Chinese Journal of Traumatology 24 (2021) 350–355

Contents lists available at ScienceDirect

Chinese Journal of Traumatology

journal homepage: http://www.elsevier.com/locate/CJTEE

Original Article

Application of machine learning to predict the outcome of pediatric traumatic brain injury

Thara Tunthanathip*, Thakul Oearsakul

Division of Neurosurgery, Department of Surgery, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, 90110, Thailand

ARTICLE INFO

Article history:
Received 26 August 2020
Received in revised form
23 May 2021
Accepted 2 June 2021
Available online 8 June 2021

Keywords:
Pediatrics
Traumatic brain injury
Machine learning
Support vector machine
Random forest
Logistic regression

ABSTRACT

Purpose: Traumatic brain injury (TBI) generally causes mortality and disability, particularly in children. Machine learning (ML) is a computer algorithm, applied as a clinical prediction tool. The present study aims to assess the predictability of ML for the functional outcomes of pediatric TBI.

Methods: A retrospective cohort study was performed targeting children with TBI who were admitted to the trauma center of southern Thailand between January 2009 and July 2020. The patient was excluded if he/she (1) did not undergo a CT scan of the brain, (2) died within the first 24 h, (3) had unavailable complete medical records during admission, or (4) was unable to provide updated outcomes. Clinical and radiologic characteristics were collected such as vital signs, Glasgow coma scale score, and characteristics of intracranial injuries. The functional outcome was assessed using the King's Outcome Scale for Childhood Head Injury, which was thus dichotomized into favourable outcomes and unfavourable outcomes: good recovery and moderate disability were categorized as the former, whereas death, vegetative state, and severe disability were categorized as the latter. The prognostic factors were estimated using traditional binary logistic regression. By data splitting, 70% of data were used for training the ML models and the remaining 30% were used for testing the ML models. The supervised algorithms including support vector machines, neural networks, random forest, logistic regression, naive Bayes and k-nearest neighbor were performed for training of the ML models. Therefore, the ML models were tested for the predictive performances by the testing datasets.

Results: There were 828 patients in the cohort. The median age was 72 months (interquartile range 104.7 months, range 2–179 months). Road traffic accident was the most common mechanism of injury, accounting for 68.7%. At hospital discharge, favourable outcomes were achieved in 97.0% of patients, while the mortality rate was 2.2%. Glasgow coma scale score, hypotension, pupillary light reflex, and subarachnoid haemorrhage were associated with TBI outcomes following traditional binary logistic regression; hence, the 4 prognostic factors were used for building ML models and testing performance. The support vector machine model had the best performance for predicting pediatric TBI outcomes: sensitivity 0.95, specificity 0.60, positive predicted value 0.99, negative predictive value 1.0; accuracy 0.94, and area under the receiver operating characteristic curve 0.78.

Conclusion: The ML algorithms of the present study have a high sensitivity; therefore they have the potential to be screening tools for predicting functional outcomes and counselling prognosis in general practice of pediatric TBIs.

Introduction

Traumatic brain injury (TBI) in children is a public health problem.[1] From previous studies in Thailand, 37.5%–65.3% of pediatric TBIs were injured from road traffic accidents[2,3] and 22.4% of the wounded was motorcycle drivers.[2] The mortality rate of pediatric TBIs of all ranges of severity has been reported at 3.2%–5.2%, while 0.3%–0.8% of those were found as a vegetative state and

* Corresponding author.
E-mail address: tsus4@hotmail.com (T. Tunthanathip).
Peer review under responsibility of Chinese Medical Association.

https://doi.org/10.1016/j.cjtee.2021.06.003

severe disability.[4,5] More than half of severe pediatric TBIs died within one year after injuries, and only 4.6%–8.7% had a good recovery.[6]

Post-traumatic sequelae and long-term consequences of severe pediatric TBIs have been recognized to incur a productivity loss for the nation.[7] Prediction of functional outcomes among the injured children is one of the important processes to evaluate future economic burden. In general, the clinical outcome is predicted from independent variables by traditional statistical methods. Nowadays, machine learning (ML) algorithms have been used as clinical prediction tools in the literature.[7,8] Prior studies have performed ML for classifying and predicting clinical outcomes such as neuro-oncology,[9,10] TBI,[11] stroke,[12] and postoperative complications.[13]

