Using an artificial neural network to predict traumatic brain injury

Andrew T. Hale, MS,1,2 David P. Stonko, MS, MD,2,3 Jaims Lim, MD,2
Oscar D. Guillamondegui, MD, MPH,2,3,6 Chevis N. Shannon, PhD,2,6,7 and
Mayur B. Patel, MD, MPH2,4–6

1Vanderbilt University School of Medicine, Medical Scientist Training Program; 2Vanderbilt University School of Medicine, Nashville, Tennessee; 3Johns Hopkins Medical Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland; 4Division of Trauma, Emergency General Surgery, and Surgical Critical Care, Departments of Surgery and Hearing and Speech Sciences, Section of Surgical Sciences, Vanderbilt University Medical Center; 5Center for Health Services Research, Vanderbilt Brain Institute, Vanderbilt University Medical Center and Geiatric Research, Education and Clinical Center Service, Surgical Service, Department of Veterans Affairs Medical Center, Tennessee Valley Health Care System; 6Department of Neurosurgery, Vanderbilt University Medical Center and Division of Pediatric Neurosurgery, Monroe Carell Jr. Children’s Hospital of Vanderbilt University; and 7Surgical Outcomes Center for Kids, Monroe Carell Jr. Children’s Hospital of Vanderbilt University, Nashville, Tennessee

OBJECTIVE Pediatric traumatic brain injury (TBI) is common, but not all injuries require hospitalization. A computational tool for ruling in patients who will have a clinically relevant TBI (CRTBI) would be valuable, providing an evidence-based way to safely discharge children who are at low risk for a CRTBI. The authors hypothesized that an artificial neural network (ANN) trained on clinical and radiologist-interpreted imaging metrics could provide a tool for identifying patients likely to suffer from a CRTBI.

METHODS The authors used the prospectively collected, publicly available, multicenter Pediatric Emergency Care Applied Research Network (PECARN) TBI data set. All patients under the age of 18 years with TBI and admission head CT imaging data were included. The authors constructed an ANN using clinical and radiologist-interpreted imaging metrics in order to predict a CRTBI, as previously defined by PECARN: 1) neurosurgical procedure, 2) intubation > 24 hours as direct result of the head trauma, 3) hospitalization ≥ 48 hours and evidence of TBI on a CT scan, or 4) death due to TBI.

RESULTS Among 12,902 patients included in this study, 480 were diagnosed with CRTBI. The authors’ ANN had a sensitivity of 99.73% with precision of 98.19%, accuracy of 97.98%, negative predictive value of 91.23%, false-negative rate of 0.0027%, and specificity for CRTBI of 60.47%. The area under the receiver operating characteristic curve was 0.9907.

CONCLUSIONS The authors are the first to utilize artificial intelligence to predict a CRTBI in a clinically meaningful manner, using radiologist-interpreted CT information, in order to identify pediatric patients likely to suffer from a CRTBI. This proof-of-concept study lays the groundwork for future studies incorporating iterations of this algorithm directly into the electronic medical record for real-time, data-driven predictive assistance to physicians.

https://thejns.org/doi/abs/10.3171/2018.8.PEDS18370

KEYWORDS TBI; pediatrics; machine learning; artificial intelligence; trauma

Traumatic brain injury (TBI) affects thousands of children in the United States every year. Despite the large number of children who experience TBI, only a small percentage actually require hospitalization or prolonged surveillance. However, identifying which patients do require monitoring compared with those who can be safely discharged from the emergency department remains an important unanswered question. Thus, creation of a tool for identifying patients at risk for clinically relevant TBI (CRTBI) could provide an evidence-based...
mechanism for early safe discharge and potentially reduce unnecessary healthcare expenditures.

