Advancing the care for traumatic brain injury: summary results from the IMPACT studies and perspectives on future research

Andrew I.R. Maas, MD, PhD1, Gordon D. Murray, PhD2, Bob Roozenbeek, MD, PhD3,4, Hester F. Lingsma, PhD4, Isabella Butcher, PhD2, Gillian S. McHugh, PhD2, James Weir, MSc2, Juan Lu, MD, PhD5, and Ewout W. Steyerberg, PhD4 on behalf of the International Mission on Prognosis Analysis of Clinical Trials in Traumatic Brain Injury (IMPACT) Study Group

1Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium 2Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK 3Department of Neurology, Erasmus MC - University Medical Center Rotterdam, Rotterdam, The Netherlands 4Department of Public Health, Erasmus MC- University Medical Center Rotterdam, Rotterdam, The Netherlands 5Department of Epidemiology and Community Health, Virginia Commonwealth University, Richmond, Virginia, USA

Abstract

Research in Traumatic Brain Injury (TBI) is challenging because of many differences between patients. Advances in basic science have failed to translate into successful clinical treatments and the evidence underpinning guideline recommendations is low. Clinical research has been hampered by lack of standardized data collection, limited multidisciplinary collaboration and by insensitive approaches to classification and efficacy analyses. Multidisciplinary collaborations are now being fostered. Approaches for dealing with heterogeneity have been developed by the IMPACT study group. These can increase statistical power in clinical trials by up to 50% and are also relevant to other heterogeneous neurological diseases, such as stroke and subarachnoid hemorrhage. Rather than trying to limit heterogeneity, we may also be able to exploit it by analyzing differences in treatment and outcome between countries and centers in comparative effectiveness designs. This concept offers an additional research approach with great potential to advance the care in TBI.

Introduction

Traumatic Brain Injury (TBI) is a serious public health problem with an estimated annual incidence of up to 500/100,000 in the US and Europe.1,2 A recent population based study from New Zealand reported an annual incidence of 790/100,000.3 In low/middle income
countries most injuries result from road traffic incidents; in high income countries falls are now a more frequent cause of TBI and commonly occur in older patients. TBI constitutes a major cause of death and disability, leading to great personal suffering to victims and relatives and huge direct and indirect costs to society. In the US these annual costs are estimated at over 76.5 billion USD. Despite the magnitude of the socio-economic and medical problem posed by TBI, the strength of evidence underpinning treatment recommendations is low. Since the first publication of the guidelines on management of severe TBI in 1996 strong evidence in support of treatment recommendations has not been forthcoming. Conventional approaches to clinical TBI research have been reductionist, attempting to isolate out one single factor for treatment. These approaches have ignored the heterogeneity of TBI as a disease in terms of causes, pathophysiology, treatment and outcome. This heterogeneity makes research in TBI particularly challenging, and may partly explain why many randomized clinical trials (RCTs) have not shown statistically significant results. This aspect is being addressed in two ways: First, by applying novel methodology for dealing with the heterogeneity of TBI. The International Mission on Prognosis and Analysis of randomized Controlled Trials in TBI (IMPACT) study has provided novel methodology for dealing with the heterogeneity of TBI, offering the potential to increase statistical efficiency by up to 50%. Second, rather than dealing with heterogeneity in care paths, treatment and outcome, we can exploit it by employing Comparative Effectiveness Research (CER) paradigms. CER is the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers and policymakers to make informed decisions that will improve health care at both the individual and population levels. CER is broadly defined and may include pragmatic clinical trials. One approach to CER is to make use of existing differences in treatment and outcome between countries and centers with the aim of identifying best practices. Modern computational techniques and the availability of robust risk adjustment models facilitate such approaches, offering the potential to acquire high quality evidence in observational studies with greater generalizability. In this report we summarize the results of the IMPACT studies and discuss the potential of CER to provide evidence in support of care paths and treatment recommendations in TBI. A discussion of this potential is currently pertinent as large observational studies in TBI will soon be initiated in the context of an international collaboration established by the European Commission, the US National Institute of Neurological Disorders and Stroke (NIH-NINDS) and the Canadian Institutes of Health Research (CIHR).

The IMPACT studies

The IMPACT studies, funded by NIH-NINDS, were initiated in 2003. Over a 10-year period this international, multidisciplinary study group addressed methodological issues to improve the design and analysis of clinical trials in TBI. The IMPACT investigators were initially granted access to 11 large datasets of clinical trials (n=8) and observational studies (n=3) conducted in North America and Europe. Permission for accessing the datasets was obtained from principal investigators and where appropriate sponsoring companies. Over the course of the project, additional studies were added and collaborations were developed with the MRC CRASH trial investigators and the TARN registry. The available datasets were used to test and validate new approaches to trial design and analysis. As the project developed, three main directions of research evolved: 1) Standardization of data collection; 2) Prognostic analysis and development of prognostic models; 3) Improving the design and analysis of randomized clinical trials.
Standardization of data collection