The ML approach, particularly supervised ML, has also been used to predict clinical outcomes following TBI. Random forest has been reported as a valuable method for predicting moderate to severe TBIs based on quantitative electroencephalography measurements and clinical parameters, which has an area under the receiver operating characteristic curve (AUC) of 0.94–0.81 according to the study of Haveman et al.[14] In addition, the performance of mortality prediction in TBIs reported that the models of random forest, support vector machines, neural networks, and gradient boosting machines have an AUC of 0.81 (95% *CI* 0.78–0.84), 0.81 (95% *CI* 0.79–0.84), 0.82 (95% *CI* 0.79–0.84), and 0.83 (95% *CI* 0.81–0.86), respectively.[15] Amorim et al.[11] used various ML algorithms to predict the mortality of TBI patients and found that the naive Bayes algorithm had the best predictive value of mortality (AUC = 0.906), while random forest had an AUC of 0.880. The concordance of the results demonstrated that naive Bayes had the best performance for predicting clinical outcomes, but an AUC equal to 0.76 was not a good performance.[13] In the face of this gap, we aimed to evaluate the predictive performance of ML algorithms for clinical outcomes. Also, the present study is based on the functional outcomes of pediatric TBI.

Methods

Study design and study population

The authors retrospectively reviewed the medical records of patients with TBI admitted to the trauma center of southern Thailand, who was younger than 15 years between January 2009 and July 2020. In detail, all TBI patients had been prospectively registered in the traumatic database of our hospital. The exclusion criteria were patients who did not undergo a CT scan of the brain, died within the first 24 h, had unavailable complete medical records during admission and those who were unable to provide the updated outcomes.

Various clinical, radiological, treatment and outcome variables were collected for analysis. Glasgow coma scale (GCS) score after resuscitation was used to classify TBI severity: mild TBI (GCS score 13–15), moderate TBI (GCS score 9–12), and severe TBI (GCS score 3–8).[5] In addition, secondary systemic injuries such as hypotension and bradycardia were defined according to the age period.[16,17] For radiological findings, the characteristics of intracranial injuries, midline shift, and obliteration of basal cistern were reviewed by two neurosurgeons. Based on Vieira et al.,[18] the diffuse axonal injury was defined where patients had a GCS score of ≤8 following resuscitation and CT scan or MRI showed signs of diffuse axonal injury.

The functional outcome of the present study was assessed using King's Outcome Scale for Childhood Head Injury (KOSCHI).[19] The KOSCHI classification was collected at hospital discharge and every follow-up visit. For patients who we were unable to follow up face by face, we evaluated the patients' outcomes by telephone. Patients who could not be contacted were excluded. For the dependent variable, the KOSCHI classification was dichotomized into favourable outcomes and unfavourable outcomes for binary proposes. Good recovery and moderate disability were categorized as favourable outcomes, whereas death, vegetative state, and severe disability were categorized into the unfavourable group.

The study was performed with the approval of the human research ethics committee (REC.63-373-10-1). Because the present study was a retrospective cohort study design, informed consent was not performed. However, the patients' identification numbers were encoded before analysis.

Statistical analysis

Clinical characteristics were calculated from descriptive data. The median with interquartile range (IQR) was used for continuous variables. The prognostic variables, which were used for building the ML models, were analyzed using the traditional binary logistic regression. Initially, clinical and imaging variables were estimated by univariate analysis and the variables that had *p* values less than 0.01 were regarded as candidate variables. In multivariable analysis, these candidate variables were estimated whether they should be included as the prognostic factors using the Akaike information criterion. Thereafter the qualified prognostic factors were used for building ML models. In this part, statistical analysis was performed using the R version 3.6.2 software (R Foundation, Vienna, Austria).

ML

By randomly splitting data, 70% of the total data were used to train the ML models, while the remaining 30% were used for testing the models' performance. Data preprocessing included dichotomization of independent variables into binary classifiers. Because patients who had incomplete data were excluded, missing data management was not performed in the present study.

Supervised algorithms with 10-fold cross-validation including support vector machines, neural networks, naive Bayes, logistic regression, and k-nearest neighbor were used for training the models from the training dataset. The parameters of each algorithm were also optimized by the grid-search method. Therefore, a confusion matrix was constructed to describe the performance of ML models for which the true values of the outcome were known. To test the ML models, we built web-based applications of various algorithms using stream lit and deployed via the Heroku platform. (Heroku, California, USA).

The performance of each algorithm included sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, F1-score, and average precision. Furthermore, the receiver operating characteristic curve and AUC were created. An AUC of ≥0.7 was acceptable discrimination, ≥ 0.8 good discrimination, and ≥0.9 excellent discrimination.[20] ML was performed using Python version 3.7.7 with Scikit-learn 0.23.2. (Python Software Foundation, Washington, USA).

