Abstract

Objectives—To assess the prognostic value of summary measures of secondary physiological insults in addition to baseline clinical variables for patients with traumatic brain injury.

Methods—A series of 110 patients with traumatic brain injury had data on intracranial pressure (ICP), arterial blood pressure (ABP), cerebral perfusion pressure (CPP), arterial O₂ saturation (SaO₂), temperature in °C (Temp), and heart rate in beats/min (HRT) monitored and recorded every minute. Secondary insults were defined according to the Edinburgh University secondary insult grading system. The prognostic value of summary measures of these secondary insults was assessed by adding them to a prognostic model for survival at 1 year after controlling for baseline clinical variables using a previously validated model.

Results—Of the eight secondary insults measured, only ICP added significantly to the prediction of survival in the first 72 hours after injury. The particular type of summary measure did not seem to influence the results. After the addition of ICP to the model, none of the other secondary insult measures could improve the predictive power of the model significantly.

Conclusions—Early intracranial hypertension is confirmed as a sign of poor prognosis in patients with traumatic brain injury, even after controlling for baseline clinical variables. The value or otherwise of treating such secondary insults, however, can only be definitively established in the context of prospective randomised controlled trials. The specific pathophysiological evolution of secondary insults is still the subject of much research, and a clear understanding will be necessary before the development of specific treatments is feasible.

(J Neurol Neurosurg Psychiatry 1999;66:26–31)

Keywords: traumatic brain injury; prognosis; survival; secondary insults; adults

Traumatic brain injury is a major cause of death and disability among a predominantly young population. Recent estimates of incidence are roughly 10 million cases/year worldwide.¹ Current understanding of the pathophysiology of traumatic brain injury stresses the importance not only of the primary injury, but also of the secondary processes occurring after injury which may lead to cerebral hypoxia and ischaemia.² Episodes of physiological derangement, often called secondary insults, occur frequently, often go unrecorded by standard clinical monitoring procedures,³ and correlate well with poor functional outcome.⁴ The automatic collection, processing, and subsequent analysis of physiological data during the acute treatment phase of patients with traumatic brain injury is a highly complex problem, combining aspects of standard prospective study methodology with that of observational epidemiology, and the technical and logistical problems associated with very large volumes of data.⁵ ⁶ The technological advancement of monitoring technology in intensive care units is such that standard bedside monitoring equipment used in the unit can now be readily equipped to record raw minute by minute (or even more frequent) values of a wide range of physiological indices. The processing, validation, analysis, and interpretation of such data, however, is far from trivial,⁷ and the ability to collect data must be matched by an understanding of their limitations.

In this paper, for a sample of patients with traumatic brain injury admitted to the intensive care unit, we assessed the prognostic value of several summary measures of secondary insults in the first 72 hours after injury. Baseline prognostic models, developed on a larger dataset, were used to adjust the comparisons, and allowed us to measure the “added value” of the physiological monitoring data to baseline clinical predictions. Four different secondary insult summary measures were calculated and compared, and final models giving the improved predictions at 24, 48, and 72 hours after injury are described.

Methods

MONITORING AND PATIENT POPULATION

From 1 January 1989 to 16 July 1991, all patients with traumatic brain injury admitted to a regional neurosurgical centre at the Royal Infirmary, Edinburgh were eligible for the study. Inclusion criteria were age 14 or more, admission to the intensive care unit, and an admission or last known Glasgow coma score (GCS) of 12 or less, or of 13–15 with an overall injury severity score (ISS)³ of 16 or more.