The Pediatric Emergency Care Applied Research Network (PECARN) is a consortium of 25 hospitals that developed a decision-making score based on head CT findings. Numerious studies have independently published on these data in an effort to develop predictive metrics to guide treatment of children with TBI; however, none have used artificial neural networks (ANNs). ANNs are a type of machine-learning algorithm that have been widely used in conventional medicine. They are often more useful than conventional statistical methods because: 1) ANNs can take any number of input variables and predict any number of outcomes; 2) they are capable of improving their predictive ability over time as they are exposed to new data; 3) they benefit from internal validation and testing; and 4) they tend to have stronger discriminant ability compared to conventional statistics.

Leveraging this technology, we created a model that combines clinical and radiologist-interpreted CT data to predict whether or not a pediatric patient will experience a CRTBI. We quantify the accuracy and error of this algorithm and provide an open-source software package to enable prediction generation and validation. We expand on previous PECARN predictive studies by utilizing a combination of demographic, clinical, and radiologist-interpreted CT data to investigate CRTBIs in pediatric patients using an ANN. For the present study we hypothesized that we could train an ANN on clinical and radiographic data to identify which pediatric TBI patients with head CT scans are at risk for a CRTBI.

**Methods**

**Study Population**

This study utilized the prospective PECARN study of children with CRTBI, as described previously. The PECARN TBI study enrolled patients under the age of 18 who experienced nonpenetrating (i.e., blunt) head trauma who presented to the emergency department between 2004 and 2006 and had admission head CT imaging classification. All data analyzed in this study were de-identified, and our study was approved by the Vanderbilt University Institutional Review Board. We included patients who had complete data available for all variables of interest and thus did not impute any missing variables. Head CT scanning was performed in 14,969 patients, 12,902 of whom had complete imaging information.

**Analysis and Included Variables**

Descriptive statistics including Pearson correlation and t-test were used to evaluate the normally distributed cohort variables. Statistical significance was set a priori at p < 0.05. The input variables included in our ANN were as follows: 1) mechanism of injury (e.g., motor vehicle collision, pedestrian struck by a moving vehicle, bicycle rider struck by an automobile, bicycle collision or fall from bicycle, other wheeled transport crash, fall to ground from standing/walking/running, walked or ran into stationary object, fall from an elevation, fall down stairs, sports, assault, object struck head [accidental], and other etiology of injury); 2) severity of injury mechanism (low [e.g., fall from ground level and walked/ran into a stationary object], moderate [any other mechanism], high [e.g., motor vehicle collision with patient ejection, death of another passenger, or rollover; pediatric or bicyclist without helmet struck by motor vehicle; and fall > 5 feet for patients 2 years and older, or falls of > 3 feet for those < 2 years]); 3) loss of consciousness; 4) Glasgow Coma Scale score at presentation; 5) age; and 6) sex.

Our outcome of interest was a CRTBI, a composite of several variables as defined by the PECARN investigators. The CRTBI variables consisted of any of the following: 1) neurosurgical procedure (e.g., dural repair for cerebrospinal fluid leak, fracture elevation, hematoma drainage, intracranial pressure monitor placement, lobectomy, tissue debridement, ventriculostomy, and “other” neurosurgical procedure), 2) intubated > 24 hours as a direct result of the head trauma, 3) hospitalization ≥ 48 hours and evidence of TBI on a head CT scan, and 4) death due to TBI.

**ANN Analysis**

We trained an ANN using offline MATLAB R2016b (version 9.1.0.441655) on a 64-bit MacBook Pro running OS 10.11.6. We randomly partitioned patients into 3 groups in order to provide holdout validation on our large data set: 70% were for training the ANN; 15% were for validating the ANN; and 15% were for subsequent final testing of the ANN. The ANN had not been exposed to any of the final test patients until after the model was finished training and validating. A two-layer, feed-forward ANN with 11 sigmoid hidden and softmax output neurons was trained using the scaled conjugate gradient back-propagation method on the dedicated partition. We tabulated confusion tables and statistics on the testing partition, as well as for the entire data set. We assessed the predictive ability of the model rigorously with various numerical measures of accuracy, precision, and error.