Merging individual patient data for analysis across the constituent studies of the IMPACT database proved to be a major challenge. Not only did data field names and codings differ, but the structures of the datasets were complex and documentation poor. This experience highlighted the importance of consensus on a basic set of core variables to be collected in TBI studies, with agreement on appropriate definitions, field names and coding. Obtaining such agreement and consensus would expedite study designs, facilitate individual patient data analysis across studies, and also reduce costs to funding agencies and pharmaceutical companies. From this perspective the IMPACT investigators initiated a process of standardization of data collection in TBI studies. This process was taken forward in the context of an international and interagency initiative towards “an integrated approach to research and psychological health and traumatic brain injury” in the US. This initiative proposed common data elements (CDEs) and included proposals for definitions and coding of demographics, basic clinical data, biomarkers, neuroimaging, and outcome. The original intent was to focus on the most important variables for characterizing TBI populations, including established prognostic indicators. The process was consensus driven with multidisciplinary input from a broad range of experts, from emergency medicine to rehabilitation and late outpatient care. The Working Group on Demographics and Clinical Assessments recognized that the required level of detail for coding a variable may vary with the aim of a specific study. Thus, up to three versions for coding data elements were developed: a basic, an advanced, and an extended format with the greatest level of detail. The coding of these variables was such that more detailed coding could always be collapsed into the basic version, thus enabling comparisons across studies. In a second phase, the CDEs have been refined. Recently version 2 was released. This revised version also addressed variables for epidemiologic, post-acute care and outcome research. As a consequence, the number of variables has been expanded substantially resulting in a less user-friendly presentation. Regrettably, the initial focus on the most important elements has partly been lost. A broad discussion would appear appropriate as to whether the CDEs should be focused on the most important variables to be collected in all studies, or that a more inclusive approach should be taken in the context of a data dictionary. Version 2 of the CDEs represents a hybrid format of these approaches in which the most relevant variables for the different domains of TBI research are designated as core or basic. The TBI CDE effort is an evolving process and recent studies that have implemented the CDEs, such as TRACK-TBI, will likely provide data to resolve some of these issues for the anticipated Version 3. This revision should be informed by experience and evidence, rather than being based on consensus.

Prognostic analysis and development of prognostic models

Differences in patient population (case mix) may confound comparison of results between studies. In randomized controlled trials, the process of randomization seeks to achieve balance between treatment arms. However, an imbalance in cumulative prognostic risk between treatment arms may occur, despite only minor differences in individual characteristics. Quantification of the initial prognostic risk is therefore highly relevant. Many previous studies have reported on associations between predictors and outcome after TBI, but most have focused on univariate analyses in small sample sizes. The few studies that integrated predictors into a prognostic model to predict outcome on an individual patient basis had many methodological shortcomings, in particular the lack of external validation. The IMPACT database allowed for extensive prognostic analyses on large numbers (n>8000). Unique features included a systematic approach to the adjusted analyses of predictors, non-linear analysis of continuous predictors, and proportional odds analysis of the Glasgow Outcome Scale, rather than using a dichotomized analysis such as survival or unfavorable versus favorable outcome. Table 1 updates the Murray et al 2007.
overview of the prognostic strength of the most important predictors for outcome in TBI in larger numbers based on the 2013 version of the IMPACT database. The importance of adjustment for other predictors in multivariable analysis is well illustrated by the prognostic strength of cause of injury. Cause of injury was strongly associated with outcome in univariate analysis, with falls having a statistically significant higher risk of poorer outcome than road traffic incidents or other causes. However, after adjustment for age and other predictors in multivariable analysis, the association between cause of injury, specifically the occurrence of a fall, and outcome as demonstrated in univariate analysis is no longer found. Thus, the effect of cause of injury is confounded by the older age of patients sustaining falls. The analysis of blood pressure highlights the importance of continuous non-linear analysis of continuous predictors: both lower and higher blood pressures were related to poorer outcome, in a U-shaped relation (Figure 1). This relation would not have been observed if one cut off for blood pressure had been chosen.

The most important predictors of outcome were included in three prognostic logistic regression models of increasing complexity: a Core model based on demographics and injury severity; an Extended model additionally including CT information and second insults; and a Lab model additionally including glucose and Hemoglobin values. Predictions from the models for individual patients can be obtained online. Specifically, we note that the simple Core model (including age, motor score and pupillary reactivity) contains most of the prognostic information. These models were initially validated internally and externally in collaboration with the CRASH trial collaborators, and thereafter in various other datasets (Table 2). Although missingness of data should be avoided as far as possible in high quality studies, it does occur and has to be dealt with. Multiple imputation was used for dealing with missing covariates. Such imputation is considered more efficient than complete case analysis in which cases with incomplete data are dropped. Discrimination was assessed by the area under the receiver operating characteristic curve (AUC) and was typically around 0.7 to 0.8 in ten external validation studies. The variability in discriminatory performance was primarily related to variation in the case-mix in the validation sets, with better performance in more heterogeneous observational studies such as TARN and POCON (Table 2). This extensive validation illustrates the robustness of the IMPACT models and their generalizability across various settings. Nevertheless, limitations in the use and interpretation of the IMPACT models should be acknowledged:

- As in any prognostic model, the output of the calculation remains a probability estimate with an inherent degree of uncertainty. Thus, particular care should be taken in interpreting prognostic estimates in individual patients.

- The focus of the IMPACT prognostic analysis was on establishment of the baseline prognostic risk and the studies did not include dynamic predictions, including new information as it becomes available over the course of the disease process.

- The IMPACT studies were limited by the selection and detail of predictors which had been collected in previous studies. Some, possibly relevant, predictors could not be analyzed in adequate detail due to relatively low number of patients in which these had been collected. Examples include details on coagulation status and the presence and severity of extracranial injuries.

- The IMPACT dataset did not include patients with mild traumatic brain injury, and the IMPACT models are consequently not valid for mild TBI. Despite the fact that up to 95% of TBIs are mild, only one prognostic model has been developed specifically for mild TBI.
- A further possible selection bias may have been introduced by the lack of population based studies. This risk is however considered low as the focus of the IMPACT prognostic analysis was more on patients with severe and moderate TBI. Population based studies are more relevant to mild TBI as the majority of these patients are not seen in the hospital setting.

