Automated Analysis and Annotation of Datasets from Patients with Severe Traumatic Brain Injury: an Initial Case Study

Derek Sleeman, Samuel Cauvin
Department of Computing Science
University of Aberdeen
d.sleeman@abdn.ac.uk,
s.cauvin.10@aberdeen.ac.uk

Jonathan Rhodes
Intensive Care Unit
Western General Hospital
Edinburgh EH4 2XU
jrhodes1@staffmail.ed.ac.uk

Treatment protocols for patients who have suffered traumatic brain injury (TBI) specify that hypoxia should be avoided and specifically that brain oxygen tension (PbtO₂) should be maintained above a particular level (20mmHg). Results from several specialized Neuro ICUs world-wide suggest that such guidelines are not achieved in at least 24% of patients (Shafi et al. 2014). Many physiological and therapeutic factors can influence PbtO₂. Furthermore, clinical staff in ICUs have many calls on their time and may not be able to direct sufficient attention to the management of a single parameter in one of several patients. We believe that automated analysis of complex data sets could be a useful step to developing software capable of assisting in the better maintenance of PbtO₂ in these complex patients.

A "manual" analysis of 5 patients' complete temporal records showed that certain distinct rules including a number of important descriptors (CPP, PbtO₂, PaO₂ & the correlation coefficient between CPP and PbtO₂) with their values segmented into discrete ranges, cover 98.2% (SD 1.6) of the available dataset once the rules have been "fine-tuned". This study was then validated in a second set of five patients, in which 92.5% (SD 12.5) of the data set was covered without additional "tuning" of the rules. Moreover, it was noted that these patterns, once some account was taken of noise, occur in "blocks". As a result of these observations we developed a correlation module for the existing Temporal Discovery workbench to replicate the "manual" analysis. Using the dataset for the 10 patients, we have now obtained effectively perfect agreement between the manual analysis and that produced by the workbench. Subsequently, the expert provided clinical actions which correspond to each of the (48) patterns potentially created in this study, and so the correlation module is now able to produce for each time-point (and each time "block") a pattern, and the correct clinical action.

Further work includes enhancing how the correlation module deals with noise, evaluating the approach across a much larger dataset, and evaluating the effectiveness of the module's recommendations in clinical settings.

Traumatic Brain Injury, ICU, patient monitoring, Hypoxia, correlation, workbench, data analysis

Note: In this paper, the terms patterns and rules are used interchangeably.

fundamental to the maintenance of oxygen delivery (The Brain Trauma Foundation 2007).

1. INTRODUCTION

This project has developed a sophisticated system for the detection and management of adverse brain oxygenation following traumatic brain injury (TBI). We believe this is an important first stage in the development of computer based decision support in the neuro-intensive care unit (NICU). Injury to the brain is often progressive. In part this is due to the brain being more vulnerable to hypoxia following TBI (Jenkins et al. 1989). Patient management focuses on maintaining oxygen delivery to the brain, and so the control of intracranial pressure (ICP) and cerebral perfusion pressure (CPP) are

However, this does not ensure that hypoxia is prevented (Rhodes et al. 2016) or outcomes improved (Andrews et al. 2015). Care which focuses extensively on cranial pressures (Jones et al. 1994) has been recognised to have significant limitations which have led to the development of multimodality monitoring, (Oddo et al. 2012). These include the direct measurement of cerebral oxygen tension (PbtO₂), which is standard in the ICU at Edinburgh's Western General Hospital, providing additional clinical information which is complimentary to Intracranial Pressure (ICP) and cerebral perfusion pressure (CPP) data (Flynn et al. 2015).

PbtO₂ values associated with adverse outcome after severe TBI have been determined in several studies and meta-analysis (Stiefel et al. 2005). The consensus in the literature is that PbtO₂ should be maintained at or above 20 mmHg (Spiotta et al. 2010). A low PbtO₂ can usually be improved by manipulations which either directly or indirectly modify CPP, ICP, arterial carbon dioxide tension (PaCO₂), blood haemoglobin concentration and inspired oxygen concentration (FiO₂). Data from several comparative series and a systematic review, recommend targeting PbtO₂ to improve patient outcome after severe TBI (Stiefel et al. 2005; Nangunoori et al. 2012).