Results

Clinical and radiological characteristics

A total of 828 patients were included in the present cohort. Median patient age was 72 months (IQR 104.7 months, range 2–179 months). There were 567 males and 261 females. The mechanisms of TBIs included motorcycle crashes (51.0%), falls (25.7%), pedestrians injured in traffic accidents (9.1%), objects striking the head (2.9%), vehicle accidents (8.7%) and bicycle accidents (2.7%). After resuscitation, 87.4% of the cases were of mild TBI (GCS 13–15), while moderate and severe TBI were found in 6.0% and 6.5% of patients, respectively. A summary of the demographic data is presented in Table 1. Long bone fractures and maxillofacial injuries were the most common associated injuries, whereas coagulopathy was found in 0.4% in the present cohort. Additionally, the post-traumatic seizure was observed in 5% of cases (Table 1).

From the radiological findings, intracranial injuries were observed as follows: skull fracture (14.6%), basilar skull fracture (4.8%), epidural hematoma (8.3%), subdural hematoma (11.7%), contusion (7.6%), brainstem contusion (0.5%), subarachnoid haemorrhage (5.8%), and primary intraventricular haemorrhage (1.6%). Moreover, a diffuse axonal injury was diagnosed in 2.9% of cases. Midline shift of 5 mm or more was observed in 1.4%, where the obliterated basal cistern was in 2.7% of total cases. The radiological findings and outcomes are revealed in Table 2.

Table 1
Baseline and clinical characteristics of the 828 pediatric patients with traumatic brain injury.

Variables	*n* (%)
Age (years)	
<2	119 (14.4)
≥2	709 (85.6)
Age (years)	
<5	335 (40.5)
≥5	493 (59.5)
Median age (IQR) (months)	72 (104.7)
Gender	
Male	567 (68.5)
Female	261 (31.5)
Mechanism of injury	
Motorcycle crash	422 (51.0)
Fall	213 (25.7)
Pedestrians injured	75 (9.1)
Vehicle crash	72 (8.7)
Object striking the head	24 (2.9)
Bicycle accident	22 (2.7)
Mean injury severity score (SD)	21.5 (7.9)
Association of injuries	
Long bone fracture	165 (19.9)
Maxillofacial injury	82 (9.9)
Orbital injury	44 (5.3)
Lung contusion	39 (4.7)
Liver injury	30 (3.6)
Spine fracture	30 (3.6)
Kidney, ureter, and bladder injury	21 (2.5)
Splenic injury	18 (2.2)
Pelvic fracture	17 (2.0)
Bowel injury	8 (0.9)
Retroperitoneal injury	7 (0.8)
Comorbidities	
Scalp hematoma/laceration	487 (58.8)
Loss of consciousness	315 (38.0)
Amnesia	225 (27.2)
Vomiting	210 (25.4)
Seizure	41 (5.0)
Hypotension episode	34 (4.1)
Bleeding per ear/nose	22 (2.7)
Hemiparesis	21 (2.5)
Bradycardia	6 (0.7)
Coagulopathy	3 (0.4)
Glasgow coma scale score	
13-15	724 (87.4)
9-12	50 (6.0)
3-8	54 (6.5)
Pupillary light reflex	
Fixed pupils both eyes	24 (2.9)
Fixed pupil one eyes	16 (1.9)
React pupils both eyes	788 (95.2)

Table 2
Imaging characteristics, treatment and outcomes of the 828 pediatric patients with traumatic brain injury.

Variables	*n* (%)
Skull fracture	121 (14.6)
Linear skull fracture	70 (8.5)
Simple depressed skull fracture	20 (2.4)
Compound depressed skull fracture	19 (2.3)
Diastatic skull fracture	12 (1.4)
Intracranial injuries	
Basilar skull fracture	40 (4.8)
Epidural hematoma	69 (8.3)
Subdural hematoma	97 (11.7)
Contusion	63 (7.6)
Brainstem contusion	4 (0.5)
Subarachnoid haemorrhage	48 (5.8)
Intraventricular haemorrhage	13 (1.6)
Diffuse axonal injury	24 (2.9)
Basal cistern	
Obliteration	22 (2.7)
Patency	806 (97.3)
Mean midline shift (mm) (SD)	0.02 (0.1)
Midline shift (mm)	
<5	816 (98.6)
≥5	12 (1.4)
Surgical treatment	20 (2.4)
Craniotomy with clot removal	27 (3.3)
Decompressive craniectomy with clot removal	11 (1.3)
Intracranial monitoring insertion	2 (0.2)
Hospital-discharge KOSCHI outcome	
Death	18 (2.2)
Vegetative stage	2 (0.2)
Severe disability	6 (0.7)
Moderate disability	13 (1.6)
Good recovery	789 (95.3)
Six-month KOSCHI outcome	
Death	18 (2.2)
Vegetative stage	2 (0.2)
Severe disability	3 (0.4)
Moderate disability	15 (1.8)
Good recovery	790 (95.4)

Abbreviation: KOSCHI: King's Outcome Scale for Childhood Head Injury.