**Improving the design and analysis of randomized controlled trials**

In TBI, nearly all RCTs have attempted to decrease heterogeneity by applying strict enrolment criteria or by targeting patients with an intermediate prognosis at randomization. A contrasting approach was followed in the CRASH mega trial, where large numbers would overcome problems caused by prognostic heterogeneity, and increase generalizability. The relative efficacy of these approaches was explored systematically in the IMPACT project. Simulation studies showed that exclusion of patients with an extreme prognosis – by the use of strict enrollment criteria or by prognostic targeting – indeed increases statistical power. However, as a result of this strict selection many patients are excluded from study participation, which limits the generalizability of the results. Moreover, strict selection will reduce the recruitment rate, thereby prolonging study duration. Beneficial effects in terms of increased statistical power of strict selection need to be balanced against adverse effects on recruitment. Assuming a uniform treatment effect, this balance was unfavorable for the execution of studies with such restrictive enrollment criteria.

Despite the fact that conditional estimation of treatment effects, also by non-normal regression models such as logistic or Cox regression models, is more powerful than unadjusted estimation, relatively few RCTs have dealt with the problem of heterogeneity by adjusting the treatment effect for important predictors of outcome using covariate adjustment. Simulation studies with TBI trial data showed that covariate adjustment for seven strong predictors of outcome increase statistical efficiency up to 30% in more heterogeneous populations of observational surveys and up to 16% increase in trial populations that initially used stricter enrollment criteria. In a re-analysis of the CRASH trial data, covariate adjustment for age, GCS motor score and pupillary reactivity reduced the required sample size by 21% to obtain the same statistical power compared to the unadjusted analysis. Covariate adjustment should be pre-specified and include established strong predictors for outcome. We note that risk adjustment models may need to be updated as newer prognostic information becomes available.

When testing efficacy of a new therapy the aim is to prove that the new treatment yields better results than placebo treatment or conventional management, in other words that patients will have a better outcome than expected. The primary outcome measure in the majority of RCTs for TBI is the Glasgow Outcome Scale (GOS) or its extended version (GOSE). It is common practice to dichotomize this scale into a favorable versus an unfavorable outcome. However, the practice of dichotomization is clinically unattractive and statistically inefficient. The prognosis of patients at the extreme ends of the outcome distribution can be so good that they will almost inevitably achieve a favorable outcome, even without the benefits of an effective therapeutic intervention, or so poor that it is unlikely that even an effective intervention would improve their outcome to such an extent that it would move from being unfavorable to favorable. Moreover, focusing only on one specific split of the outcome scale ignores the fact that other transitions of the outcome are clinically relevant. We considered two novel approaches to ordinal efficacy analysis: the sliding dichotomy approach and proportional odds regression (Figure 2). With the sliding dichotomy approach, the point of the dichotomy is differentiated according to the baseline prognostic risk. For patients with a poor prognosis, survival may be most relevant, while in those with a good prognosis any outcome worse than good recovery may be considered.
unfavorable. Proportional odds regression considers all possible ways in which the ordinal scale can be dichotomized, assuming that the odds ratio for a better versus worse outcome is identical wherever the scale is dichotomized (the ‘proportional odds assumption’). Conceptually, the model combines all potential splits to estimate an overall effect measure: the common odds ratio. This common odds ratio can be interpreted as the odds for a shift in outcome across the full ordinal scale.\textsuperscript{43}

Simulation studies with the IMPACT database showed that ordinal approaches to the efficacy analysis reduced required sample sizes by 23 to 30\%, compared to the traditional dichotomized analysis.\textsuperscript{21} These gains were consistent across studies and, remarkably, also remained if the proportional odds assumption was violated. Applying covariate adjustment together with ordinal analysis reduced the sample size requirements by up to 50\%. These findings were confirmed when testing these approaches on the CRASH trial\textsuperscript{44}; Combining ordinal analysis with covariate adjustment increased statistical efficiency by approximately 50\% and statistically significant effects were already present after enrolment of approximately 50\% of the population\textsuperscript{44}. The combined results of the simulation studies and empirical proof of effectiveness of the IMPACT recommendations in CRASH provide strong support to incorporate these approaches in the design of new clinical trials in the field of TBI (Panel 1).

**Perspectives on future research**

The IMPACT studies illustrate how international and multidisciplinary collaboration can accelerate research and how methodological research can lead directly to improved clinical research. The recent institution of the International Initiative for Traumatic Brain Injury Research (InTBIR) as a collaboration between funding agencies (EC, NIH-NINDS and CIHR) represents a milestone accomplishment and provides a platform for global collaboration in TBI research\textsuperscript{45,46}.

**Standards for data collection and prognostic research**

The concepts developed by the IMPACT study group are being taken forward in TBI research. Use of common data elements is currently required in all observational studies and trials on TBI funded by NIH-NINDS. A recent call by the European Commission also mandated use of core common data elements.\textsuperscript{47} This adoption of CDEs by funding agencies may be expected to facilitate comparisons between studies, meta-analyses of individual patient data across studies and, importantly, will reduce costs for designing Case Report Forms for new studies. From a global perspective, the generalizability for use of the CDEs across different settings and empirical experience in their use should form the basis for further refinements where compromises may need to be made between international generalizability and a national/local focus. The differentiation of coding into three levels of detail may provide opportunities for harmonizing these two perspectives. Importantly, the CDEs should be presented in a user-friendly format.