In clinical practice, it is difficult to maintain a patient's PbtO₂ above this threshold consistently. The management of patients with severe TBI is multi-faceted and occurs in a very busy environment. Furthermore, the quantity of information available from a number of monitors adds to the difficulties in interpreting and responding appropriately to the data (Citerio et al. 2015). Paradoxically this abundance of data has the potential to support the development of computer algorithms for the detection and prediction of (adverse) events. This is particularly important when the principal aim of the treatment is the prevention of secondary insults. In similar complex and failure-intolerant environments, such as the aviation industry, the use of technologies to monitor and interpret data has had a significant impact on safety (Sabatini 2006). We believe that the development of intelligent data analysis systems and the subsequent development of decision support tools will be useful in guiding complex care (such as the management of PbtO₂), and thereby enhancing outcomes.

Overview of the Paper: section 2 outlines the manual analyses carried out on the Traumatic Brain Injury dataset; section 3 discusses software developed to support the analyses of such temporal datasets (Temporal Discovery Workbench & its correlation module); section 4 discusses a case study of a TBI dataset using both manual analysis and the workbench's correlation module; and section 5 discusses the study and planned further work.

2. MANUAL "PILOT STUDY WITH TBI DATASETS"

Patients with severe TBI (GCS≤7) are admitted to Edinburgh's Western General Hospital NICU and are managed in accordance with Brain Trauma Foundation guidelines (The Brain Trauma Foundation 2007). Since May 2010 patients receiving ICP measurement also have PbtO₂ monitored.

A simple treatment protocol has been developed to guide the management of PbtO₂. This includes steps to optimise PaO₂, PaCO₂, ICP, CPP, haemoglobin and cardiac output. Minute by minute physiological data including PbtO₂, CPP and ICP are recorded by a bedside computer.

Cumulative Hypoxia and Management Failure.

Despite the introduction of protocol guided management to attain a PbtO₂ ≥ 20mmHg our current practice often fails to achieve this therapeutic target. In 25 patients monitored for up to 5 days post injury, PbtO₂ was <20mmHg for 51.9% (12.7 to 64.8) of the monitored time (Rhodes et al. 2016). This high incidence of treatment failure occurred despite staff being aware of the treatment protocol. It is our experience that following a patient review, such as occurs during a clinical ward round, the PbtO₂ can often be improved, simply by adherence to the existing protocol. However, compliance with such guideline based management in the NICU may generally be difficult to achieve on a day-to-day basis. On the other hand, a simplified analysis of the data set which only considered high/low dichotomisation of CPP and PaO₂ revealed that when PbtO₂ was low (54,654 minutes of data), a corrective action that might increase PbtO₂ (such as optimising CPP or correcting a low PaO₂) was found for 72.5% (SD 16.4) of the time-points. This suggests that compliance could have been 85.7% (IE 48.1% + .725 * 51.9%).

Similarly, a retrospective study of severe TBI patients managed in 11 North American level-1 trauma centres reported that the target value for CPP was not achieved in 24% of patients (Shafi et al. 2014). We believe that the current practice of passive monitoring and infrequent expert reviews misses the potential of the increasing volume of physiological data available to clinicians. However, the amount of data available can be overwhelming to many, requires very specific expertise to interpret, which might only be available from a small number of individuals in a therapeutic unit. These difficulties are compounded by the round the clock needs of TBI patients and the competing demands for clinicians' attention in the critical care environment.

Development of PbtO₂ data interpretation.

As a prelude to developing procedures for how PbtO₂ data might be processed, archived ("off-line") multi-modality data sets from 5 patients with TBI were retrospectively assessed (by JR). Recordings for individual patients, which passed certain acceptability criteria, ranged from 30 to 391 hours. Using Excel logic functions such as IF, IF AND etc., it was possible to define a set of "rules" based on ranges for PbtO₂, CPP, PaO₂ and the 3 hour retrospective correlation coefficient between

PbtO₂ and CPP (referred to as ORx). By refining these ranges after each data set was analyzed, the overall fit of the rules to the data was increased; further, during this process, rules were refined so that only a single rule applied to each time-point.

