit by allowing adjustment for case mix.[22]

Future research

The differences found between the prognostic models for low-middle and high income countries are important. Although most of the burden of trauma occurs in low-middle income countries, most research takes place in high income countries.[3] A recent systematic review found that few prognostic models for traumatic brain injury were developed in low-middle income countries.[7] More research is therefore needed to obtain reliable data from these settings. An improved understanding of the differences between these regions might also clarify the mechanisms of predictors that are not immediately obvious when we analyse a homogeneous population. Because our models were developed with data from a clinical trial, and validated only in patients from high income countries, further prospective validation in independent cohorts is needed to strengthen the generalisability of the models. Future research could also evaluate different ways, or formats, for presenting the models to physicians; their use in clinical practice; and whether ultimately they have any

WHAT IS ALREADY KNOWN ON THIS TOPIC

Traumatic brain injury is a leading cause of death and disability worldwide with most cases occurring in low-middle income countries

Prognostic models may improve predictions of outcome and help in clinical research

Many prognostic models have been published but methodological quality is generally poor, sample sizes small, and only a few models have included patients from low-middle income countries

WHAT THIS STUDY ADDS

In a basic model prognostic indicators included age, Glasgow coma scale, pupil reactivity, and the presence of major extracranial injury

In a CT model additional indicators were the presence of petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleeding, mid-line shift, and non-evacuated haematoma

The strength of predictors of outcomes varies according to whether patients are from high or low-middle income countries

These prognostic models, that include simple variables, are available on the internet (www.crash2.lshtm.ac.uk/)

impact on the management and outcomes of patients with traumatic brain injury.