The IMPACT prognosis studies have been instrumental in developing and setting standards for prognostic research in TBI. The predictive effects of many known prognostic variables have been confirmed in much larger numbers than before and novel predictors were identified. The development of the IMPACT prognostic models for severe and moderate TBI has provided opportunities for summarizing the baseline prognostic risks in study populations and can be used as robust risk adjustment models. They have already been widely adopted in TBI research.\textsuperscript{48} A major limitation is that these models were not developed for mild TBI. The CRASH prognostic models\textsuperscript{49} included patients with milder injuries in the development and may consequently have broader generalizability across the spectrum of severity. Prognostic models should never be considered final, but will require
continuous updating and evaluation/validation. In this context the added value of new predictors should be taken into consideration. New candidate predictors would include gene signature, biomarkers, coagulation parameters, and the presence of systemic injuries. The application of prognostic models in TBI is broader than clinical trial design and may be used towards classification and benchmarking the quality of health care delivery in TBI. 48

Dealing with heterogeneity in clinical trials

The IMPACT recommendations for trial design have also been widely adopted. The Pharmos Dexanabinol trial 50 was one of the first in TBI to use covariate adjustment and proportional odds approaches in the efficacy analysis. Since then, many completed and ongoing studies have adopted (parts of) the IMPACT recommendations. These studies include the DECRA trial on decompressive craniectomy in patients with diffuse traumatic brain injury 51 and the ongoing EuroTherm (therapeutic hypothermia for TBI), SyNAPSe® and PROTECT III trials on progesterone for severe TBI.

Prognostic heterogeneity of patient populations does not only apply to TBI, but also to other neurological diseases (e.g., ischemic stroke, subarachnoid hemorrhage (SAH), intracerebral hemorrhage and the Guillain-Barré syndrome) and to other fields of medicine. 52,53,37,54 The methodology to deal with heterogeneity in randomized controlled trials is essential for all these fields. In SAH, the example of IMPACT is being followed in the establishment of the Subarachnoid Hemorrhage International Trialists (SAHIT) data repository aiming to optimize the design and analysis of phase III trials in aneurysmal SAH. 55 As in TBI, most large randomized controlled trials in acute ischemic stroke have been neutral. 56 Stroke and TBI populations both have substantial prognostic heterogeneity in the population and an ordinal outcome measure is generally used (modified Rankin Scale), which is often dichotomized for the analysis. 57,58 A recent paper advises tailoring the approach to the analysis of the treatment effect to each individual trial, based on how the intervention under study is most likely to modify the distribution of outcomes, recognizing that ordinal analysis should be preferred over the traditional dichotomy, and covariate adjustment used. 59 SCAST 60, IST-3 61 and INTERACT2 62 were three stroke trials where the results were only significant when ordinal analysis was used. Several acute stroke trials were published that have used different aspects of the methodology consistent with the IMPACT recommendations; 63,64,65,66,67 Other ongoing stroke trials, such as STASH (statins for subarachnoid hemorrhage), EuroHYP (hypothermia for acute ischemic stroke), MR CLEAN (endovascular treatment for acute ischemic stroke) and others, have planned to adopt (parts) of the recommendations.

Exploiting heterogeneity in Comparative Effectiveness Research

We should recognize that many other issues (including deficiencies in pre-clinical studies and early clinical work up as well as uncertainty on time windows and dosing) than the heterogeneity addressed by the IMPACT studies have contributed to the disappointments in clinical TBI trials. These aspects have not been addressed by the IMPACT study group, but have been previously reviewed in detail. 68,69,70 Neither have the IMPACT studies addressed the problem of heterogeneity related to mechanism. Early mechanistic endpoints, which can serve as intermediate outcomes in TBI trials, are still lacking. We further recognize that it will be impossible to mount a sufficient number of adequately powered clinical trials to address all existing uncertainties in the management of TBI. Major advances in the care for TBI patients have not come from clinical trials, but rather from observational studies and guideline developments. 7 Randomized controlled trials are not the only source of high quality evidence to support practice recommendations. Alternative designs may be considered in a Comparative Effectiveness Research (CER) framework. CER is not new to TBI. Studies that compared treatment and outcome between centers in the 1980s would
There are several features of TBI that would favor CER approaches:

- There are large between centre differences and between country differences in management and outcome. In the IMPACT studies, analyzing 9578 moderate and severe TBI patients from 265 centers, we found a 3.3 fold difference in the odds of having unfavorable outcome at 6 months between very good and very poor centers (2.5 vs. 97.5 percentile) after adjustment for chance effects and for differences in case-mix. In CRASH, differences were even larger with a 6.6-fold between-centre difference in 14-day mortality and a 15-fold difference between countries.

- Robust risk adjustment models, specific for TBI, are available to adjust for differences in major prognostic factors (case-mix).

- Advanced statistical methods, including random effect models, have been field tested for TBI.

CER offers opportunities to exploit the existing heterogeneity and differences between countries, centers and patients in TBI to identify best practices. This approach requires high quality contemporaneous observational data, for which calls were recently published in the EU FP7 program and US NIH programs. These calls seek to establish high quality contemporary observational data as a basis for CER studies. This is important as, although methodology for CER studies is available, large-scale observational studies on TBI date back at least 20 years. These calls will lead to research to better characterize TBI as a disease, and identify the most effective clinical interventions for managing TBI. Better characterization will facilitate Personalized Medicine approaches as recently advocated by the National Academy of Science. Phenotypic heterogeneity may interact critically with genetics and exploring this will require studies in large patient numbers. Novel information will come forward on disease processes, treatment, outcome and prognosis in TBI, whilst the establishment of bio-repositories on neuro-imaging, genetics and biomarkers will ensure opportunities for future research including legacy research. Data sharing policies will need to be established to encourage academic productivity and to accelerate TBI research. Initiatives, currently being developed within the context of InTBIR, are summarized in panel 2. These initiatives have each a different focus and, although “stand alone” analysis is expected to yield important contributions, the major benefit will most likely result from integration of analyses across studies: InTBIR is more than the sum of its parts.