Two stand alone PC based data recording systems⁸ were available in the intensive care unit for the duration of the study. As there were
eight bed spaces in the unit, it was possible that both systems were in use when a new eligible patient was admitted. This, and the obvious practical problems associated with such a technologically advanced system (at the time of the study), meant that only one third (124/372) of eligible patients actually had their physiological monitoring recorded. The six most consistent and reliable monitoring channels recorded were intracranial pressure (ICP), arterial blood pressure (ABP), cerebral perfusion pressure (CPP), arterial O2 saturation (SaO2), core temperature in °C (Temp), and heart rate in beats/ min (HRT). These gave rise to eight potential types of secondary insult; raised ICP, hypertension, reduced CPP, low SaO2, pyrexia, bradycardia, and tachycardia. Patients were only monitored on a particular channel if it was clinically indicated, this decision being made on the basis of conventional clinical wisdom in the late 1980s. Thus, not all patients were monitored on all channels. The criteria for ICP monitoring included all patients who required emergency evacuation of a haematoma, all patients requiring mechanical ventilation, and any other patients in which the responsible clinician thought that there was a risk of raised ICP.

In general we have taken the term “secondary insults” to refer to episodes during the acute care of patients with traumatic brain injury in the intensive care unit during which a physiological index—for example, temperature, was within a range defined as “abnormal”. Specifically we have used the thresholds and indices defined by the Edinburgh University secondary insult grades (EUSIG), as discussed by Jones et al1 and shown in table 1.

Patients were treated according to a protocol that emphasised resuscitation of blood pressure, oxygenation, and ICP reduction. Therapeutic goals were a minimum mean ABP of 70 mm Hg and minimum mean CPP of 60 mm Hg. CPP therapy was “ICP based” with an emphasis on ICP reduction and followed a regimen described previously.10

### OUTCOME AND BASELINE PROGNOSTIC MODEL

Outcome assessment was performed at 6, 12, and 24 months after injury, using a combination of structured interview and test battery administered by a neuropsychologist, a 60 item questionnaire sent to the patients’ closest relative or primary carer,1 and a letter to the patients’ general practitioner. For this study, survival at 1 year was taken as the primary outcome.

An externally validated baseline clinical prognostic model for predicting survival was constructed on the larger dataset of all 372 eligible patients with moderate or severe head injury. This model predicts survival on the basis of age, GCS on admission, pupil reactivity, ISS score, and the presence of a visible haematoma on CT. It has been validated, both internally and on an external dataset, and is described in detail elsewhere.11 The linear predictor LP (log odds of the predicted probability of survival) was used in all models as an adjusting factor for initial prognosis, and was calculated as follows:

\[
LP = 1.735 - 0.121 \times (AGE_{60}) + 0.270 \times (ISS) - 0.512 \times (GCS score) - 0.061 \times (ISS) - 1.784 \times (Pupil reaction) - 0.512 \times (If only one reactive pupil) - 1.784 \times (If neither pupil reacting)
\]

where \(AGE_{60}\) is zero if the patient was less than 50 years of age, and \((age=50)\) otherwise. If the patient had two reactive pupils and a visible haematoma on CT, then only age, GCS, and ISS contribute to the prediction.

Similar but not identical methods were used by Lemeshow et al15 to incorporate baseline predictive models into predictions at later times for general intensive care unit patients.

### SECONDARY INSULTS

Careful consideration of how to summarise the insult data was essential. Previous work used total insult time as a potential covariate,4 but this ignored the fact that patients vary greatly in the duration of monitoring, and there may be interruptions for trips to the operating theatre or imaging departments. More severely injured patients may spend longer in the intensive care unit, and so, even if secondary insults were unrelated to outcome, we would find a greater burden of insult in these patients, who of course have a poorer prognosis, and we may then have erroneously inferred that secondary insults were associated with poorer outcome.

For a fair comparison insults were quantified within fixed time intervals rather than over the whole of the stay in the intensive care unit. The initiating event and primary insult was the traumatic impact itself and so this was taken to be time zero, and secondary insult summaries were calculated over periods of 24, 48, and 72 hours from this point. In addition to the periods of secondary insult, there were occasions when, due to standard intensive care unit procedures, equipment failure or patient transfer, recording of the physiological indices was

---

**Table 1** Definition of secondary insult grades. Each derangement must be sustained for at least 5 minutes to be deemed a secondary insult, with the exception of pyrexia, which must last at least 60 minutes.

