J Neurosurg 113:585–590, 2010

Use of an artificial neural network to predict head injury outcome

Clinical article

Anand I. Rughani, M.D.,[1] Travis M. Dumont, M.D.,[1] Zhenyu Lu, M.S.,[2] Josh Bongard, Ph.D.,[2] Michael A. Horgan, M.D.,[1] Paul L. Penar, M.D.,[1] and Bruce I. Tranmer, M.D.[1]

[1]Division of Neurosurgery and [2]Department of Computer Science, University of Vermont, Burlington, Vermont

Object. The authors describe the artificial neural network (ANN) as an innovative and powerful modeling tool that can be increasingly applied to develop predictive models in neurosurgery. They aimed to demonstrate the utility of an ANN in predicting survival following traumatic brain injury and compare its predictive ability with that of regression models and clinicians.

Methods. The authors designed an ANN to predict in-hospital survival following traumatic brain injury. The model was generated with 11 clinical inputs and a single output. Using a subset of the National Trauma Database, the authors "trained" the model to predict outcome by providing the model with patients for whom 11 clinical inputs were paired with known outcomes, which allowed the ANN to "learn" the relevant relationships that predict outcome. The model was tested against actual outcomes in a novel subset of 100 patients derived from the same database. For comparison with traditional forms of modeling, 2 regression models were developed using the same training set and were evaluated on the same testing set. Lastly, the authors used the same 100-patient testing set to evaluate 5 neurosurgery residents and 4 neurosurgery staff physicians on their ability to predict survival on the basis of the same 11 data points that were provided to the ANN. The ANN was compared with the clinicians and the regression models in terms of accuracy, sensitivity, specificity, and discrimination.

Results. Compared with regression models, the ANN was more accurate ($p < 0.001$), more sensitive ($p < 0.001$), as specific ($p = 0.260$), and more discriminating ($p < 0.001$). There was no difference between the neurosurgery residents and staff physicians, and all clinicians were pooled to compare with the 5 best neural networks. The ANNs were more accurate ($p < 0.0001$), more sensitive ($p < 0.0001$), as specific ($p = 0.743$), and more discriminating ($p < 0.0001$) than the clinicians.

Conclusions. When given the same limited clinical information, the ANN significantly outperformed regression models and clinicians on multiple performance measures. While this paradigm certainly does not adequately reflect a real clinical scenario, this form of modeling could ultimately serve as a useful clinical decision support tool. As the model evolves to include more complex clinical variables, the performance gap over clinicians and logistic regression models will persist or, ideally, further increase. *(DOI: 10.3171/2009.11.JNS09857)*

Key Words • **artificial neural network** • **outcome prediction** • **head trauma** • **traumatic brain injury**

Each year 1.5 million Americans sustain a TBI,[18] leading to more than 50,000 deaths and 290,000 hospitalizations with outcomes that are far from uniform. Accurately predicting TBI outcomes is a significant clinical challenge and has implications for factors extending from treatment decisions to family counseling.[15] A recent survey of physicians has revealed that only 37% believe that they accurately assess prognosis.[17] Of 60 survey respondents, 88% said that a more accurate model would change the way they report prognosis to a relative, and 67% said it would alter patient care.

Artificial neural networks, computational models based on the presumed architecture of the human brain, offer a unique way to model complex systems. Although many types of ANNs exist, one of the most common structures consists of an interconnected group of nodes in multiple layers, in which input nodes and output nodes have clinical correlates. Hidden nodes, which connect to inputs and outputs, allow nonlinear interactions among the input variables and do not have real-world correlates (Fig. 1). The nodes are connected by links, and "each" link has an associated weight. This network is "trained" by exposure to inputs paired with known outputs, and "learning" occurs when the weights between nodes are modified according to feedback. The computational power of an ANN is derived from the distributed nature of connections.[1,4] Once a model is trained it can then be tested against novel records to predict outputs.

