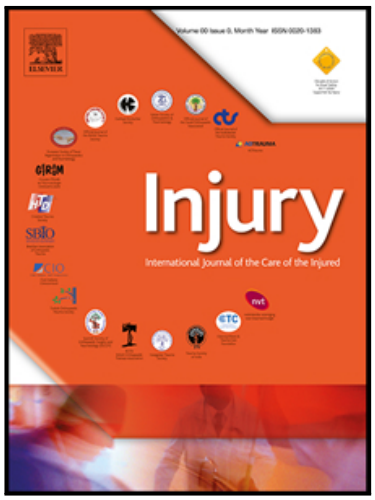

Predictive model for early functional outcomes following acute care after traumatic brain injuries: A machine learning-based development and validation study

Meng Zhang , Moning Guo , Zihao Wang , Haimin Liu , Xue Bai , Shengnan Cui , Xiaopeng Guo , Lu Gao , Lingling Gao , Aimin Liao , Bing Xing , Yi Wang

PII: S0020-1383(23)00006-2
DOI: https://doi.org/10.1016/j.injury.2023.01.004
Reference: JINJ 10566

To appear in: *Injury*

Accepted date: 2 January 2023

Please cite this article as: Meng Zhang , Moning Guo , Zihao Wang , Haimin Liu , Xue Bai , Shengnan Cui , Xiaopeng Guo , Lu Gao , Lingling Gao , Aimin Liao , Bing Xing , Yi Wang , Predictive model for early functional outcomes following acute care after traumatic brain injuries: A machine learning-based development and validation study, *Injury* (2023), doi: https://doi.org/10.1016/j.injury.2023.01.004

Predictive model for early functional outcomes following acute care after traumatic brain injuries: A machine learning-based development and validation study

Meng Zhang[a,b], Moning Guo[c,d], Zihao Wang[e], Haimin Liu[a,b], Xue Bai[a,b], Shengnan Cui[a,b], Xiaopeng Guo[f], Lu Gao[e], Lingling Gao[g], Aimin Liao[a,b], Bing Xing[f*], and Yi Wang[a,b*]

[a] Department of Medical Records, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100730, China

[b] National Center for Quality Control of Medical Records, Beijing 100730, China

[c] Beijing Municipal Health Big Data and Policy Research Center, Beijing 100034, China

[d] Beijing Institute of Hospital Management, Beijing 100034, China

[e] Department of Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100730, China

[f] Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100730, China

[g] Peking University Clinical Research Institute, Beijing 100191, China

Authors' email addresses and ORCIDs:

Meng Zhang: Email: zhangmeng97@pumch.cn, ORCID: 0000-0003-1112-9822

Moning Guo: Email: Guomoning@wjw.beijing.gov.cn

Zihao Wang: Email: wangzihaopumc@163.com, ORCID: 0000-0002-3961-0100

Haimin Liu: Email: liuhm@pumch.cn

Xue Bai: Email: 0601snow@sina.com, ORCID: 0000-0002-1483-6832

Shengnan Cui: Email: cuishengnan_86@126.com

Xiaopeng Guo: Email: guoxiaopeng_pumch@163.com, ORCID: 0000-0002-4183-5560

Lu Gao: Email: gaoluemail@yeah.net, ORCID: 0000-0003-1632-5101

Lingling Gao: Email: gaolinger@foxmail.com

Aimin Liao: Email: liaoaiminxh@163.com

Bing Xing: Email: xingbingemail@aliyun.com, ORCID: 0000-0002-3864-5168

Yi Wang: Email: wypumch@163.com, ORCID: 0000-0003-3484-3738

***Corresponding author**

Bing Xing, Address: No. 1 Shuaifuyuan, Dongcheng District, Beijing 100730, China; Email: xingbingemail@aliyun.com; Tel.: +86-10-69151898; Fax: +86-10-65124875

Yi Wang, Address: No. 1 Shuaifuyuan, Dongcheng District, Beijing 100730, China; Email: wypumch@163.com; Tel.: +86-10-69151883; Fax: +86-10-65124875

Highlights

- A random forest model for early functional outcomes following acute care after traumatic brain injury (TBI) was developed and validated.
- The BI score at admission, age, use of nonsurgical treatment, neurosurgery status, and mCCI score were the top 5 prognostic predictors.
- The predictive model is generalisable by using hospital discharge abstract data.
- The model can inform decision-making regarding TBI patient management and facilitate health care quality assessment and resource allocation.

Abstract

Introduction: Few studies on early functional outcomes following acute care after traumatic brain injury (TBI) are available. The aim of this study was to develop and validate a predictive model for functional outcomes at discharge for TBI patients using machine learning methods.

