
Predicting patient outcomes via neural network estimation of discharge APACHE scores for traumatic brain injury

Cindy Crump^{1,2,*}, Christine Tsien Silvers^{1,3}, Bruce Wilson⁴, Loretta Schlachta-Fairchild⁵, Jeffrey Scott Ashley⁶

¹AFrame Digital, Inc., Vienna, VA, USA

²The Center for Study of Chronic Illness and Disability, George Mason University, Fairfax, VA, USA

³Children's Hospital Informatics Program, Boston, MA, USA

⁴Barron Associates, Charlottesville, VA, USA

⁵iTelehealth, Inc., Cocoa Beach, FL, USA

⁶Center for Nursing Science and Clinical Inquiry, Walter Reed National Military Medical Center, Bethesda, MD, USA

Email address:

ccrump@afamedigital.com (C. Crump), tsien@alum.mit.edu (C. T. Silvers), bruce@wilson.org (B. Wilson),

lschlachta@itelehealthinc.com (L. Schlachta-Fairchild), jeffrey.s.ashley.mil@mail.mil (J. S. Ashley)

To cite this article:

Cindy Crump, Christine Tsien Silvers, Bruce Wilson, Loretta Schlachta-Fairchild, Jeffrey Scott Ashley. Predicting Patient Outcomes Via Neural Network Estimation of Discharge APACHE Scores for Traumatic Brain Injury. *American Journal of Health Research*.

Vol. 2, No. 6, 2014, pp. 361-365. doi: 10.11648/j.ajhr.20140206.17

Abstract: It is highly desirable to be able to predict the likely outcome of critical patients admitted to the intensive care unit (ICU) for traumatic brain injury (TBI). Vital signs, laboratory values, and clinical assessments from throughout a patient's ICU stay were collected retrospectively in an IRB-approved protocol from a Level I Trauma-Military Medical Center in the Southwest. ICU patients were included if they had been admitted for TBI during a five-year period ending in October 2007. Data were collected for 139 ICU patients with TBI. Admission and discharge APACHE IV scores were then derived from the collected data for each patient. A static back propagation neural network was developed to predict a patient's ICU outcome vis-a-vis discharge APACHE IV scores. The resulting network, trained using leave-one-out methodology, was able to predict the discharge APACHE score on average within 12.9% of the actual score.

Keywords: APACHE Score, Intensive Care Unit, Neural Network, Patient Outcome, Prediction, Traumatic Brain Injury

1. Introduction

The ability to accurately predict a traumatic brain injury (TBI) patient's outcome after admission to the intensive care unit (ICU) is important for a variety of reasons, including improved understanding by and realistic goal-setting for the patient, family and caregivers; improved resource utilization for the hospital; and better ability to set uniform standards for judging the institutional performance of trauma centers and intensive care units. Patient severity scores are furthermore useful for clinical research, for example, as inclusion criteria or for comparison between study and control groups. The goal of the present study was to determine the feasibility of developing an artificial neural network (ANN) model for predicting discharge APACHE IV (or simply 'APACHE') severity scores in ICU patients admitted with TBI based upon admission APACHE scores and clinical data available from

the first ten days of each patient's stay in the ICU.

2. Background

Predicting mortality based upon trauma scoring systems, such as the Major Trauma Outcome Study (MTOS), the Trauma Audit Research Network (TARN), the Base Excess Injury Severity Scale (BISS), the Trauma and Injury Severity Score (TRISS), and Acute Physiology and Chronic Health Evaluation (APACHE), have been described [1-5]. Other models, such as artificial neural networks and multivariate logistic regression, have also been attempted [6-8]. These systems, however, look at some subset of pre-hospital data, emergency department data, and data from the first 24 hours in the ICU. It could be useful for care providers as well as family to have a better sense of a TBI patient's predicted severity level at the end of his/her ICU stay, based upon additional

clinical data as they become available in the ICU beyond the first 24 hours.