Abbreviations used in this paper: ANN = artificial neural network; AUROC = area under the receiver operating characteristic curve; ED = emergency department; GCS = Glasgow Coma Scale; GOS = Glasgow Outcome Scale; ICP = intracranial pressure; NTDB = National Trauma Data Bank; TBI = traumatic brain injury.

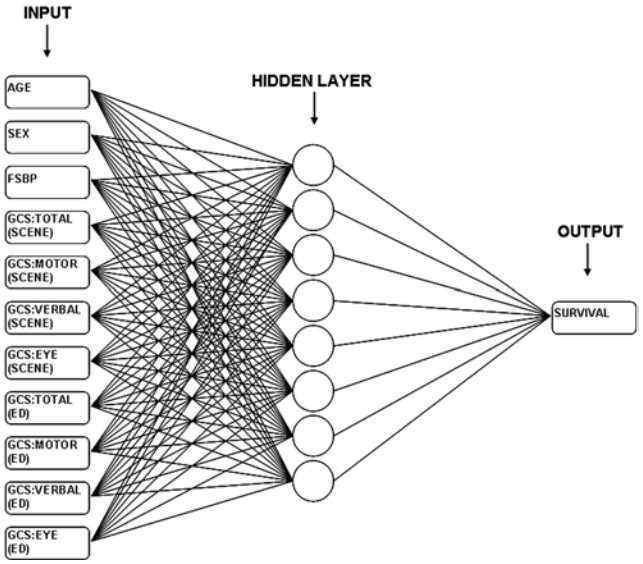

Fig. 1. A schematic representation showing our model with 11 input nodes, 8 nodes in a single hidden layer, and a single output node, which represents in-hospital survival.

Compared with traditional forms of modeling such as linear regression analysis, ANNs provide certain advantages. The most significant is the ability to allow for nonlinear interactions among variables.[1] Other advantages of ANNs are the fact that no a priori knowledge of the relationships between variables is necessary and that ANNs are capable of readily handling missing data points.[18]

We hypothesized that an ANN could be used to predict survival in TBI with results similar to those of a clinician or regression model when exposed to the same data. We trained an ANN by using a large national database and tested the network against clinicians and 2 logistic regression models for comparison.

Methods

The data set was derived from the NTDB, version 6.2, which is a national registry maintained by the American College of Surgeons and contains nearly 3,000,000 records assembled from 712 hospitals from 2002 to 2007. This data set is not a population data set and relies on the participation of individual hospitals, which provide data. Variability in the data collection and interhospital variability are limitations of this type of data set. Furthermore, NTDB data points are collected by the individual reporting hospitals and are verified by the American College of Surgeons for logical consistency and completeness but not for accuracy. The main advantage of this data set is its size, which allows sufficient training of the ANN. The data set we used was limited to cases of head trauma only by selecting records with a positive head CT. This strategy limited the data set to > 200,000 records. The 11 input variables included in our model were age, sex, total GCS score at the scene of injury, individual components of the GCS score at the scene of injury (eye, verbal, and motor score), total score and individual components of the GCS in the ED, and first systolic blood pressure measured in the ED. Variables were selected on the basis of known or expected influence on outcome, including age,[11] sex,[6] and hypotension.[3] It should be emphasized that several variables crucial to outcome prediction are absent from the NTDB, such as pupillary reactivity and specific radiographic findings. The solitary output variable we examined was in-hospital survival. All incomplete records were excluded, which further limited the data set to 7769 records.

A schematic representation of our model is featured in Fig. 1, and the detailed training algorithm is described in the *Appendix*. Unlike traditional back-propagation models, training in our model relied on a technique called "informative sampling" that we have described in a feed-forward 3-layer neural network.[12] This technique allows a small but very informative subset of the entire training set to be selectively emphasized by the algorithm during the training phase. In short, training consists of pairing the 11 inputs with the known output and allowing the algorithm to adjust the strength of the different connections between variables. The strength of these connections, which is defined by mathematical functions, can vary with repeated exposure to data and is adjusted according to feedback.