KOSCHI categories at hospital-discharge included death (2.2%), vegetative state (0.2%), severe disability (0.7%), moderate disability (1.6%), and good recovery (95.3%). At 6-month follow-up, the rates of death, vegetative state, severe disability, moderate disability, and good recovery were 2.2%, 0.2%, 0.4%, 1.8%, 95.4%, respectively. Therefore, the dichotomized KOSCHI categories comprised of 97.0% favourable and 3.0% unfavourable outcomes after hospital discharge (Table 2).

When clinical and imaging characteristics were estimated for the prognostic variables using the traditional binary logistic regression, GCS score, hypotension, pupillary light reflex, and subarachnoid haemorrhage were found to be prognostic variables in the final model, as shown in Table 3. Hence, these prognostic variables were used for supervised learning.

ML

After splitting the data, 580 patients were considered suitable for the training dataset and were used to construct the ML models. The parameters of the ML models were optimized using the grid search method. In detail, the support vector machines model was optimized with a linear kernel, scale grama and the regularization parameter (C parameter) of 1.0, while the optimized logistic regression model was C parameter of 0.10 with “l2” penalization. Optimization of the neural networks model comprised of two hidden layers, “identity” activation, “lbfgs” “solver”, “invasling” learning and an alpha of 0.0001. For k-nearest neighbor, the model

Table 3
Traditional binary logistic regression for favourable outcome.

Variables	Univariate analysis		Multivariable analysis	
	Odds ratio (95% *CI*)	*p* value	Odds ratio (95% *CI*)	*p* value
Age group (years)				
<2	Ref			
≥2	1.12 (0.21–5.95)	0.89		
Age group (years)				
<5	Ref			
≥5	1.86 (0.73–4.72)	0.19		
Road traffic injury				
No	Ref			
Yes	0.96 (0.41–2.26)	0.93		
Comorbidities				
Loss of consciousness[a]	0.65 (0.29–1.45)	0.30		
Vomiting[a]	4.02 (0.94–17.19)	0.06		
Hemiparesis[a]	0.14 (003–0.36)	<0.001		
Scalp injury[a]	0.79 (0.34–1.82)	0.59		
Bleeding per nose/ear[a]	0.17 (0.04–0.64)	0.009		
Seizure[a]	0.58 (0.13–2.58)	0.48		
Hypotension	0.01 (0.004–0.02)	<0.001	0.05 (0.009–0.37)	0.003
Bradycardia	0.02 (0.005–0.14)	<0.001		
Injury severity score	0.14 (0.14–2.36)	0.56		
Glasgow coma scale score	2.04 (1.68–2.49)	<0.001	1.69 (1.30–2.19)	<0.001
Pupillary light reflex				
Fixed both eyes	Ref		Ref	
React one eye	5.0 (1.23–20.30)	0.02	0.23 (0.01–4.05)	0.32
React both eyes	21.7 (6.86–68.7)	<0.001	9.46 (1.28–69.92)	0.02
Intracranial injuries				
Skull fracture[a]	0.23 (0.10–0.53)	0.001		
Basilar skull fracture[a]	0.14 (0.05–0.37)	<0.001		
Epidural hematoma[a]	0.21 (0.08–0.53)	0.001		
Subdural hematoma[a]	0.10 (0.04–0.23)	<0.001		
Contusion[a]	0.15 (0.06–0.37)	<0.001		
Brainstem contusion[a]	0.02 (0.004–0.21)	0.001		
Intraventricular haemorrhage[a]	0.04 (0.01–0.13)	<0.001		
Subarachnoid haemorrhage[a]	0.04 (0.01–0.09)	<0.001	0.05 (0.01–0.29)	0.001
Diffuse axonal injury[a]	0.02 (0.01–0.06)	<0.001		
Midline shift (mm)				
<0.5	Ref			
≥0.5	0.01 (0.005–0.05)	<0.001		
Basal cistern				
Obliteration	Ref			
Patent	5.17 (0.13–7.17)	0.98		

[a] Data only show "yes group" while reference groups (no group) are hidden.

was also optimized with 5 neighbors and uniform weight function. Additionally, the optimized random forest model comprised 5 maximum depths of the tree with 1000 trees in the forest.