<table>
<thead>
<tr>
<th>Insult category</th>
<th>Grade 1</th>
<th>Grade 2</th>
<th>Grade 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raised ICP</td>
<td>mean ≥ 20 mm Hg</td>
<td>mean ≥ 30 mm Hg</td>
<td>mean ≥ 40 mm Hg</td>
</tr>
<tr>
<td>Hypotension</td>
<td>mean ≤ 70 mm Hg or systolic ≤ 90 mm Hg</td>
<td>mean ≤ 55 mm Hg or systolic ≤ 50 mm Hg</td>
<td>mean ≤ 40 mm Hg or systolic ≤ 30 mm Hg</td>
</tr>
<tr>
<td>Hypertension</td>
<td>mean ≥ 110 mm Hg or systolic ≥ 160 mm Hg</td>
<td>mean ≥ 130 mm Hg or systolic ≥ 190 mm Hg</td>
<td>mean ≥ 150 mm Hg or systolic ≥ 220 mm Hg</td>
</tr>
<tr>
<td>Low CPP</td>
<td>mean ≤ 60 mm Hg</td>
<td>mean ≤ 50 mm Hg</td>
<td>mean ≤ 50 mm Hg</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>SaO2 ≤ 85%</td>
<td>SaO2 ≤ 80%</td>
<td>SaO2 ≤ 80%</td>
</tr>
<tr>
<td>Pyrexia</td>
<td>Temp. ≥ 38 °C</td>
<td>Temp ≥ 39°C</td>
<td>Temp ≥ 40°C</td>
</tr>
<tr>
<td>Tachycardia</td>
<td>HRT ≥ 120 bpm</td>
<td>HRT ≥ 135 bpm</td>
<td>HRT ≥ 150 bpm</td>
</tr>
<tr>
<td>Bradycardia</td>
<td>HRT ≤ 50 bpm</td>
<td>HRT ≤ 40 bpm</td>
<td>HRT ≤ 30 bpm</td>
</tr>
</tbody>
</table>
not possible. This resulted in periods when nothing was known about the status of the patient. Subtracting these periods from the total monitoring time gave what we have termed “good monitoring time”, which was an upper limit on the maximum amount of secondary insult which would have been possible to record.

The summary measures examined, in increasing order of complexity, were:

(A) the presence or absence of any grade of insult (1–3) in an insult category.
(B) The total time (in minutes) of insult of grade 1 or worse in an insult category.
(C) The total time spent at grade 1 insult or worse expressed as a proportion of good monitoring time (GMT).
(D) A weighted sum of the proportion of GMT spent at each grade of insult, with weights equal to 1, 2, and 3 for grades 1, 2, and 3 respectively.

These summary measures were considered as candidate variables in a predictive model of survival over and above the linear predictor from the previously described baseline model. The crucial difference between this work and previous studies is that we attempted to adjust for variations in the severity and initial prognosis of patients before assessing the importance of secondary insults. The linear predictor variable (LP) measured what was known and predictable when the patient was first admitted to the intensive care unit, whereas the secondary insult summaries carry additional data on the evolution of the patient’s condition after this point. Our interest was in assessing the value of such additional information.

Rather than try to choose between the four types of insult summary (A-D above) automatically, each family of summary measures was treated separately and comparisons between them were based on global measures of model performance, simplicity and practical considerations. For each of the three epochs (cumulatively over the first 24, 48, and 72 hours after injury), the four summary measures were calculated for the eight categories of secondary insult for the entire epoch. Patients who were not monitored on a particular channel were assumed to have no insult on the corresponding insult categories. This assumption is justifiable because clinical monitoring was only initiated when it was deemed necessary, and hence where possible derangement were clinically expected.