**Results**

In this study, we included 12,902 patients of whom 63% were male and the average age was 7.99 ± 5.91 years (Table 1). Of the 12,902 patients, 480 suffered a CRTBI.
TABLE 1. Characteristics of pediatric TBI patients studied using an ANN

<table>
<thead>
<tr>
<th>Variable</th>
<th>Non-CRTBI (n = 12,422)</th>
<th>CRTBI (n = 480)</th>
<th>All Patients (n = 12,902)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean ± SD)</td>
<td>8.00 ± 5.92</td>
<td>7.88 ± 5.67</td>
<td>7.00 ± 5.91</td>
<td>0.648</td>
</tr>
<tr>
<td>Sex ratio (M/F)</td>
<td>1.71</td>
<td>1.89</td>
<td>1.72</td>
<td>0.305</td>
</tr>
<tr>
<td>Severity of injury (no.)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low</td>
<td>1807</td>
<td>25</td>
<td>1832</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>7798</td>
<td>243</td>
<td>8041</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>2817</td>
<td>212</td>
<td>3029</td>
<td></td>
</tr>
<tr>
<td>Loss of consciousness</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No</td>
<td>7928</td>
<td>173</td>
<td>8101</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>3220</td>
<td>239</td>
<td>3459</td>
<td></td>
</tr>
<tr>
<td>NOS</td>
<td>1274</td>
<td>68</td>
<td>1342</td>
<td></td>
</tr>
<tr>
<td>GCS score total</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

GCS = Glasgow Coma Scale; NOS = not otherwise specified.

Our outcome of interest was CRTBI, a composite of several variables as defined by the PECARN investigators: 1) neurosurgical procedure; 2) intubation > 24 hours as a direct result of the head trauma; 3) hospitalization ≥ 48 hours; and 4) death due to TBI.

* Univariate statistical significance examined using the t-test or Pearson’s chi-square test.

from age and sex, all other clinical and imaging variables had a univariate association with CRTBI (Table 2).

The ANN has a sensitivity of 99.73% and a negative predictive value (NPV) of 91.23% for CRTBI in the testing cohort (Table 3). When the data used for testing were combined with the remaining 85% of data, which the network was trained and validated on, the sensitivity remained very high at 99.54% and NPV of 84.38% (Table 3). Determination of specificity was much lower at 60.47% for testing and 64.17% for the entire data set. We included other statistical measures of the ANN binary classifier (Table 3). A pictorial representation of the ANN constructed here is shown in Fig. 1.

Receiver operator characteristic (ROC) curves for both test patients and the entire data set were calculated (Fig. 2). The area under the ROC curve (AUC) was 0.9907 for prediction of a CRTBI in test patients and 0.9790 for the entire data set.

Discussion

We constructed and validated an ANN, a machine-learning computational algorithm, to predict CRTBIs in children using clinical and imaging data. This platform has apparent clinical utility for the inexperienced pediatric emergency care provider for assigning admission for children with TBI given its very high sensitivity for CRTBI, which, while it has a low prevalence (<5%), can have serious consequences, such as future intracranial procedure(s), respiratory failure, prolonged hospitalization, and/or death. This would be the first study, to our knowledge, aiming to predict TBI of any type in any patient population.

Predictive outcome and prognostication models are
becoming increasingly important across medicine and surgery (e.g., CHA2DS2-VASc score,11,12 APACHE,17 SOFA15), and modeling techniques have evolved over the years. These models have classically relied on logistic regression or conventional statistics to generate predictions, and often they use fewer input variables that are manually entered. More recently, ANNs have been shown to robustly predict complications, outcomes, and prognosis among numerous fields,3,13,31,34–36 including TBI.8,18,26,32,37,38 Thus, an ANN tool yielding predictive information concerning CRTBI would be helpful and provide an evidence-based mechanism for treating these patients.