We trained 30 ANNs simultaneously, which allowed us to produce a single prediction by averaging the 30 individual predictions. This approach also allowed us to evaluate each ANN individually for statistical purposes. In this way the 5 most accurate ANNs were identified and pooled for analysis. The models were compared with clinicians, which consisted of 5 neurosurgery residents and 4 neurosurgery staff physicians, using the same testing set.

Traditional outcome modeling relies heavily on the use of logistic regression models, such as logistic model trees[7] and Bayesian logistic regression methods.[12] For the sake of comparing our ANN with logistic models, we developed a logistic model tree and a Bayesian model using the same data points that the ANN was trained on. This approach provided us with 2 separate regression models that could be evaluated against the ANN using the same testing set.

A testing set of 100 novel patients was generated from the NTDB. Because of an over-representation of patients with a total GCS score of 15, our testing set excluded those patients to create a more heterogeneous sample. When we initially used a testing set with GCS scores of 3–15, we found that because of the over-representation of patients with a GCS score of 15, the performance of both the ANN and clinicians was similar given that the high survival rate was easily predicted. Thus, the testing set we used excluded all patients with a GCS score of 15 only for the purpose of testing. The distribution of ED total GCS scores from our 100-patient testing set is represented in Fig. 2. Of note, the models were trained on a group of patients that had GCS scores of 3–15. Both the models and the clinicians were supplied with the same data set, which exposed the clinicians to only 11 inputs, and were asked to predict in-hospital survival. The clinicians were given a table, which consisted of the 100 patients in each row and the 11 clinical variables in each column. Clinical predictions were made at one sitting and marked on the table, with no real-time feedback on performance.

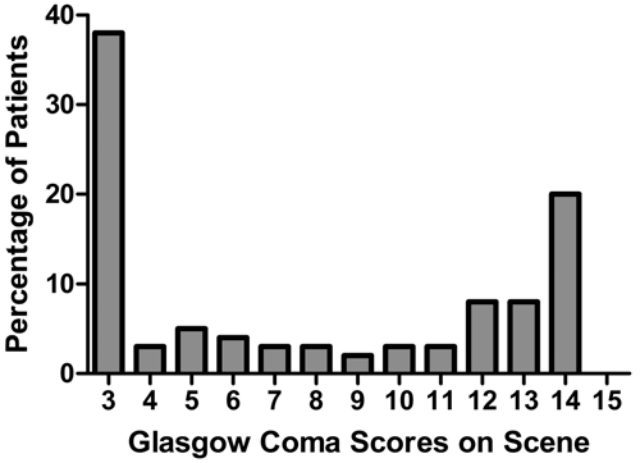

Fig. 2. Bar graph showing the frequency of GCS scores in the testing set of 100 patients. All patients with a GCS score of 15 have been eliminated to produce a more heterogeneous testing set.

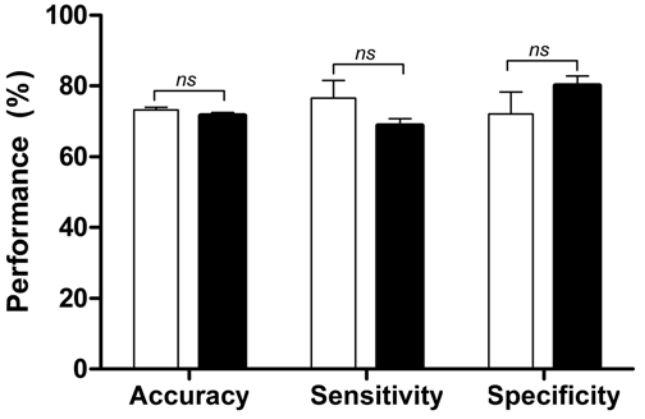

Fig. 3. Bar graph comparing the performance of 5 neurosurgery residents with that of 4 staff physicians in predicting in-hospital survival of 100 patients with TBI. *White bars* indicate residents; *black bars*, staff physicians. ns = not significant.

