BMJ Open

Use of machine learning to develop a prehospital-stage prediction tool for traumatic brain injury

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055918
Article Type:	Original research
Date Submitted by the Author:	02-Aug-2021
Complete List of Authors:	Choi, Yeong Ho; Seoul National University Hospital, Emergency Department Park, Jeong Ho; Seoul National University Hospital, Emergency Department Hong, Ki Jeong; Seoul National University College of Medicine, Emergency Medicine; Seoul National University Seoul Metropolitan Government Boramae Medical Center, Emergency Medicine Ro, Young Sun; Seoul National University Hospital, Emergency Department Song, Kyoung Jun; Seoul Metropolitan Boramae Hospital, Department of Emergency Medicine Shin, Sang Do; Seoul National University Hospital, Department of Emergency Medicine
Keywords:	ACCIDENT & EMERGENCY MEDICINE, Neurological injury < NEUROLOGY, Trauma management < ORTHOPAEDIC & TRAUMA SURGERY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title page

1. Title

Use of machine learning to develop a prehospital stage prediction tool for traumatic brain injury

2. Authors

Yeong Ho Choi, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: d2uk87@gmail.com

Jeong Ho Park, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: timthe@gmail.com

Ki Jeong Hong, MD, PhD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital

Biomedical Research Institute, Seoul, Korea

E-mail: emkjhong@gmail.com

Young Sun Ro, MD, DrPH

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital

Biomedical Research Institute, Seoul, Korea

E-mail: ro.youngsun@gmail.com

Kyoung Jun Song, MD, PhD

Department of Emergency Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital

Biomedical Research Institute, Seoul, Korea

E-mail: skciva@gmail.com

Sang Do Shin, MD, PhD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital

Biomedical Research Institute, Seoul, Korea

E-mail: shinsangdo@gmail.com

3. Address correspondence and requests for reprints: Jeong Ho Park, MD

Address: Seoul National University Hospital, 101 Daehak-Ro, Jongno-Gu, Seoul 03080, Korea

Phone: +82-2-2072-1800

FAX: +82-2-741-7855

E-mail: timthe@gmail.com

Abstract

Objectives: Predicting diagnosis and prognosis of traumatic brain injury (TBI) at the prehospital stage is challenging; however, using comprehensive prehospital information and machine learning may improve the performance of the predictive model. We developed and tested predictive models for TBI that use machine learning algorithms using information that can be obtained in the prehospital stage.

Design: This was a multi-center retrospective study.

Setting and participants: This study was conducted at three tertiary academic emergency departments (EDs) located in an urban area.of South Korea. The data from adult patients with severe trauma who were assessed by emergency medical service (EMS) providers and transported to three participating hospitals between 2014 to 2018 were analyzed.

Results: We developed and tested five machine learning algorithms—logistic regression analyses, extreme gradient boosting, support vector machine, random forest, and elastic net (EN)—to predict TBI, TBI with intracranial hemorrhage or injury (TBI-I), TBI with emergency department or admission result of admission or transferred (TBI-ND), and TBI with emergency department or admission result of death (TBI-D). Of the 1,169 patients in the development cohort, TBI, TBI-I, TBI-ND, and TBI-D was 24.0%, 21.5%, 21.3%, and 3.7%, respectively. The EN model yielded an AUROC of 0.799 for TBI, 0.844 for TBI-I, 0.811 for TBI-ND, and 0.871 for TBI-D. The EN model also yielded the highest specificity, and significant reclassification improvement. Variables related to loss of consciousness, Glasgow Coma Scale, and light reflex were the three most important variables to predict all outcomes.

Conclusion: Our results inform the diagnosis and prognosis of TBI. Machine learning models resulted in significant performance improvement over that with logistic regression analyses, and the best performing model was EN.

Keywords: brain injuries; traumatic; outcome; prognosis; machine learning.

Strengths and limitations of this study

- By using high dimensional prehospital data, we developed and validated prediction models for the diagnosis and prognosis of traumatic brain injury using machine learning algorithms among patients with severe trauma, identified by emergency medical service providers.
- Machine learning models showed acceptable-to-excellent discrimination performance (AUROCs were 0.799–0.871 according to outcomes in the best-performing model). When identifying 80% of target patients with traumatic brain injury, the false positive rate was almost 19.7–39.0%.
- We used retrospective analysis of electronically collected prehospital data. We treated missing status as a separate category for our analysis, however, there could be different reasons for missing data.
- External validation for other areas should be conducted to generalize the developed prediction model.

Introduction

Traumatic brain injury (TBI) is a significant health burden worldwide.[1 2] It is the leading cause of mortality and disability among young individuals.[3] Patients with TBI are vulnerable to hypoxia and hypotension in the early period of their course and these insults are associated with poor outcomes.[4-6] Prehospital assessment and management of patients with TBI is important,[7 8] as early prediction of TBI and correcting hypoxia and hypotension during the prehospital stage could be beneficial.[9 10] However, the identification of TBI can often be challenging in the prehospital area.[7] Vulnerable patients, including the elderly or patients who take medications like anti-platelet or anticoagulant drugs, often have TBI owing to low energy insults.[11] Prehospital clinical signs are also reported to have poor sensitivity for raised intracranial pressure following TBI.[12]

Several prediction models to target patients with TBI have been reported.[13-15] However, most incorporated information that is available only in the hospital, such as laboratory results or image findings.[13 14 16] In addition, most previous prediction models focused on the outcomes of patients with TBI, not the identification of TBI. Previously, predictors of older adult patients with TBI who required transport to a trauma center were identified. However, this was consensus-based; therefore, there is a lack of clinical data.[17] Accurate prehospital prediction of TBI and its severity could prevent delays to definite care for patients with TBI. Most emergency medical service (EMS) providers collect various information including demographics, past medical history, circumstances of the trauma, and clinical signs including vital signs; but those variables have not been evaluated together as predictors of TBI and its severity. Using a variety of prehospital information, and adapting newly emerging machine learning algorithms for predicting diagnosis, disposition, and outcome of TBI, might improve the accuracy of identification of TBI and its severity.[18]

The aim of this study was to develop and test prediction models for the diagnosis and prognosis of TBI using prehospital information and machine learning algorithms among patients with severe trauma. We hypothesized that incorporating prehospital information could achieve acceptable performance in predicting TBI, and machine learning algorithms could contribute to performance improvement.

Materials and Methods

Study design and settings

This was a multi-center retrospective study conducted at three tertiary academic emergency departments (EDs) located in an urban area (Seoul and Bundang) of South Korea. These EDs received 50,000–90,000 visits annually. We adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement on reporting predictive models.[19]

The EMS system in South Korea is operated by the National Fire Agency. The EMS level is considered intermediate, as EMS providers can perform bleeding control, spinal motion restriction, immobilization and splintage, advanced airway management, and administer fluid intravenously. As only physicians can declare death in South Korea, EMS providers cannot stop resuscitation and must transport all patients including those in cardiac arrest to the ED. For all EMS transport, EMS providers record an ambulance run-sheet by law. Since 2012, the National Fire Agency adapted the United States Centers for Disease Control and Prevention of the United States field triage decision scheme to evaluate patients with trauma,[20] and they developed an EMS severe trauma in-depth registry. For said patients, EMS providers evaluate whether patients met trauma center transport criteria in the field triage decision scheme. If they did, the in-depth registry should be recorded, and EMS

transport protocol recommends that patients are transferred to a near regional trauma center; but it is not mandatory.

The Ministry of Health and Welfare designated three ED levels according to the resources and functional requirements; level 1 (n = 36) and level 2 (n = 118) EDs have more resources and better facilities for emergency care and must be staffed by emergency physicians 24 hours a day/365 days a year; whereas level 3 EDs (n = 248) can be staffed by general physicians. In accordance with the EMS Act, all EDs participated annually in a nationwide functional performance evaluation program, which was administered by the Ministry of Health and Welfare. The three participating hospitals in this study were all level 1 EDs that can perform acute trauma care for patients with TBI 24 hours a day/365 days a year—including emergency neurosurgical operation and angiographic interventions.

Data source

We used an EMS ambulance run-sheet, EMS trauma in-depth registry, and ED administrative database. The EMS database information, including ambulance run-sheet and trauma in-depth registry, was collected electronically by EMS providers using tablets. The EMS record review for each severe trauma has been performed by EMS medical directors of each fire department since 2012. The ED administrative database contains patients' demographic characteristics, route of visit, time of visit, and diagnosis and disposition. We merged the EMS database with the ED administrative database based on patients' arrival time, age, and sex.

Study population

We included adult (age ≥ 15) EMS users who were transported to participating hospitals with severe trauma from January 1, 2014 to December 31, 2018. Severe trauma was assessed by

EMS providers and defined as patients who fulfilled trauma center transport criteria (physiologic criteria, anatomic criteria, mechanism of injury criteria, or special patients or system consideration criteria) in the field triage decision scheme.[21] Patients were excluded if they had out-of-hospital cardiac arrest or their main cause of EMS call was medical or nontraumatic injury including choking, drowning, fire, flame, heat, cold, poisoning, chemical, sexual assault, weather, or natural disaster. Patients with an unknown outcome were also excluded.

Outcome measure

The primary outcome measure was the diagnosis of TBI. TBI diagnosis was defined as patients whose diagnostic code, according to the International Statistical Classification of Diseases and Related Health Problems (ICD-10), was between S06.0 and S06.9.[22 23] The ED administrative database has two types of primary diagnostic codes: the final diagnostic codes at ED discharge and at hospital discharge. We extracted up to 20 codes for each. We defined the diagnostic code as positive for TBI if a confirmative diagnostic code was found in any level of the discharge record. A secondary outcome measure was the diagnosis of TBI with intracranial hemorrhage or injury (TBI-I), defined as ED discharge or hospital discharge diagnosis ICD-10 code S06.1–S06.9. Concussion (ICD-10 code with S06.0) was excluded in TBI-I. A tertiary outcome was TBI with non-discharge (TBI-ND). Non-discharge was defined as patients whose ED discharge disposition included admission, transfer, or death. Quaternary outcome measure was TBI with death (TBI-D). Death was defined as patients whose ED discharge disposition or hospital discharge disposition was death.

Variables and preprocessing

We collected patients' demographic data, circumstances of trauma, chief complaints, EMS

vital sign assessment, EMS management and hospital outcomes. The detailed descriptions of each variable are described in Supplementary Table 1. Categorical variables were preprocessed with the one-hot encoding (dummy variable encoding) method. Continuous variables were divided into four quantiles and unknown or missing values were categorized as a fifth category. One-hot encoding was also applied to discretized continuous variables. Preprocessing measures including discretization results of continuous variables are presented in Supplementary Table 1.

Model development

We developed prediction models for outcomes by using five machine learning algorithms: logistic regression analyses (LR), extreme gradient boost (XGB), random forest (RF), support vector machine (SVM), and elastic net (EN). The LR algorithm was chosen as baseline comparison algorithm because it is widely used in the medical field and has been used for previous prediction model development in TBI studies.[14 24] Backward stepwise LR was selected for feature selection. The other four algorithms were selected based on their ability to model nonlinear associations, their relative ease of implementation, and their general acceptance in the machine learning community.[25-29] All algorithms have a method to calculate the probability of the outcome occurring and algorithms other than LR need hyperparameter tuning for proper training and prediction.

The study population was split into training cohorts that included development, validation, and test cohorts. The development cohort included a training cohort from which each of the machine learning prediction models were derived and a validation cohort in which the prediction models were applied to adjust the hyperparameters of the algorithm. The test cohort was used for the final evaluation of the performance of the prediction models. Chronological split was used for data split. Patients enrolled from January 1, 2014 to

December 31, 2016 were used as the training cohort; patients from January 1, 2017 to December 31, 2017 were used as the validation cohort; and patients from January 1, 2018 to December 31, 2018 were used as the test cohort. Hyperparameter tuning using validation data was conducted by, first, a random search within 10,000 randomly generated hyperparameters; then, grid search hyperparameters chosen around from random search with five candidates per each hyperparameter. Finally, hyperparameter with best area under receiver-operation curve (AUROC) in validation cohorts were selected. Test data were separated during training and tuning processes and used to measure algorithm performance.