The final set of rules has the following form:

- R1: CPP≤110mmHg, PbtO₂ <40mmHg, ORx≥0.4
- R2: CPP≥100mmHg, PbtO₂≥15mmHg, ORx <0.4
- R3: CPP<100mmHg, PbtO₂>9.0mmHg, ORx<0.4, PaO₂>9kPa
- R4: CPP>110mmHg, PbtO₂<25mmHg, ORx ≥0.5
- R5: CPP>70mmHg, PbtO₂>40mmHg, ORx>0.4, PaO₂ >10kPa

Overall this set of 5 rules describes 98.2% (SD 1.6) of the data for which a valid correlation coefficient is present. These descriptive rules were subsequently validated against a second set of 5 patients, which contained between 52 and 188 hours of data. With no further adjustment of the ranges used, these rules described 92.5% (SD 12.7) of the data. This exercise demonstrated that large sets of data can be described by a simple rule set where each time-point¹ is "covered" by only a single rule. A further important observation, is that an individual rule can apply to continuous data sequences which vary between several minutes to several hours in length. This finding, that stable relationships exist between variables over time, encouraged the development of a software system to detect the correlations which exist between PbtO₂ and the domain's other significant clinical descriptors.

3. OVERVIEW OF TEMPORAL DISCOVERY WORKBENCH

We have been addressing the general scenario in which an unusual event, E, happens at time-point, T, and we aim to predict this event by analyzing trends and absolute values in the several descriptors recorded in the time-period prior to E.

To help this analysis, it is likely we will also have datasets involving the same descriptors in which the event, E, does not occur. The first domain in which we have applied this approach is the detection of myocardial damage from the temporal records recorded for ICU patients by the ICU's patient management system (PMS).

The PMS records, at least hourly, standard physiological data (e.g. Heart Rate (HR), Blood

Pressure (MAP), Blood Oxygen Saturation (SpO₂), Administered Oxygen (FiO₂)), drugs administered, and treatments such as dialysis. The presence / absence of myocardial damage is confirmed by the level of Troponin in the patient's blood; this test is usually performed every 72 hours and when requested by a clinician.

To address this class of problems, we decided to implement a Workbench (Temporal Discovery Workbench: TDWB) as we believe this provides a great deal of flexibility. Although we may have a clear idea of an initial project's objectives we do not know in advance the range of applications we might encounter and thus we do not know the detailed nature of the analyses which domain experts might wish to carry out on their datasets.

Workbenches (WBs) generally present their user with options at each stage in the analysis and allow the analyst (sometimes with guidance) to decide the data display mode or analysis package to be used and with what descriptors. It is essential that WBs provide user-friendly interfaces, and they are modular in construction, so that functionality not envisaged at the initial design can be subsequently added if needed.

The original initial version of TDWB, which was implemented in Java, contained 2 principal components, namely:

- The Data Handling module to load and check the validity of various datasets to be analyzed. (Note TDWB at present is only able to load pre-existing datasets that is, it is currently unable to process patient real-time datasets.)
- A module which can create/discover from the "approved" datasets, patterns which ideally only occur before a positive special event and in none of the segments which contain negative special event markers.

Some details of these modules are given in the following subsections.

3.1 Data Handling Module

Temporal datasets are presented to TDWB as CSV files, and must contain a column called Time-point (containing data of the following form: [DD:MM:YYYY; hh:mm:ss], and a column called "Special Event" which can only contain the strings: Positive, Negative or blank. Additionally, the file can contain as many other column headings as required by the domain. So in the case of the ICU domain this is likely to include: variables such as HR, Mean (or MAP), FiO₂, SpO₂, together with drugs information. Each column is typed to help TDWB spot data errors; currently only the following

¹ Each time-point in this dataset last for one minute, and contains the relevant descriptor-value pairs for that time-point.

data types are accepted: "Timepoint", "Int" (integer), "Real" and "String". The data associated with a particular time point is held as a separate record; each record is terminated by a new line; and files are terminated by a special terminator.