For a particular epoch and each summary measure, the eight secondary insult categories were assessed for predictive power using a stepwise backwards elimination algorithm which was constrained to always contain the case mix adjusting variable LP. Thus the algorithm began with all secondary insult variables in the model, then eliminated those which did not contribute significantly.

The four resulting models were summarised using the model deviance, the quadratic or Brier score, the error rate, and the area under the receiver operating characteristic (ROC) curve. The deviance is analogous to the sum of squared errors in the linear model and is a good measure of global fit. Similarly, the average quadratic or Brier score is the sum of the squared differences between the observed outcome (0 or 1) and the predicted probability. In both cases smaller values are associated with better fitting models. Both the error rate and the area under the ROC curve assess discrimination (how well the model prediction classified survivors and non-survivors). The first is based on classifying all those with a predicted probability >0.5 as a survivor, all those with a predicted probability <0.5 as a non-survivor, and counting the number of misclassifications as a percentage of the total. The second is an extension of this process, and is exactly equal to the concordance statistic c, a generally applicable measure of predictive discrimination. It varies between 0 and 1, with 0.5 being equivalent to random guessing and 1 being perfect prediction.

As a final modelling exercise, the 72 hour insult summary measures were broken down into contributions from each of the three 24 hour periods, and a stepwise forwards selection algorithm employed to determine which insults in which epoch were most predictive.

Not all patients had arrived in the intensive care unit at 24 hours after injury, and some had already been admitted and died before this point. Models were calculated only on those patients with any monitoring starting before 24 hours who were still alive at 24 hours. Thus we modelled the conditional probability of survival at 1 year, given that the patient was alive and being monitored in the intensive care unit at 24 hours. Although this is a complex patient population, it is intuitively appealing and practical, in that these models can then be applied to any patient being monitored in the intensive care unit who is still alive 24 hours after injury. Similar approaches were taken for the 48 and 72 hour periods.

Models containing secondary insult summary measures were externally validated using 140 patients accrued to an almost identical study from 17 July 1991 to 30 April 1996. This study took place under an altered neurosurgical emergency service, but the patients were subject to the same inclusion criteria and thought to be broadly similar. Each model containing a significant secondary insult covariate was refitted in this new dataset, including the linear predictor LP, to protect the study against false positive results. Thus rather than adjusting the model selection procedure to explicitly account for the multiple tests used, an external dataset was used to validate any potentially significant predictors of survival.

Results
A total of 124 patients were enrolled in the study, as discussed by Jones et al. The precise time of injury was not known in 10 patients, who were excluded from this study. Two patients were lost to follow up at 1 year, and two did not begin to be monitored until after 72 hours, leaving 110 patients. This emphasised the advantage gained by using a baseline prognostic model developed on a larger set of patients rather than this very small group. We
have achieved much better precision, and not rejected candidate baseline clinical predictors purely because of low power, by including in our modelling process all patients with a moderate or severe head injury, not just the 33% of patients actually monitored in the intensive care unit. Table 2 shows the demographic, injury, survival, and crude secondary insult characteristics of the patients. Of the 23 deaths, two were within 24 hours of injury, two between 24 and 48 hours, two between 48 and 72 hours, a further three were within the first week, with the remaining 14 occurring between 1 week and 1 year after injury. Thus most deaths occurred after the particular period of interest. No patients were in a persistent vegetative state at 12 months.

The number of patients who were actually monitored on each of the six potential channels is shown in Table 3. The channels which caused most concern over the assumption that “no monitoring implied no insult” were ICP and CPP (of which course required an ICP monitor to be inserted), as the other four channels had monitoring rates of 85% or greater. Given that ICP monitors were inserted on the basis of perceived risk of intracranial hypertension, the assumption that patients without monitoring on these channels had no insults seems justifiable. The number of patients who experienced at least one secondary insult on each of the various channels is also shown in Table 2, demonstrating that nearly all patients had some hypotensive or hypertensive insult, and about half experienced raised ICP and reduced CPP at some point.