CRASH trial collaborators by country (number of patients randomised)
Albania (41 patients): Fatos Olldashi (national coordinator), Itan Muzha (Central Military University Hospital National Trauma Centre, 35); Nikolin Filipi (University Hospital "Mother Teresa" Tirana, 6).
Argentina (359 patients): Roberto Lede, Pablo Copertari, Carolina Traverso, Alejandro Copertari (IAMBE—national and regional coordinators for southern Latin America); Enrique Alfredo Vergara, Carolina Montenegro, Roberto Ruiz de Huidobro, Pantaleón Saladino (Hospital San Bernardo, 106); Karina Surt, José Cialzeta, Silvio Lazzeri (Hospital Escuela Jose de San Martin, 52); Gustavo Piñero, Fabiana Ciccioli (Hospital Municipal "Dr Leonidas Lucero," 37); Walter Videtta, María Fernanda Barboza (Hospital Dr Ramón Carrillo, 35); Silvana Svampa, Victor Sciuto (Hospital Castro Rendon, 28); Gustavo Domeniconi, Marcelo Bustamante (Hospital Zonal General De Agudos "Heroes de Malvinas," 27); Maximiliano Waschbusch (Policlinico Sofia T de Santamarina. 20); María Paula Gullo (Hospital Municipal Dr Hector J D'Agnillo, 17); Daniel Alberto Drago (Hospital Nacional Profesor Alejandro Posadas, 11); Juan Carlos Arjona Linares (Hospital Español de Mendoza, 10); Luis Camputaro (Hospital Italiano, 10); Gustavo Tróccoli (Hospital "Dr José Penna," 5); Hernán Galimberti (Hospital Aleman, 1).
Australia (13 patients): Mandy Tallott (Gold Coast Hospital, 13).
Austria (21 patients): Christian Eybner, Walter Buchinger (Waldviertelklinikum Standort Horn, 17); Sylvia Fitzal (Wilhelminenspital der Stadt Wien, 4).
Belgium (403 patients): Guy Mazairac (national coordinator), Véronique Oleffe, Thierry Grollinger, Philippe Delvaux, Laurent Carlier (Centre Hospitalier Regional de Namur, 356); Veronique Braet (AZ Klina Hospital, 34); Jean-Marie Jacques (Hospital of Jolimont, 11); Danielle de Knoop (Clinique Saint-Luc, 2).
Brazil (119 patients): Luiz Nasi (national coordinator), Humberto Kukhuyn Choi, Mara Schmitt (Hospital de Pronto Socorro de Porto Alegre, 113); André Gentil (Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, 5); Flavio Nacul (Clínica São Vicente, 1).
Chile (3 patients): Pedro Bedoya Barrios (Hospital Regional Copiapo, 3).
China (87 patients): Chen Xinkang, Lin Shao Hua, Huang Han Tian (Zhongshan City People's Hospital, 79); Cai Xiaodong (Sheng Zheng Second People's Hospital, 8).
Colombia (832 patients): Wilson Gualteros, Alvaro Ardila Otero (Hospital Universitario San Jorge, 216); Miguel Arango (national coordinator, regional coordinator for northern Latin America and Caribbean), Juan Ciro, Hector Jaramillo (Gloria GarciaClinica Las Americas, 199); Ignacio Gonzalez, Carolina Gomez (Hospital General de Medellin, 119); Arturo Arias, Marco Fonseca, Carlos Mora (Hospital Erasmo Meoz, 90); Edgar Giovanni Luna Cabrera, José Luis Betancurth, Porfirio Muñoz (Hospital Departamental de Nariño, 51); Jesus Alberto Quiñónez, Maria Esther Gonzalez Castillo (Hospital San Andres, 37); Orlando Lopez (Hospital Federico Lleras, 31); Rafael Perez Yepes, Diana Leon Cuellar, Gerson Paez (Hospital El Tunal, 24); Hernán Delgado Chaves, Pablo Emilio Ordoñez (Hospital Civil de Ipiales, 21); Ricardo Plata, Martha Pineda (Hospital Universitario del Valle, 15); Libardo Enrique Pulido (Hospital Regional de Duitama, 12); John Sergio Velez Jaramillo (Hospital Timothy Britton, 12); Carlos Rebolledo (Organización Clinica General del Norte, 5).
Costa Rica (20 patients): Oscar Palma (Hospital México, 20).
Cuba (404 patients): Caridad Soler (national coordinator); Irene Pastrana, Raul Falero (Hospital Abel Santamaria Cuadrado, 77); Mario Domínguez Perera, Agustín Arocha García, Raydel Oliva (Hospital Universitario "Arnaldo Milián Castro," 55); Hubiel López Delgado (Hospital Provincial Docente "Manuel Ascunce Domenech," 43); Aida Madrazo Carnero, Boris Leyva López (Hospital VI Lenin, 42); Angel Lacerda Gallardo, Amarilys Ortega Morales (Hospital General de Morón, 40); Humberto Lezcano (Hospital General Universitario "Carlos Manuel de Cèspedes," 38); Marcos Iraola Ferrer (Hospital Universitario "Dr Gustavo Aldereguia Lima," 37); Irene Zamalea Bess, Gladys Rivas Canino (Hospital Miguel Enriquez, 36); Ernesto Miguel Piferrer Ruiz (Hospital Clínico-Quirúrgico Docente "Saturnino Lora," 32); Orlando Garcia Cruz (Centro de Investigaciones Médico-Quirúrgicas, 4).
Czech Republic (961 patients): Petr Svoboda (national coordinator), Ilona Kantorová, Jiří Ochmann, Peter Scheer, Ladislav Kozumplík, Jitka Maršová (Research Institute for Special Surgery and Trauma, 852); Karel Edelmann (Masaryk Hospital, 41); Ivan Chytra, Roman Bosman (Charles University Hospital, Plzen, 35); Hana Andrejsová (University Hospital Hradec Kralove, 15); Jan Pachl (Hospital Kralovske Vinohgrady, 9); Jan Bürger (Hospital Pribram, 7); Filip Kramar (Univerzity Karlovy Neurochirurgicka Klinika, 2).
Ecuador (258 patients): Mario Izurieta Ulloa (national coordinator), Luis Gonzalez, Alberto Daccach, Antonio Ortega, Stenio Cevallos (Hospital Luis Vernaza, 202); Boris Zurita Cueva (Hospital de la Policia Guayaquil, 16); Marcelo Ochoa (Hospital Jose Carrasco Arteaga, 11); Jaime Velásquez Tapia (Hospital Naval, 11); Jimmy Hurtado (Clínica Central, 8); Miguel Chung Sang Wong (Hospital Militar de Guayaquil, 5); Roberto Santos (Hospital Regional del IESS "Dr Teodoro Maldonado Carbo," 5).
Egypt (775 patients): Hussein Khamis (national coordinator), Abdul Hamid Abaza, Abdalla Fekry, Salah El Kordy, Tarek Shawky (Mataria Teaching Hospital, 364); Hesham El-Sayed (national coordinator), Nabil Khalil, Nader Negm, Salem Fisal (Suez Canal University, 180); Mamdouh Alamin, Hany Shokry (Aswan Teaching Hospital, 160); Ahmed Yahia Elhusseny, Atif Radwan, Magdi Rashid (Zagazig University Hospital, 71).
Georgia (56 patients): Tamar Gogichaisvili (national coordinator), George Ingorokva, Nikoloz Gongadze (Neurosurgery Department of Tbilisi State Medical University, 55); Alexander Otarashvili (Tbilisi 4th Hospital, 1).
Germany (27 patients): Waltraud Kleist (Ernst Moritz Arndt University, 14); Mathias Kalkum (Kreiskrankenhaus Tirschenreuth, 8); Peter Ulrich (Klinikum Offenbach, 5).
Ghana (7 patients): Nii Andrews (Narh-Bita Hospital, 7).
Greece (20 patients): George Nakos (University Hospital of Ioannina, 8); Antonios Karavelis (University General Hospital of Larissa, 5); George Archontakis (Chania General Hospital "St George," 4); Pavlos Myrianthefs (KAT Hospital of Athens, 3).
India (973 patients): Yadram Yadav, Sharda Yadav, R Khatri, Arvind Baghel (NSCB Medical College, 177); Mazhar Husain (national coordinator for north India), Deepak Jha (King George Medical College, 105); Wu Hoong Chhang, Manohar Dhandhania, Choden Fonning (North Bengal Neuro Research Centre, 65); S N Iyengar, Sanjay Gupta (G R Medical College, 51); R R Ravi, K S Bopiah, Ajay Herur (Medical Trust Hospital Kochi, 51); N K Venkataramana (national coordinator for south India), A Satish (Manipal Hospital, 50); K Bhavadasan, Raymond Morris, Ramesh S (Medical College Hospital Trivandrum, 50); A Satish (Abhaya Hospital, 42); Yashbir Dewan, Yashpal Singh (Christian Medical College, 36); Rajesh Bhagchandani, Sanjana Bhagchandani (Apex Hospital Bhopal, 32); Vijaya Ushanath Sethurayar (Meenakshi Mission Hospital and Research Centre, 32); Sojan Ipe, G Sreekumar (MOSC Medical College Hospital, 32); Manas Panigrahi, Agasti Reddy (Nizam's Institute of Medical Sciences, 28); Varinder Khosla, Sunil Gupta (Postgraduate Institute of Medical Education and Research, 28); Haroon Pillay, Nisha Thomas (Baby Memorial Hospital, 25); Krishnamurthy Sridhar, Bobby Jose (V H S Hospital, 22); Nadakkavvkakan Kurian (Jubilee Mission Hospital, 20); Shanti Praharaj, Shibu Pillai (National Institute of Mental Health and Neurosciences, 17); Ramana (Care Hospital (16); Sanjay Gupta, Smita Gupta (Sri Sai Hospital, 16); Dilip Kiyawat (Hirabi Cowasji Jehangir Medical Research Institute, 15); Kishor Maheshwari (Maheshwari Orthopaedic Hospital, 13); Dilip Panikar (Amrita Institute of Medical Sciences, 11); Jayant Chawla (Hartej Maternity and Nursing Home, 7); Satyanarayana Shenoy, Annaswamy Raja (Kasturba Medical College and Hospital, 7); Yeshomati Rupayana (Choitram Hospital and Research Centre, 6); Suryanarayan Reddy (Gowri Gopal Superspeciality Hospital, 6); Nelanuthala Mohan (Apex Hospital Visakhapatnam, 3); Shailesh Kelkar (Central India Institute of Medical Sciences, 3); Yadram Yadav (Marble City Hospital and Research Centre, 3); Jayant Chawla (Government Medical College Amritsar, 1); Mukesh Johri (Johri Hospital, 1); Yadram Yadav (National Hospital Jabalpur, 1).
Indonesia (238 patients): Nyoman Golden (national coordinator), Sri Maliawan (Sanglah General Hospital, 222); Achmad Fauzi, Umar Farouk (Sidoarjo General Hospital, 14).
Iran (233 patients): Esmaeel Fakharian, Amir Aramesh (Naghavi University Hospital, 110); Maasoumeh Eghtedari, Farhad Ahmadzadeh, Alireza Gholami (Fatemeh Zahra Hospital, 85); Maasoumeh Eghtedari, Farhad Ahmadzadeh (Social Security Hospital, 38).
Ireland (113 patients): Patrick Plunkett, Catherine Redican, Geraldine McMahon (St James's Hospital, 113).
Italy (9 patients): Maria Giuseppina Annetta (Università Cattolica del Sacro Cuore, 4); Homère Mouchaty (Università di Firenze, 3); Eros Bruzzone (Ospedale San Martino, 2).
Ivory Coast (3 patients): Béatrice Harding (CHU de Cocody, 3).
Kenya (2 patients): Mahmood Qureshi (Aga Khan Hospital, 2).
Malaysia (176 patients): Zamzuri Idris, Jafri Abdullah NC, Ghazaime Ghazali, Abdul Rahman Izaini Ghani (Hospital University Science Malaysia, 162); Fadzli Cheah (Ipoh Specialist Hospital, 14).
Mexico (17 patients): Alfredo Cabrera (national coordinator); José Luis Mejía González (Hospital General Regional No 1, 12); José Luis Mejía