Almost all the ML models had excellent performances after training. Support vector machines, neural networks, and logistic regression were the best models with high sensitivity and PPV. Thus, details of performances from the training data are shown in Table 4. Moreover, a confusion matrix of the ML models from the training dataset is shown via https://pedtbi-train.herokuapp.com.

For testing the ML models, AUC of all the algorithms decreased. However, the support vector machines model and the neural networks model still had AUC in acceptable levels at 0.78 and 0.72, while the AUC of other ML models dropped to less than 0.7, as shown in Fig. 1. Moreover, support vector machines had a sensitivity of 0.95, specificity 0.60, PPV 0.99, NPV 0.21, and accuracy 0.94, whereas the neural networks model had the sensitivity of 0.84, specificity 0.60, PPV 0.99, NPV 0.07 and accuracy 0.83. In addition, various ML models were built in the web-based applications for testing further unseen data (https://pedtbi-home.herokuapp.com).

Discussion

A favourable outcome was reported to be 84.2%–99.2% in pediatric TBI patients in the literature.[19,21] Hawley et al.[19] studied the

Table 4
Performances of each algorithm to predict outcome using training data ($n = 580$).

Algorithm	Sensitivity	Specificity	PPV	NPV	Accuracy score[a] (SD)	F1-score[a] (SD)	Mean AUC[a] SD
Support vector machines	0.99 (0.98–1.00)	0.71 (0.37–1.00)	0.99 (0.98–1.00)	0.71 (0.37–1.00)	0.90 (0.254)	0.50 (0.33)	0.98 (0.04)
Neural networks	0.99 (0.98–1.00)	0.71 (0.37–1.00)	0.99 (0.98–1.00)	0.83 (0.53–1.00)	0.93 (0.14)	0.61 (0.30)	0.99 (0.02)
Logistic regression	0.99 (0.98–1.00)	0.57 (0.20–0.93)	0.98 (0.97–1.00)	0.66 (0.28–1.00)	0.93 (0.14)	0.61 (0.30)	0.98 (0.03)
k-nearest neighbor	0.98 (0.97–1.00)	0.57 (0.20–0.93)	0.98 (0.97–1.00)	0.57 (0.20–0.93)	0.93 (0.15)	0.47 (0.38)	0.89 (0.16)
Naive Bayes	0.95 (0.93–0.98)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.95 (0.93–0.98)	0.89 (0.27)	0.75 (0.28)	0.95 (0.11)
Random forest	0.99 (0.98–1.00)	0.85 (0.59–1.00)	0.99 (0.98–1.00)	0.85 (0.59–1.00)	0.91 (0.21)	0.55 (0.36)	0.97 (0.08)

Abbreviation: AUC: area under the curve, NPV: negative predictive value, PPV: positive predictive value.
[a] 10-fold cross-validation.

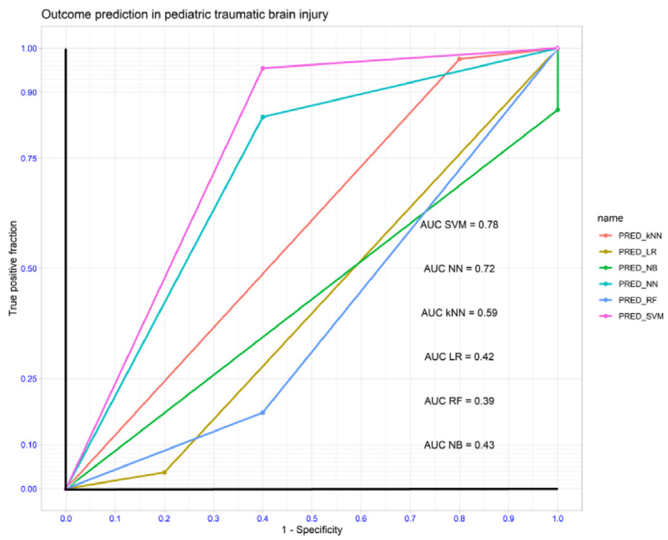

Fig. 1. Receiver operating characteristic curve and AUC of various algorithms to predict the outcomes following pediatric traumatic brain injury from the testing dataset. AUC: area under the curve, kNN: k-nearest neighbor, LR: Logistic regression, NB: naive Bayes, NN; Neural networks, RF: random forest, SVM: support vector machine.