"Input of Data files" loads patient (CSV) files, performs various checks on the dataset (including: type-checking of elements, that temporal records are correctly ordered, check length of gaps between time-points), provides options for extrapolation of missing time-points etc.; allows the analyst to select from all the descriptors in the CSV file which should be included in the current analysis / study; and set ranges for the selected descriptors. For example, the expert decided that SpO2 should have the following 5 ranges: L4 (Low-4), L3 (Low-3), L2 (Low-2), L1 (Low-1) & N (Normal). Multiple patient datasets can be loaded. There are also facilities to display the datasets in different formats: raw/original, cleaned-up (i.e., when extra and missing elements/time-points are dealt with), continuous data with a predefined set of ranges for each descriptor, and discrete where the names of the ranges are displayed. Once these processes have been successfully completed, the analyst is given the opportunity to save this information to a project file so that the "set up" work does not need to be done again.

3.2 Pattern Matching & Discovery Module

Once datasets have been loaded and checked by the previous module, it is then possible to use this module to create patterns which in a sense "explain" the special positive event which occurs as part of segment's last time-point. Given space limitations we are unable to explain this process here, but we refer the reader to earlier papers for details (Sleeman et al, 2015).

However, we are providing a summary of a study which we have been able to carry out with the TDWB in the domain of Myocardial Damage. "We have shown that the sets of patterns produced by TDWB generally have better "coverage", than those produced by the original expert-derived model. We then investigated whether some of the TDWB-created patterns might not be clinically acceptable.

Recently we ran a pilot study in which we asked a single clinician to evaluate the patterns produced by TDWB, and to say whether they were acceptable, and why. This further information has now been implemented in TDWB; the resulting set of filtered patterns still has better coverage than the initial set of "manual" patterns." (Sleeman et al, 2015).

3.3 Enhancements of TDWB: the Correlation Module

Subsequently we have developed the Correlation module, which takes the checked temporal datasets and attempts to show for each time-point whether some variables are correlated, and additionally reports for each time-point the ranges for a pre-specified number of descriptors, as well as the correlation between the descriptors PbtO2 and CPP. Optionally this module can also report the appropriate clinical actions to be taken (where the domain expert will have provided the system with the appropriate clinical actions for each of the possible "patterns"). In the next section we provide a detailed example.

Figure 1: Screenshot of the Correlation Parameters Dialogue.

4. INITIAL RESULTS OF TDWB'S CORRELATION MODULE

The previous section and section 2 have outlined the task to be addressed by the Correlation module. The dataset provides data for a number of patients; the data for each patient consists of a number of time-points which generally have values provided for all the descriptors / variables; the descriptors include PbtO2, CPP, ICP, PaO2, & PaCO2. These datasets have values for descriptors reported every minute. (Data for 10 patients were processed in this study, mean length 9583 (SD 7658) minutes. Note these figures exclude the initial periods when the PbtO2 readings are unstable.) TDWB's data handling module allows the analyst to decide which descriptors to include in the analysis; their ranges are also set up in this module.

Further, the Correlation module allows the analyst to specify the ranges to be used for the correlation coefficient and the pair of descriptors that the correlation coefficient is to be calculated for (Figure 1); to be consistent with the manual analysis we have chosen to calculate the correlation coefficient between PbtO2 & CPP (referred to as ORx). The ranges which the analyst provided, as a result of a knowledge acquisition session, for this study are:

CPP Low[0:70]; Med[70: 110], High [110:151]
PaO2 Low[0: 13.0]; High[13.0: 100.0]
PbtO2 Low [0:9.0]; Med[9.0: 20.0];
High[20.0: 40.0]; VHigh[40.0: 101]
ORx Low[-1: 0.4]; High[0.4: 1.0]

Before the correlation coefficient can be calculated for each of the time-points in all the datasets the remaining parameters in the Correlation Parameters Dialogue (Figure 1) have to be provided.

The values input for this analysis are shown in that figure. Here we discuss the values provided for the more important parameters; as noted earlier the Correlation is to be calculated between descriptors PbtO2 & CPP and so these are input as the Primary & Secondary variables. Parameter 3 allows us to specify that the values of the other variables are to be reported even if they are not correlated with the primary & secondary variables. (So selecting this options means that the Correlation & all the selected descriptors are reported.) Parameter 5 says that the first 122 time-points are to be ignored (as the PbtO2 sensor is not stable during this period). Parameter 6 specifies the length of the correlation window (but this window might effectively contain less elements if noise is detected in that range). TDWB provides several ways in which noisy elements can be defined and handled. For details of these and the other parameters, see the TDWB User Manual (Sleeman & Cauvin, 2016).