Patients and methods: In this retrospective study, data from 5281 TBI patients admitted for acute care who were identified in the Beijing hospital discharge abstract database were analysed. Data from 4181 patients in 52 tertiary hospitals were used for model derivation and internal validation. Data from 1100 patients in 21 secondary hospitals were used for external validation. A poor outcome was defined as a Barthel Index (BI) score ≤ 60 at discharge. Logistic regression, XGBoost, random forest, decision tree, and back propagation neural network models were used to fit classification models. Performance was evaluated by the area under the receiver operating characteristic curve (AUC), the area under the precision-recall curve (AP), calibration plots, sensitivity/recall, specificity, positive predictive value (PPV)/precision, negative predictive value (NPV) and F1-score.

Results: Compared to the other models, the random forest model demonstrated superior performance in internal validation (AUC of 0.856, AP of 0.786, and F1-score of 0.724) and external validation (AUC of 0.779, AP of 0.630, and F1-score of 0.604). The sensitivity/recall, specificity, PPV/precision, and NPV of the model were 71.8%, 69.2%, 52.2%, and 84.0%, respectively, in external validation. The BI score at admission, age, use of nonsurgical treatment, neurosurgery status, and modified Charlson Comorbidity Index were identified as the top 5 predictors for functional outcome at discharge.

Conclusions: We established a random forest model that performed well in predicting early functional outcomes following acute care after TBI. The model has utility for informing

decision-making regarding patient management and discharge planning and for facilitating health care quality assessment and resource allocation for TBI treatment.

Keywords: traumatic brain injury; early functional outcome; machine learning; predictive model; hospital discharge abstract data

Introduction

Traumatic brain injury (TBI) causes temporary or permanent cognitive, behavioural, emotional, and physical impairments, which affect a patient's activities, participation in society and quality of life [1]. TBI leads to increasingly substantial burdens for individuals and society, making it a global public health priority [2,3]. It is estimated that the annual incidence of TBI is 369 per 100 000 people and that 27.08 million new patients suffer from TBI annually worldwide [3]. TBI is severely disabling, resulting in permanent disability in approximately 40% of TBI survivors [4].

Early prognosis after TBI is important for clinical decision-making, conducting research, assessing the quality of health care, and allocating medical resources, and it would be helpful in preparing patients and their relatives for expected outcomes [5,6]. The functional status of patients at discharge from acute hospital care is associated with subsequent rehabilitation arrangements and long-term physical, psychological, and employment outcomes [7-10]. Functional outcomes after TBI are of concern to multiple stakeholders, including patients and their relatives, health care providers, and policy-makers. A predictive model for predicting early functional outcomes for TBI patients is useful for informing decision-making regarding patient management and discharge planning and facilitating health care quality assessments and resource allocation for TBI treatment. Few studies are available on functional outcomes following acute care after TBI. De Guise et al. [6] developed a predictive model for cognitive functional outcomes of patients with TBI in a trauma centre. Wada et al. [11] established an ICD-10-based disability predictive index for patients admitted to hospitals with trauma. However, there have been no studies on predicting physical function following acute care after TBI.

Hospital administrative data could serve as a cost-effective data source with good

availability for health care research and quality improvement [12-16]. This study aimed to develop and validate a predictive model for early functional outcomes in TBI survivors by using the Beijing municipal hospital discharge abstract database (DAD).

This study was approved by the Institutional Review Board. Patient consent was waived for this retrospective study because it used unidentified patient data.

Patients and methods

Eligibility criteria

TBI patients were defined as those with a principal diagnosis of intracranial injury (World Health Organization (WHO) International Classification of Diseases (ICD)-10 codes: S06.0-S06.9). TBI patients who had a Barthel index (BI) score ≤ 60 at admission were included in this study. Patients who were under 14 years old, admitted for rehabilitation, readmitted for previous injuries, or died during hospitalization were excluded.

Data source

The Beijing hospital DAD is an administrative database that routinely collects hospital discharge abstract data from 107 (74.3%) secondary and 87 (75.0%) tertiary medical institutions in Beijing. Information on eligible TBI patients who were admitted to secondary and tertiary medical institutions from 1 January to 31 December 2017 was accessed from the DAD. Data from 4181 TBI patients from 52 tertiary hospitals and 1100 patients from 21 secondary hospitals were analysed in this study.