The predictions were compared with known actual outcomes, as recorded in the NTDB. Both the models and the clinicians were evaluated on accuracy, sensitivity, and specificity. In this context, "sensitivity" refers to the ability to correctly predict survival, whereas "specificity" refers to the ability to correctly predict death. The AUROC, which is a plot of the true-positive fraction against the false-positive fraction, provides a measure of discrimination.[8,9] The AUROC would be 1 for any model that performs without any errors and would be 0.5 if that model operates at chance. On each measure, the groups were statistically compared using a 1-tailed Student t-test.

Results

The training set was composed of 7769 patients, 72% male, with a mean age of 39.1 years, a mean total on-scene GCS score of 8.3 (eye = 2.3, verbal = 2.4, and motor = 3.6), and a mean total ED GCS score of 8.5 (eye = 2.4, verbal = 2.4, and motor = 3.8). The testing set consisted of 100 novel patients, 74% male, with a mean age of 37.1 years, a mean total on-scene GCS score of 7.8 (eye = 2.2, verbal = 2.3, and motor = 3.4), and a mean total ED GCS score of 7.6 (eye = 2.1, verbal = 2.2, and motor = 3.3). The in-hospital survival rate in this group was 75%.

The clinicians consisted of 5 neurosurgery residents with an average of 4 years of postgraduate training and 5 neurosurgery staff physicians with an average of 11 years of active surgical practice including the care of trauma patients. When the level of training or years of practice was analyzed individually, there was no effect of level of training, or years of practice on the accuracy, sensitivity, specificity, or AUROC. Residents were compared with the staff physicians (Fig. 3) and showed no differences in accuracy (73.2 vs 71.5%), sensitivity (76.6 vs 68.7%), specificity (72.0 vs 80.0%), or AUROC (0.74 vs 0.74), with $p > 0.05$ on all comparisons.

The 5 most accurate ANNs were pooled and compared with all clinicians (Fig. 4). Accuracy was higher for the ANNs than the clinicians (87.8 vs 72.4%, $p < 0.0001$), sensitivity was higher for the ANNs than the clinicians (98.6 vs 73.1%, $p < 0.0001$), and specificity was no different for the ANNs compared with the clinicians (74.1 vs 75.5%, $p = 0.40$). Discrimination was better for the ANNs (Fig. 5A), as measured by the AUROC (0.86 for the ANNs vs 0.74 for the clinicians, $p < 0.0001$).

The same 5 ANNs were pooled and compared with the 2 different regression models (Fig. 4). Accuracy was higher for the ANNs than the regression models (87.8 vs 78.0%, $p < 0.001$), sensitivity was higher for the ANNs than the regression models (98.6 vs 78.7%, $p < 0.001$), specificity was similar (74.1 vs 76.0%, $p = 0.260$), and the AUROC was 0.86 for the ANNs compared with 0.77 for the regression models ($p < 0.001$), as represented in Fig. 5B.

Discussion

We used a large database to train and test a neural network in outcome prediction in TBI. This model was