Statistical analysis

The demographic findings and outcomes of the study population were described in this study. Additionally, the baseline characteristics of the training cohort and the validation cohort were compared. The continuous variables were compared by using Student's T-test or the Wilcoxon rank sum test, and the categorical variables were compared by using the chi-squared test or the Fisher exact test, as appropriate.

We assessed discrimination performance by comparing the AUROC for each model in the test cohort. We considered an AUROC of 0.5 as no discrimination, 0.7 to 0.8 as acceptable, 0.8 to 0.9 as excellent, and more than 0.9 is considered outstanding.[30] Area under the precision-recall curve (AUPRC) was assessed for each model in the test cohort. We assessed the calibration power by using the Hosmer–Lemeshow test, the scaled Brier score, and a calibration plot in the test cohort.[31] For the delineation of test characteristics, the sensitivity, specificity, and positive and negative predictive values with 95% CIs were determined using a cutoff probability at a sensitivity of 80%. Given that poor sensitivity of clinical predictors for TBI in previous studies,[12 32] and almost 75% sensitivity level for other severe disease prediction in prehospital settings,[33 34] we thought that 80% sensitivity was an

appropriate target for our prediction model. We calculated false positive rate as 1 – specificity. The added prognostic power of each prediction model compared to the LR model was also evaluated by continuous net reclassification index (NRI). NRI is a statistical method to quantify how well a new model correctly reclassifies the study population with the other models. Details of NRI are described elsewhere.[35]

By using a model-specific metric, the variable importance of each model was assessed, except for the SVM algorithm. The variable importance was determined by the coefficient effect sizes for the LR model. The XGB and RF models were ranked by variable importance on the selection frequency of the variable as a decision node. The absolute value of the coefficients corresponding to the tuned model were used for the measurement of variable importance in the EN algorithm.[36] To compare the variable importance of each prediction models efficiently, top 5 variables of each model was presented.

All statistical analyses were performed with R Statistical Software (version 4.0.1; R Foundation for Statistical Computing, Vienna, Austria). Packages included caret, e1071, xgboost, randomForest, and glmnet for the analysis of the machine learning algorithms.

No patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

Result

Demographic findings

Among the 157,134 EMS users transported to three hospitals from 2014 to 2018, 1,169 patients were included in the final analysis (Figure 1). Patients were split into 2 datasets: data from 2014 to 2017, consisting of 867 patients (74.2%) in the development cohort; and the remaining data from 2018 consisting of 302 patients (25.8%) in the test cohort (Figure 1). Among the development cohort, data from 2014 to 2016—consisting of 661 patients—were used as the training cohort, and 2017 data—consisting of 206 patients—were used as the validation cohort in the model.

Table 1 shows key demographic findings of the development and test cohorts. Median (IQR) age was 52 years (35–66) in the development cohort and 56 years (40–69) in the test cohort. Traffic accident was most common mechanism of trauma (43.3% for the development cohort and 41.4% for the test cohort). The proportion of patients with alert mental status was 58.1% for the development cohort and 69.5% in the test cohort. Overall, TBI, TBI-I, TBI-ND, TBI-D occurred in 215 (24.8%), 195 (22.5%), 192 (22.1%), and 32 (3.7%) in the development cohort; and 66 (21.9%), 56 (18.5%), 57 (18.9%), and 11 (3.6%) in the test cohort. All demographic characteristics of the development and test cohorts are described in Supplementary Table 2.

Main analysis

The discrimination and NRI of the prediction models on the test cohort are presented in Table 2. The AUROC for outcomes were 0.770–0.806 for TBI, 0.820–0.844 for TBI-I, 0.767–0.811 for TBI-ND, and 0.664–0.889 for TBI-D (Table 2 and Supplementary Figure 1). Compared to LR, XGB performed significantly well in predicting TBI, and RF and EN performed well in

predicting TBI-ND and TBI-D. EN model generally performed well on all outcomes. The AUROC of the EN model for outcomes were 0.799 (95% CI: 0.732–0.867), 0.844 (95% CI: 0.779–0.910), 0.811 (95% CI: 0.741–0.882), and 0.871 (95% CI: 0.764–0.978) for TBI, TBI-I, TBI-ND, and TBI-D, respectively. Machine learning models generally resulted in significant reclassification improvement compared to LR for TBI, TBI-I, and TBI-ND. For prediction TBI-D, AUROC difference, and reclassification improvement compared to LR was non-significant in all machine learning models. The precision-recall curve is shown in Supplementary Figure 2. AUPRC were 0.479–0.564 for TBI, 0.469–0.606 for TBI-I, 0.477–0.551 for TBI-ND and 0.094–0.140 for TBI-D. EN model showed highest AUPRC among all prediction models. Supplementary Figure 3 shows the calibration plot of prediction models according to outcomes. All prediction models generally showed poor calibration. Given the high AUROC and AUPRC among prediction models, and reclassification improvement compared to LR, we determined EN as a best-performing prediction model in our analysis.

Using cutoff of 80% sensitivity, specificity was 47.5–68.2% for TBI, 71.1–81.3% for TBI-I, 46.1–74.3% for TBI-ND, and 42.6-–.0 for TBI-D. EN showed the highest specificity and PPV among all outcomes. False positive rate (1 – specificity) was almost 19.7–39.0% according to outcomes in the EN model. The 95% CI of specificity of the EN model was not overlapped with LR in TBI, TBI-ND, and TBI-D predictions. NPV was almost 89–99% for all outcomes in the prediction models (Table 3).

Table 4 shows the top 5 variable importance of prediction models according to outcomes. Variables related to patients' symptom of loss of consciousness, Glasgow Coma Scale component, and light reflex were the three most important variables to predict all outcomes. Compared to other outcomes, the difference between variable importance for TBI-D was prominent, and the mechanism of injury, heart rate, and age showed the highest importance for predicting TBI-D.

Discussion

By using prehospital data from EMS users visiting three teaching hospitals, we developed and validated prediction models for the diagnosis and prognosis of TBI using machine learning algorithms among patients with severe trauma, identified by EMS providers in South Korea. We found that 24% of patients were diagnosed with TBI, 22% showed intracranial injury, 21% could not be discharged from the ED with a TBI diagnosis, and 4% showed TBI-related death. Machine learning models showed acceptable-to-excellent discrimination performance (AUROCs were 0.799–0.871 according to outcomes in the best-performing EN model). When identifying 80% of target patients with TBI, the false positive rate was almost 19.7–39.0%. Consciousness status related variables ranging from patients' symptom to EMS providers' assessment showed the highest importance for predicting all outcomes. This study adds considerably to the understanding of prehospital prediction performance of TBI among patients with severe trauma. Use of comprehensive prehospital information and certain machine learning approaches led to increased performance with a diminished false positive rate compared to those of the traditional statistical model.

Several studies reported that EMS providers' assessment using prehospital information is effective for the identification of patients with severe trauma who require direct transport to a trauma center.[37-39] Because TBI accounts for a significant portion of patients with severe trauma,[38] and the majority of patients have poor access to trauma centers,[40] identification of TBI among patients with severe trauma by EMS providers could contribute to proper prehospital management and destination hospital decisions.[6] However, prehospital identification of TBI is challenging.[41] Prehospital clinical signs showed poor predictive performance for differentiating patients with TBI. [12], and previous prediction models related to TBI mostly focused on TBI outcomes.[13 14 16] One study reported the predictors for mild TBI with persistent symptoms; but a single-center case-control study

design and ED-based model development lacks applicability to prehospital settings.[32] In this study, we developed and tested TBI prediction models that used prehospital information, and we found acceptable discrimination power for the prediction of diagnosis and prognosis of TBI. Uniquely, we incorporated various demographic variables, trauma circumstances, patients' complaints, and EMS assessment information in the prediction models, and we adapted the machine learning algorithms.

When using a cutoff for 80% sensitivity for TBI detection, the false positive rate was 19.7−39.0% (Table 2). Those false positive rate levels are plausible for detecting severe diseases in EMS settings. A previous study reported a 26% of false positive rate of EMS triage for myocardial infarction with a sensitivity of 74% and 50% of false positive rate of EMS recognition of stroke in sensitivity of 74%.[33 34] Considering the prevalence of outcomes (24% in TBI, 22% in TBI-I, 21% in TBI-ND, and 4% in TBI-D; Table 1), there would be 16, 9, 12, and 67 false-positive patients for every 10 patients that are accurately identified as TBI, TBI-I, TBI-ND, and TBI-D, respectively. Because of the low prevalence of TBI-D, a similar specificity of the prediction model for outcomes resulted in a very low positive predictive value and a high proportion of false positive cases, which suggested the limited applicability of prediction models for TBI-D in prehospital settings.

Consciousness-status-related variables ranging from patients' complaints to EMS assessment showed the highest importance regardless of models and outcomes in our study. Consciousness status is closely associated with head trauma. Head trauma can result in structural brain injury or physiological disruption of brain function, which could result in altered mental status.[42] Mental status is also associated with TBI severity, [43] and its association with TBI outcomes have been reported.[13 14 16] History taking and physical examination for altered mental status is key to early diagnosis and proper management of TBI in prehospital settings.[44]

We adapted machine learning algorithms for the prediction of TBI-related outcomes and found an improvement in discrimination and an increase in specificity with the same sensitivity thresholds. However, the LR model also showed acceptable or similar performance compared to machine learning models, according to the outcomes. In clinical prediction models, a previous systematic review reported no performance benefit of the machine learning model over LR.[45] The previous study stated that machine learning models tend to show high performance with a strong signal-to-noise ratio problem like gaming, image recognition. However, clinical prediction problems often result in a poor signal-to-noise ratio.[45 46] If we could use unstructured data, which has strong signal-to-noise ratio like continuous vital sign monitoring data or audiovisual data for patients' appearance, machine learning models might perform better than LR models. In addition, if we analyzed more patient data, the performance improvement of machine models might be elucidated.

Precise assessment in prehospital field could contribute to improved patient-related outcomes. High demand of EMS call and response,[47] disparity in accessibility to definitive care capable hospitals according to regions,[40] and the importance of timely management in acute disease care are the chief reasons behind the necessity for the accurate assessment of EMS providers. Although information acquisition and processing is quite difficult in prehospital areas, various instruments and information systems could attribute to diminish those problems. Complex data acquisition like mobile CT or other unstructured data[48 49], information sharing through telemedicine,[50] and decision support tools in prehospital environments[51] could contribute to the accurate assessment of EMS providers. More information acquisition and real-time processing of those data could improve the clinical prediction models in prehospital areas, which could lead to the improvement of patients' safety and outcomes.

Our study had several limitations. First, our data were collected at three teaching hospitals in urban areas of South Korea. Therefore, external validation for other areas should be conducted to generalize the developed prediction model. Second, we used retrospective analysis of electronically collected prehospital and hospital data. There might be various information loss and missing data. We treated missing status as a separate category for our analysis;[52] however, there could be different reasons for missing data. Third, there is a possibility that the prediction model was overfitted or underfitted. To minimize this issue, we rigorously searched hyperparameters and carefully chose hyperparameters according to the performance in independent validation cohorts. Lastly, this study was performed in an intermediate-service-level EMS system. The generalization of our study findings to different EMS settings should be made with caution.

In conclusion, we presented data on TBI among patients with severe trauma assessed by EMS providers, and our results inform the development of prediction models for the diagnosis and prognosis of TBI in our population. We used various information that can be obtained in prehospital settings and showed acceptable outcome performance. The consistent importance of consciousness-status-related variables emphasizes the importance of assessment and monitoring of consciousness status in prehospital areas. Although prospective, and implementation studies are needed for TBI prediction in prehospital areas, our study outlined a novel method for the precise assessment of EMS providers using a machine-learning-based prediction model. Further collection of various types of patient-related data would contribute to the enhanced performance of the clinical prediction model in prehospital settings.