The maximum number of patterns which can be produced in this analysis is the product of the number of CPP ranges, the number of PaO2 ranges, the number of PbtO2 ranges, and the number of ORx ranges, which from the information above we know is 48 (i.e., $3 * 2 * 4 * 2$). Additionally, the expert clinician has provided a list of the clinical actions (including "NO ACTION") which he believes should be performed when the conditions specified in the (48) patterns apply.

This information has been encoded and provided to TDWB's Correlation module which means that as well as calculating, for each time-point, the appropriate pattern, the system is also able to recommend the associated Clinical Action.

Here are several examples of Clinical actions and their conditions:

- IF PbtO2 [High] THEN NO-ACTION
- IF PbtO2[Low] and CPP[Med] and ORx[High] and PaO2[High] THEN Increase-CPP-by-10mmHG
- IF PbtO2[Low] and CPP[High] and ORx[High] and PaO2[High] THEN Increase-Oxygen-Delivery

In this study, we focussed initially on the dataset from a single patient to ensure that the manual analysis and that of TDWB's correlation module give the same results for both the patterns / rules and the clinical actions recommended. These comparisons highlighted a number of differences which have been systematically removed, including:

- Whereas the ranges set for the descriptors by TDWB systematically cover the whole "space", there were gaps in the ranges used in the manual analysis, as well as overlapping ranges.
- Also the ranges used in the manual analysis tended to be different from those adopted for the Correlation module. For example, whereas the correlation module uses just 2 ranges for ORx (low & high), we note that the rules reported for the manual analysis in section 2 use 3 as the upper range is subdivided at 0.5.
- Additionally, in the process of this detailed checking, we found a number of programming and conceptual errors in the correlation module, and situations where the manual analysis had been wrongly applied.

Once these issues had been addressed then the results for the manual and the correlation module were effectively identical for the first patient. We then processed the data from the remaining 9 patients, and again a high level of agreement (99.9%) between the 2 sets of patterns and proposed clinical actions was achieved. Note we expect that NULL patterns and NULL actions will be reported for the first ~300 minutes of each patient dataset as the PbtO2 sensor takes about 120 minutes to stabilize, and then a further 180 minutes which is the size chosen in this analysis for the correlation window (see figure 1).)

Table1 shows some of the "data blocks" / sequences which TDWB detects in the dataset for Patient-17, and shows the corresponding clinical actions. Note, in transition periods, usually several clinical actions are suggested.

Table 1: Part of the analysis reported by the correlation module for Patient-17 (with some interpretation).

Range	Clinical Actions	Comments
1-342	None	PbtO2 values stabilizing, correlation window of 180, & then some TPs with PbtO2 ≥ 20 .
343-353	Incr PaO2 to $>13\text{KPa}$	
354-357	Various	Transition period
358-387	Incr CPP to $\geq 70\text{mmHg}$	
388-417	Incr CPP by 10mmHg	
418-421	Various	Transition period
422-475	Incr CPP to $\geq 70\text{mmHg}$	
...		
1351-1498	Incr_Oxygen_Delivery	

5. DISCUSSION & FURTHER WORK

- Using a deterministic method the TDWB Correlation module is able to generate Patterns and Clinical Actions which are essentially the same as those produced by the manual analysis in a very high percentage of the situations, (99.9%). And of course TDWB is able to process these datasets much more quickly and consistently, thus allowing the analyst to modify parameter values systematically (say the range limits for descriptors) to determine the effects.
- Currently, smoothing operations within the TDWB are very basic: the analyst can specify the number of time-points in a given correlation window that can have a NULL value. If less than this number is present in a correlation window, then a correlation, and thus a rule/pattern is reported for that time-point.

Further Work.

- We plan to run standard training and test studies. For example, the domain expert will annotate the 10 patient datasets with the clinical actions they think are appropriate. This annotated set will then be randomly split into a training set (2/3) and a test set (1/3). A machine learning algorithm such as a decision tree algorithm (Witten & Frank 2005) will then be used to create a decision tree from the training dataset, extract a ruleset from that decision tree,

and then run the test set against the "extracted" rule set.