The following data were extracted from the database: demographic information; ICD-10 codes for primary diagnoses; the first 10 secondary diagnoses of TBI; ICD-9 Clinical Modification Volume 3 (ICD-9-CM-3) codes for surgeries and procedures; functional status at admission and at discharge, as assessed by the Barthel index (BI) [17]; duration of loss of consciousness (LOC) after the injury; ICU admission; mechanical ventilation use; duration of ventilation; and length of hospital stay (LOS). Individual scores for 10 activities, including feeding, bathing, grooming, dressing, bowel control, bladder control, toileting, transferring from a bed to a chair, walking and stair climbing, were evaluated and documented by a trained nurse on the first day of admission and on the day of discharge. The BI score was calculated by summing the 10 individual scores. Duration of LOC after the injury refers to the total time of a state in which the patient lacks normal awareness of self and the surrounding environment after the injury, which was calculated by summing the duration of LOC prior to hospital and that during hospitalization.

A total of 4716 patients were missing data on the duration of ventilation. However, all the patients were missing data on the duration of ventilation because they did not use mechanical ventilation during hospitalization. Thus, the missing values were recoded to 0 in the preprocessing of the data.

Outcome measures

A poor outcome in this study was defined as a BI score ≤ 60 at discharge, which was used to describe severe physical disability in previous studies [11,18,19]. The BI assessment was performed for all patients by trained nurses at discharge.

Candidate predictor variables

A number of candidate predictor variables were considered in this study, including age, sex, comorbidities, injury nature, injury severity, multiple injuries, BI score at admission, duration of LOC, type of treatment, ICU admission, mechanical ventilation use, duration of ventilation, and LOS.

A modified Charlson Comorbidity Index (mCCI) was used to measure comorbidities by weighting up to 10 secondary diagnoses in the study population. The mCCI score was generated by summing the weighted values for each comorbid condition as described by Bouamra et al. [20]. The mCCI score was input as a continuous variable for model derivation.

The nature of the TBI was divided into 9 groups: concussion (ICD-10 code: S06.0), traumatic cerebral oedema (ICD-10 code: S06.1), diffuse brain injury (ICD-10 code: S06.2), focal brain injury (ICD-10 code: S06.3), epidural haemorrhage (ICD-10 code: S06.4), traumatic subdural haemorrhage (ICD-10 code: S06.5), traumatic subarachnoid haemorrhage (ICD-10 code: S06.6), intracranial injury with prolonged coma (ICD-10 code: S06.7), other intracranial injuries (ICD-10 code: S06.8), and unspecified intracranial injury (ICD-10 code: S06.9).

The International Classification of Diseases-based Injury Severity Score (ICISS) was used to estimate the severity of injuries. Survival risk ratios (SRRs) were assigned to each ICD-10 code at the three- or four-digit level by using the reference dataset of previously reported diagnosis-specific SRRs [21]. The ICISS was generated as the SRR of the single worst injury. The ICISS was categorized into two groups by a cut-off value of 0.85 [22].

Multiple injuries were defined as the concurrence of TBI and injuries to the thorax, abdomen/lower back/pelvis, upper extremities, lower extremities, or spine/spinal cord.

The type of treatment was divided into 4 groups: neurosurgery (ICD-9-CM-3 codes: 01-05), orthopaedic surgery (ICD-9-CM-3 codes: 76-84), other surgery (ICD-9-CM-3 codes excluding 01-05 and 76-84) and nonsurgical treatment (without ICD-9-CM-3 codes).

The other candidate predictors, including age, sex, BI score at admission, duration of LOC, ICU admission, mechanical ventilation use, duration of ventilation and LOS, were taken directly from the DAD.

Model derivation and validation

The cohort from the tertiary hospitals was randomly split by a ratio of 8:2 into derivation (79.9%, 3342) and internal validation (21.1%, 839) datasets. The cohort from the secondary hospitals was used for external validation. The derivation dataset was used to determine the optimized parameters of the predictive model. The internal and external validation datasets were used to evaluate the performance of the trained models. One-hot encoding was used to encode categorical variables. Age, mCCI score, BI score at admission, duration of LOC, duration of ventilation, and LOS were input as continuous variables.

Logistic regression, XGBoost, random forest, decision tree, and back propagation (BP) neural network classification models were used. As information on LOS is available only at discharge, we input it for the derivation of the model to assess its effect on the outcome but excluded it in the final model. For each model, a grid search method was used to adjust the parameters and improve the generalization performance of the models. The optimal parameters for each model are shown in Supplemental Digital Content Table S1. In the random forest model, Gini importance (mean decrease in impurity) was used to calculate each feature importance.

Table 2. Performance metrics of the candidate models in the validation cohorts.