One approach to predicting an ICU patient's severity level at ICU discharge (or death, in some cases) could be to use both admission data and additional clinical data as they become available on subsequent days in the ICU to serve as inputs into artificial neural networks (or simply neural networks) for developing a prediction model. Artificial neural networks have arisen over the last fifty years as an important means of using machines to model complex physiological and biological learning processes as represented in brain function [9]. Machine learning techniques have been shown in a variety of medical research settings to deliver statistically significant results with strong predictive capabilities for complex medical diagnoses in cardiac cases, trauma, pain, sepsis, and a host of other clinical conditions [10, 11]. Brause demonstrated in 2001 that ANNs can outperform human diagnostic capabilities and that about 50% of diagnoses by medical professionals are incorrect [12]. This is due to the fact that humans cannot analyze complex data informally without making errors and tiring easily. Ishak notes that research in the use of neural networks in medicine has been accelerating since the mid-1990s given that relevant patient data can be overlooked by even the most experienced specialists in the field [13]. His estimate is that most of the research employing neural networks has an observed accuracy of between 70% and 80%. Furthermore, he notes a number of important studies in use of neural networks in clinical medicine to predict survival and death in patients with cancer, trauma, and acute myocardial infarction, as well as those in long-term care. The results of these studies compare very well with estimates of experienced physicians and objective condition scoring systems such as APACHE II [13].

ANNs can support nonlinear, complex, multivariable data relationships. ANNs, however, are mechanistic and not transparent (that is, they are "black box" models of the domain). Furthermore, they require a target outcome to train the model. Therefore, they do not allow for the data to uncover previously unknown relationships, nor do they allow for establishment of conditional probabilities that can be easily translated into a clinical picture based on prior knowledge associated with the population-based patient data. Nonetheless, ANNs are used today in a number of relevant clinical applications and are explored here for prediction of outcome severity in ICU patients with TBI.

3. Methods

Physiological vital signs, clinical laboratory values, and clinical assessments from throughout each patient's hospital stay, along with outcomes ('Discharged' versus 'Deceased' and length of stay) were collected retrospectively as part of an IRB-approved protocol for patients meeting inclusion criteria who were admitted to the ICU at a Level I Trauma-Military Medical Center that also participates in the civilian city-wide trauma program, in the southwest. Patients were included in this analysis if their diagnosis during the five-year period

ending in October 2007 included traumatic brain injury. Vital sign values collected include the following: systolic, diastolic, and mean blood pressures from an arterial line; systolic, diastolic, and mean blood pressures from a noninvasive blood pressure cuff; heart rate; body temperature; respiratory rate; and oxygen saturation. Laboratory values collected include the following: white blood cell count, hemoglobin, hematocrit, sodium, potassium, chloride, bicarbonate, blood urea nitrogen, creatinine, glucose, prothrombin time, and partial thromboplastin time. Clinical assessments collected include Glasgow Coma Scale score (eye, verbal, and motor) and level of consciousness. Outcomes collected include length of stay (LOS) in the ICU, and whether or not the patient died during the ICU stay. These values were a convenience sample, selected based on their potential availability and ease of automatic evaluation. Values for the different clinical measurement types (i.e., parameters) were collected at different frequencies, dependent upon routine clinical practice. For example, heart rate may have been recorded multiple times within a given hour, while a laboratory value such as sodium might only have been collected once or twice in a day.

Admission and discharge APACHE IV scores (with a possible range from 0 to 299) were calculated by a research nurse independent of the authors for each ICU TBI patient. In order to try to predict the discharge APACHE scores, a static ANN was used. The inputs to this network consisted of clinical data from the collected ICU dataset. To minimize the effects of missing data values and varying frequencies of measurement recordings, daily averages of each of the clinical values were used as inputs to the system. If a measurement type had no values recorded at all for a given patient on a particular day (such that a value for that day's "daily average" was therefore unavailable), then that patient was not included in further analysis. To furthermore ensure that there was a sufficient amount of longitudinal ICU data for each patient and that outcome prediction was not made using data too close in time to the date of ICU discharge or death, only patients who had a length of stay in the ICU of at least 20 days and who had clinical data available for the first 20 days were used in development of this APACHE score prediction system. For each of those patients, the patient's preprocessed clinical data for the first ten days of ICU stay, along with the admission APACHE score, were then used to predict that patient's ultimate discharge APACHE score.