González (Hospital General de Queretaro, 4); Jorge Loría-Castellanos (Hospital General Regional No 25, 1).
New Zealand (43 patients): Suzanne Jackson, Robyn Hutchinson (Dunedin Hospital, 43).
Nigeria (180 patients): Edward Komolafe (national coordinator), Augustine Adeolu, Morenikeji Komolafe (Obafemi Awolowo University Teaching Hospitals, 77); Olusanya Adeyemi-Doro, Femi Bankole (Lagos University Teaching Hospital, 43); Bello Shehu, Victoria Danlami (Usmanu Danfodiyo University Teaching Hospital, 36); Olugbenga Odebode (University of Ilorin Teaching Hospital, 15); Kehinde Oluwadiya (Lautech Teaching Hospital, 7); Ahmed Sanni (Lagos State University Teaching Hospital, 1); Herb Giebel (Seventh Day Adventist Hospital, 1); Sushil Kumar (St Stephen's Hospital, 1).
Pakistan (17 patients): Rashid Jooma (Jinnah Postgraduate Medical Centre, 17).
Panama (7 patients): Jose Edmundo Mezquita (Complejo Hospitalario M A Guerrero, 7).
Paraguay (10 patients): Carlos Ortiz Ovelar (Instituto de Prevision Social, 10).
Peru (8 patients): Marco Gonzales Portillo (Hospital Nacional "Dos de Mayo," 6); Diana Rodriguez (national coordinator) (Hospital Nacional Arzobispo Loayza, 2).
Romania (319 patients): Laura Balica (national coordinator), Bogdan Oprita, Mircea Sklerniacof, Luiza Steflea, Laura Bandut (Spitalul Clinic de Urgenőtùa Bucuresö ti, 282); Adam Danil, Remus Iliescu (Sfantum Pantelimon Hospital, 28); Jean Ciurea (Prof Dr D Bagdasar Clinical Emergency Hospital, 9).
Saudi Arabia (32 patients): Abdelazeem El-Dawlatly, Sherif Alwatidy (King Khalid University Hospital, 24); Walid Al-Yafi, Megahid El-Dawlatly (King Khalid National Guard Hospital, 8).
Serbia (23 patients): Ranka Krunic-Protic, Vesna Janosevic (Klinicki Centar Srbije, 23).
Singapore (23 patients): James Tan (national coordinator) (National Neuroscience Institute, 21); Charles Seah (Changi General Hospital, 2).
Slovakia (179 patients): Štefan Trenkler (national coordinator), Matuš Humenansky, Tatiana Stajančová (Reiman Hospital, 71); Ivan Schwendt, Anton Laincz (NsP Poprad, 39); Zeman Julius, Stano Maros (Nemocnica Bojnice, 25); Jozef Firment (FNsp Kosice, 12); Maria Cifraničova (Ns P Trebisov, 11); Beata Sániová (Faculty Hospital in Martin, 10); Karol Kalig (NsP Ruzinov, 4); Soňa Medekova (NsP Nové Zámky, 3); Radovan Wiszt (NsP Liptovsky Mikulas, 2); NsP F D Roosevelt, 1); Ivan Mačuga (NsP Zilina, 1).
South Africa (366 patients): Bennie Hartzenberg (national coordinator), Grant du Plessis, Zelda Houlie (Tygerberg Academic Hospital, 307); Narendra Nathoo, Sipho Khumalo (Wentworth Hospital, 57); Ralph Tracey (Curamed Kloof Hospital, 1).
Spain (259 patients): Angeles Muñoz-Sánchez (national coordinator), Francisco Murillo-Cabezas NC, Juan Flores-Cordero, Dolores Rincón-Ferrari (Hospital Universitario Virgen del Rocio, 133); Martin Rubi, Lopez Caler (Hospital Torrecárdenas, 37); Maite Misis del Campo, Luisa Bordejé Laguna (Hospital Universitario Germans Trias i Pujol, 32); Juan Manuel Nava (Hospital Mútua de Terrassa, 20); Miguel Arruego Minguillón (Hospital Universitario de Girona Dr Josep Trueta, 12); Alfonso Muñoz Lopez (Hospital Carlos Haya, 10); Luis Ramos-Gómez (Hospital General de La Palma, 6); Victoria de la Torre-Prados (Hospital Universitario Virgen de la Victoria, 5); Romero Pellejero (Hospital General Yagüe, 4).
Sri Lanka (132 patients): Véronique Laloë (national coordinator), Bernhard Mandrella, Suganthan (Batticaloa General Hospital-Médecins Sans Frontières, 84); Sunil Perera (National Hospital of Sri Lanka, 39); Véronique Laloë, Kanapathipillai Mahendran (Point-Pedro Base Hospital, 9).
Switzerland (160 patients): Reto Stocker (national coordinator), Silke Ludwig (national coordinator) (University Hospital of Zurich, 133); Heinz Zimmermann (University Hospital Bern, 15); Urs Denzler (Kantonsspital Schaffhausen, 12).
Thailand (579 patients): Surakrant Yutthakasemsunt (national coordinator), Warawut Kittiwattanagul, Parnumas Piyavechvirat, Pojana Tapsai, Ajchara Namuang-jan (Khon Kaen Regional Hospital, 535); Upapat Chantapimpa (Chiangrai Prachanuko Hospital, 12); Chanothai Watanachai, Pusit Subsompon (Rayong Hospital, 11); Wipurat Pussanakawatin, Pensri Khunjan (Krabi Hospital, 10); Sakchai Tangchitvittaya, Somsak Nilapong (Suratthani Hospital, 8); Tanagorn Klangsang, Wibul Taechakosol (Roi Et Hospital, 2); Atirat Srinat (Lampang Hospital, 1).
Tunisia (63 patients): Zouheir Jerbi (national coordinator), Nebiha Borsali-Falfoul, Monia Rezgui (Hospital Habib Thameur, 63).
Turkey (2 patients): Nahit Cakar (Istanbul Medical Faculty, 2).
Uganda (43 patients): Hussein Ssenyonjo, Olive Kobusingye (Makerere Medical School, 43).
UK (1391 patients): Gabrielle Lomas, David Yates, Fiona Lecky (Hope Hospital, 209); Anthony Bleetman, Alan Baldwin, Emma Jenkinson, Shiela Pantrini (Birmingham Heartlands Hospital, 123); James Stewart, Nasreen Contractor, Trudy Roberts, Jim Butler (North Manchester General Hospital, 85); Alan Pinto, Diane Lee (Royal Albert Edward Infirmary, 83); Nigel Brayley, Karly Robbshaw, Clare Dix (Colchester General Hospital, 79); Sarah Graham, Sue Pye (Whiston Hospital, 69); Marcus Green, Annie Kellins (Selly Oak Hospital, 61); Chris Moulton, Barbara Fogg (Royal Bolton Hospital, 51); Rowland Cottingham, Sam Funnell, Utham Shanker (Eastbourne District General Hospital, 50); Claire Summers, Louise Malek (Trafford General Hospital, 41); Rowland Cottingham (national coordinator), Christopher Ashcroft, Jacky Powell (Royal Sussex County Hospital, 38); Steve Moore, Stephanie Buckley (Countess of Chester Hospital, 36); Mandy Grocutt, Steve Chambers (Worthing Hospital, 34); Amanda Morrice, Helen Marshall (Medway Maritime Hospital, 29); Julia Harris, Wendy Matthews, Jane Tippet (Chelsea and Westminster Hospital, 28); Simon Mardell, Fiona MacMillan, Anita Shaw (Furness General Hospital, 27); Pramod Luthra, Gill Dixon (Royal Oldham Hospital, 26); Mohammed Ahmed, John Butler, Mike Young (Stepping Hill Hospital, 26); Sue Mason, Ian Loveday (Northern General Hospital, 25); Christine Clark, Sam Taylor (Blackburn Royal Infirmary, 23); Paul Wilson (Cheltenham General Hospital, 23); Kassim Ali, Stuart Greenwood (Fairfield General Hospital, 23); Martin White, Rosa Perez (Queen Elizabeth the Queen Mother Hospital, 21); Sam Eljamel (Ninewells Hospital and Medical School, 19); Jonathan Wasserberg, Helen Shale (Queen Elizabeth Hospital Birmingham, 18); Colin Read, John McCarron (Russell's Hall Hospital, 18); Aaron Pennell (Princess Alexandra Hospital, 16); Gautam Ray (Princess Royal Hospital, 14); John Thurston, Emma Brown (Darent Valley Hospital, 13); Lawrence Jaffey, Michael Graves (Royal Liverpool University Hospital, 12); Richard Bailey, Nancy Loveridge (Chesterfield and North Derbyshire Royal Hospital, 10); Geraint Evans, Shirleen Hughes, Major Kafeel Ahmed (Withybush General Hospital, 10); Jeremy Richardson, Claire Gallagher (Aberdeen Royal Infirmary, 8); Titus Odedun, Karen Lees (Ormskirk and District General Hospital, 8); David Foley, Nick Payne (Queen Mary's Hospital, 8); Alan Pennycook, Carl Griffiths (Arrowe Park Hospital, 6); David Moore, Denise Byrne (City Hospital Birmingham, 5); Sunil Dasan (St Helier Hospital, 4); Ashis Banerjee, Steve McGuinness (Whittington Hospital, 4); Claude Chikhani (Doncaster Royal Infirmary, 2); Nigel Zoltie, Ian Barlow (Leeds General Infirmary, 2); Ian Stell (Bromley Hospital, 1); William Hulse, Jacqueline Crossley (Harrogate District Hospital, 1); Laurence Watkins (Institute of Neurology, 1); Balu Dorani (Queen Elizabeth Hospital Gateshead, 1).
Vietnam (2 patients)—Truong Van Viet (Cho Ray Hospital, 2).