Author Contribution Statement

YHC and JH Park designed and developed the study, analysed and interpreted the data, and drafted the initial manuscript. KJH, YSR, KJS and SDS were involved in the acquisition of data, the development of the research question and assisted with analysis and interpretation of data. All authors revised the drafts for intellectual content and edited the manuscript. All authors reviewed and approved the final draft.

Funding

This study was supported by grant No. '04-2019-0680' from the Seoul National University Hospital Research Fund.

Competing Interests

There are no conflicts of interest for all authors in this study.

Patients consent

Not required

Data availability statement

No data are available. We do not have ethics approval to share data.

Ethical statements

This study complied with the Declaration of Helsinki, and its protocol was approved by the Institutional Review Board of the Seoul National University Hospital with a waiver of informed consent (IRB No: E-2006-004-1128).

References

1. Maas AI, Stocchetti N, Bullock R. Moderate and severe traumatic brain injury in adults. *The Lancet Neurology* 2008;7(8):728-41.
2. Hsia RY, Markowitz AJ, Lin F, et al. Ten-year trends in traumatic brain injury: a retrospective cohort study of California emergency department and hospital revisits and readmissions. *BMJ Open* 2018;8(12):e022297. doi: 10.1136/bmjopen-2018-022297 [published Online First: 2018/12/16]
3. Finfer SR, Cohen J. Severe traumatic brain injury. *Resuscitation* 2001;48(1):77-90. doi: 10.1016/s0300-9572(00)00321-x [published Online First: 2001/02/13]
4. DeWitt DS, Jenkins LW, Prough DS. Enhanced vulnerability to secondary ischemic insults after experimental traumatic brain injury. *New Horiz* 1995;3(3):376-83. [published Online First: 1995/08/01]
5. McHugh GS, Engel DC, Butcher I, et al. Prognostic value of secondary insults in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;24(2):287-93. doi: 10.1089/neu.2006.0031 [published Online First: 2007/03/23]
6. Spaite DW, Bobrow BJ, Keim SM, et al. Association of Statewide Implementation of the Prehospital Traumatic Brain Injury Treatment Guidelines With Patient Survival Following Traumatic Brain Injury: The Excellence in Prehospital Injury Care (EPIC) Study. *JAMA Surg* 2019;154(7):e191152. doi: 10.1001/jamasurg.2019.1152 [published

Online First: 2019/05/09]

7. Pelieu I, Kull C, Walder B. Prehospital and Emergency Care in Adult Patients with Acute Traumatic Brain Injury. *Med Sci (Basel)* 2019;7(1) doi: 10.3390/medsci7010012 [published Online First: 2019/01/24]

8. Goldberg SA, Rojanasarntikul D, Jagoda A. The prehospital management of traumatic brain injury. *Handb Clin Neurol* 2015;127:367-78. doi: 10.1016/B978-0-444-52892-6.00023-4 [published Online First: 2015/02/24]

9. Chi JH, Knudson MM, Vassar MJ, et al. Prehospital hypoxia affects outcome in patients with traumatic brain injury: a prospective multicenter study. *J Trauma* 2006;61(5):1134-41. doi: 10.1097/01.ta.0000196644.64653.d8 [published Online First: 2006/11/14]

10. Barton CW, Hemphill JC, Morabito D, et al. A novel method of evaluating the impact of secondary brain insults on functional outcomes in traumatic brain-injured patients. *Acad Emerg Med* 2005;12(1):1-6. doi: 10.1197/j.aem.2004.08.043 [published Online First: 2005/01/07]

11. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 2012;61(1):1-20.

12. Ter Avest E, Taylor S, Wilson M, et al. Prehospital clinical signs are a poor predictor of raised intracranial pressure following traumatic brain injury. *Emerg Med J* 2021;38(1):21-26. doi: 10.1136/emermed-2020-209635 [published Online First: 2020/09/20]

13. Collaborators MCT, Perel P, Arango M, et al. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;336(7641):425-9. doi: 10.1136/bmj.39461.643438.25 [published Online First:

2008/02/14]

14. Steyerberg EW, Mushkudiani N, Perel P, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med* 2008;5(8):e165; discussion e65. doi: 10.1371/journal.pmed.0050165 [published Online First: 2008/08/08]

15. Gozt AK, Hellewell SC, Thorne J, et al. Predicting outcome following mild traumatic brain injury: protocol for the longitudinal, prospective, observational Concussion Recovery (CREST) cohort study. *BMJ Open* 2021;11(5):e046460. doi: 10.1136/bmjopen-2020-046460 [published Online First: 2021/05/15]

16. Miller PR, Chang MC, Hoth JJ, et al. Predicting Mortality and Independence at Discharge in the Aging Traumatic Brain Injury Population Using Data Available at Admission. *J Am Coll Surg* 2017;224(4):680-85. doi: 10.1016/j.jamcollsurg.2016.12.053 [published Online First: 2017/03/07]

17. Wasserman EB, Shah MN, Jones CM, et al. Identification of a neurologic scale that optimizes EMS detection of older adult traumatic brain injury patients who require transport to a trauma center. *Prehosp Emerg Care* 2015;19(2):202-12. doi: 10.3109/10903127.2014.959225 [published Online First: 2014/10/08]

18. Hale AT, Stonko DP, Brown A, et al. Machine-learning analysis outperforms conventional statistical models and CT classification systems in predicting 6-month outcomes in pediatric patients sustaining traumatic brain injury. *Neurosurgical Focus FOC* 2018;45(5):E2. doi: 10.3171/2018.8.Focus17773

19. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Ann Intern Med* 2015;162(10):735-6. doi: 10.7326/L15-5093-2 [published Online First: 2015/05/20]

20. Sasser SM, Hunt RC, Sullivent EE, et al. Guidelines for field triage of injured patients:

recommendations of the National Expert Panel on Field Triage. 2009

21. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *MMWR Recomm Rep* 2012;61(RR-1):1-20. [published Online First: 2012/01/13]

22. Andelic N, Anke A, Skandsen T, et al. Incidence of hospital-admitted severe traumatic brain injury and in-hospital fatality in Norway: a national cohort study. *Neuroepidemiology* 2012;38(4):259-67. doi: 10.1159/000338032 [published Online First: 2012/06/09]

23. Ro YS, Shin SD, Holmes JF, et al. Comparison of clinical performance of cranial computed tomography rules in patients with minor head injury: a multicenter prospective study. *Acad Emerg Med* 2011;18(6):597-604. doi: 10.1111/j.1553-2712.2011.01094.x [published Online First: 2011/06/17]

24. Gang MC, Hong KJ, Shin SD, et al. New prehospital scoring system for traumatic brain injury to predict mortality and severe disability using motor Glasgow Coma Scale, hypotension, and hypoxia: a nationwide observational study. *Clin Exp Emerg Med* 2019;6(2):152-59. doi: 10.15441/ceem.18.027 [published Online First: 2019/07/03]

25. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: Association for Computing Machinery, 2016:785–94.

26. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67(2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x

27. Dong Q, Taylor RA, Moore CL, et al. Predicting urinary tract infections in the emergency department with machine learning. *Plos One* 2018;13(3) doi: 10.1371/journal.pone.0194085

28. Breiman RF. Vaccines as tools for advancing more than public health: perspectives of a former director of the National Vaccine Program office. *Clin Infect Dis* 2001;32(2):283-8. doi: 10.1086/318461 [published Online First: 2001/02/15]

29. Hearst MA, Dumais ST, Osuna E, et al. Support vector machines. *IEEE Intelligent Systems and their Applications* 1998;13(4):18-28. doi: 10.1109/5254.708428

30. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression: John Wiley & Sons 2013.

31. Fenlon C, O'Grady L, Doherty ML, et al. A discussion of calibration techniques for evaluating binary and categorical predictive models. *Prev Vet Med* 2018;149:107-14. doi: 10.1016/j.prevetmed.2017.11.018 [published Online First: 2018/01/02]

32. Wojcik SM. Predicting mild traumatic brain injury patients at risk of persistent symptoms in the Emergency Department. *Brain Inj* 2014;28(4):422-30. doi: 10.3109/02699052.2014.884241 [published Online First: 2014/02/26]

33. Oostema JA, Konen J, Chassee T, et al. Clinical predictors of accurate prehospital stroke recognition. *Stroke* 2015;46(6):1513-7. doi: 10.1161/STROKEAHA.115.008650 [published Online First: 2015/04/30]

34. Swan PY, Nighswonger B, Boswell GL, et al. Factors associated with false-positive emergency medical services triage for percutaneous coronary intervention. *West J Emerg Med* 2009;10(4):208-12. [published Online First: 2010/01/05]

35. Park JH, Shin SD, Song KJ, et al. Prediction of good neurological recovery after out-of-hospital cardiac arrest: A machine learning analysis. *Resuscitation* 2019;142:127-35. doi: 10.1016/j.resuscitation.2019.07.020 [published Online First: 2019/07/31]

36. Weng SF, Reps J, Kai J, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;12(4):e0174944. doi: 10.1371/journal.pone.0174944 [published Online First: 2017/04/05]

37. Esposito TJ, Offner PJ, Jurkovich GJ, et al. Do prehospital trauma center triage criteria identify major trauma victims? *Arch Surg* 1995;130(2):171-6. doi: 10.1001/archsurg.1995.01430020061010 [published Online First: 1995/02/01]

38. Ocak G, Sturms LM, Hoogeveen JM, et al. Prehospital identification of major trauma patients. *Langenbecks Arch Surg* 2009;394(2):285-92. doi: 10.1007/s00423-008-0340-4 [published Online First: 2008/06/27]

39. Fries GR, McCalla G, Levitt MA, et al. A prospective comparison of paramedic judgment and the trauma triage rule in the prehospital setting. *Ann Emerg Med* 1994;24(5):885-9. doi: 10.1016/s0196-0644(94)70207-1 [published Online First: 1994/11/01]

40. Branas CC, MacKenzie EJ, Williams JC, et al. Access to trauma centers in the United States. *JAMA* 2005;293(21):2626-33. doi: 10.1001/jama.293.21.2626 [published Online First: 2005/06/02]

41. Whiting MD, Dengler BA, Rodriguez CL, et al. Prehospital Detection of Life-Threatening Intracranial Pathology: An Unmet Need for Severe TBI in Austere, Rural, and Remote Areas. *Front Neurol* 2020;11:599268. doi: 10.3389/fneur.2020.599268 [published Online First: 2020/11/17]

42. Management of Concussion/m TBIWG. VA/DoD Clinical Practice Guideline for Management of Concussion/Mild Traumatic Brain Injury. *J Rehabil Res Dev* 2009;46(6):CP1-68. [published Online First: 2010/01/30]

43. Grote S, Bocker W, Mutschler W, et al. Diagnostic value of the Glasgow Coma Scale for traumatic brain injury in 18,002 patients with severe multiple injuries. *J Neurotrauma* 2011;28(4):527-34. doi: 10.1089/neu.2010.1433 [published Online First: 2011/01/27]

44. Badjatia N, Carney N, Crocco TJ, et al. Guidelines for prehospital management of traumatic brain injury 2nd edition. *Prehosp Emerg Care* 2008;12 Suppl 1:S1-52. doi: 10.1080/10903120701732052 [published Online First: 2008/09/06]

45. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019;110:12-22. doi: 10.1016/j.jclinepi.2019.02.004 [published Online First: 2019/02/15]

46. Ennis M, Hinton G, Naylor D, et al. A comparison of statistical learning methods on the Gusto database. *Stat Med* 1998;17(21):2501-8. doi: 10.1002/(sici)1097-0258(19981115)17:21<2501::aid-sim938>3.0.co;2-m [published Online First: 1998/11/20]

47. Crowe RP, Bower JK, Cash RE, et al. Association of Burnout with Workforce-Reducing Factors among EMS Professionals. *Prehosp Emerg Care* 2018;22(2):229-36. doi: 10.1080/10903127.2017.1356411 [published Online First: 2017/08/26]

48. Hov MR, Zakariassen E, Lindner T, et al. Interpretation of Brain CT Scans in the Field by Critical Care Physicians in a Mobile Stroke Unit. *J Neuroimaging* 2018;28(1):106-11. doi: 10.1111/jon.12458 [published Online First: 2017/08/03]

49. Nakada TA, Masunaga N, Nakao S, et al. Development of a prehospital vital signs chart sharing system. *Am J Emerg Med* 2016;34(1):88-92. doi: 10.1016/j.ajem.2015.09.048 [published Online First: 2015/10/29]

50. Kim Y, Groombridge C, Romero L, et al. Decision Support Capabilities of Telemedicine in Emergency Prehospital Care: Systematic Review. *J Med Internet Res* 2020;22(12):e18959. doi: 10.2196/18959 [published Online First: 2020/12/09]

51. Reisner AT, Khitrov MY, Chen L, et al. Development and validation of a portable platform for deploying decision-support algorithms in prehospital settings. *Appl Clin Inform* 2013;4(3):392-402. doi: 10.4338/ACI-2013-04-RA-0023 [published Online First: 2013/10/25]

52. Maslove DM, Podchiyska T, Lowe HJ. Discretization of continuous features in clinical

603 datasets. *J Am Med Inform Assoc* 2013;20(3):544-53. doi: 10.1136/amiajnl-2012-
604 000929 [published Online First: 2012/10/13]

605

Figure legends

Figure 1. Population flow. EMS, emergency medical service; OHCA, out-of-hospital cardiac arrest; TBI, traumatic brain injury.