The static back propagation neural network used to perform the APACHE score prediction consisted of three layers: an input layer of 90 'log-sig' neurons, a second layer of 25 'log' neurons, and an output layer of one linear neuron. The network was trained using resilient back propagation. This uses the error between the desired output of the network and the actual output of the network to systematically adjust the individual neuron weights in each layer. The process is repeated for each member of the training set sequentially until a desired error is achieved or a predetermined number of sequences is reached. The neural network used to perform the discharge APACHE score prediction is shown in Fig. 1.

In order to train the neural network, a "leave one out"

methodology was used. This means that one patient's data are removed from the overall dataset and all other patients' data are used to train the neural network. The removed patient's data (previously unseen to the network) are then processed by the system and the predicted discharge APACHE score is compared to the actual discharge APACHE score of that patient. The weights of the neural network are reset to their initial values and the removed patient's data are added back into the overall dataset. The process is then repeated for each patient's data until all patients have been processed. Each of the predicted APACHE scores and their associated actual APACHE scores can then be compared in order to quantify the effectiveness of the neural network classifier.

4. Results

Data were collected from the ICU stays of 139 patients diagnosed with TBI. Of those 139 patients, 132 were eventually discharged from the ICU, while 7 died. Fifty-four patients had an ICU length of stay of fewer than or equal to 30 days, 54 patients had an ICU LOS of 31-60 days, 29 patients had an ICU LOS of more than 60 days, and two patients had an unknown ICU LOS. Daily averages, when possible, were computed for each of the 26 clinical values collected. Twenty-seven patients had sufficient data spanning all 26 clinical values for each of the ICU days of interest to be included in neural network model analysis and testing.

Table 1. Results of discharge APACHE IV score prediction with a neural network applied to the dataset of intensive care unit (ICU) patients with traumatic brain injury (TBI) who had an ICU length of stay of at least 20 days and who had clinical data available for all 26 collected parameters for at least 20 days. Patient numbers correspond to unique identifiers assigned to each patient in the ICU TBI dataset.

Patient Number	Predicted Score	Actual Score
104	22.76	25
105	34.57	57
1167	-2.9	67
1538	52.89	62
1677	58.54	71
2110	40.06	9
235	46.04	5
244	29.35	9
2470	19.7	7
248	30.84	23
3217	33.72	78
3245	39.38	64
3299	111.35	4
3625	68.33	10
373	36.16	70
4148	82.4	54
4356	38.88	14
471	61.29	13
5435	26.03	59
5538	121.22	14
6066	52.79	131
6106	60.4	14
6395	7.76	14
651	47.84	152
7	54.66	39
7981	28.36	17

The results of the neural network testing are shown in Table 1 and Fig. 2. The average APACHE score difference between the actual and predicted values is 38.51 out of 299 points, or 12.9% of the possible range of scores.

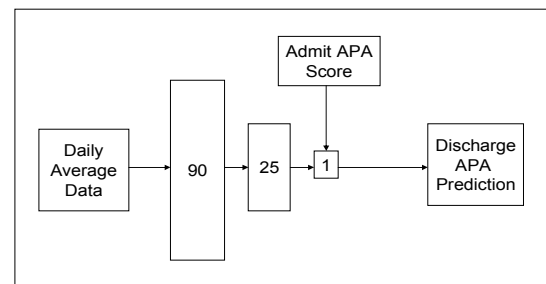


Figure 1. Neural network architecture for prediction of discharge APACHE IV scores in ICU patients with traumatic brain injury [APA = APACHE].