These initiatives and the establishment of InTBIR illustrate a shift in TBI research towards international and multidisciplinary collaborations, which bridge the traditional disconnection between acute and post-acute research. We further note a shift from current reductionistic approaches in clinical research towards broader approaches with greater generalizability. It should be recognized, however, that RCTs remain the preferred approach for evaluating efficacy of novel treatment approaches. We hope that the design of such trials will be influenced by the results from the IMPACT studies, such as to optimize assessment of treatment effects.

**Summary and conclusions**

The landscape of TBI research is changing. Broad based, sustainable, multidisciplinary and international approaches are required to address the complexity of TBI. Large international collaborations of not only researchers but also funding agencies are currently being implemented, as exemplified by InTBIR. Disconnects in research between acute and post-acute care settings, between research in milder TBI and more severe injuries requiring hospital admission are being repaired. The disappointing results of most clinical trials in TBI...
have led to a reappraisal of pre-clinical work-up and clinical trial methodology. Weaknesses of previous pre-clinical studies include poor design, use of experimental models and paradigms that fail to incorporate key elements germane to human severe TBI, insufficient pre-clinical testing before proceeding into clinical trials, insufficient attention for studies on brain penetration and pharmacokinetics in clinical studies, as well as publication bias for positive results. Methodological challenges in clinical trials, in particular challenges posed by the inherent heterogeneity of the patient population, have been addressed in the IMPACT studies. These have resulted in recommendations which have the potential to increase statistical efficiency by up to 50%. This is a major advance offering better chances for demonstrating efficacy of new treatments in the context of randomized controlled trials. The IMPACT recommendations have broad applicability and principles of the recommendations are also being taken over in the fields of stroke and subarachnoid hemorrhage. International consensus on standardization of data collection is being sought in the development of common data elements which will facilitate comparability between studies and meta-analyses of individual patient data across studies in large numbers. Reductionist approaches originating from traditional research paradigms in which single factors are isolated and targeted are slowly being replaced by more holistic approaches more representative of the clinical situation. In this context, new dimensions are being added to TBI research in the form of comparative effectiveness approaches. The potential of these approaches is now being recognized by funding agencies as evidenced by recent calls in the European Commission FP7 and NIH programs. Improved clinical trial methodology and exploiting the heterogeneity of TBI in the context of comparative effectiveness research holds great promise for advancing the care for traumatic brain injury.

Acknowledgments

The authors gratefully acknowledge the collaboration within the IMPACT study group from which this manuscript originates. Further, the authors wish to express their gratitude for the administrative support in preparing this manuscript provided by Véronique De Keyser.

This work was supported by the NIH/National Institute of Neurological Disorders and Stroke (NS-042691).

References

6. Manley GT, Maas AI. Traumatic Brain Injury: An international knowledge-based approach. JAMA. 2013 Accepted for publication.


Panel: Search strategy and selection criteria

References for this Personal View were identified through searches of PubMed, by use of (combinations of) the search terms “traumatic brain injury”, “prognostic models”, “heterogeneity”, “clinical trial design”, “clinical trial analysis”, “comparative effectiveness research” and other appropriate terms up to May, 2013. Papers were also identified from the authors’ own files and from references cited in relevant articles. We considered only publications written in English. The final reference list was generated on the basis of relevance to the topics covered in this Personal View.
Panel 1. Recommendations for design and analysis of randomized controlled trials for traumatic brain injury.\textsuperscript{8}