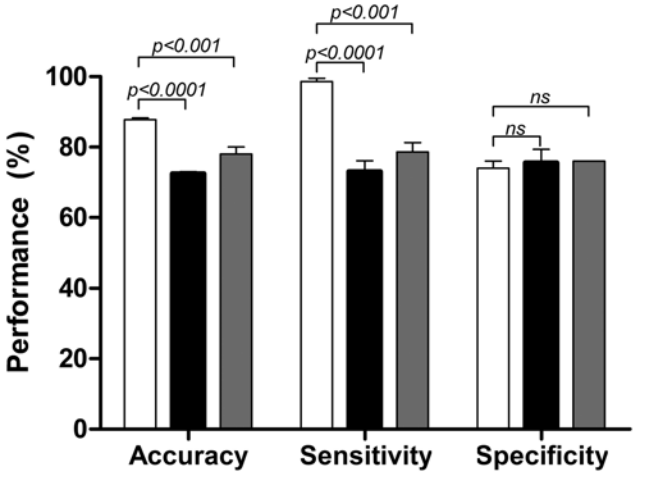

Fig. 4. Bar graph comparing performance of 9 clinicians with the 5 most accurate ANNs in predicting in-hospital survival of 100 patients with TBI. *White bars* indicate ANN; *black bars*, clinicians; *gray bars*, regression models.

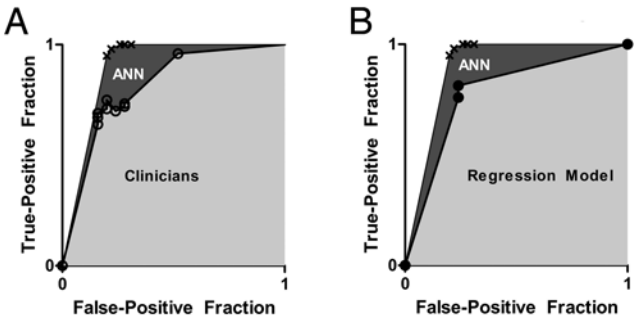

Fig. 5. Graphs demonstrating a mean AUROC for the ANNs of 0.86 compared with 0.74 for the clinicians **(A)** and 0.77 for the regression models **(B)**. The AUROC gives an indication of how well a model discriminates among different outcomes. A model with an AUROC of 0.5 suggests that the model operates at the level of chance, and a model with an AUROC of 1.0 discriminates outcomes perfectly.

tested against actual outcomes, and the neural network performance was compared with clinicians and 2 logistic regression models created with the same inputs. We showed that by using a limited number of clinical inputs and a simple outcome measure, the ANN performance surpasses that of clinicians and logistic regression models using the same limited variables.

While ANNs are not a novel form of modeling, they remain underutilized in the field of neurosurgery. With varying results, 5 studies have been focused on the role of ANNs in predicting the outcome of TBI.

Lang and colleagues[13] were the first group to demonstrate that an ANN was as accurate, sensitive, and specific as standard logistic regression in predicting 6-month survival following severe head injury. More recently, Segal and colleagues[18] have compared an ANN with a multiple regression model in predicting several different functional outcome scores at 1 year after TBI. They used sophisticated linear models, which performed with the same accuracy as an ANN. Another group[5] actually showed that an ANN was less accurate than linear regression in predicting survival, although it was marginally better at discriminating outcomes. While they did not compare ANN with other forms of modeling, Hsu and colleagues[10] did show that an ANN could predict the GOS score at 12 months or more after injury with 75.8% accuracy. Most recently, Pang and colleagues[16] compared an ANN to 4 other types of models, all of which outperformed the ANN in accurately predicting the GOS score at 6 months after severe TBI.

In short, the use of ANNs in TBI outcome prediction has shown mixed results, probably because of variations in modeling techniques, patient demographics, and the clinical variables examined. We believe that our results support the continued use and refinement of ANNs in modeling TBI outcomes. Our current model has a limited number of inputs, which is primarily a function of the data set used, and nonetheless outperforms clinicians in accuracy, sensitivity, and discrimination when using the same testing set. We predict that as the number of model inputs is expanded and the predictions are made more specific, this discrepancy will further increase.

We believe that the advantage of this model over previous ANN models derives from a unique aspect of the training algorithm known as "informative sampling." The model takes an active approach to the training algorithm;[2,14,19] that is, new patient records are periodically selected from the database for further training of the ANN, based on the ability of a particular patient record to induce disagreement among the ANN's predicted outcomes. Essentially, during the training phase, the model preferentially selects patients who are likely to have an ambiguous outcome instead of a patient who, for example, has a GCS score of 15 and is clearly predicted to survive. By selectively emphasizing those borderline cases, the model is better able to discriminate outcomes during the testing phase.