outcome of 526 children with TBI and found that the rates of good recovery and moderate disability were 51.3% and 47.9%, respectively; while another prior study[21] reported a good recovery of 94% and moderate disability of 1.7% among 948 pediatric TBI patients. In the present cohort, we found an overall favourable outcome of 96.9%. The severity of TBIs, which is categorized by the GCS score, is well known as a TBI predictor, including TBIs in children. Bredy et al.[21] studied 315 children with TBIs and reported that severe TBIs were associated with mortality or disability compared with moderate and mild TBIs (odds ratio (OR) 2.55, 95% *CI* 1.96–4.52). Additionally, a previous study used the traditional binary logistic regression for identifying prognostic factors significantly associated with unfavourable outcomes as follows: GCS score 3–8, hemiparesis, pupillary light reflex, hypotension, basilar skull fracture, epidural hematoma, subdural hematoma, and motorcycle crash mechanism.[5]

When traditional binary logistic regression was performed with a backward elimination procedure, the prognostic factors were GCS score, hypotension, pupillary light reflex, and subarachnoid haemorrhage. These clinical characteristics have been reported in previous studies as prognostic factors. Hypotension meaningfully increased the mortality and poor outcome in pediatric TBIs, which could be explained by the systematic hypoperfusion directly effecting the global cerebral ischemia.[22,23] Low GCS scores and poor pupillary light reflex have been associated with poor outcomes and low survival rate[24,25] because these factors indicate the severity of TBI and brainstem dysfunction.[25] In addition, subarachnoid haemorrhage is one of the prognostic factors in the present study. However, this result differed from a prior study by Hochstadter et al.[26] who studied 171 pediatric severe TBIs and reported that subarachnoid haemorrhage was not significantly associated with mortality in multivariable analysis. Nevertheless, there is a lack of studies that mention the relation of imaging findings with prognosis in pediatric TBIs. This topic needs further studies to explore the controversial results.

For ML-based prediction, prognostic factors were used for developing the model with various algorithms. We observed the overfitting performances during the model training, therefore the application of the model with the testing datasets resolved the problem. The support vector machines algorithm was achieved and validated with the test dataset. From a literature review, various ML algorithms have been performed as clinical prediction tools. Gravesteijn et al.[15] studied the prognostication of TBI patients using the algorithms of random forest, support vector machines, neural networks, and gradient boosting machines and the AUCs of all of them were at a good level. Moreover, Amorim et al.[11] built the ML-based model for mortality of TBIs in Brazil and found that AUCs of naive Bayes, Bayesian generalized linear model, random forest, and penalized discriminant analysis were of high values of 0.906, 0.881, 0.880, and 0.880, respectively.

In the present study, the support vector machines and neural networks algorithms had acceptable performances in predicting the functional outcome following pediatric TBIs, with a particularly high sensitivity. For implication, these ML algorithms may be involved in general practice as the screening tools for counselling parents about prognosis. Stromberg et al.[27] performed the ML-based prediction for employment following moderate to severe TBIs. AUCs of the decision tree model was 0.77 (95% *CI* 0.74–0.80) and 0.72 (95% *CI* 0.68–0.76) for 1-year and 5-year outcomes, respectively. Moreover, the predictability of surgical site infection in neurological operation using various ML algorithms and the naive Bayes algorithm had the highest AUC of 0.76, sensitivity 0.63, specificity 0.87, PPV 0.29, NPV 0.96, and accuracy 0.86.[13]

ML algorithms had varying performances for predicting clinical outcomes. The characteristics of the study population possibly explain this. The type of predictors such as continuous variables and categorized variables (GCS score and severity of TBIs) affects the predictability of ML models.[13] In addition, a number of predictors were associated with high performance.[28,29] More significant predictors and more high-dimensional data increased the performance of ML.[30,31] Moreover, because several ML algorithms are flexible from hyperparameters adjustment, the performance of each algorithm was considerably dependent on its hyperparameters.[15]

Limitations of the present study are acknowledged that study design was a retrospective cohort study which might have led to bias from confounding factors. The multivariable analysis helps to control this bias, but the number of parameters for building the ML model is limited after adjustment with multivariable analysis. Additionally, the imbalance among the severity of TBI was probably why performance reached an excellent level and high sensitivity. Because a limited number of moderate and severe TBI was observed in the present cohort, a multicenter-study may be able to resolve this problem for improving the predictive performance. Also, the predictability of the ML models of the present study needs to be confirmed by external validation with unseen data in the future. Alternatively, a nomogram is one of the clinical prediction tools that have been performed for predicting the outcome of various diseases.[32–34] Therefore, researches should be explored in the future with a comparison of predictability between ML-based models and nomogram.