- Details of the major baseline prognostic characteristics should be provided in every report on a TBI study; in trials they should be differentiated per treatment group. We also advocate the reporting of a summary of the baseline prognostic risk as determined by validated prognostic models.
- Inclusion criteria should be as broad as is compatible with the current understanding of the mechanisms of action of the intervention being evaluated. This will maximize recruitment rates and enhance the generalizability of the results.
- The statistical analysis should incorporate (prespecified) covariate adjustment to mitigate the effects of heterogeneity.
- The statistical analysis should use an ordinal approach, based on either sliding dichotomy or proportional odds methodology.
<table>
<thead>
<tr>
<th>Acronym/short title</th>
<th>Institution and Principal Investigator</th>
<th>Project</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>European Commission(EU)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CENTER-TBI</td>
<td>Antwerp University Hospital (A. Maas) University of Cambridge (D. Menon)</td>
<td>Collaborative European NeuroTrauma Effectiveness Research in TBI</td>
</tr>
<tr>
<td>CREACTIVE</td>
<td>IRCCS - Istituto di Ricerche Farmacologiche Mario Negri Milano (G. Bertolini)</td>
<td>Collaborative Research on Acute Traumatic Brain Injury in Intensive Care Medicine in Europe</td>
</tr>
<tr>
<td><strong>NIH (US)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAPT trial</td>
<td>Children’s hospital of Pittsburgh (M. Bell)</td>
<td>Approaches and Decisions for Acute Pediatric TBI</td>
</tr>
<tr>
<td>TRACK-TBI</td>
<td>University of California (G. Manley)</td>
<td>Transforming Research and Clinical Knowledge in Traumatic Brain Injury</td>
</tr>
<tr>
<td><strong>CIHR (Canada)</strong>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safe to Play</td>
<td>Hotchkiss Brain Institute, University of Calgary (C. Emery)</td>
<td>A longitudinal research program to establish best practice in the prevention, early diagnosis and management of sport-related concussion in youth ice hockey players</td>
</tr>
<tr>
<td>Innovation through the use of common data</td>
<td>McGill University (I. Gagnon)</td>
<td>Generating innovation through the use of common data: Improving the diagnosis and treatment of pediatric and adolescent mild traumatic brain injury in Canada.</td>
</tr>
<tr>
<td>Play Game</td>
<td>University of Calgary (K. Barlow)</td>
<td>Post-concussion syndrome Affecting Youth: GABAergic effects of Melatonin</td>
</tr>
<tr>
<td>Post concussion problems in pediatric TBI</td>
<td>Children’s Hospital of Eastern Ontario, University of Ottawa (R. Zemek)</td>
<td>Predicting Persistent Postconcussive Problems in Pediatrics (5P)</td>
</tr>
<tr>
<td>&quot;NeuroCare” as Innovation in Intervention</td>
<td>University of Toronto (M. Keightley)</td>
<td>A Neurophysiological Approach to Determine Readiness for Return to Activity</td>
</tr>
<tr>
<td><strong>15 Catalyst Grants (1-year duration)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>3 New Post-doctoral awards</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Co-funding partners of the Canadian Institutes of Health Research (CIHR) for the Team Grants are: Fonds de recherche du Québec Santé; Hotchkiss Brain Institute; Ontario Brain Institute; Ontario Neurotrauma Foundation.
Figure 1. Relationship between systolic blood pressure and outcome (n=8172)
X axis: systolic blood pressure at admission
Y axis: linear predictor of unfavorable outcome (GOS 1–3). A higher linear predictor corresponds to a higher probability of unfavorable outcome.
Figure 2.
Graphical illustration of three approaches to the analysis of the Glasgow Outcome Scale as the primary outcome measure of randomized controlled trials for traumatic brain injury. The traditional approach (A) to the efficacy analysis in a clinical traumatic brain injury trial is to dichotomize the Glasgow Outcome Scale (GOS) into unfavorable (Dead, Vegetative State, Severe Disability) versus favorable outcome (Moderate Disability, Good Recovery). The proportions of patients with an unfavorable outcome in the treatment and placebo groups are compared, by calculating an odds ratio with logistic regression analysis. With the sliding dichotomy approach (B), the study population is first subdivided in (for example three) equally large prognostic risk groups. For each of the risk groups, the point of the dichotomy of the GOS is based on the baseline prognostic risk (e.g., for patients in the good prognostic risk group, only Good Recovery is considered a favorable outcome). A pooled odds ratio is calculated, which can be interpreted as the summary measure for having a better outcome than expected. With the proportional odds approach (C) the population is not subdivided. The proportional odds model considers every possible way the GOS can be dichotomized, assuming that the odds ratio for a better versus a worse outcome is similar wherever the GOS is dichotomized (the proportional odds assumption). The common odds ratio can be interpreted as a summary measure for the shift in outcome across the full GOS.

Abbreviations: D= Dead, VS= Vegetative State, SD= Severe Disability, MD= Moderate Disability, GR= Good Recovery, OR= odds ratio.
Table 1
Overview of the prognostic strength of 28 predictors in TBI, expressed as proportional odds ratios for 6 month GOS.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reference / Category</th>
<th>UNIVARIATE ANALYSIS</th>
<th>ADJUSTED ANALYSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Odds ratio (95% CI)</td>
<td>N</td>
</tr>
<tr>
<td>Predictors included in IMPACT model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>46 vs 22</td>
<td>11022</td>
<td>2,09 (1.96 – 2.22)</td>
</tr>
<tr>
<td>GCS motor score</td>
<td>Localises/ obeys</td>
<td>11383</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>4,34 (3.03 – 6.23)</td>
<td>3.31 (2.52 – 4.34)</td>
</tr>
<tr>
<td></td>
<td>Extension</td>
<td>6.75 (5.27 – 8.63)</td>
<td>5.13 (4.10 – 6.42)</td>
</tr>
<tr>
<td></td>
<td>Abnormal flexion</td>
<td>3.43 (2.73 – 4.30)</td>
<td>2.84 (2.34 – 3.44)</td>
</tr>
<tr>
<td></td>
<td>Normal flexion</td>
<td>1.71 (1.46 – 2.02)</td>
<td>1.59 (1.34 – 1.90)</td>
</tr>
<tr>
<td></td>
<td>Missing/untestable</td>
<td>2.29 (1.73 – 3.01)</td>
<td>2.02 (1.58 – 2.58)</td>
</tr>
<tr>
<td>Pupil response</td>
<td>Both reacting</td>
<td>9830</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>One reacting</td>
<td>2.69 (2.37 – 3.06)</td>
<td>2.2 (1.90 – 2.56)</td>
</tr>
<tr>
<td></td>
<td>Neither reacting</td>
<td>6.62 (4.85 – 9.06)</td>
<td>4.33 (3.45 – 5.44)</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>No</td>
<td>8195</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Suspected/definite</td>
<td>1.91 (1.56 – 2.35)</td>
<td>1.48 (1.27 – 1.72)</td>
</tr>
<tr>
<td>Hypotension</td>
<td>No</td>
<td>9191</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Suspected/definite</td>
<td>2.44 (1.96 – 3.03)</td>
<td>1.86 (1.54 – 2.24)</td>
</tr>
<tr>
<td>CT class</td>
<td>No visible pathology/diffuse</td>
<td>6407</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Swelling/shift</td>
<td>2.68 (2.08 – 3.46)</td>
<td>2.32 (1.84 – 2.92)</td>
</tr>
<tr>
<td></td>
<td>Mass lesion</td>
<td>2.36 (1.93 – 2.90)</td>
<td>1.63 (1.42 – 1.88)</td>
</tr>
<tr>
<td>tSAH</td>
<td>No</td>
<td>9184</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>2.69 (2.42 – 2.99)</td>
<td>2 (1.84 – 2.18)</td>
</tr>
<tr>
<td>Haemoglobin</td>
<td>14.2 vs 10.9 Hb (g/dL)</td>
<td>4957</td>
<td>0.69 (0.62 – 0.77)</td>
</tr>
<tr>
<td>Glucose</td>
<td>10.2 vs 6.6 (mMol/l)</td>
<td>5889</td>
<td>1.64 (1.52 – 1.77)</td>
</tr>
</tbody>
</table>