Our model was trained on patients with mild, moderate, and severe TBI. When tested on all patients with TBI, the accuracy of the ANN and the clinicians was the same (unpublished data), because of the extremely high survival rates when patients with GCS scores of 15 are included. To create greater heterogeneity in the testing set, all patients with a GCS score of 15 were eliminated. While we have shown that our model can readily generalize to include mild, moderate, and severe TBI, the model fails to outperform experts on most mild TBI cases, when survival is unambiguous.

Our ANN suffers from several limitations. First, the inputs do not include pupillary examination, brainstem reflexes, reference to comorbidities or multiple injuries, CT findings, procedures performed, or intubation status, primarily because of the nature of the database we used for both training and testing. It speaks to the utility of the ANN; that is, despite such limited data points, the ANN's performance surpasses that of both regression models and clinicians. We believe that the inclusion of such data points will allow a further increase in accuracy and discrimination over regression models, and we are developing a database that includes such variables.

A related limitation when comparing a model to a clinician is the failure to capture the myriad data points that are present in a clinical encounter. An individual patient's medical record certainly fails to capture many of these variables. A large database fails to capture even more of the clinical nuance specific to each patient, including such important factors as elective withdrawal of care. Furthermore, a database such as the NTDB suffers from interhospital variability in data gathering and reporting. When these various biases are reduced they will likely decrease the outperformance of the ANN compared with clinicians when a prospective comparison is made.

Accurately predicting outcome in head injury will ultimately require predicting more than just survival. Of course, survival remains the basic question that emerges early in the care of severely brain-injured patients, but it is certainly not the only question that clinicians face. Clinically meaningful predictions will also come in the form of predicting clinical course, such as increased ICP, need for ICP monitoring, need for surgical treatment, duration of hospitalization, duration of intensive care unit stay, discharge disposition, and short- and long-term functional status. Adequately predicting these outcomes with any form of modeling, including ANNs, will require large

and sophisticated data sets with detailed clinical as well as long-term follow-up information. In the current study, the NTDB did not allow for these predictions. Although training ANNs requires a large number of cases and determining a priori the number of patients required is exceedingly difficult to do, we are currently building a data set that we hope will allow more sophisticated prognostication than the current model.

Conclusions

We have demonstrated that with the same data an ANN can consistently outperform a linear regression model. Our work serves principally to generate support for the continued use of ANNs as predictive modeling techniques in neurosurgery. It would be premature to characterize our current model as a clinically useful predictive model. Future directions for work on this model will include the development of a more detailed database that includes CT and pupillary examination findings as well as more detailed outcome data. As our model evolves, we hope that its performance will continue to improve and that it will eventually have utility as an adjunctive clinical decision tool.

Appendix

The ANN is a standard feed-forward neural network with 3 layers (input, hidden, and output). The input layer has 16 nodes; the hidden layer, 8 nodes; and the output layer, 1 binary node (that is, 1 for fatal and 0 for nonfatal patient outcome). An illustration of the structure of the ANN is provided in Fig. 1. For each link between nodes, there is an associated weight. For each input layer, the value is its corresponding feature value from the current data point. For each node in the hidden layer, the value is calculated as the sum of each linked weight multiplied by the weight's corresponding input node. The resulting sum is passed through an activation function. For each node in the output layer, the value is calculated like the nodes in the hidden layer, with the input values from the hidden nodes. The behavior of the ANN can be summarized by the following equation:

$$o = h(\sum_j w_{jk} \times h(\sum_i w_{ij} \times f_i + \theta_j) + \theta_k)$$