Contributors: The writing committee comprised Pablo Perel (Chair), Miguel Arango, Tim Clayton, Phil Edwards, Edward Komolafe, Stuart Poccock, Ian Roberts, Haleema Shakur, Ewout Steyerberg, and Surakrant Yutthakasemsunt

Funding: The MRC CRASH trial was funded by the UK Medical Research Council.

Competing interests: None declared.

Ethical approval: All MRC CRASH collaborators obtained local ethics or research committee approval.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Bruns J Jr, Hauser WA. The epidemiology of traumatic brain injury: a review. *Epilepsia* 2003;44(suppl 10):2-10.
2 Fleminger S, Ponsford J. Long term outcome after traumatic brain injury. *BMJ* 2005;331:1419-20.
3 Hofman K, Primack A, Keusch G, Hrynkow S. Addressing the growing burden of trauma and injury in low- and middle-income countries. *Am J Public Health* 2005;95:13-7.
4 Perel P, Wasserberg J, Ravi RR, Shakur H, Edwards P, Roberts I. Prognosis following head injury: a survey of doctors from developing and developed countries. *J Eval Clin Pract* 2007;13:464-5.
5 Lee KL, Pryor DB, Harrell FE Jr, Califf RM, Behar VS, Floyd WL, et al. Predicting outcome in coronary disease. Statistical models versus expert clinicians. *Am J Med* 1986;80:553-60.
6 Murray LS, Teasdale GM, Murray GD, Jennett B, Miller JD, Pickard JD, et al. Does prediction of outcome alter patient management? *Lancet* 1993;341:1487-91.
7 Perel P, Edwards P, Wentz R, Roberts I. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* 2006;6:38.
8 CRASH Trial Collaborators. Effect of intravenous corticosteroids on death within 14 days in 10 008 adults with clinically significant head