Other Predictors
### Variable | Reference / Category | UNIVARIATE ANALYSIS | ADJUSTED ANALYSIS
--- | --- | --- | ---
**Gender** |  |  |  
Male | 10379 | 1 (reference) | 10021 | 1 (reference)  
Female | 1,02 (0.94 – 1.10) |
**Race** |  |  |  
Caucasian | 7075 | 1 (reference) | 7070 | 1 (reference)  
Black | 1,37 (1.16 – 1.61) |
Asian | 1,21 (0.91 – 1.62) |
Other | 1,1 (0.90 – 1.34) |
**Education** |  |  |  
0–8 years | 2589 | 1 (reference) | 2555 | 1 (reference)  
9–12 years | 0,82 (0.64 – 1.06) |
Over 12 years | 0,74 (0.56 – 0.98) |
**Cause of injury** |  |  |  
Fall | 11363 | 1 (reference) | 11005 | 1 (reference)  
Road traffic incident | 0,71 (0.64 – 0.80) |
Assault | 0,68 (0.56 – 0.81) |
Work–related | 0,94 (0.75 – 1.18) |
Sports/recreation | 0,47 (0.32 – 0.68) |
Other | 0,85 (0.71 – 1.01) |
**GCS eye score** |  |  |  
Pain/sound/spontaneous | 11383 | 1 (reference) | 10637 | 1 (reference)  
None | 2,69 (2.18 – 3.31) |
Missing/untestable | 2,3 (1.71 – 3.09) |
**GCS verbal score** |  |  |  
Sounds-orientated | 11383 | 1 (reference) | 10637 | 1 (reference)  
None | 2,43 (1.95 – 3.02) |
Missing/untestable | 2,6 (2.22 – 3.04) |
**Systolic BP** | 85–150 mm Hg |  |  
<120 mm Hg | 8172 | 1 (reference) | 8168 | 1 (reference)  
120–150 mm Hg | 1,48 (1.29 – 1.70) |
>150 mm Hg | 1,37 (1.19 – 1.57) |
**Mean arterial BP** | 85–110 mm Hg |  |  
<85 mm Hg | 8250 | 1 (reference) | 8245 | 1 (reference)  
85–110 mm Hg | 1,28 (1.13 – 1.45) |
>110 mm Hg | 1,35 (1.16 – 1.58) |
### UNIVARIATE ANALYSIS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Reference / Category</th>
<th>N</th>
<th>Odds ratio (95% CI)</th>
<th>N</th>
<th>Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothermia</td>
<td>No</td>
<td>5034</td>
<td>1 (reference)</td>
<td>5017</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Suspected/definite</td>
<td></td>
<td>2.01 (1.47 – 2.75)</td>
<td>1.59</td>
<td>(1.18 – 2.14)</td>
</tr>
<tr>
<td>Cisterns</td>
<td>Present</td>
<td>5233</td>
<td>1 (reference)</td>
<td>5229</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Compressed/absent</td>
<td></td>
<td>2.49 (2.05 – 3.04)</td>
<td>1.95</td>
<td>(1.68 – 2.25)</td>
</tr>
<tr>
<td>Shift</td>
<td>No</td>
<td>5546</td>
<td>1 (reference)</td>
<td>5542</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>1–5 mm</td>
<td></td>
<td>1.33 (1.11 – 1.59)</td>
<td>1.3</td>
<td>(1.12 – 1.52)</td>
</tr>
<tr>
<td></td>
<td>&gt; 5 mm</td>
<td></td>
<td>2.16 (1.68 – 2.77)</td>
<td>1.37</td>
<td>(1.06 – 1.77)</td>
</tr>
<tr>
<td>EDH</td>
<td>No</td>
<td>9872</td>
<td>1 (reference)</td>
<td>9531</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
<td>0.65 (0.58 – 0.72)</td>
<td>0.67</td>
<td>(0.59 – 0.75)</td>
</tr>
<tr>
<td>SDH</td>
<td>No</td>
<td>9881</td>
<td>1 (reference)</td>
<td>9540</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
<td>2.26 (1.85 – 2.75)</td>
<td>1.4</td>
<td>(1.24 – 1.57)</td>
</tr>
<tr>
<td>Contusion</td>
<td>No</td>
<td>8953</td>
<td>1 (reference)</td>
<td>8761</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td></td>
<td>1.39 (1.16 – 1.67)</td>
<td>1.39</td>
<td>(1.19 – 1.61)</td>
</tr>
<tr>
<td>Sodium</td>
<td>137–142 mmol/L</td>
<td>6335</td>
<td>1 (reference)</td>
<td>6331</td>
<td>1 (reference)</td>
</tr>
<tr>
<td></td>
<td>&lt;137 mmol/L</td>
<td></td>
<td>1.38 (1.23 – 1.55)</td>
<td>1.16</td>
<td>(0.97 – 1.39)</td>
</tr>
<tr>
<td></td>
<td>&gt;142 mmol/L</td>
<td></td>
<td>1.15 (0.99 – 1.34)</td>
<td>1.18</td>
<td>(1.03 – 1.35)</td>
</tr>
<tr>
<td>pH</td>
<td>7.45 vs 7.32</td>
<td>4268</td>
<td>0.81 (0.75 – 0.88)</td>
<td>4264</td>
<td>0.84 (0.78 – 0.92)</td>
</tr>
<tr>
<td>Haematocrit</td>
<td>42.0 vs 32.6</td>
<td>1914</td>
<td>0.71 (0.60 – 0.83)</td>
<td>0.83</td>
<td>(0.71 – 0.97)</td>
</tr>
<tr>
<td>Platelets</td>
<td>253 vs 154 (10^9/L)</td>
<td>2466</td>
<td>0.76 (0.64 – 0.92)</td>
<td>2466</td>
<td>0.85 (0.71 – 1.02)</td>
</tr>
<tr>
<td>Prothrombin time</td>
<td>14.4 vs 12.1/sec</td>
<td>1032</td>
<td>1.47 (1.13 – 1.92)</td>
<td>1032</td>
<td>1.57 (1.28 – 1.92)</td>
</tr>
</tbody>
</table>