In summary, we found that ML algorithms have a high sensitivity and PPV; therefore, ML models had the potential performance as a screening tool for predicting the functional outcome in pediatric TBIs for real-world implication.

Transparency declaration

This research was a part of a retrospective cohort study that will be published elsewhere, whereas this study focused on application of machine learning to predict the outcome of pediatric traumatic brain injury.

Funding

Nil.

Ethical statement

The study has been approved by the human research ethics committee (REC.63-373-10-1).

Declaration of competing interest

All the authors declare no competing interest.

References

1. Phuenpathom N, Tiensuwan M, Ratanalert S, et al. The changing pattern of head injury in Thailand. *J Clin Neurosci*. 2000;7:223–225.
2. Tunthanathip T, Duangsuwan J, Wattanakitrungroj N, et al. Clinical nomogram predicting intracranial injury in pediatric traumatic brain injury. *J Pediatr Neurosci*. 2020;15:409–415. https://doi.org/10.4103/jpn.JPN_11_20.
3. Taweesomboonyat C, Kaewborisutsakul A, Tunthanathip T, et al. Necessity of in-hospital neurological observation for mild traumatic brain injury patients with negative computed tomography brain scans. *JHSMR*. 2020;38:267–274. https://doi.org/10.31584/jhsmr.2020746.
4. Chaitanya K, Addanki A, Karambelkar R, et al. Traumatic brain injury in Indian children. *Childs Nerv Syst*. 2018;34:1119–1123. https://doi.org/10.1007/s00381-018-3784-z.
5. Tunthanathip T, Phuenpathom N. Impact of road traffic injury to pediatric traumatic brain injury in Southern Thailand. *J Neurosci Rural Pract*. 2017;8:601–608. https://doi.org/10.4103/jnrp.jnrp_381_17.
6. Fulkerson DH, White IK, Rees JM, et al. Analysis of long-term (median 10.5 years) outcomes in children presenting with traumatic brain injury and an initial Glasgow Coma Scale score of 3 or 4. *J Neurosurg Pediatr*. 2015;16:410–419.
7. Kavosi Z, Jafari A, Hatam N, et al. The economic burden of traumatic brain injury due to fatal traffic accidents in Shahid Rajaei trauma hospital, Shiraz, Iran. *Arch Trauma Res*. 2015;4, e22594. https://doi.org/10.5812/atr.22594.
8. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. *Nat Methods*. 2018;15:233–234. https://doi.org/10.1038/nmeth.4642.
9. Valdebenito J, Medina F. Machine learning approaches to study glioblastoma: a review of the last decade of applications. *Cancer Rep (Hoboken)*. 2019;2:e1226. https://doi.org/10.1002/cnr2.1226.
10. Papp L, Pötsch N, Grahovac M, et al. Glioma survival prediction with combined analysis of in vivo 11C-MET PET features, ex vivo features, and patient features by supervised machine learning. *J Nucl Med*. 2018;59:892–899. https://doi.org/10.2967/jnumed.117.202267.
11. Amorim RL, Oliveira LM, Malbouisson LM, et al. Prediction of early tbi mortality using a machine learning approach in a LMIC population. *Front Neurol. Front Neurol.* 2019;10:1366. https://doi.org/10.3389/fneur.2019.01366.
12. Heo J, Yoon JG, Park H, et al. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke*. 2019;50:1263–1265. https://doi.org/10.1161/STROKEAHA.118.024293.
13. Tunthanathip T, Sae-Heng S, Oearsakul T, et al. Machine learning applications for the prediction of surgical site infection in neurological operations. *Neurosurg Focus*. 2019;47:E7. https://doi.org/10.3171/2019.5.FOCUS19241.
14. Haveman ME, Van Putten M, Hom HW, et al. Predicting outcome in patients with moderate to severe traumatic brain injury using electroencephalography. *Crit Care*. 2019;23:401. https://doi.org/10.1186/s13054-019-2656-6.
15. Gravesteijn BY, Nieboer D, Ercole A, et al. Machine learning algorithms performed no better than regression models for prognostication in traumatic brain injury. *J Clin Epidemiol*. 2020;122:95–107. https://doi.org/10.1016/j.jclinepi.2020.03.005.