610 Table 1. Key characteristics of the development and test cohorts.

	n (%) or Median (IQR)			
	Total	Development cohort	Test cohort	P
Total	N = 1169	n = 867	n = 302	
Demographics				
Age, years	53 (36–66)	52 (35–66)	56 (40–69)	< 0.01
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Job, unemployed	299 (25.6)	197 (22.7)	102 (33.8)	< 0.01
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	< 0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	< 0.01
Circumstances of trauma				
Location, road/highway	444 (38.0)	326 (37.6)	118 (39.1)	0.65
Season, summer	336 (28.7)	253 (29.2)	83 (27.5)	0.57
Weekday, weekend	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time, 6 p.m. to midnight	361 (30.9)	265 (30.6)	96 (31.8)	0.69
Mechanism of injury, TA	500 (42.8)	375 (43.3)	125 (41.4)	0.57
Chief complaint				
Fracture/abrasion/laceration	302 (25.8)	204 (23.5)	98 (32.5)	< 0.01
EMS vital sign assessment				
SBP, mmHg*	130 (109–150)	130 (104–146)	131 (115–150)	< 0.01
DBP, mmHg*	80 (70–91)	80 (69–90)	80 (70–92)	0.21
RR, mmHg*	18 (16–20)	18 (16–20)	18 (16–20)	0.33
HR, /min*	86 (75–99)	86 (74–99)	86 (76–100)	0.40
SpO2, %*	98 (95–99)	98 (95–99)	98 (96–99)	0.67
AVPU scale, Alert	714 (61.1)	504 (58.1)	210 (69.5)	< 0.01
EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	0.95

611 IQR, interquartile range; TA, traffic accident; SBP, systolic blood pressure; DBP, diastolic
612 blood pressure; RR, respiratory rate; ED, emergency department; TBI, traumatic brain injury.

Table 2. Discrimination and reclassification of prediction models for outcomes on test cohort.

Outcome	Model	AUROC (95% CI)	p[a]	NRI (95% CI)	p[b]	AUPRC
TBI						
	LR	0.770 (0.698, 0.841)	NA	NA	NA	0.492
	XGB	0.809 (0.743, 0.876)	0.04	0.689 (0.427, 0.951)	< 0.01	0.552
	SVM	0.776 (0.708, 0.844)	0.77	0.339 (0.072, 0.607)	0.01	0.479
	RF	0.800 (0.735, 0.865)	0.13	0.308 (0.047, 0.569)	0.02	0.532
	EN	0.799 (0.732, 0.867)	0.06	0.698 (0.441, 0.954)	< 0.01	0.564
TBI-I						
	LR	0.820 (0.751, 0.890)	NA	NA	NA	0.551
	XGB	0.838 (0.775, 0.901)	0.28	0.539 (0.258, 0.821)	< 0.01	0.554
	SVM	0.812 (0.748, 0.875)	0.66	0.729 (0.464, 0.994)	< 0.01	0.469
	RF	0.836 (0.772, 0.899)	0.38	0.333 (0.058, 0.607)	0.02	0.552
	EN	0.844 (0.779, 0.910)	0.15	1.093 (0.845, 1.342)	< 0.01	0.606
TBI-ND						
	LR	0.767 (0.690, 0.844)	NA	NA	NA	0.482
	XGB	0.800 (0.727, 0.873)	0.07	0.605 (0.326, 0.884)	< 0.01	0.496
	SVM	0.778 (0.704, 0.852)	0.56	0.285 (-0.001, 0.572)	0.05	0.477
	RF	0.809 (0.739, 0.880)	0.03	0.194 (-0.059, 0.448)	0.13	0.535
	EN	0.811 (0.741, 0.882)	0.02	0.768 (0.496, 1.039)	< 0.01	0.551
TBI-D						
	LR	0.664 (0.490, 0.838)	NA	NA	NA	0.138
	XGB	0.714 (0.512, 0.917)	0.64	-0.026 (-0.605, 0.553)	0.93	0.094
	SVM	0.814 (0.718, 0.910)	0.09	0.209 (-0.325, 0.742)	0.44	0.140
	RF	0.889 (0.801, 0.976)	< 0.01	-0.204 (-0.742, 0.334)	0.46	0.196
	EN	0.871 (0.764, 0.978)	0.01	0.119 (-0.415, 0.654)	0.66	0.293

[a]Comparing the AUROC and the logistic regression model.

[b]Comparing the NRI and the logistic regression model.

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; NRI, net reclassification index; AUPRC, area under precision-recall curve; TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support vector machine; RF, random forest; EN, elastic net

626 Table 3. Test characteristics of prediction models for outcomes on test cohort.

Outcome	Model	Specificity (95% CI)	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)	Cutoff
TBI						
	LR	47.5 (40.9, 54.0)	80.3 (68.7, 89.1)	29.9 (23.3, 37.3)	89.6 (82.9, 94.3)	0.136
	XGB	72.5 (66.3, 78.1)	80.3 (68.7, 89.1)	44.9 (35.7, 54.3)	92.9 (88.2, 96.2)	0.268
	SVM	64.8 (58.4, 70.9)	80.3 (68.7, 89.1)	39.0 (30.7, 47.7)	92.2 (87.0, 95.8)	0.191
	RF	68.2 (61.9, 74.1)	80.3 (68.7, 89.1)	41.4 (32.8, 50.4)	92.5 (87.6, 96.0)	0.185
	EN	61.0 (54.5, 67.3)	80.3 (68.7, 89.1)	36.6 (28.7, 44.9)	91.7 (86.3, 95.5)	0.205
TBI-I						
	LR	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.164
	XGB	74.0 (68.0, 79.4)	80.4 (67.6, 89.8)	41.3 (31.9, 51.1)	94.3 (90.0, 97.1)	0.143
	SVM	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.172
	RF	76.0 (70.2, 81.2)	80.4 (67.6, 89.8)	43.3 (33.6, 53.3)	94.4 (90.3, 97.2)	0.205
	EN	81.3 (75.9, 86.0)	80.4 (67.6, 89.8)	49.5 (38.8, 60.1)	94.8 (90.9, 97.4)	0.204
TBI-ND						
	LR	46.1 (39.8, 52.6)	80.7 (68.1, 90.0)	25.8 (19.6, 32.9)	91.1 (84.7, 95.5)	0.090
	XGB	66.5 (60.2, 72.4)	80.7 (68.1, 90.0)	35.9 (27.7, 44.9)	93.7 (89.0, 96.8)	0.242
	SVM	59.2 (52.7, 65.4)	80.7 (68.1, 90.0)	31.5 (24.1, 39.7)	92.9 (87.7, 96.4)	0.147
	RF	60.4 (54.0, 66.6)	80.7 (68.1, 90.0)	32.2 (24.6, 40.5)	93.1 (88.0, 96.5)	0.138
	EN	74.3 (68.3, 79.6)	80.7 (68.1, 90.0)	42.2 (32.8, 52.0)	94.3 (90.0, 97.1)	0.201
TBI-D						
	LR	42.6 (36.9, 48.5)	81.8 (48.2, 97.7)	5.1 (2.4, 9.5)	98.4 (94.4, 99.8)	0.005
	XGB	57.7 (51.8, 63.5)	81.8 (48.2, 97.7)	6.8 (3.2, 12.5)	98.8 (95.8, 99.9)	0.002
	SVM	74.2 (68.8, 79.2)	81.8 (48.2, 97.7)	10.7 (5.0, 19.4)	99.1 (96.7, 99.9)	0.039
	RF	74.9 (69.5, 79.8)	81.8 (48.2, 97.7)	11.0 (5.1, 19.8)	99.1 (96.8, 99.9)	0.005
	EN	79.0 (73.9, 83.6)	81.8 (48.2, 97.7)	12.9 (6.1, 23.0)	99.1 (96.9, 99.9)	0.033

627 TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-
628 ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with
629 death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support
630 vector machine; RF, random forest; EN, elastic net.
631

Table 4. Top 5 important variables for outcomes in descending order using model specific metrics

Outcome	Rank	LR	XGB	RF	EN
TBI					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Motor, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	4	Light reflex	Other mechanism	Light reflex	GCS, Eye, 1
	5	GCS, Motor, 1	GCS, Verbal, 2	GCS, Motor, 1	GCS, Verbal, 1
TBI-I					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	GCS, Eye, 1
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	Loss of consciousness
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	Light reflex	GCS, Verbal, 1
	5	GCS, Motor, 1	Other mechanism	GCS, Motor, 1	Light reflex
TBI-ND					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	GCS, Verbal, 2	GCS, Verbal, 1
	5	GCS, Motor, 1	GCS, Motor, 1	GCS, Motor, 4	Light reflex
TBI-D					
	1	Loss of consciousness	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	2	GCS, Verbal, 1	Oxygen saturation<96%	Light reflex	GCS, Verbal, 1
	3	GCS, Eye, 1	Fall mechanism	Loss of consciousness	Loss of consciousness
	4	Light reflex	Afternoon	GCS, Eye, 1	Age over 80
	5	GCS, Motor, 1	Light reflex	GCS, Motor, 1	HR 87-99

TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression; XGB, extreme gradient boosting; RF, random forest; EN, elastic net; GCS, Glasgow coma scale; HR, heart rate.

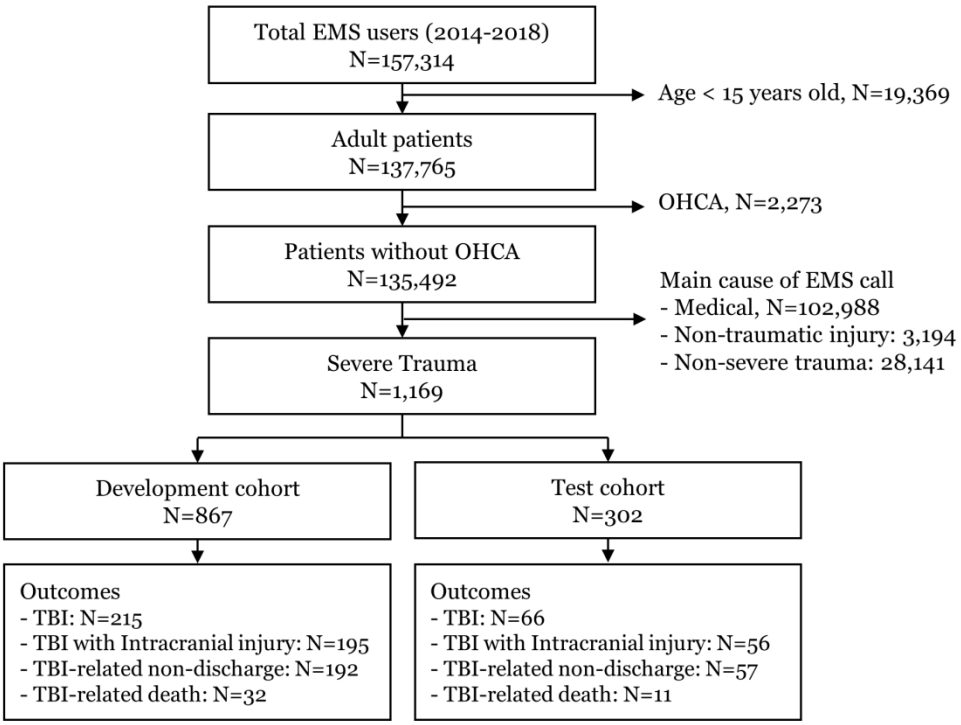


Figure 1

165x119mm (300 x 300 DPI)

Supplementary Table 1. List of analyzed variables.