Summary measures for each of the baseline (containing only LP) and the final models selected by the stepwise algorithm are shown in Table 4, and are discussed separately for each epoch.

MODELS FOR 24 HOURS AFTER INJURY

Ninety two of the 110 patients started to be monitored within 24 hours of injury, but two died before this point, leaving 90 patients for analysis. The linear predictor LP was a highly significant predictor of outcome. No patients experienced any bradycardic secondary insult in this period, so this variable was omitted from the stepwise selection algorithm. For each of the four types of insult summary, only the ICP category remained in the final model, as shown in the first five rows of Table 4. None of the other insult categories improved on the prediction once baseline prognosis and ICP were taken into account. There was some evidence that the more complex measures of insult burden gave slightly better models, in that using the weighted sum (D) resulted in the smallest model deviance and Brier score, and the largest

In Table 4, the comparison of models selected by backwards elimination algorithm from all potential secondary insult categories.
area under the ROC curve, although the differences were marginal.

MODELS FOR 48 HOURS POSTINJURY
There were 105 patients still alive and being monitored at 48 hours postinjury. Once again, as shown in the centre five rows of table 4, only information on ICP insult significantly improved prediction over and above the initial prognosis at admission. In this case, however, it seemed that the simple measure of total time that ICP was above 20 mm Hg (B) was slightly more informative than the weighted sum measure (D), and there was doubt over the value of the crudest measure, the simple presence or absence of any secondary insult.

MODELS FOR 72 HOURS POSTINJURY
One hundred and four of the 110 patients were alive and being monitored at 72 hours postinjury, and the results of the variable selection procedure are shown in the final five rows of table 4. In this case, none of the dichotomous indicators of any secondary insult (A) seemed to improve on the baseline model. For the other three types of insult summary, ICP was again the only category selected, although the differences between the models were very small.

FINAL MODEL
Finally, the 72 hour summaries were each broken down into their components, from 0 to 24 hours, 24 to 48 hours, and 48 to 72 hours. The 24 potential predictors within each type of summary measure were assessed for predictive power using a forward selection algorithm to add them to a model already containing the linear predictor LP. The patient sample now contained 86 patients who began monitoring within 24 hours of injury and were still alive at 72 hours postinjury.

Table 5 summarises the final selected model within each type of summary measure. For the crudest insult summary (A) the presence of any ICP insult within the first 24 hours was the only indication which added significantly to the predictive model. For the other three summaries, however, the period from 24 to 48 hours was most informative, with very little to choose between the three types, although the total time with ICP insult within this period seemed to be marginally more predictive.

This model was compared in this set of patients to the model containing the total time with ICP insult from 0 to 48 hours, as this was thought to be simpler to calculate, with the results shown in the final row of table 5.

VALIDATION
Each model from table 4 was refitted on data from the validation set of 140 similar patients admitted during a subsequent period after alteration of the emergency neurosurgery service. In every case for summaries B, C, and D, and for each of the three epochs, the ICP covariate was a significant predictor in addition to the linear predictor (p<0.01 in every case). For the type A summaries, however, neither the 24 hour (p=0.152), the 48 hour (p=0.386) nor the 72 hour (p=0.456) observation of an ICP insult was predictive of survival. The external measures of model adequacy—namely, the Brier score, the error rate, and the area under the ROC curve—were obviously higher than in the original dataset, but not high enough to suggest that the models were fitted poorly. With the exception of the 24 hour type A result these findings were entirely consistent with what was found with the original dataset, and suggested that these results can be replicated in other patient samples.