where o is the value of a node in the output layer, i denotes a node in the input layer, j denotes a node in the hidden layer, k denotes the node in the output layer, w_{ij} is the weight from node i to j, w_{jk} is the weight from node j to k, f_i is the value of the ith feature of an input data point, Θj is a constant called the bias of node j, Θk is the bias of node k, and h is the sigmoid activation function. The value of the output node is a real value between 0 and 1. In this paper, when $o \leq 0.5$, the output of the ANN is set to 0, and when $o > 0.5$ the output is 1. An output of 1 represents survival, and 0 nonsurvival. The algorithm starts by creating a population of 30 ANNs. The weights of the neural networks are randomly selected real numbers between −1 and 1. The data set is divided into subsets of size p. The exploration phase is run before the estimation phase at the outset of the algorithm. The first data point to the pth data point is supplied to the current set of candidate models (the random 30 ANNs) sequentially. The predictions of the 30 models on 1 data point are recorded, and the difference between the number of survival predictions and fatal predictions is then calculated for that point. The data point with the lowest such value is chosen and added into the training set with its corresponding label. The exploration phase is followed by the estimation phase. At the outset of the first pass through the estimation phase, the random ANNs are trained on the single training datum using the following equation:

$$fitness = (c_{nf} \div t_{nf}) \times (c_f \div t_f)$$

where c_{nf} is the number of correct predictions of nonfatal patients, t_{nf} is the total number of nonfatal patients, c_f is the number of correct predictions of fatal patients, and t_f is the total number of fatal patients. The training is done by applying a mutation operator on each ANN. The fitness of the parent ANN and the child ANN is calculated. If the child ANN is better, the parent model is replaced with it; otherwise, the parent is retained. One generation is evolved each time the estimation phase is run. This process instantiates a parallel hill climbing algorithm. The mutation operator takes a random weight of the ANN and changes it to a random real value between −1 and 1. The fitness function is suggested specifically for highly unbalanced data sets (for example, when survival is much more likely than nonsurvival). A simple fitness function would mislead the models to only output the majority label (that is, survival) regardless of the input. The fitness function used here shapes the ANNs to produce correct predictions on data from both the majority (survival) and minority (nonsurvival) classes. The algorithm then returns to the exploration phase. This time, the $(p + 1)$th data point to the $2p$th data point are supplied to the models that were evolved in the estimation phase. One data point is selected based on model disagreement and added into the training set (which now contains 2 data points). The estimation phase is run again after the exploration phase on the updated training set. One round of the algorithm consists of a single run of the exploration phase followed by a single run of the estimation phase. These steps are repeated until the performance of the model reaches a plateau with subsequent iterations.

Disclosure

This work is partially supported by the National Science Foundation Grant No. EPS-0701410 (J.B. and Z.L.).

Author contributions to the study and manuscript preparation include the following. Conception and design: AI Rughani, J Bongard. Analysis and interpretation of data: AI Rughani, TM Dumont, Z Lu. Drafting the article: AI Rughani, PL Penar. Critically revising the article: AI Rughani, MA Horgan, PL Penar. Reviewed final version of the manuscript and approved it for submission: AI Rughani, TM Dumont, Z Lu, J Bongard, MA Horgan, PL Penar, BI Tranmer. Statistical analysis: AI Rughani, TM Dumont, Z Lu. Administrative/technical/material support: J Bongard, MA Horgan, BI Tranmer. Study supervision: J Bongard, BI Tranmer. Other (programming of model): Z Lu, J Bongard.

Acknowledgments

The authors would like to thank Alan Howard for assistance with statistical analysis. Committee on Trauma, American College of Surgeons, National Trauma Data Bank, version 6.2, Chicago, IL, 2007. The content reproduced from the NTDB remains the full and exclusive copyrighted property of the American College of Surgeons. The American College of Surgeons is not responsible for any claims arising from works based on the original data, text, tables, or figures.