injury (MRC CRASH trial): randomised placebo-controlled trial. *Lancet* 2004;364:1321-8.
9 CRASH Trial Collaborators. Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months. *Lancet* 2005;365:1957-9.
10 Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet* 1975;i:480-4.
11 World Bank. *World development indicators* . Washington, DC: World Bank, 2006.
12 Brain Trauma Foundation (US) AAoNS. *Management and prognosis of severe traumatic brain injury* . New York: 2000. www.braintrauma.org/site/DocServer/Prognosis_Guidelines_for_web.pdf?docID=241.
13 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361-87.
14 Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515-24.
15 Maas AI, Marmarou A, Murray GD, Teasdale SG, Steyerberg EW. Prognosis and clinical trial design in traumatic brain injury: the IMPACT study. *J Neurotrauma* 2007;24:232-8.
16 Hukkelhoven CW, Steyerberg EW, Rampen AJ, Farace E, Habbema JD, Marshall AF, et al. Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. *J Neurosurg* 2003;99:666-73.
17 Signorini DF, Andrews PJD, Jones PA, Wardlaw JM, Miller JD. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *J Neurol Neurosurg Psychiatry* 1999;66:20-5.
18 Maas AI, Hukkelhoven CW, Marshall LF, Steyerberg EW. 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* 2005;57:1173-82.
19 Wardlaw JM, Easton VJ, Statham P. Which CT features help predict outcome after head injury? *J Neurol Neurosurg Psychiatry* 2002;72:188-92.
20 Hukkelhoven CW, Steyerberg EW, Farace E, Habbema JD, Marshall LF, Maas AI. Regional differences in patient characteristics, case management, and outcomes in traumatic brain injury: experience from the tirilazad trials. *J Neurosurg* 2002;97:549-57.
21 Hukkelhoven CW, Steyerberg EW, Habbema JD, Farace E, Marmarou A, Murray GD, et al. Predicting outcome after traumatic brain injury: development and validation of a prognostic score based on admission characteristics. *J Neurotrauma* 2005;22:1025-39.
22 Altman DG. Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables. *BMJ* 2001;323:224-8.

Accepted: 5 December 2007

BMC Medical Research Methodology

BioMed Central

Research article

Open Access

Use of the Oxford Handicap Scale at hospital discharge to predict Glasgow Outcome Scale at 6 months in patients with traumatic brain injury

Pablo Perel*, Phil Edwards, Haleema Shakur and Ian Roberts

Address: Department of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK

Email: Pablo Perel* - pablo.perel@lshtm.ac.uk; Phil Edwards - phil.edwards@lshtm.ac.uk; Haleema Shakur - haleema.shakur@lshtm.ac.uk; Ian Roberts - ian.roberts@lshtm.ac.uk

* Corresponding author

Published: 6 November 2008

BMC Medical Research Methodology 2008, **8**:72 doi:10.1186/1471-2288-8-72

Received: 24 June 2008
Accepted: 6 November 2008

This article is available from: http://www.biomedcentral.com/1471-2288/8/72

Abstract

Background: Traumatic brain injury (TBI) is an important cause of acquired disability. In evaluating the effectiveness of clinical interventions for TBI it is important to measure disability accurately. The Glasgow Outcome Scale (GOS) is the most widely used outcome measure in randomised controlled trials (RCTs) in TBI patients. However GOS measurement is generally collected at 6 months after discharge when loss to follow up could have occurred. The objectives of this study were to evaluate the association and predictive validity between a simple disability scale at hospital discharge, the Oxford Handicap Scale (OHS), and the GOS at 6 months among TBI patients.

Methods: The study was a secondary analysis of a randomised clinical trial among TBI patients (MRC CRASH Trial). A Spearman correlation was estimated to evaluate the association between the OHS and GOS. The validity of different dichotomies of the OHS for predicting GOS at 6 months was assessed by calculating sensitivity, specificity and the C statistic. Uni and multivariate logistic regression models were fitted including OHS as explanatory variable. For each model we analysed its discrimination and calibration.

Results: We found that the OHS is highly correlated with GOS at 6 months (spearman correlation 0.75) with evidence of a linear relationship between the two scales. The OHS dichotomy that separates patients with severe dependency or death showed the greatest discrimination (C statistic: 84.3). Among survivors at hospital discharge the OHS showed a very good discrimination (C statistic 0.78) and excellent calibration when used to predict GOS outcome at 6 months.

Conclusion: We have shown that the OHS, a simple disability scale available at hospital discharge can predict disability accurately, according to the GOS, at 6 months. OHS could be used to improve the design and analysis of clinical trials in TBI patients and may also provide a valuable clinical tool for physicians to improve communication with patients and relatives when assessing a patient's prognosis at hospital discharge.

Trial Registration Number: ISRCTN74459797

Background

Traumatic brain injury (TBI) is an important cause of acquired disability. In evaluating the effectiveness of clinical interventions for TBI it is important to measure disability accurately. The Glasgow Outcome Scale (GOS) is the most widely used outcome measure in randomised controlled trials (RCTs) in TBI patients.[1] However, because the GOS assesses how well patients function in their daily social interactions, it is only applicable after the patient has been discharged from hospital.

Loss to follow up after hospital discharge is a common problem in clinical trials in TBI and some amount of missing data is often unavoidable.[2]. If an early outcome measure was available that could predict long term disability, it could be valuable for dealing with missing data, and might potentially be used as a surrogate outcome.

The MRC CRASH Trial was a large, randomised placebo controlled trial of the effects of a 48-hour infusion of corticosteroids on death and disability, among 10,008 adults.[3] Using data from this cohort of patients we have previously identified hospital admission variables that accurately predict 6 months GOS.[4] This cohort also presents an opportunity to evaluate the predictive validity of an early disability outcome measure for TBI patients. A modified version of the Oxford Handicap Scale (OHS) was completed at hospital discharge and the GOS was completed at 6 months after injury. The OHS, which was originally developed for stroke patients, comprises six categories: no symptoms, minor symptoms, minor handicap, moderate handicap, moderately severe handicap, and severe handicap.[5] In the MRC CRASH Trial a modified form of the OHS was used in which moderate handicap and moderately severe handicap were combined. Although the OHS has been previously used in brain injury trials, its association with GOS at 6 months in TBI patients has not been previously reported.[5]

The aim of this paper is to describe the association between an early disability outcome (OHS), and a 6 months disability outcome (GOS). Specifically the objectives were to:

1) Evaluate the correlation between OHS at hospital discharge and GOS at 6 months

2) Evaluate different dichotomies of the OHS at hospital discharge in predicting GOS at 6 months

3) Evaluate the extent to which OHS at hospital discharge predicts GOS at 6 months in survivors

Methods

Potential predictor

The OHS (table 1) was assessed at 14 days, hospital discharge or death (whichever occur first).