Discussion
The aim of this work was to assess the predictive value for survival of recording and summarising secondary insults over and above simple baseline data, and we have shown that in patients managed with an intention to treat raised ICP,16 17 raised during the first 24, 48, and 72 hours, was independently predictive of mortality. Other indices recorded, such as CPP, did not add predictive value over and above ICP in this dataset, but it is important to note that in observational studies in which patients are managed with an intention to prevent and treat insults, we can infer very little about the causative relations between secondary insults and outcome. The correlation coefficients between the summary measures of insult for ICP and CPP for each summary measure and epoch ranged from 0.48 to 0.75, suggesting that although these variables are, as expected, related, the correlation is not overwhelming. The fact that episodes of raised ICP are predictive of death does not necessarily imply that more aggressive treatment to reduce or prevent these episodes will improve survival. Similarly, the absence of any predictive power of low CPP or hypoxia does not imply that derangements in these physiological variables should go uncorrected. These questions can only be answered definitively within the context of a prospective randomised clinical trial of adequate size. As a rough guide, if the management protocol for a patient in a new intensive care unit were to reduce mortality over standard practice from 20% to 15% (a dramatic relative improvement of 25%), then a trial with 90% power and 5% type I error rate would require about 2500 patients.

It is still possible, however, to speculate on the clinical and pathological interpretation of our results. The most important predictor of survival in the 72 hour postinjury period was found to be raised ICP, within all epochs and however crudely it was measured, with the notable exception of the very simplest “presence or absence of ICP insult” measure. This is
Early secondary insults after traumatic brain injury

encouraging when the largely subjective (but physiologically well founded) choice of the original thresholds for secondary insult is considered. It should also be noted that treatment thresholds for ICP were 25 mm Hg for the first 24 hours of intensive care unit care and thereafter 30 mm Hg, thus above the grade one ICP insult threshold. The fact that the period from 24 to 48 hours was found to be more predictive than the 0–24 hour period may have been due to the variability in the beginning of monitoring leading to increased noise in the earlier period.

Although we have attempted to adjust for differences in severity and prognosis using a previously validated model, a potential criticism of this process could be the fact that we have taken a very crude summary of the CT imaging information (haematoma present or absent), which ignores the potential for signs of raised ICP to be noted on the original scan. Whereas this may detract from our findings, we think that given the substantial variations from place to place in such things as the time of CT, the technology used, and the experience of the reporting radiologist, the use of a more robust measure which recorded only haematomas deemed relevant enough to be reported in the standard way was a sensible approach which gave a generalisable result. We are, however, returning to the original CT for both the initial and validation set of patients described here and extracting more detailed imaging data. The replication or otherwise of these results after inclusion of CT data on raised ICP signs will allow further speculation and perhaps clarification as to the particular mechanisms of action, and will be reported in a future publication.

There is a great deal of research into the biochemical and pathophysiological processes at work locally within the injured brain, but the fact remains that current monitoring and intervention is based on systemic or global brain variables. Moreover, there may also be a burden of undetected or unreported systemic hypoxia and or hypotension before admission to hospital. Although our data may only confirm early research and the opinion of many clinicians that early intracranial hypertension is a sign of poor prognosis, it is inevitable that as monitoring equipment develops, ever greater quantities of this kind of data will be collected. It is imperative, if we are not to be overwhelmed by the sheer volume, that careful plans of analysis are determined beforehand, as this work shows that these observational data are vulnerable to many sources of bias.

In summary we have shown that prediction of survival for patients with head injury in the intensive care unit can be improved with the addition to the model of the total time that ICP is above 20 mm Hg in the periods up to 24, 48, and 72 hours after injury. Measurement of other categories of systemic physiological variables may be important for patient management, but do not add to predicative ability in this situation.

This work is a small part of a large ongoing project to study the mechanisms and pathology of traumatic brain injury begun by the late Professor J D Miller. We are very grateful to all the staff who aided in the laborious task of collecting and verifying the vast quantity of monitoring data—namely, J Corrie, M Dearden, J Findlay, T Howells, J Kerr, EG McKeaning, S Midgley, MJ Souter, and J Tocher. Funding was provided by the MRC, under grants S8609195, S86125498, G9301975, G9301951, and G980852.

References