References

1. Baxt WG: Application of artificial neural networks to clinical medicine. **Lancet 346:**1135–1138, 1995
2. Bongard J, Lipson H: Nonlinear system identification using coevolution of models and tests. **IEEE Trans Evol Comput 9:**361–384, 2005
3. Chesnut RM, Marshall SB, Piek J, Blunt BA, Klauber MR, Marshall LF: Early and late systemic hypotension as a frequent and fundamental source of cerebral ischemia following severe brain injury in the Traumatic Coma Data Bank. **Acta Neurochir Suppl (Wien) 59:**121–125, 1993

4. Cross SS, Harrison RF, Kennedy RL: Introduction to neural networks. **Lancet 346:**1075–1079, 1995
5. Eftekhar B, Mohammad K, Ardebili HE, Ghodsi M, Ketabchi E: 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 5:**3, 2005 (Abstract)
6. Farace E, Alves WM: Do women fare worse: a metaanalysis of gender differences in traumatic brain injury outcome. **J Neurosurg 93:**539–545, 2000
7. Genkin A, Lewis D, Madigan D: Large-scale Bayesian logistic regression for text categorization. **Technometrics 49:**291–304, 2007
8. Hagen MD: Test Characteristics. How good is that test? **Prim Care 22:**213–233, 1995
9. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. **Radiology 143:**29–36, 1982
10. Hsu MH, Li YC, Chiu WT, Yen JC: Outcome prediction after moderate and severe head injury using an artificial neural network. **Stud Health Technol Inform 116:**241–245, 2005
11. Hukkelhoven CW, Steyerberg EW, Rampen AJ, Farace E, Habbema JD, Marshall LF, et al: Patient age and outcome following severe traumatic brain injury: an analysis of 5600 patients. **J Neurosurg 99:**666–673, 2003
12. Landwehr N, Hall M, Frank E: Logistic Model Trees. **Machine Learning 59:**161–205, 2005
13. Lang EW, Pitts LH, Damron SL, Rutledge R: Outcome after severe head injury: an analysis of prediction based upon comparison of neural network versus logistic regression analysis. **Neurol Res 19:**274–280, 1997
14. Lu Z, Rughani AI, Tranmer BI, Bongard J: Informative sampling for large unbalanced data sets, in **Proceedings of the 2008 GECCO conference companion on Genetic and evolutionary computation.** New York: ACM Press, 2008, pp 2047–2054
15. Machado SG, Murray GD, Teasdale GM: Evaluation of designs for clinical trials of neuroprotective agents in head injury. **J Neurotrauma 16:**1131–1138, 1999
16. Pang BC, Kuralmani V, Joshi R, Hongli Y, Lee KK, Ang BT, et al: Hybrid outcome prediction model for severe traumatic brain injury. **J Neurotrauma 24:**136–146, 2007
17. Perel P, Wasserberg J, Ravi RR, Shakur H, Edwards P, Roberts I: Prognosis following head injury: a survey of doctors from developing and developed countries. **J Eval Clin Pract 13:** 464–465, 2007
18. Segal ME, Goodman PH, Goldstein R, Hauck W, Whyte J, Graham JW, et al: The accuracy of artificial neural networks in predicting long-term outcome after traumatic brain injury. **J Head Trauma Rehabil 21:**298–314, 2006
19. Seung HS, Opper M, Sompolinsky H: Query by Committee, in **Proceedings of the Fifth Annual ACM Conference on Computational Learning Theory.** New York: ACM Press, 1992, pp 287–294

Manuscript submitted May 31, 2009.
Accepted November 17, 2009.
Preliminary results from this project have been presented in part in abstract form at the 2008 Annual Meeting of the Congress of Neurological Surgeons in Orlando, Florida.
Please include this information when citing this paper: published online December 18, 2009; DOI: 10.3171/2009.11.JNS09857.
Address correspondence to: Anand I. Rughani, M.D., 111 Colchester Avenue, Fletcher 5, Burlington, Vermont 05401. email: anand.rughani@vtmednet.org.