		Internal Validation Cohort (n = 839)	External Validation Cohort (n = 1100)	Internal & External Validation Cohort (n = 1939)
XGBoost	AUC	0.855 (0.828-0.882)	0.770 (0.741-0.800)	0.808 (0.787-0.828)
	AP	0.795 (0.750-0.840)	0.618 (0.567-0.669)	0.695 (0.660-0.730)
	Sensitivity/Recall	0.667 (0.615-0.719)	0.615 (0.564-0.666)	0.640 (0.603-0.677)
	Specificity	0.875 (0.847-0.903)	0.786 (0.757-0.815)	0.823 (0.802-0.844)
	PPV/Precision	0.759 (0.708-0.810)	0.574 (0.524-0.624)	0.652 (0.615-0.689)
	NPV	0.816 (0.784-0.848)	0.814 (0.786-0.842)	0.815 (0.794-0.836)
	F1-score	0.710 (0.679-0.741)	0.594 (0.565-0.623)	0.646 (0.625-0.667)
Random Forest	AUC	0.856 (0.830-0.883)	0.779 (0.750-0.808)	0.813 (0.793-0.833)
	AP	0.786 (0.740-0.832)	0.630 (0.579-0.681)	0.695 (0.660-0.730)
	Sensitivity/Recall	0.747 (0.699-	0.718 (0.671-	0.732 (0.698-

		0.795)	0.765)	0.766)
	Specificity	0.812 (0.779-0.845)	0.692 (0.659-0.725)	0.741 (0.717-0.765)
	PPV/Precision	0.702 (0.653-0.751)	0.522 (0.477-0.567)	0.595 (0.561-0.629)
	NPV	0.844 (0.812-0.876)	0.840 (0.811-0.869)	0.842 (0.821-0.863)
	F1-score	0.724 (0.694-0.754)	0.604 (0.575-0.633)	0.656 (0.635-0.677)
Decision Tree	AUC	0.722 (0.685-0.759)	0.652 (0.615-0.687)	0.683 (0.657-0.709)
	AP	0.557 (0.502-0.612)	0.423 (0.371-0.475)	0.480 (0.442-0.518)
	Sensitivity/Recall	0.657 (0.604-0.710)	0.581 (0.529-0.633)	0.617 (0.580-0.654)
	Specificity	0.778 (0.743-0.813)	0.718 (0.686-0.750)	0.743 (0.719-0.767)
	PPV/Precision	0.637 (0.584-0.690)	0.492 (0.444-0.540)	0.555 (0.519-0.591)
	NPV	0.793 (0.758-0.828)	0.785 (0.754-0.816)	0.789 (0.766-0.812)
	F1-score	0.647 (0.615-0.679)	0.533 (0.504-0.562)	0.584 (0.562-0.606)

BP Neural Network	AUC	0.847 (0.819-0.874)	0.764 (0.734-0.796)	0.800 (0.779-0.820)
	AP	0.776 (0.730-0.822)	0.593 (0.542-0.644)	0.670 (0.634-0.706)
	Sensitivity/Recall	0.686 (0.635-0.737)	0.618 (0.567-0.669)	0.65 (0.614-0.686)
	Specificity	0.880 (0.852-0.908)	0.780 (0.750-0.810)	0.821 (0.800-0.842)
	PPV/Precision	0.773 (0.724-0.822)	0.568 (0.518-0.618)	0.654 (0.618-0.690)
	NPV	0.826 (0.795-0.857)	0.813 (0.784-0.842)	0.819 (0.798-0.840)
	F1-score	0.727 (0.697-0.757)	0.592 (0.563-0.621)	0.652 (0.631-0.673)
Logistic Regression	AUC	0.842 (0.814-0.870)	0.777 (0.748-0.806)	0.805 (0.785-0.825)
	AP	0.771 (0.724-0.818)	0.627 (0.576-0.678)	0.689 (0.654-0.724)
	Sensitivity/Recall	0.779 (0.733-0.825)	0.752 (0.707-0.797)	0.765 (0.733-0.797)
	Specificity	0.755 (0.718-0.792)	0.649 (0.615-0.683)	0.693 (0.668-0.718)
	PPV/Precision	0.653 (0.605-	0.501 (0.458-	0.564 (0.532-

	0.701)	0.544)	0.596)
NPV	0.852 (0.820-0.884)	0.848 (0.819-0.877)	0.850 (0.828-0.872)
F1-score	0.711 (0.680-0.742)	0.601 (0.572-0.630)	0.649 (0.628-0.670)

AUC, area under the receiver operating characteristic curve; AP, area under the precision-recall curve; PPV, positive predictive value; NPV, negative predictive value.