ANNs are computational constructs used to interpret the maximum number of combinations of data in complex systems, such as making medical diagnoses in neurosurgery,36,37 where many competing factors influence outcome.2 During training of the ANN, random “weights” are assigned to each input variable, compared against every variable in the model, and are then used to predict the strength of correlation with the outcome of interest (Fig. 1). While there is no maximum number of variables that can be included in an ANN, addition of irrelevant variables will not make the data prediction any stronger.41 Thus, we chose to rationally design the ANN described here by only including variables that had previously been shown by univariate statistical analyses to be significantly associated with CRTBI.

We trained an ANN on data collected from PECARN and successfully developed a very sensitive (sensitivity 99.73%, AUC 0.9907) tool for identifying CRTBI (NPV 91%) in children. We optimized the ANN for sensitivity over specificity to conservatively identify patients likely to be diagnosed with a CRTBI. Future iterations of this ANN with additional variables and data not available through PECARN could be similarly leveraged to optimize specificity, thereby safely ruling out disease. However, since PECARN is a group consisting of 25 hospitals with data collected prospectively, these data most accurately reflect the epidemiological and treatment diversity seen across North America for pediatric TBI. Importantly, the number of variables included in the predictive ANN algorithm can be greatly increased compared with prior risk-calculation tools due to the overwhelmingly computational superiority of machine-learning over conventional statistical approaches, which are limited by degrees of freedom.29,31,37,42 Our intent was to provide software that allows for real-world incorporation of data into a standalone application or an electronic medical record (EMR). Future applications could self-collect these clinical data (i.e., our published algorithm depends on collected and interpreted data) and therefore would integrate all necessary input data from EMR systems and provide results for the clinician on the ground. Another strength of this study is its external generalizability, as we did not further divide our cohort based on any further head-injury patterns (e.g., subdural, epidural, intraparenchymal hematoma, shear) often seen in the literature, as we wanted to reflect the full spectrum of pediatric patients with all severities and pathoanatomical types of TBI presenting to any emergency department.

In the our opinion, future iterations of ANN-based predictive modeling should be centered around 3 guiding principles: 1) prospective data collection leading to real-time updates and refinement of the algorithm, 2) direct linkage of ANN models to the EMRs, and 3) increase in the granularity of data available for training the ANN—for instance, using image-based processing. First, compared to traditional statistical approaches that require new analyses to be performed each time new data are added, ANNs can be constantly updated, providing real-time, up-to-date information and quantitative evidence. ANNs could be designed to use national, regional, or even provider-specific data. Second, directly linking ANNs to the EMR would provide streamlined data collection and up-to-date predictive capabilities based on the most current evidence. Lastly, ANNs could be trained directly on the CT images themselves, leading to a quicker diagnosis, prognosis, and better utilization of hospital resources. Although we lay the groundwork for incorporating machine-learning into evaluation of children with TBI, this study is not without limitations. First, because machine-learning algorithms are computational constructs that are not familiar to most physicians, these models can be seen as foreign and/or unproven entities.6 However, as the importance of utilizing “big data” increases, utilizing artificial intelligence and machine learning will inevitably be tools used going forward.6,24 Second, despite the very large number of total patients, the number of patients in each individual subset of CRTBI was low. However, with additional data, we believe we can create more sophisticated models with higher specificity in the future, providing even better data on who can be safely discharged without risk for readmission. Furthermore, we were not able to incorporate standardized metrics observed during the patient’s physical examination, details that are difficult to quantify and capture. Thus, while algorithm-based decision tools can be useful in guiding the physician’s decision, these constructs absolutely do not replace the

| TABLE 3. Confusion table statistics: testing results of an ANN on pediatric TBI patients |
|---------------------------------|-----------------|-----------------|
| Measure                  | Test Group Patients | All Patients |
| Sensitivity              | 0.9973           | 0.9954         |
| Specificity              | 0.6047           | 0.6417         |
| Precision                | 0.9819           | 0.9863         |
| NPV                      | 0.9123           | 0.8438         |
| False-positive rate      | 0.3953           | 0.3583         |
| False discovery rate     | 0.0181           | 0.0137         |
| Accuracy                 | 0.0027           | 0.0046         |
| F1 score                 | 0.9798           | 0.9823         |
| Matthews correlation coefficient | 0.7337         | 0.7272         |
| Area under ROC curve     | 0.9907           | 0.9790         |