Variables that have previously been reported to be associated with the outcome were considered as potential confounders and included in an adjusted model: age, Glasgow Coma Scale (GCS) at randomization, pupil reactivity, whether the patient sustained a major extra cranial injury and computerised tomography (CT) scan results.[4]

Table 1: Original Oxford Handicap Scale and OHS used in the MRC CRASH Trial

Original OHS	Modified OHS used in CRASH
Categories	**Categories**
No symptoms	No symptoms
Minor symptoms that do not interfere with lifestyle	Minor symptoms
Minor handicap, symptoms that lead to some restriction in lifestyle but do not interfere with the patient's capacity to look after himself	Some restriction in lifestyle but independent
Moderate handicap, symptoms that significantly restrict lifestyle and prevent totally independent existence	Dependent but not requiring constant attention
Moderately severe handicap, symptoms that clearly prevent independent existence though not needing constant attention	
Severe handicap, totally dependent patient requiring constant attention night and day	Fully dependent requiring attention day and night
	Death

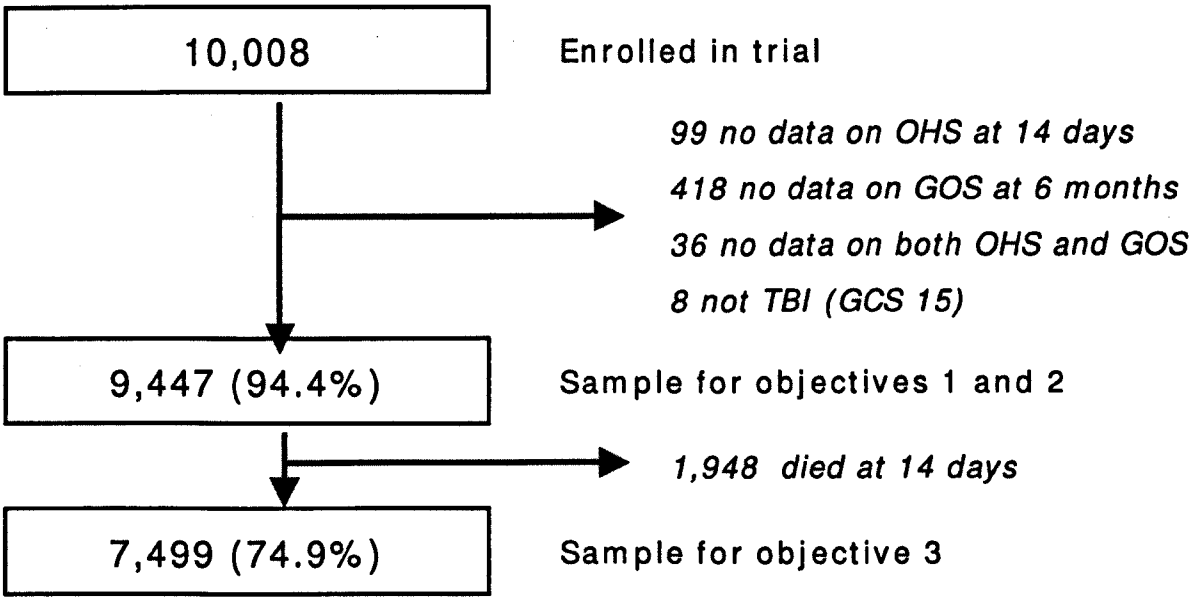


Figure 1
Flowchart of patients used in the analysis.

Outcome

The outcome was GOS at 6 months. The GOS comprises five categories: death, persistent vegetative state, severe disability, moderate disability and good recovery.[6] GOS was dichotomised for analysis in the CRASH Trial into favourable outcome (good recovery or moderate disability) and unfavourable outcome (severe disability, persistent vegetative state or death). We created two further dichotomies: good recovery versus other outcomes, and survival versus death.

The sample of patients

The MRC CRASH trial was a large international double-blind randomised placebo-controlled trial of the effect of early administration of a 48-h infusion of a corticosteroid (methylprednisolone) on the risk of death and disability after TBI. The characteristics of the patients randomised, and results of the trial have already been reported in detail.[3,7] Briefly, adults (aged 16 years or older) with a head injury and a GCS of 14 or less were randomly allocated to commence either a 48 hour infusion of methylprednisolone or matching placebo within eight hours of injury; patients from 239 hospitals in 48 countries were randomised. All collaborating MRC CRASH investigators were required to secure local ethics or research committee approval before recruitment could begin. Patients with clinically significant head injury are unable to give valid informed consent. Local ethics committees set consent procedures for participating hospitals. Some allowed consent waiver and others consent from a legal representative. We always adhered to these requirements.

Of 10,008 study participants enrolled in the MRC CRASH Trial, 99 (1%) had missing data on the OHS, 418 (4.2%) had missing data on the GOS at 6 months, and 36 (0.3%) had missing data for both OHS and GOS. A further 8 patients were excluded from analysis as they had a Glasgow Coma Scale (GCS) score of 15 at randomisation. Analysis for objectives 1 and 2 were therefore performed using data for 9,447 (94.4%) patients (figure 1). For the third objective (predictive validity of OHS among survivors), the 1,948 patients who died within 14 days of admission were excluded and the analysis was based on data for the remaining 7,499 patients (figure 1).

Analysis

Objective 1

A cross-tabulation between the OHS and GOS categories was performed. Their relation was assessed with the Spearman rank correlation index.

Objective 2

The validity of the different dichotomies of the OHS for predicting GOS at 6 months was assessed by calculating sensitivity, specificity and the c statistic (an equivalent concept to area under the receiver operator characteristic curve).

Objective 3

A logistic regression model was first fitted including only OHS as explanatory variables (model 1). A second model was then fitted that also included demographic and clinical variables (model 2). Finally, a third model was fitted that included all variables in model 2, plus CT scan variables. All the demographic, clinical and CT variables have been previously reported as being independently associated with unfavourable outcome at 6 months.[4] For each model we analysed its discrimination using the c statistic and calibration (graphically and with the Hosmer-Lemeshow test).

We then estimated the positive predictive value (with 95% confidence intervals) of each OHS category for GOS at 6 months.

Results

General characteristics of the population

Table 2 shows the characteristics of the sample included in the analysis. At 14 days 1,863 (19%) were dependent and 1,948 patients had died (21%). At 6 months, 3,525 (37.3%) patients were severely disabled or had died. Most deaths (84%) occurred within the first 14 days. OHS at 14 days and GOS at 6 months were highly correlated (Spearman rank correlation coefficient 0.75) and they showed a linear relationship (figure 2).

OHS for predicting 6 months outcome

Five dichotomies of the OHS were considered (Table 3).