5. Discussion

As can be seen in Table 1 and Fig. 2, some predictions are very close to the patient's actual discharge APACHE score. The overall accuracy of the system, however, is influenced by the presence of several large outliers. The majority of these outliers occur when final APACHE scores are very high (i.e., the patient is doing very poorly) compared to the average APACHE score of the patients.

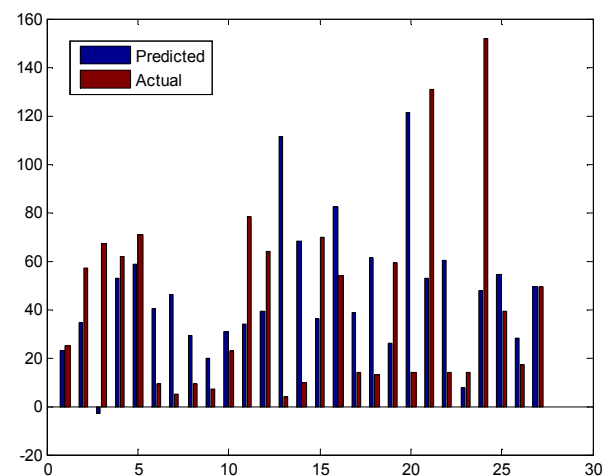


Figure 2. Results of discharge APACHE IV score prediction with a neural network applied to the dataset of intensive care unit patients with traumatic brain injury. Vertical axis represents discharge APACHE IV scores. Horizontal axis represents each of the 27 patients used in this analysis.

A limitation in this analysis is that the ICU data collected had many missing values. The dataset provided to the neural network system still had missing data points despite the fact that a daily average of each clinical value of interest was used as input into the neural network. If more time-granular data (i.e., more frequent than a daily average) of better quality (i.e., with fewer missing values) could be provided to the system, a more accurate prediction of the discharge APACHE score might be possible.

Alternatively, a more sophisticated method for handling missing values could be investigated to try to improve results.

Additional modification to the system, such as further preprocessing of inputs and more sophisticated network techniques, could also enable a more accurate prediction model. From the original 139 ICU patients with traumatic brain injury, only 27 patients had adequate amounts of data available to be useful for this application of neural network training. Clearly, the inability to better handle missing data during neural network development was a large limitation to making good use of the full ICU TBI dataset.

Neural networks were chosen for this study because of their familiarity, though typically they are used to *classify* a new case. In this domain, for example, a classifier would decide between two choices, e.g., predicting whether an ICU patient will likely get discharged from the ICU, or will likely die. Future work should explore the use of various regression methods for predicting values of a numerical nature. A more accurate prediction system may also be possible using other types of machine learning classifiers, such as Bayesian networks and dynamic Bayesian networks, which might be better suited to handling missing data values and disparate types of time-series input data streams [14-19].

6. Conclusion

The current results show that prediction of discharge APACHE IV scores for intensive care unit patients admitted with traumatic brain injury is indeed feasible using a neural network that considers clinical measurements from the first ten days of a patient's ICU admission.

Acknowledgments

The views expressed herein are those of the authors and do not reflect the official policy or position of Brooke Army Medical Center, the U.S. Army Medical Department, the U.S. Army Office of the Surgeon General, the Department of the Army, the Department of Defense, the U.S. Government, or the Defense Advanced Research Projects Agency (DARPA). The authors would like to acknowledge and thank DARPA for financial support via DARPA research contract W31P4Q-06-C-0016.

The authors would like to thank Minhyung Lee, PhD, and Michael Roan, PhD, of Virginia Tech for their significant contributions to this project. The authors would also like to thank the research nurse who performed the APACHE IV score calculations. The authors would furthermore like to thank the staff and patients at the Trauma-Military Medical Center from where these data were collected, without whom this study could not have been performed.