- More sophisticated smoothing functions based on statistical and ML approaches to be developed for subsequent versions of the Correlation Module.

We believe that a software platform could identify low PbtO2 and alert clinical staff to corrective actions with greater efficiency than current practice. This would reduce the incidence of low PbtO2 and be associated with improvement in clinical outcomes.

6. ACKNOWLEDGEMENT

TDWB was implemented by Samuel Cauvin and Michael Gibson (Univ of Aberdeen) with financial support from the Univ of Aberdeen Development Trust. Helpful discussions on the project with Dr Laura Moss (Univ of Aberdeen & GGC Health Board), Dr Martin Shaw (GGC Health Board), & Dr Wamberto Vasconcelos (Univ of Aberdeen).

7. REFERENCES

- Andrews, P.J.D. et al., 2015. Hypothermia for Intracranial Hypertension after Traumatic Brain Injury. The New England journal of medicine, 373(25), pp.2403-12.
- Citerio, G. et al., 2015. Data collection and interpretation. Neurocritical care, 22(3), pp.360-8.
- Flynn, L.M.C., Rhodes, J. & Andrews, P.J.D., 2015. Therapeutic Hypothermia Reduces Intracranial Pressure and Partial Brain Oxygen Tension in Patients with Severe Traumatic Brain Injury: Preliminary Data from the Eurotherm3235 Trial. Therapeutic hypothermia and temperature management, 5(3), pp.143-51.
- Jenkins, L.W. et al., 1989. Increased vulnerability of the mildly traumatized rat brain to cerebral ischemia: the use of controlled secondary ischemia as a research tool to identify common or different mechanisms contri. Brain Research, 477(1-2), pp.211-224.
- Jones, P.A. et al., 1994. Measuring the burden of secondary insults in head-injured patients during intensive care. J.Neurosurg.Anesthesiol., 6(0898-4921), pp.4-14.
- Nangunoori, R. et al., 2012. Brain tissue oxygen-based therapy and outcome after severe

- traumatic brain injury: a systematic literature review. *Neurocritical care*, 17(1), pp.131-8.
- Oddo, M., Villa, F. & Citerio, G., 2012. Brain multimodality monitoring: an update. *Current opinion in critical care*, 18(2), pp.111-8.
- Rhodes, J.K., Chandrasekaran, S. & Andrews, P.J., 2016. Early Changes in Brain Oxygen Tension May Predict Outcome Following Severe Traumatic Brain Injury. *Acta Neurochir.Suppl* 2016; 122: 9-16.
- Sabatini, N., 2006. Downward Pressure on the Accident Rate. Speech by FAA Associate Administrator for Aviation Safety, to the International Society of Air Safety Investigators. Available at: <http://www.chinaaviationdaily.com/news/0/456.html>.
- Shafi, S. et al., 2014. Suboptimal compliance with evidence-based guidelines in patients with traumatic brain injuries. *Journal of neurosurgery*, 120(3), pp.773-7.
- Sleeman, D. et al., 2015. Experts Assess Patterns produced by a Temporal Discovery Workbench. In *Proceedings of the SciKn Workshop*, held at the KCAP-15 Conference, held at Palisades, NY, October 2015.
- Sleeman, D. & Cauvin, S.R. (2016). *TDWB User Manual*. Department of Computer Science, The University of Aberdeen.
- Spiotta, A.M. et al., 2010. Brain tissue oxygen-directed management and outcome in patients with severe traumatic brain injury. *J.Neurosurg.*, 113(3), pp.571-580.
- Stiefel, M.F. et al., 2005. Reduced mortality rate in patients with severe traumatic brain injury treated with brain tissue oxygen monitoring. *J.Neurosurg.*, 103(5), pp.805-811.
- The Brain Trauma Foundation, 2007. *Guidelines for the Management of Severe Traumatic Brain Injury* 3rd edition. *J.Neurotrauma.*, 24(S1), pp.1-106.
- Witten, I.H., Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd Ed. Morgan Kaufmann 2005.