16. Wetzel RC. Anesthesia for pediatric trauma. In: Stane JK, Grande CM, eds. *Trauma Anesthesia*. Baltimore: William & Wilkins; 1991:312.
17. O'Rourke PP, Crone RK. The respiratory system. In: *Gregory GA. Pediatric Anesthesia*. second ed. New York: Churchill Livingstone; 1989.
18. Vieira RC, Paiva WS, de Oliveira DV, et al. Diffuse axonal injury: epidemiology, outcome and associated risk factors. *Front Neurol*. 2016;7:178. https://doi.org/10.3389/fneur.2016.00178.
19. Hawley CA, Ward AB, Magnay AR, et al. Outcomes following childhood head injury: a population study. *J Neurol Neurosurg Psychiatry*. 2004;75:737–742. https://doi.org/10.1136/jnnp.2003.020651.
20. Roelen CA, Bültmann U, van Rhenen W, et al. External validation of two prediction models identifying employees at risk of high sickness absence: cohort study with 1-year follow-up. *BMC Publ Health*. 2013;13:105. https://doi.org/10.1186/1471-2458-13-105.
21. Bedry T, Tadele H. Pattern and outcome of pediatric traumatic brain injury at hawassa university comprehensive specialized hospital, southern Ethiopia: observational cross-sectional study. *Emerg Med Int*. 2020;2020:1965231. https://doi.org/10.1155/2020/1965231.
22. Pigula FA, Wald SL, Shackford SR, et al. The effect of hypotension and hypoxia on children with severe head injuries. *J Pediatr Surg*. 1993;28:310–314. https://doi.org/10.1016/0022-3468(93)90223-8.
23. Suttipongkaset P, Chaikittisilpa N, Vavilala MS, et al. Blood pressure thresholds and mortality in pediatric traumatic brain injury. *Pediatrics*. 2018;142, e20180594. https://doi.org/10.1542/peds.2018-0594.
24. Tunthanathip T, Udomwitthayaphiban S. Development and validation of a nomogram for predicting the mortality after penetrating traumatic brain injury. *Bull Emerg Trauma*. 2019;7:347–354. https://doi.org/10.29252/beat-070402.
25. Vathanalaoha K, Oearsakul T, Tunthanathip T. Predictive factors of survival and 6-month favorable outcome of very severe head trauma patients; a historical cohort study. *Emerge*. 2017;5:e24.
26. Hochstadter E, Stewart TC, Alharfi IM, et al. Subarachnoid hemorrhage prevalence and its association with short-term outcome in pediatric severe traumatic brain injury. *Neurocritical Care*. 2014;21:505–513. https://doi.org/10.1007/s12028-014-9986-7.
27. Stromberg KA, Agyemang AA, Graham KM, et al. Using decision tree methodology to predict employment after moderate to severe traumatic brain injury. *J Head Trauma Rehabil*. 2019;34:E64–E74. https://doi.org/10.1097/HTR.0000000000000438.
28. Delahanty RJ, Kaufman D, Jones SS. Development and evaluation of an automated machine learning algorithm for in-hospital mortality risk adjustment among critical care patients. *Crit Care Med*. 2018;46:e481–e488. https://doi.org/10.1097/CCM.0000000000003011.
29. Desautels T, Das R, Calvert J, et al. Prediction of early unplanned intensive care unit readmission in a UK tertiary care hospital: a cross-sectional machine learning approach. *BMJ Open*. 2017;7, e017199. https://doi.org/10.1136/bmjopen-2017-017199.
30. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med*. 2018;1:18. https://doi.org/10.1038/s41746-018-0029-1.
31. Beam AL, Kohane IS. Big data and machine learning in health care. *J Am Med Assoc*. 2018;319:1317–1318. https://doi.org/10.1001/jama.2017.18391.
32. Parmontree P, Tunthanathip T, Doungngern T, et al. Predictive risk factors for early seizures in traumatic brain injury. *J Neurosci Rural Pract*. 2019;10:582–587. https://doi.org/10.1055/s-0039-1700791.
33. Tunthanathip T, Ratanalert S, Sae-Heng S, et al. Prognostic factors and nomogram predicting survival in diffuse astrocytoma. *J Neurosci Rural Pract*. 2020;11:135–143. https://doi.org/10.1055/s-0039-3403446.
34. Tunthanathip T, Oearsakul T. Machine learning approaches for prognostication of newly diagnosed glioblastoma. *Int J Nutr Pharmacol Neurol Dis*. 2021;11:57–63. https://doi.org/10.4103/ijnpnd.ijnpnd_93_20.