References

- [1] J.E. Zimmerman, A.A. Kramer, D.S. McNair, F.M. Malila, "Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients," *Crit Care Med*, vol. 34, 2006, pp. 1297-1310.
- [2] M.A. de Jongh, M.H. Verhofstad, L.P. Leenen, "Accuracy of different survival prediction models in a trauma population," *Br J Surg.*, vol. 97, Epub 2010 Aug 19, pp. 1805-13.
- [3] S.C. Lim, A.C.K. Fok, Y.Y. Ong, "Patient outcome and intensive care resource allocation using APACHE II," *Singapore Med J.*, vol. 37, 1996, pp. 488-491.
- [4] F.H. Millham, W.W. LaMorte, "Factors associated with mortality in trauma: re-evaluation of the TRISS method using the National Trauma Data Bank," *J Trauma*, vol. 56, 2004, pp. 1090-1096.
- [5] M.J. Vassar, F.R. Lewis Jr., J.A. Chambers, *et al*, "Prediction of outcome in intensive care unit trauma patients: a multicenter study of Acute Physiology and Chronic Health Evaluation (APACHE), Trauma and Injury Severity Score (TRISS), and a 24-hour intensive care unit (ICU) point system," *J Trauma*, vol. 47, 1999, pp. 324-9.
- [6] S.M. DiRusso, T. Sullivan, C. Holly, *et al*, "An artificial neural network as a model for prediction of survival in trauma patients: validation for a regional trauma area," *J Trauma*, vol. 49, 2000, pp. 212-223.
- [7] B. Eftekhari, K. Mohammad, H.E. Ardebili, *et al*, "Comparison of artificial neural network and logistic regression models for prediction of mortality in head trauma based on initial clinical data," *BMC Med Inform Decis Mak.*, vol 5, 2005.
- [8] C.M. Ryan, D.A. Schoenfeld, W.P. Thorpe, *et al*, "Objective estimates of the probability of death from burn injuries," *N Engl J Med.*, vol. 338, 1998, pp. 362-6.
- [9] A. Knoll, M. De Kamps, "Roadmap of neuro-IT development," Ver 1.3: April 19, 2004, <http://www.neuro-it.net/NeuroIT/Roadmap/RoadmapVersions/Roadmapv1.3>.
- [10] M. Sordo, "Introduction to neural networks in healthcare," *Open Clinical*, October, 2002.
- [11] ANNIMAB-1 Abstracts: Clinical diagnosis and medical decision support. www.phil.gu.se/ann/abstracts4.html.
- [12] R.W. Brause, "Medical analysis and diagnosis by neural networks, Lecture Notes in Computer Science, J.W. Goethe University, Frankfurt, Germany, 2001.
- [13] W. Ishaket *al*, "The potential of neural networks in medical applications," July 2004, <http://www.generation5.org/content/2004/NNAppMed.asp>.
- [14] C. J. Langmead, "Generalized queries and Bayesian statistical model checking in dynamic Bayesian networks: application to personalized medicine," *Proceedings of The 8th Annual International Conference on Computational Systems Bioinformatics (CSB)*, 2009, pp. 201-212.
- [15] E. Castillo, J.M. Gutierrez, A.S. Hadi, "Learning Bayesian Networks," in: *Expert Systems and Probabilistic Network Models (Monographs in Computer Science)*. New York, NY: Springer-Verlag, 1997, pp. 481-528.
- [16] Z. Ghahramani. "Learning dynamic Bayesian networks," *Lecture Notes in Computer Science*, vol. 1387, 1997, pp. 168-197.
- [17] F.V. Jensen, T.D. Nielsen, "Bayesian networks and decision graphs," *Information Science and Statistics series (2nd ed.)*. New York, NY; Springer Science + Business Media, LLC; 2007.

- [18] J. Pearl, S. Russell, "Bayesian networks," In M.A. Arbib, Handbook of Brain Theory and Neural Networks. Cambridge, MA: MIT Press, 2002, pp. 157-160.