Variables	Descriptions	Type of raw data	Category	Preprocessing
Gender	Sex of the patients	Binary	Male, Female	
Age	Age of patients	Continuous	15-39 years, 40-59 years, 60-79 years, and 80- years	Discretization and one hot encoding
Job	Job of patients	Categorical	Unemployed, Student/Housewife; Office/Commercial/Service workers; Industrial/Agricultural/Fishery/Miner worker; Others	One hot encoding Missing data were classified into others
Diabetes	History of diabetes mellitus	Binary	Yes, No	Missing data were classified into no
Hypertension	History of hypertension	Binary	Yes, No	Missing data were classified into no
Location of injury	Location of injury	Categorical	home/residential area/medical facility/school/gym; Road/highway; Off-road traffic area; Others	One hot encoding Missing data were classified into others
Season	Season when injury occurred	Categorical	Spring, Summer, Fall, Winter	One hot encoding
Weekend	Whether Injury occurred on weekday or weekend	Binary	Weekday, Weekend	
Daytime	When injury was occurred	Categorical	Night (Midnight to 5AM), Morning (6AM to 11AM), Afternoon (Midday to 5PM), Evening (6PM to 11PM)	One hot encoding Missing time were imputed using EMS call time
Mechanism of injury	Mechanism of injury	Categorical	Slip down, Fall down, Traffic accident, Other	One hot encoding Missing data were classified into others
Glasgow coma scale eye	Eye element of Glasgow coma scale	Categorical	1;2;3;4;Unknown	One hot encoding
Glasgow coma scale Verbal	Verbal element of Glasgow coma scale	Categorical	1;2;3;4;5;Unknown	One hot encoding
Glasgow coma scale Motor	Motor element of Glasgow coma scale	Categorical	1;2;3;4;5;6;Unknown	One hot encoding
Light Reflex any Abnormal	Any abnormality of light reflex on any side	Categorical	No, Yes, Unknown	One hot encoding Missing data were classified into unknown

Systolic blood pressure	Systolic blood pressure	Continuous	-107 mmHg, 108-130 mmHg, 131-145 mmHg, 146- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Diastolic blood pressure	Diastolic blood pressure	Continuous	-69 mmHg, 70-80 mmHg, 81-91 mmHg, 92- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Heart rate	Heart rate	Continuous	-74/min, 75-86/min, 87-99/min, 100-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Respiratory rate	Respiratory rate	Continuous	-16/min, 17-18/min, 19-20/min, 21-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Oxygen saturation	Oxygen saturation	Continuous	-95%, 96-98%, 99%, 100%, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Body temperature	Body temperature	Continuous	-36℃, 36.1-36.3℃, 36.4-36.8℃, 36.9-℃, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Chest pain or abdominal pain	Symptom of chest pain or abdominal pain	Binary	Yes, No	

Fracture, abrasion, or laceration	Symptom of fracture, abrasion, or laceration	Binary	Yes, No	
Loss of consciousness	Symptom of loss of consciousness	Binary	Yes, No	
Dyspnea	Symptom of dyspnea	Binary	Yes, No	
Nose bleeding	Symptom of nose bleeding	Binary	Yes, No	
Nausea or vomiting	Symptom of nausea or vomiting	Binary	Yes, No	
Headache, paralysis or dizziness	Symptom of headache, paralysis or dizziness	Binary	Yes, No	

Supplementary Table 2. Demographic characteristics of development and test cohorts

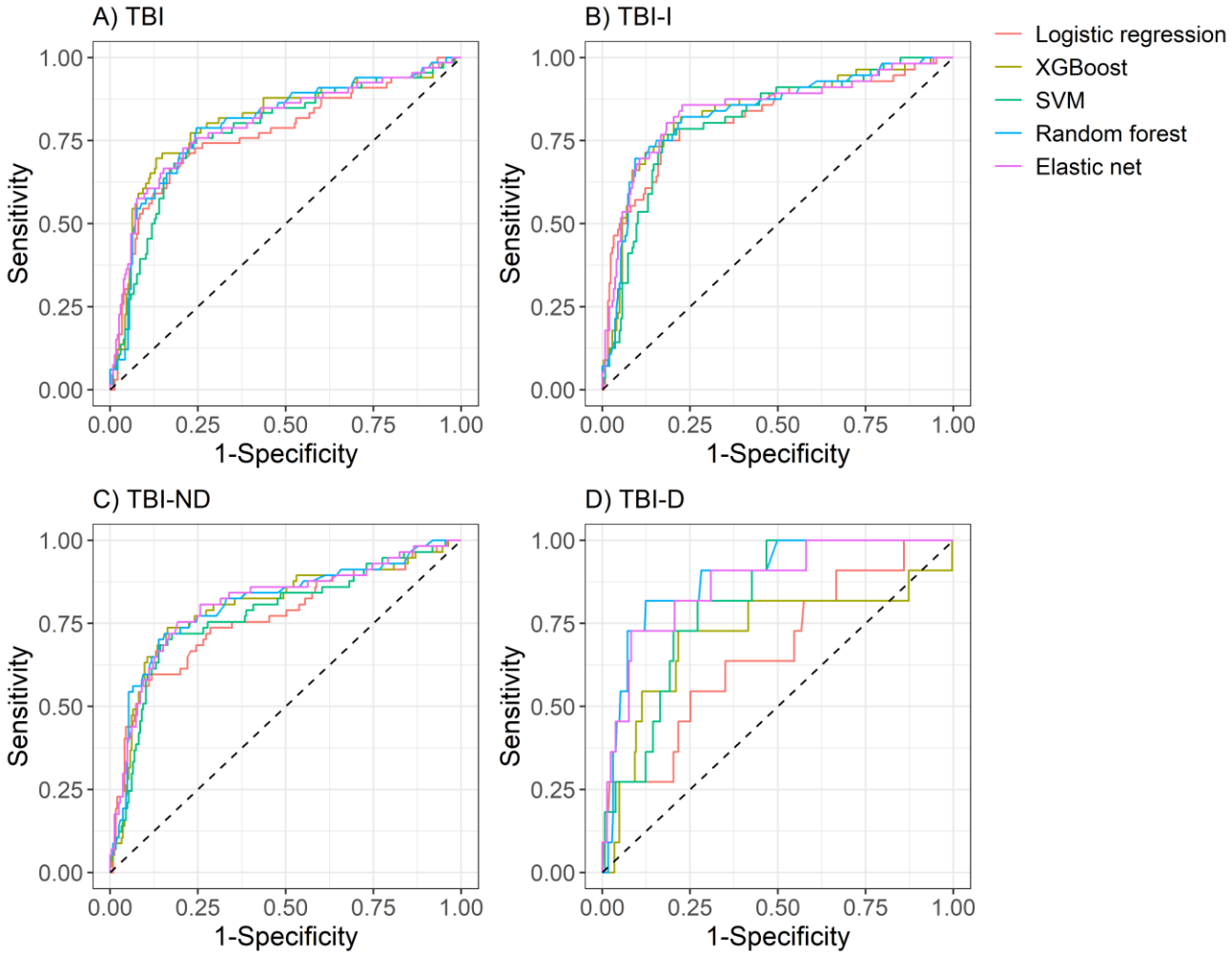
	N (%) or Median (IQR)			
Characteristics	Total	Development	Test	P-value
Total	1169	867	302	
Demographics				
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Age, years	53 (36-66)	52 (35-66)	56 (40-69)	<0.01
Job of patients				<0.01
Unemployed	299 (25.6)	197 (22.7)	102 (33.8)	
Student/Housewife	161 (13.8)	129 (14.9)	32 (10.6)	
Office/Commercial/Service worker	283 (24.2)	176 (20.3)	107 (35.4)	
Industrial/Agricultural/Fishery/Minery worker	36 (3.1)	25 (2.9)	11 (3.6)	
Others	390 (33.4)	340 (39.2)	50 (16.6)	
Past medical history				
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	<0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	<0.01
Circumstances of Trauma				
Location of trauma				0.52
Residential/Nursing/Education/Exercise facility	303 (25.9)	218 (25.1)	85 (28.1)	
Road/Highway	444 (38.0)	326 (37.6)	118 (39.1)	
Off-road traffic area	181 (15.5)	140 (16.1)	41 (13.6)	
Others	241 (20.6)	183 (21.1)	58 (19.2)	
Season of trauma				<0.01
Spring	249 (21.3)	167 (19.3)	82 (27.2)	
Summer	336 (28.7)	253 (29.2)	83 (27.5)	
Fall	304 (26.0)	242 (27.9)	62 (20.5)	
Winter	280 (24.0)	205 (23.6)	75 (24.8)	
Weekday	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time of trauma				0.83
6A-MD	281 (24.0)	206 (23.8)	75 (24.8)	
MD-6P	266 (22.8)	203 (23.4)	63 (20.9)	
6P-MN	361 (30.9)	265 (30.6)	96 (31.8)	
MN-6A	261 (22.3)	193 (22.3)	68 (22.5)	
Mechanism of Trauma				0.60
Traffic accident	500 (42.8)	375 (43.3)	125 (41.4)	
Slip down	325 (27.8)	232 (26.8)	93 (30.8)	
Fall down	171 (14.6)	129 (14.9)	42 (13.9)	
Others	173 (14.8)	131 (15.1)	42 (13.9)	
Chief complaint				
Altered mentality	279 (23.9)	223 (25.7)	56 (18.5)	0.01
Facture/Abrasion/Laceration	302 (25.8)	204 (23.5)	98 (32.5)	<0.01
Chest/Abdominal pain	47 (4.0)	31 (3.6)	16 (5.3)	0.19
Dyspnea	25 (2.1)	20 (2.3)	5 (1.7)	0.50

Epistaxis	44 (3.8)	30 (3.5)	14 (4.6)	0.36
Headache/Paralysis/Dizziness/Vertigo	95 (8.1)	64 (7.4)	31 (10.3)	0.11
Nausea/Vomiting	32 (2.7)	20 (2.3)	12 (4.0)	0.13
EMS Vital sign assessment				
SBP, mmHg	130 (109-150)	130 (104-146)	131 (115-150)	<0.01
Missing	65 (5.6)	56 (6.5)	9 (3.0)	0.02
DBP, mmHg	80 (70-91)	80 (69-90)	80 (70-92)	<0.01
Missing	75 (6.4)	65 (7.5)	10 (3.3)	0.01
HR, /min	86 (75-99)	86 (74-99)	86 (76-100)	<0.01
Missing	31 (2.7)	28 (3.2)	3 (1.0)	0.04
RR, /min	18 (16-20)	18 (16-20)	18 (16-20)	<0.01
Missing	36 (3.1)	33 (3.8)	3 (1.0)	0.01
SpO2, %	98 (95-99)	98 (95-99)	98 (96-99)	<0.01
Missing	38 (3.3)	33 (3.8)	5 (1.7)	0.07
Temperature, ℃	36.5 (36-36.8)	36.5 (36-36.8)	36.5 (36-36.7)	<0.01
Missing	94 (8.0)	65 (7.5)	29 (9.6)	0.25
AVPU scale				<0.01
Alert	714 (61.1)	504 (58.1)	210 (69.5)	
Verbal	168 (14.4)	136 (15.7)	32 (10.6)	
Pain	199 (17.0)	158 (18.2)	41 (13.6)	
Unresponsive	88 (7.5)	69 (8.0)	19 (6.3)	
Abnormal light reflex	165 (14.1)	132 (15.2)	33 (10.9)	<0.01
Missing	66 (5.6)	57 (6.6)	9 (3.0)	
GCS scale component				
Glasgow coma scale eye				<0.01
4	558 (47.7)	380 (43.8)	178 (58.9)	
3	128 (10.9)	109 (12.6)	19 (6.3)	
2	110 (9.4)	82 (9.5)	28 (9.3)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	
Glasgow coma scale Verbal				0.01
5	520 (44.5)	359 (41.4)	161 (53.3)	
4	118 (10.1)	88 (10.1)	30 (9.9)	
3	25 (2.1)	19 (2.2)	6 (2.0)	
2	132 (11.3)	105 (12.1)	27 (8.9)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	200 (17.1)	155 (17.9)	45 (14.9)	
Glasgow coma scale Motor				<0.01
6	499 (42.7)	333 (38.4)	166 (55.0)	
5	124 (10.6)	103 (11.9)	21 (7.0)	
4	158 (13.5)	123 (14.2)	35 (11.6)	
3	47 (4.0)	39 (4.5)	8 (2.6)	
2	17 (1.5)	15 (1.7)	2 (0.7)	
1	125 (10.7)	99 (11.4)	26 (8.6)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	