When their validity was assessed in relation to unfavourable outcome as defined by the GOS (severe disability or death), dichotomy D showed the highest discrimination (c statistic) with high specificity (Table 4).

Among survivors at hospital discharge the OHS showed a strong association with GOS at 6 months. The crude analysis showed that patients who were fully dependent at 14 days had 24 higher odds of an unfavourable outcome at 6 months. Although adjusting for known prognostic factors attenuated the strength of the association, OHS remained a strong predictor with a highly statistically significant test (Table 5). Most importantly, when considered alone,

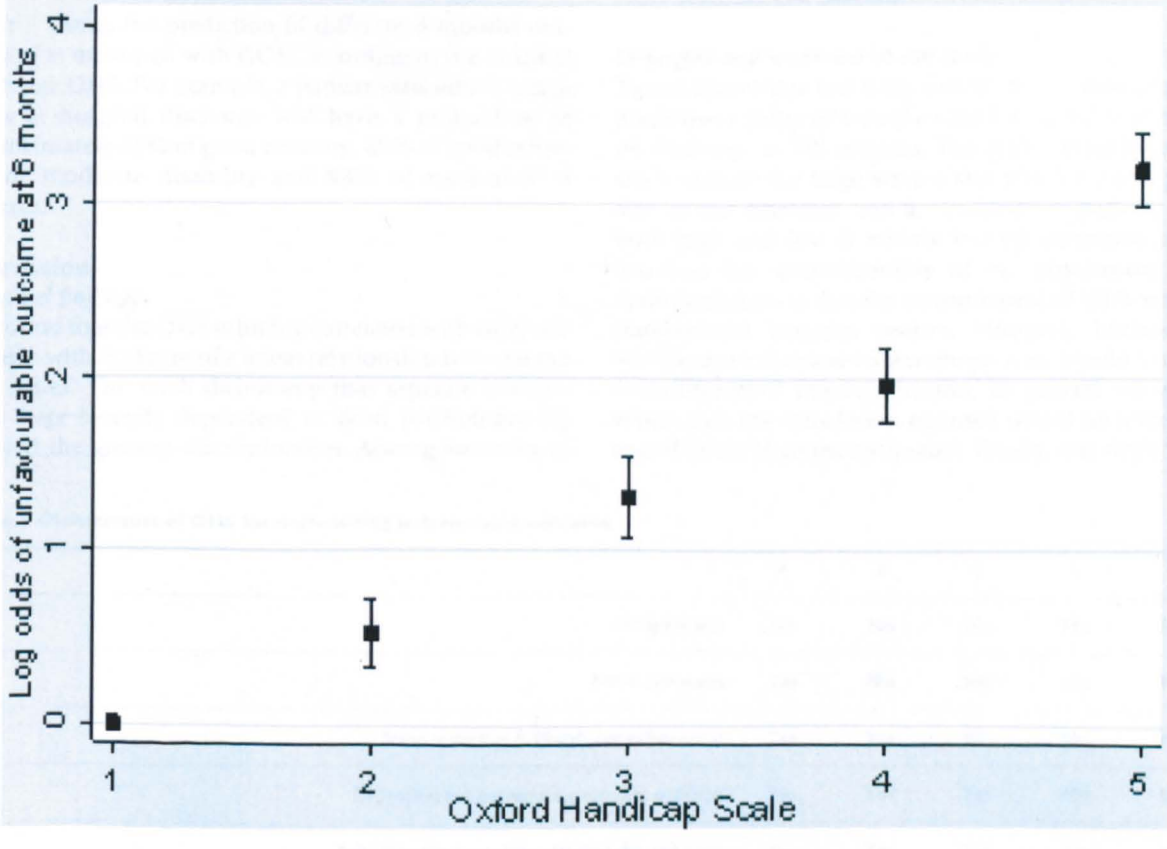


Figure 2
Relationship between Oxford Handicap Scale and unfavourable outcome (GOS) at 6 months.

Table 2: Glasgow Outcome Scale at 6 months by Oxford Handicap Scale at 14 days

	Glasgow Outcome Scale at 6 months								
Oxford Handicap Scale at 14 days	Good recovery		Moderate disability		Severe disability		Death		Total
	n	%	n	%	n	%	n	%	n
No symptoms	1,910	79	334	14	150	6	17	1	2411
Minor symptoms	1,646	67	537	22	233	9	42	2	2,458
Some restriction in lifestyle but independent	354	46	246	32	147	19	20	3	767
Dependent but not requiring constant attention	232	30	273	35	221	29	45	6	771
Fully dependent requiring attention day & night	148	14	242	22	457	42	245	22	1,092
Death	0	0	0	0	0	0	1948	100	1948
Total	4,290	45	1,632	17	1,208	13	2317	25	9,477

OHS showed very good discrimination and excellent calibration (H-L = 1) (figure 3).

Table 6 shows the prediction of different 6 months outcomes (as measured with GOS) according to the hospital discharge OHS. For example, a patient with minor symptoms at hospital discharge will have a probability of approximately 67% of good recovery, 89% of good recovery or moderate disability and 98% of survival at 6 months.

Discussion

Principal findings

We found that the OHS is highly correlated with GOS at 6 months with evidence of a linear relationship between the two scales. The OHS dichotomy that separate patients who were severely dependent or dead (dichotomy D) showed the greatest discrimination. Among survivors at hospital discharge the OHS showed a very good discrimination and excellent calibration when used to predict GOS outcome at 6 months.

Strengths and weakness of the study

To our knowledge this is the first study that evaluated the predictive validity of a simple scale for disability at hospital discharge in TBI patients. The main strengths of our study include the large sample size which ensures precision in our estimates, and the inclusion of patients from both high and low & middle income countries, which increases the generalizability of our conclusions. The main limitation is that the measurement of OHS was not standardized between centres. However, because we would expect that any measurement error would result in non-differential misclassification, in general we would expect that the association reported would be underestimated rather than overestimated. Finally, our study is the

Table 3: Dichotomies of OHS for determining unfavourable outcome

	A	B	C	D	E
No Symptoms	No	No	No	No	No
Minor Symptoms	Yes	No	No	No	No
Some restriction in lifestyle but independent	Yes	Yes	No	No	No
Dependent but not requiring constant attention	Yes	Yes	Yes	No	No
Fully dependent requiring attention day and night	Yes	Yes	Yes	Yes	No
Death	Yes	Yes	Yes	Yes	Yes

Table 4: Validity of the Oxford Handicap Scale at 14 days for Glasgow Outcome Scale at 6 months

OHS dichotomy	Sensitivity	Sensitivity	C stat
A	95.3	37.9	66.6
B	87.5	74.8	81.1
C	82.7	84.9	83.8
D	**75.2**	**93.4**	**84.3**
E	55.3	100.0	77.6

first to report this association which should therefore be examined in an external cohort of patients in order to confirm the findings.