Multivariable logistic regression analysis was performed on the association between the prognostic factor of interest, with and without adjustment for the core predictors: age, GCS motor score and pupillary reactivity. The numbers in the adjusted analysis column show availability of each covariate in adults with non-missing outcome in the overall dataset. In the adjusted analyses missing values for pupillary response were replaced by imputed values. For the continuous predictors with a linear relation to outcome, the odds ratios were scaled so that they correspond to changing from the 25th percentile to the 75th percentile (interquartile range, IQR). The IQR is reported for each continuous variable. An odds ratio > 1 for a continuous prognostic indicates that the risk of a poor outcome increases as the variable increases over the interquartile range. Abbreviations: CT, computerized tomography; tSAH, traumatic subarachnoid haemorrhage; EDH, epidural hematoma; SDH, subdural hematoma; GCS, Glasgow come scale; BP, blood pressure.
* Adjusted for GCS Motor and pupillary response;
** Adjusted for age and pupillary response;
*** Adjusted for age and GCS Motor
Overview of the discriminative ability of the IMPACT prognostic models in 9 external validation studies, expressed as the area under the receiver operating characteristic curve (AUC) for 6-month mortality and unfavorable outcome

<table>
<thead>
<tr>
<th>Dataset ***/author</th>
<th>Sample size</th>
<th>Time period</th>
<th>Study design</th>
<th>Reference validation</th>
<th>Mortality</th>
<th>Unfavourable outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Core**</td>
<td>Extended**</td>
</tr>
<tr>
<td>CRASH **</td>
<td>n=6272</td>
<td>1999–2004</td>
<td>RCT</td>
<td>Steyerberg et al. 2008 **</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>TBI-TRAC **</td>
<td>n=2513</td>
<td>2000–2009</td>
<td>Obs study</td>
<td>Roozenbeek et al., 2012a **</td>
<td>0.79 *</td>
<td>0.83 *</td>
</tr>
<tr>
<td>APOE **</td>
<td>n=404</td>
<td>1996–1999</td>
<td>Obs study</td>
<td></td>
<td>0.81</td>
<td>0.80</td>
</tr>
<tr>
<td>TARN TBI **</td>
<td>n=6874</td>
<td>1989–2009</td>
<td>Obs study</td>
<td></td>
<td>0.83</td>
<td>0.86</td>
</tr>
<tr>
<td>NABIS Hypothermia **</td>
<td>n=385</td>
<td>1994–1998</td>
<td>RCT</td>
<td>Roozenbeek et al. 2012b **</td>
<td>0.70</td>
<td>0.74</td>
</tr>
<tr>
<td>Cerestat</td>
<td>n=517</td>
<td>1996–1997</td>
<td>RCT</td>
<td></td>
<td>0.75</td>
<td>0.76</td>
</tr>
<tr>
<td>Pharmos **</td>
<td>n=856</td>
<td>2001–2004</td>
<td>RCT</td>
<td></td>
<td>0.65</td>
<td>0.71</td>
</tr>
<tr>
<td>POCON **</td>
<td>n=415</td>
<td>2008–2009</td>
<td>Obs study</td>
<td>Lingsma et al., 2012 **</td>
<td>0.85</td>
<td>0.88</td>
</tr>
<tr>
<td>Panczykowski et al. **</td>
<td>n=587</td>
<td>1994–2009</td>
<td>Obs study</td>
<td>Panczykowski et al., 2012 **</td>
<td>0.78</td>
<td>0.83</td>
</tr>
<tr>
<td>Raj et al. **</td>
<td>n=342</td>
<td>2009–2010</td>
<td>Obs study</td>
<td>Raj et al., 2013 **</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

RCT=randomized controlled trials; Obs=observational; NA=not available

* 14 day outcome

** Models:
Core: age, GCS motor score, pupillary reactivity
Extended: core + CT information (CT classification and tSAH), second insults (hypoxia and hypotension)
Lab: Extended + glucose, hemoglobin

*** Datasets:
CRASH: Corticosteroid Randomization After Significant Head Injury
TBI-TRAC: Database of the Brain Trauma Foundation in New York for tracking the treatment of severe TBI patients
APOE: Apolipoprotein E single-center observational cohort study
TARN TBI: Trauma Audit and Registry Network TBI study
NABIS: National Acute Brain Injury Study
Cerestat: Randomized controlled trial on a non-competitive NMDA antagonist