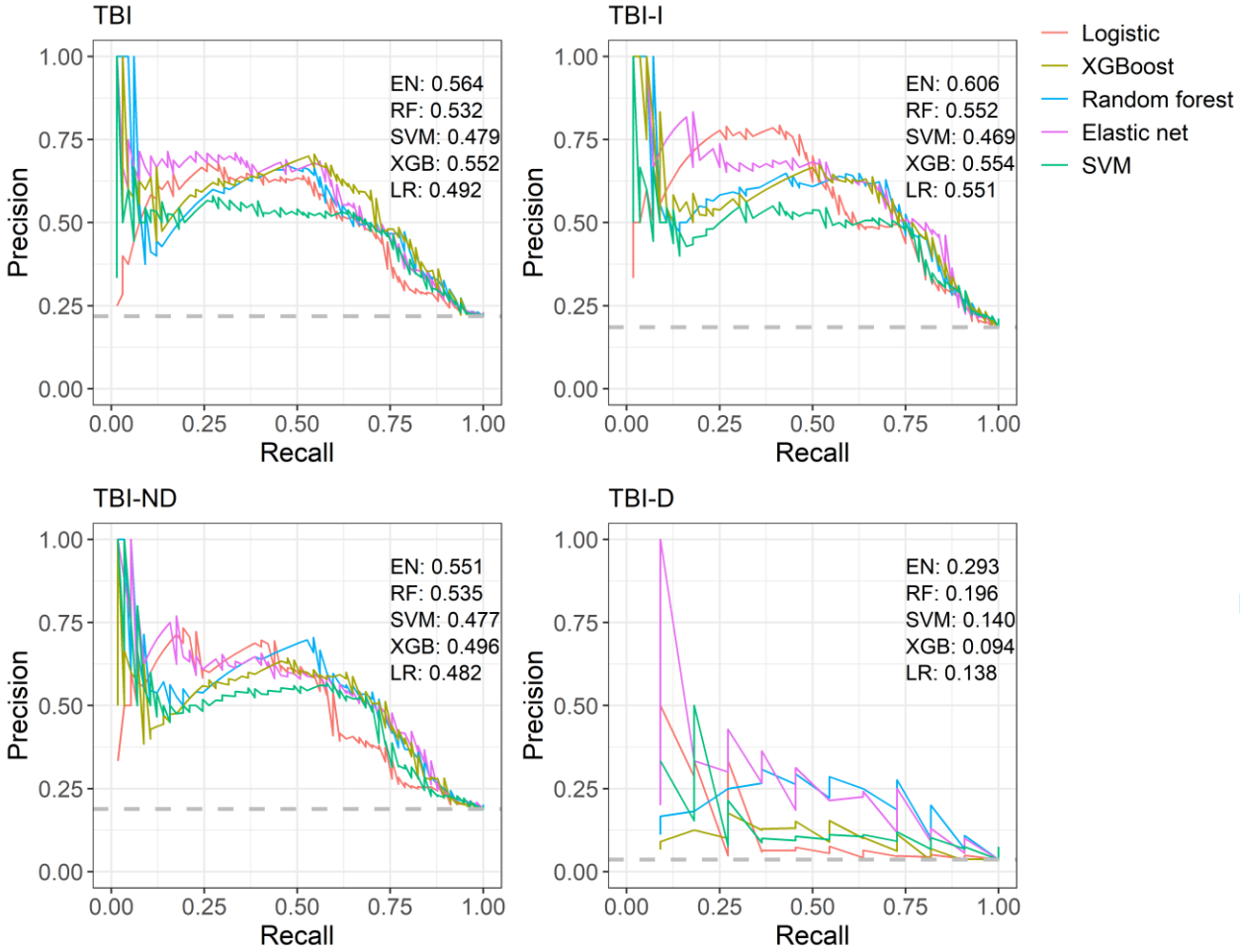
EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Advanced airway management	4 (0.3)	2 (0.2)	2 (0.7)	0.28
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
Field triage decision scheme criteria				
Physiological criteria				
SBP<90 mmHg	58 (5.0)	42 (4.8)	16 (5.3)	0.75
RR<10 or >29 /min	11 (0.9)	11 (1.3)	0 (0.0)	0.08
Non-Alert	429 (36.7)	343 (39.6)	86 (28.5)	<0.01
Anatomic criteria				
All penetrating injuries to head, neck, torso and extremities proximal to elbow or knee	34 (2.9)	23 (2.7)	11 (3.6)	0.38
Chest wall instability or deformity	4 (0.3)	4 (0.5)	0 (0.0)	0.58
Two or more proximal long bone fractures	19 (1.6)	13 (1.5)	6 (2.0)	0.60
Crush, degloved, mangled or pulseless extremity	15 (1.3)	13 (1.5)	2 (0.7)	0.38
Amputation proximal to wrist or ankle	9 (0.8)	9 (1.0)	0 (0.0)	0.12
Pelvic fractures	8 (0.7)	6 (0.7)	2 (0.7)	>0.95
Open or depressed skull fracture	17 (1.5)	9 (1.0)	8 (2.6)	0.05
Paralysis	21 (1.8)	11 (1.3)	10 (3.3)	0.02
Mechanism of injury criteria				
Fall > 6 meter	113 (9.7)	84 (9.7)	29 (9.6)	>0.95
High-risk auto crash	96 (8.2)	73 (8.4)	23 (7.6)	0.66
Auto vs pedestrian/bicyclist thrown, run over, or with significant (>30km/h) impact	119 (10.2)	83 (9.6)	36 (11.9)	0.25
Motorcycle crash > 30 km/hour	105 (9.0)	70 (8.1)	35 (11.6)	0.07
ED disposition				0.11
Discharge	320 (27.4)	241 (27.8)	79 (26.2)	
Transfer	444 (38.0)	316 (36.4)	128 (42.4)	
Admitted	366 (31.3)	276 (31.8)	90 (29.8)	
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	>0.95

IQR, interquartile range; SBP, systolic blood pressure; RR, respiratory rate; ED, emergency department; TBI, traumatic brain injury.

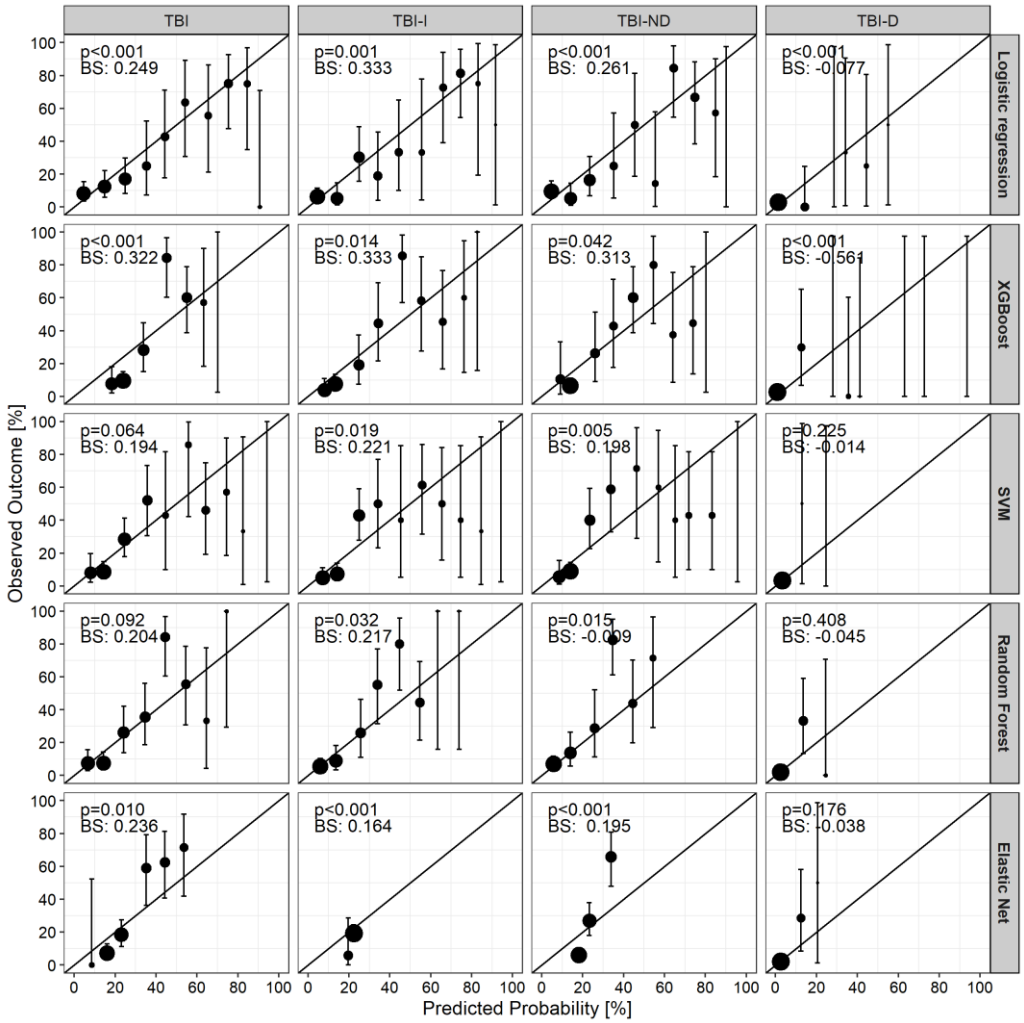
Supplementary Figure 1. Receiver operating characteristics of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death.



Supplementary Figure 2. Precision-recall curve of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; LR, logistic regression analysis; XGB, extreme gradient boosting; RF, random forest, EN, elastic net.



Supplementary Figure 3. Calibration plot of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; p, p-value of Hosmer-Lemeshow test; BS, scaled Brier score.



TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	4
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	7
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	8
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8-9
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	9
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8-9
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	N/A
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	10-11
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	11
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	8	D;V	Explain how the study size was arrived at.	14
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	11
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	11
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	11-12
	10c	V	For validation, describe how the predictions were calculated.	12-13
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	12-13
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	12
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	14
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	14
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	14
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
	15b	D	Explain how to the use the prediction model.	14-15
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	14-15
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	19
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	16-17
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	16
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Suppl
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	20

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Development and validation of a prehospital-stage prediction tool for traumatic brain injury: a multicentre retrospective cohort study in Korea

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055918.R1
Article Type:	Original research
Date Submitted by the Author:	25-Oct-2021
Complete List of Authors:	Choi, Yeong Ho; Seoul National University Hospital, Emergency Department; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Park, Jeong Ho; Seoul National University Hospital, Emergency Department; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Hong, Ki Jeong; Seoul National University Hospital, Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Ro, Young Sun; Seoul National University Hospital, Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Song, Kyoung Jun; Seoul Metropolitan Boramae Hospital, Department of Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Shin, Sang Do; Seoul National University Hospital, Department of Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Emergency medicine, Health informatics
Keywords:	ACCIDENT & EMERGENCY MEDICINE, Neurological injury < NEUROLOGY, Trauma management < ORTHOPAEDIC & TRAUMA SURGERY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title page

1. Title

Development and validation of a prehospital-stage prediction tool for traumatic brain injury: a multicentre retrospective cohort study in Korea

2. Authors

Yeong Ho Choi, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: d2uk87@gmail.com

Jeong Ho Park, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: timthe@gmail.com

Ki Jeong Hong, MD, PhD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: emkjhong@gmail.com

Young Sun Ro, MD, DrPH
Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: ro.youngsun@gmail.com

Kyoung Jun Song, MD, PhD
Department of Emergency Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: skciva@gmail.com

Sang Do Shin, MD, PhD
Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: shinsangdo@gmail.com

3. Address correspondence and requests for reprints: Jeong Ho Park, MD

Address: Seoul National University Hospital, 101 Daehak-Ro, Jongno-Gu, Seoul 03080, Korea

Phone: +82-2-2072-1800

FAX: +82-2-741-7855

E-mail: timthe@gmail.com

Abstract

Objectives: Predicting diagnosis and prognosis of traumatic brain injury (TBI) at the prehospital stage is challenging; however, using comprehensive prehospital information and machine learning may improve the performance of the predictive model. We developed and tested predictive models for TBI that use machine learning algorithms using information that can be obtained in the prehospital stage.

Design: This was a multi-center retrospective study.

Setting and participants: This study was conducted at three tertiary academic emergency departments (EDs) located in an urban area.of South Korea. The data from adult patients with severe trauma who were assessed by emergency medical service (EMS) providers and transported to three participating hospitals between 2014 to 2018 were analyzed.

Results: We developed and tested five machine learning algorithms—logistic regression analyses, extreme gradient boosting, support vector machine, random forest, and elastic net (EN)—to predict TBI, TBI with intracranial hemorrhage or injury (TBI-I), TBI with emergency department or admission result of admission or transferred (TBI-ND), and TBI with emergency department or admission result of death (TBI-D). Of the 1,169 patients in the development cohort, TBI, TBI-I, TBI-ND, and TBI-D was 24.0%, 21.5%, 21.3%, and 3.7%, respectively. The EN model yielded an AUROC of 0.799 for TBI, 0.844 for TBI-I, 0.811 for TBI-ND, and 0.871 for TBI-D. The EN model also yielded the highest specificity, and significant reclassification improvement. Variables related to loss of consciousness, Glasgow Coma Scale, and light reflex were the three most important variables to predict all outcomes.

Conclusion: Our results inform the diagnosis and prognosis of TBI. Machine learning models resulted in significant performance improvement over that with logistic regression analyses, and the best performing model was EN.

Keywords: brain injuries; traumatic; outcome; prognosis; machine learning.

Strengths and limitations of this study

- This is an original research to develop and internally validate prehospital-stage prediction models for traumatic brain injury using high dimensional prehospital information.
- Machine learning models showed acceptable-to-excellent discrimination performance.
- The retrospective observational study design could lead to certain types of bias (eg, selection bias, confounding bias).
- External validation for other areas should be conducted to generalize the developed prediction model.

Introduction

Traumatic brain injury (TBI) is a significant health burden worldwide.[1] It is the leading cause of mortality and disability among young individuals.[2] Patients with TBI are vulnerable to hypoxia and hypotension in the early period of their course and these insults are associated with poor outcomes.[3 4] Prehospital assessment and management of patients with TBI is important,[5] as early prediction of TBI and correcting hypoxia and hypotension during the prehospital stage could be beneficial.[3] However, the identification of TBI can often be challenging in the prehospital area.[5] Vulnerable patients, including the elderly or patients who take medications like anti-platelet or anticoagulant drugs, often have TBI owing to low energy insults.[6] Prehospital clinical signs are also reported to have poor sensitivity for raised intracranial pressure following TBI.[7]

Several prediction models to target patients with TBI have been reported.[8-12] However, most incorporated information that is available only in the hospital, such as laboratory results or image findings.[8 9 13] In addition, most previous prediction models focused on the outcomes of patients with TBI,[14-16] not the identification of TBI. Previously, predictors of older adult patients with TBI who required transport to a trauma center were identified. However, this was consensus-based; therefore, there is a lack of clinical data.[17] Accurate prehospital prediction of TBI and its severity could prevent delays to definite care for patients with TBI. Most emergency medical service (EMS) providers collect various information including demographics, past medical history, circumstances of the trauma, and clinical signs including vital signs; but those variables have not been evaluated together as predictors of TBI and its severity. Using a variety of prehospital information, and adapting newly emerging machine learning algorithms for predicting diagnosis, disposition, and outcome of TBI, might improve the accuracy of identification of TBI and its severity.

The aim of this study was to develop and test prediction models for the diagnosis and prognosis of TBI using prehospital information and machine learning algorithms among patients with severe trauma. We hypothesized that incorporating prehospital information could achieve acceptable performance in predicting TBI, and machine learning algorithms could contribute to performance improvement.

Materials and Methods

Study design and settings

This was a multi-center retrospective study conducted at three tertiary academic emergency departments (EDs) located in an urban area (Seoul and Bundang) of South Korea. These EDs received 50,000–90,000 visits annually and are not designated trauma centers. We adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement on reporting predictive models.[18]

The EMS system in South Korea is operated by the National Fire Agency. The EMS level is considered intermediate, as EMS providers can perform bleeding control, spinal motion restriction, immobilization and splintage, advanced airway management, and administer fluid intravenously. As only physicians can declare death in South Korea, EMS providers cannot stop resuscitation and must transport all patients including those in cardiac arrest to the ED. For all EMS transport, EMS providers record an ambulance run-sheet by law. Since 2012, the National Fire Agency adapted the United States Centers for Disease Control and Prevention of the United States field triage decision scheme to evaluate patients with trauma,[19] and they developed an EMS severe trauma in-depth registry. For said patients, EMS providers evaluate whether patients met trauma center transport criteria in the field triage decision scheme. If they did, the in-depth registry should be recorded, and EMS

transport protocol recommends that patients are transferred to a near regional trauma center; but it is not mandatory.

The Ministry of Health and Welfare designated three ED levels according to the resources and functional requirements; level 1 (n = 36) and level 2 (n = 118) EDs have more resources and better facilities for emergency care and must be staffed by emergency physicians 24 hours a day/365 days a year; whereas level 3 EDs (n = 248) can be staffed by general physicians. In accordance with the EMS Act, all EDs participated annually in a nationwide functional performance evaluation program, which was administered by the Ministry of Health and Welfare. The three participating hospitals in this study were all level 1 EDs that can perform acute trauma care for patients with TBI 24 hours a day/365 days a year—including emergency neurosurgical operation and angiographic interventions. The Ministry of Health and Welfare also designated trauma centers in Korea. Total 16 trauma centers were designated as trauma centers in 2018. Among them, 15 were Level I EDs.

Data source

We used an EMS ambulance run-sheet, EMS trauma in-depth registry, and ED administrative database. The EMS database information, including ambulance run-sheet and trauma in-depth registry, was collected electronically by EMS providers using tablets. The EMS record review for each severe trauma has been performed by EMS medical directors of each fire department since 2012. The ED administrative database contains patients' demographic characteristics, route of visit, time of visit, and diagnosis and disposition. We merged the EMS database with the ED administrative database based on patients' arrival time, age, and sex.

Study population

We included adult (age ≥ 15) EMS users who were transported to participating hospitals with severe trauma from January 1, 2014 to December 31, 2018. Severe trauma was assessed by EMS providers and defined as patients who fulfilled trauma center transport criteria (physiologic criteria, anatomic criteria, mechanism of injury criteria, or special patients or system consideration criteria) in the field triage decision scheme.[20] Patients were excluded if they had out-of-hospital cardiac arrest or their main cause of EMS call was medical or nontraumatic injury including choking, drowning, fire, flame, heat, cold, poisoning, chemical, sexual assault, weather, or natural disaster. Patients with an unknown outcome were also excluded.

Outcome measure

The primary outcome measure was the diagnosis of TBI. TBI diagnosis was defined as patients whose diagnostic code, according to the International Statistical Classification of Diseases and Related Health Problems (ICD-10), was between S06.0 and S06.9.[21 22] Although S06.7 is codes for the duration of unconscious, we included S06.7 in our study outcome according to the previous studies. [21-23] However, no patients only have S06.7 code for TBI diagnosis in our study.The ED administrative database has two types of primary diagnostic codes: the final diagnostic codes at ED discharge and at hospital discharge. We extracted up to 20 codes for each. We defined the diagnostic code as positive for TBI if a confirmative diagnostic code was found in any level of the discharge record. Because ICD 10 code is not directly linked to the severity of TBI, we further included a variety of additional outcome measures to perform analysis that take into account severity. A secondary outcome measure was TBI diagnosis with intracranial hemorrhage or injury (TBI-I), defined as TBI

patients excluding concussion (ICD 10 code with S06.0). A tertiary outcome was TBI with non-discharge (TBI-ND), defined as TBI patients excluding ED discharged patients. Because TBI-ND patients needed further management by hospitalization or transfer, we thought that this group of patients had clinically significant severity. A quaternary outcome measure was TBI with death (TBI-D), defined as TBI patients who died in ED or hospital. Because TBI-D patients are most severe group, TBI-D patients were also included in TBI-ND.

Variables and preprocessing

We collected patients' demographic data, circumstances of trauma, chief complaints, EMS vital sign assessment, EMS management and hospital outcomes. The detailed descriptions of each variable are described in Supplementary Table 1. Categorical variables were preprocessed with the one-hot encoding (dummy variable encoding) method. Continuous variables were divided into four quantiles and unknown or missing values were categorized as a fifth category. One-hot encoding was also applied to discretized continuous variables. Preprocessing measures including discretization results of continuous variables are presented in Supplementary Table 1.

Model development

We developed prediction models for outcomes by using five machine learning algorithms: traditional logistic regression analyses (LR), extreme gradient boost (XGB), random forest (RF), support vector machine (SVM), and elastic net (EN). The LR algorithm was chosen as baseline comparison algorithm because it is widely used in the medical field and has been used for previous prediction model development in TBI studies.[12] Backward stepwise LR was selected for feature selection, and we used the default parameter of stepAIC function from MASS package (version 7.3-53.1) in R for the selection. The other four algorithms were

selected based on their ability to model nonlinear associations, their relative ease of implementation, and their general acceptance in the machine learning community.[24-26] All algorithms have a method to calculate the probability of the outcome occurring and algorithms other than LR need hyperparameter tuning for proper training and prediction.

The study population was split into training cohorts that included development, validation, and test cohorts. The development cohort included a training cohort from which each of the machine learning prediction models were derived and a validation cohort in which the prediction models were applied to adjust the hyperparameters of the algorithm. The test cohort was used for the final evaluation of the performance of the prediction models. Chronological split was used for data split. Patients enrolled from January 1, 2014 to December 31, 2016 were used as the training cohort; patients from January 1, 2017 to December 31, 2017 were used as the validation cohort; and patients from January 1, 2018 to December 31, 2018 were used as the test cohort. Hyperparameter tuning using validation data was conducted by, first, a random search within 10,000 randomly generated hyperparameters; then, grid search hyperparameters chosen around from random search with five candidates per each hyperparameter. Finally, hyperparameter with best area under receiver-operation curve (AUROC) in validation cohorts were selected. Test data were separated during training and tuning processes and used to measure algorithm performance.

Statistical analysis

The demographic findings and outcomes of the study population were described in this study. Additionally, the baseline characteristics of the training cohort and the validation cohort were compared. The continuous variables were compared by using Student's T-test or the Wilcoxon rank sum test, and the categorical variables were compared by using the chi-squared test or the Fisher exact test, as appropriate.