Comparison with other studies

The incidence of unfavourable GOS outcome at 6 months in our cohort was lower in comparison to one reported in a series of TBI cohorts.[8] However, unlike ours, most of these cohorts had been restricted to severe TBI patients. The OHS has previously been used in RCTs of brain injury patients, and Bamford et al. reported good inter-observer agreement (a weighted kappa of 0.72).[5] Ours is the first study in TBI which has evaluated the relationship between OHS and GOS. Nevertheless, previous studies have shown a good agreement between the Modified Rankin Scale (the scale from which the OHS was derived) and the GOS.[9]

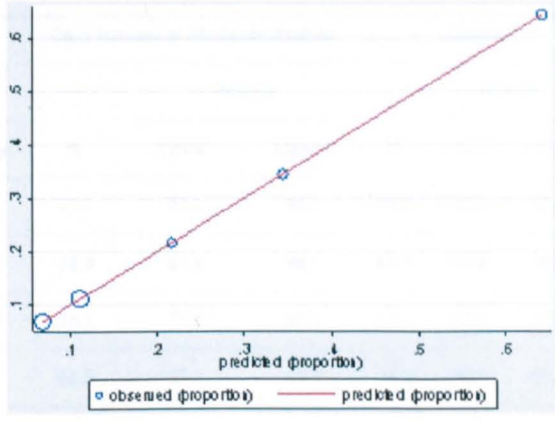


Figure 3
Calibration of model 1.

Conclusion

We have shown that OHS is strongly related and predicts accurately the GOS at 6 months. It may therefore be helpful in tackling the problem of missing data in clinical trials in TBI. It might also serve as a potential surrogate outcome measure and this application should be explored in further studies. If our findings are replicated, OHS could be a simple and useful outcome measure to use in trials in settings for which long term follow-up is problematic. Furthermore, OHS could be a useful variable to collect in rehabilitation trials in TBI patients to ensure that there is a similar distribution of disability among participants

Table 5: Association between OHS and unfavourable outcome (GOS) among survivors

	Model 1			*Model 2*			***Model 3***		
		95% CI			95% CI			**95% CI**	
OHS	**OR**	**Lower**	**Upper**	OR	Lower	Upper	**OR**	**Lower**	**Upper**
				1.0			1.0		
Minor symptoms	1.7	1.4	2.1	1.6	1.3	1.9	1.6	1.3	2.0
Some restriction in lifestyle but independent	3.7	3.0	4.7	2.7	2.1	3.4	2.7	2.1	3.5
Dependent but not requiring constant attention	7.1	5.7	8.8	4.5	3.6	5.7	4.7	3.7	6.0
Fully dependent requiring attention day & night	24.1	19.8	29.4	12.9	10.3	16.2	13.3	10.4	16.9
C statistic for the model	0.78			0.83			0.83		

Model 1: OHS
Model 2: model 1 plus GCS, pupil reactivity, major extra-cranial injury and age
Model 3: model 2 plus CT findings (petechial haemorrhages, obliteration of the third ventricle or basal cisterns, subarachnoid bleed, midline shift, non evacuated haematoma)

Table 6: Prediction of three dichotomies of GOS at 6 months according to OHS

	Good recovery			*Good recovery or Moderate disability*			*Survival*		
		95% CI			95% CI			95% CI	
OHS	%	Lower	Upper	%	Lower	Upper	%	Lower	Upper
No symptoms	79.2	77.6	80.8	93.1	92.1	94.1	99.3	99.0	99.6
Minor symptoms	**67.0**	65.1	68.8	**88.8**	87.6	90.1	**98.3**	97.8	98.8
Some restriction in lifestyle but independent	46.1	42.6	49.7	78.2	75.3	81.2	97.4	96.3	98.6
Dependent but not requiring constant attention	30.1	26.8	33.3	65.5	62.1	68.9	94.2	92.5	95.8
Fully dependent requiring attention day & night	13.6	11.5	15.6	35.8	32.9	38.6	77.6	75.1	80.0

between groups at baseline. We have also shown that, among survivors, the OHS is able to predict disability at 6 months and thus may provide a valuable clinical tool for physicians to improve communication with patients and relatives when assessing a patient's prognosis at hospital discharge.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

PP designed the study, performed the analysis and prepared the manuscript. IR conceived the study, revised and drafted the manuscript. PE and HS revised and drafted the manuscript.

Acknowledgements

The authors express their gratitude to all of the study participants and all the collaborators from the MRC CRASH Trial. We also want to thank Taemi Kawahara for helping with the tables and figures.

References

1. Bullock MR, Merchant RE, Choi SC, Gilman CB, Kreutzer JS, Marmarou A, Teasdale GM: **Outcome measures for clinical trials in neurotrauma.** *Neurosurg Focus* 2002, **13(1):**ECP1.
2. Dickinson K, Bunn F, Wentz R, Edwards P, Roberts I: **Size and quality of randomised controlled trials in head injury: review of published studies.** *BMJ* 2000, **320(7245):**1308-1311.
3. CRASH Trial Collaborators: **Final results of MRC CRASH, a randomised placebo-controlled trial of intravenous corticosteroid in adults with head injury-outcomes at 6 months.** *Lancet* 2005, **365(9475):**1957-1959.
4. CRASH Trial Collaborators: **Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients.** *BMJ* 2008, **12:**12.
5. Bamford JM, Sandercock PA, Warlow CP, Slattery J: **Interobserver agreement for the assessment of handicap in stroke patients.** *Stroke* 1989, **20(6):**828.
6. Jennett B, Bond M: **Assessment of outcome after severe brain damage.** *Lancet* 1975, **1(7905):**480-484.
7. CRASH Trial Collaborators: **Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial.** *Lancet* 2004, **364(9442):**1321-1328.
8. Maas AI, Marmarou A, Murray GD, Teasdale SG, Steyerberg EW: **Prognosis and clinical trial design in traumatic brain injury: the IMPACT study.** *J Neurotrauma* 2007, **24(2):**232-238.
9. Tilley BC, Marler J, Geller NL, Lu M, Legler J, Brott T, Lyden P, Grotta J: **Use of a global test for multiple outcomes in stroke trials with application to the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial.** *Stroke* 1996, **27(11):**2136-2142.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2288/8/72/prepub