We randomly partitioned patients into 3 groups in order to provide holdout validation on our large data set: 70% were for training the ANN; 15% were for validating the ANN; and 15% were for subsequent final testing of the ANN. Various measures of predictive ability on the test patients, as well as test patients combined with those used for training and validation, are presented as proportions.
information that can only be obtained by a trained physician. Third, each patient in our study underwent head CT scanning, an assumption in itself that our model is heavily dependent on, and the decision to order scans is not standardized across institutions and likely changes over time.

There has been extensive literature presented on the utility and safety of head CT for mild TBI in children since the time these data were collected. There have been dichotomized without providing further quantification per covariate (e.g., degree of midline shift, quanti-
tification of hemorrhage). In reality, these CT images are interpreted by a combination of emergency medicine and/or night-hawk radiologists, such that decisions would be made long before a complex research-level interpretation could be accomplished. Last, we used a single data source (PECARN) that is publicly available and has undergone rigorous quality improvement. However, we are limited by the clinical-practice standards of those years (2004–2006), including the rationale and threshold to obtain head CT scans in children. Further computational restructuring of our ANN model may also provide additional metrics for future studies that analyze the CT data directly instead of the radiologist’s interpretation.

We posit that in-hospital use of the model may actually increase the power of the algorithm as ANNs can be trained on new data, and the model has the potential to be implemented as a future online tool or packaged into the EMR system (available for download in the Supplemental Material), but this would require much of the heavy research-level classification to be performed immediately for this to be time sensitive and clinically relevant. Currently, much of the trauma registry classifications, clinical documentation, and final imaging reads are done well after clinical decisions are made and often times are only fully complete well after patient discharge or death.

Conclusions

Training an ANN model using data from PECARN, we constructed a highly sensitive tool to diagnose CRTBIs. Further iterations of this ANN may bring real-time, data-driven updates to the hands of pediatric emergency personnel in order to provide the most accurate evidence-based care and may particularly aid midlevel and/or inexperienced practitioners in small outlying or austere facilities. Immediate identification of pediatric TBI patients who are likely to require additional hospital resources allows clinical teams and hospital administrators to work synergistically to provide the best clinical care. We believe that approaches like our ANN can offer more robust and accurate predictions that can be updated prospectively in real time.

Acknowledgments

A.T.H. receives salary and tuition support through the Vanderbilt University School of Medicine Medical Scientist Training Program (ST32GM007347). This project is supported by the Vanderbilt Clinical and Translational Science Award (UL1RR024975, NIH/NCAT; to M.B.P.). M.B.P. is funded by R01 (GM120484, NIH/NIGMS) and a Vanderbilt Faculty Scholars Award.

We would also like to thank the Surgical Outcomes Center for Kids at Monroe Carell Jr. Children’s Hospital of Vanderbilt University for administrative support.

References

4. Badawy MK, Dayan PS, Tunik MG, Nadel FM, Lillis KA,
Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Hale, Stonko. Acquisition of data: Hale, Stonko, Lim. Analysis and interpretation of data: Hale, Stonko, Lim, Patel. Drafting the article: Hale. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Hale. Statistical analysis: Hale, Stonko. Study supervision: Shannon, Patel.

Supplemental Information
Online-Only Content
Supplemental material is available with the online version of the article.


Correspondence
Andrew T. Hale: Vanderbilt University School of Medicine, Nashville, TN. andrew.hale@vanderbilt.edu.