We assessed discrimination performance by comparing the AUROC for each model in the test cohort. We considered an AUROC of 0.5 as no discrimination, 0.7 to 0.8 as acceptable, 0.8 to 0.9 as excellent, and more than 0.9 is considered outstanding.[27] Area under the precision-recall curve (AUPRC) was assessed for each model in the test cohort. We assessed the calibration power by using the Hosmer–Lemeshow test, the scaled Brier score, and a calibration plot in the test cohort. For the delineation of test characteristics, the sensitivity, specificity, and positive and negative predictive values with 95% CIs were determined using a cutoff probability at a sensitivity of 80%. Given that poor sensitivity of clinical predictors for TBI in previous studies,[7] and almost 75% sensitivity level for other severe disease prediction in prehospital settings,[28 29] we thought that 80% sensitivity was an appropriate target for our prediction model. We calculated false positive rate as 1 – specificity. The added prognostic power of each prediction model compared to the LR model was also evaluated by continuous net reclassification index (NRI). NRI is a statistical method to quantify how well a new model correctly reclassifies the study population with the other models. Details of NRI are described elsewhere.[30]

By using a model-specific metric, the variable importance of each model was assessed, except for the SVM algorithm. The variable importance was determined by the coefficient effect sizes for the LR model. The XGB and RF models were ranked by variable importance on the selection frequency of the variable as a decision node. The absolute value of the coefficients corresponding to the tuned model were used for the measurement of variable importance in the EN algorithm. To compare the variable importance of each prediction models efficiently, top 5 variables of each model was presented.

All statistical analyses were performed with R Statistical Software (version 4.0.1; R Foundation for Statistical Computing, Vienna, Austria). Packages included caret, e1071, xgboost, randomForest, and glmnet for the analysis of the machine learning algorithms.

Patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

Result

Demographic findings

Among the 157,134 EMS users transported to three hospitals from 2014 to 2018, 1,169 patients were included in the final analysis (Figure 1). Patients were split into 2 datasets: data from 2014 to 2017, consisting of 867 patients (74.2%) in the development cohort; and the remaining data from 2018 consisting of 302 patients (25.8%) in the test cohort (Figure 1). Among the development cohort, data from 2014 to 2016—consisting of 661 patients—were used as the training cohort, and 2017 data—consisting of 206 patients—were used as the validation cohort in the model.

Table 1 shows key demographic findings of the development and test cohorts. Median (IQR) age was 52 years (35–66) in the development cohort and 56 years (40–69) in the test cohort. Traffic accident was most common mechanism of trauma (43.3% for the development cohort and 41.4% for the test cohort). The proportion of patients with alert mental status was

58.1% for the development cohort and 69.5% in the test cohort. Overall, TBI, TBI-I, TBI-ND, TBI-D occurred in 215 (24.8%), 195 (22.5%), 192 (22.1%), and 32 (3.7%) in the development cohort; and 66 (21.9%), 56 (18.5%), 57 (18.9%), and 11 (3.6%) in the test cohort. All demographic characteristics of the development and test cohorts are described in Supplementary Table 2.

Main analysis

The final hyperparameters of prediction models are described in Supplementary Table 3. The discrimination and NRI of the prediction models on the test cohort are presented in Table 2. The AUROC for outcomes were 0.770–0.806 for TBI, 0.820–0.844 for TBI-I, 0.767–0.811 for TBI-ND, and 0.664–0.889 for TBI-D (Table 2 and Supplementary Figure 1). Compared to LR, XGB performed significantly well in predicting TBI, and RF and EN performed well in predicting TBI-ND and TBI-D. EN model generally performed well on all outcomes. The AUROC of the EN model for outcomes were 0.799 (95% CI: 0.732–0.867), 0.844 (95% CI: 0.779–0.910), 0.811 (95% CI: 0.741–0.882), and 0.871 (95% CI: 0.764–0.978) for TBI, TBI-I, TBI-ND, and TBI-D, respectively. Machine learning models generally resulted in significant reclassification improvement compared to LR for TBI, TBI-I, and TBI-ND. For prediction TBI-D, AUROC difference, and reclassification improvement compared to LR was non-significant in all machine learning models. The precision-recall curve is shown in Supplementary Figure 2. AUPRC were 0.479–0.564 for TBI, 0.469–0.606 for TBI-I, 0.477–0.551 for TBI-ND and 0.094–0.140 for TBI-D. EN model showed highest AUPRC among all prediction models. Supplementary Figure 3 shows the calibration plot of prediction models according to outcomes. All prediction models generally showed poor calibration. Given the high AUROC and AUPRC among prediction models, and reclassification improvement

compared to LR, we determined EN as a best-performing prediction model in our analysis.

Using cutoff of 80% sensitivity, specificity was 47.5–68.2% for TBI, 71.1–81.3% for TBI-I, 46.1–74.3% for TBI-ND, and 42.6-–.0 for TBI-D. EN showed the highest specificity and PPV among all outcomes. False positive rate (1 – specificity) was almost 19.7–39.0% according to outcomes in the EN model. The 95% CI of specificity of the EN model was not overlapped with LR in TBI, TBI-ND, and TBI-D predictions. NPV was almost 89–99% for all outcomes in the prediction models (Table 3).

Table 4 shows the top 5 variable importance of prediction models according to outcomes. Variables related to patients' symptom of loss of consciousness, Glasgow Coma Scale component, and light reflex were the three most important variables to predict all outcomes. Compared to other outcomes, the difference between variable importance for TBI-D was prominent, and the mechanism of injury, heart rate, and age showed the highest importance for predicting TBI-D.

Discussion

By using prehospital data from EMS users visiting three teaching hospitals, we developed and validated prediction models for the diagnosis and prognosis of TBI using machine learning algorithms among patients with severe trauma, identified by EMS providers in South Korea. We found that 24% of patients were diagnosed with TBI, 22% showed intracranial injury, 21% could not be discharged from the ED with a TBI diagnosis, and 4% showed TBI-related death. Machine learning models showed acceptable-to-excellent discrimination performance (AUROCs were 0.799–0.871 according to outcomes in the best-performing EN model). When identifying 80% of target patients with TBI, the false positive rate was almost 19.7–39.0%. Consciousness status related variables ranging from patients' symptom to EMS

providers' assessment showed the highest importance for predicting all outcomes. This study adds considerably to the understanding of prehospital prediction performance of TBI among patients with severe trauma. Use of comprehensive prehospital information and certain machine learning approaches led to increased performance with a diminished false positive rate compared to those of the traditional statistical model.

Several studies reported that EMS providers' assessment using prehospital information is effective for the identification of patients with severe trauma who require direct transport to a trauma center.[31-33] Because TBI accounts for a significant portion of patients with severe trauma,[32] and the majority of patients have poor access to trauma centers,[34] identification of TBI among patients with severe trauma by EMS providers could contribute to proper prehospital management and destination hospital decisions.[3] However, prehospital identification of TBI is challenging.[35] Prehospital clinical signs showed poor predictive performance for differentiating patients with TBI. [7], and previous prediction models related to TBI mostly focused on TBI outcomes.[8 9 13] One study reported the predictors for mild TBI with persistent symptoms; but a single-center case-control study design and ED-based model development lacks applicability to prehospital settings.[36] In this study, we developed and tested TBI prediction models that used prehospital information, and we found acceptable discrimination power for the prediction of diagnosis and prognosis of TBI. Uniquely, we incorporated various demographic variables, trauma circumstances, patients' complaints, and EMS assessment information in the prediction models, and we adapted the machine learning algorithms.

When using a cutoff for 80% sensitivity for TBI detection, the false positive rate was 19.7–39.0% (Table 2). Those false positive rate levels are plausible for detecting severe diseases in EMS settings. A previous study reported a 26% of false positive rate of EMS

triage for myocardial infarction with a sensitivity of 74% and 50% of false positive rate of EMS recognition of stroke in sensitivity of 74%.[28 29] Considering the prevalence of outcomes (24% in TBI, 22% in TBI-I, 21% in TBI-ND, and 4% in TBI-D; Table 1), there would be 16, 9, 12, and 67 false-positive patients for every 10 patients that are accurately identified as TBI, TBI-I, TBI-ND, and TBI-D, respectively (Supplementary Table 4). Because of the low prevalence of TBI-D, a similar specificity of the prediction model for outcomes resulted in a very low positive predictive value and a high proportion of false positive cases, which suggested the limited applicability of prediction models for TBI-D in prehospital settings.

Consciousness-status-related variables ranging from patients' complaints to EMS assessment showed the highest importance regardless of models and outcomes in our study. Consciousness status is closely associated with head trauma. Head trauma can result in structural brain injury or physiological disruption of brain function, which could result in altered mental status.[37] Mental status is also associated with TBI severity, [38] and its association with TBI outcomes have been reported.[8 9 13] History taking and physical examination for altered mental status is key to early diagnosis and proper management of TBI in prehospital settings.[39]

We adapted machine learning algorithms for the prediction of TBI-related outcomes and found an improvement in discrimination and an increase in specificity with the same sensitivity thresholds. However, the LR model also showed acceptable or similar performance compared to machine learning models, according to the outcomes. In clinical prediction models, a previous systematic review reported no performance benefit of the machine learning model over LR.[40] The previous study stated that machine learning models tend to show high performance with a strong signal-to-noise ratio problem like gaming, image recognition. However, clinical prediction problems often result in a poor signal-to-

noise ratio.[40] If we could use unstructured data, which has strong signal-to-noise ratio like continuous vital sign monitoring data or audiovisual data for patients' appearance, machine learning models might perform better than LR models. In addition, if we analyzed more patient data, the performance improvement of machine models might be elucidated.

Precise assessment in prehospital field could contribute to improved patient-related outcomes. High demand of EMS call and response, disparity in accessibility to definitive care capable hospitals according to regions,[34] and the importance of timely management in acute disease care are the chief reasons behind the necessity for the accurate assessment of EMS providers. Although information acquisition and processing is quite difficult in prehospital areas, various instruments and information systems could attribute to diminish those problems. Complex data acquisition like mobile CT or other unstructured data[41], information sharing through telemedicine,[42] and decision support tools in prehospital environments[43] could contribute to the accurate assessment of EMS providers. More information acquisition and real-time processing of those data could improve the clinical prediction models in prehospital areas, which could lead to the improvement of patients' safety and outcomes.

Our study had several limitations. First, our data were collected at three teaching hospitals in urban areas of South Korea. Therefore, external validation for other areas should be conducted to generalize the developed prediction model. Second, we used retrospective analysis of electronically collected prehospital and hospital data. There might be various information loss and missing data. We treated missing status as a separate category for our analysis;[44] however, there could be different reasons for missing data. Third, there is a possibility that the prediction model was overfitted or underfitted. The use of large number of predictors also can contribute to overfitting. To minimize this issue, we rigorously searched hyperparameters and carefully chose hyperparameters according to the performance in

independent validation cohorts. Fourth, we selected our study population using trauma center transport criteria for EMS providers in Korea. Although those criteria are based on the field triage decision scheme which is the most widely used prehospital trauma triage protocol,[6] extrapolation to another EMS setting or general trauma patients would be limited. Fifth, Abbreviated Injury Scale (AIS) codes were not used to identify our study outcome because of a lack of information. To compensate for this limitation, we further identified TBI-I, TBI-ND, and TBI-D patients to consider severity. However, different definitions of clinical severity, including ICU admission or emergency operation, might be possible. Lastly, this study was performed in an intermediate-service-level EMS system. The generalization of our study findings to different EMS settings should be made with caution.

In conclusion, we presented data on TBI among patients with severe trauma assessed by EMS providers, and our results inform the development of prediction models for the diagnosis and prognosis of TBI in our population. We used various information that can be obtained in prehospital settings and showed acceptable outcome performance. The consistent importance of consciousness-status-related variables emphasizes the importance of assessment and monitoring of consciousness status in prehospital areas. Although prospective, and implementation studies are needed for TBI prediction in prehospital areas, our study outlined a novel method for the precise assessment of EMS providers using a machine-learning-based prediction model. Further collection of various types of patient-related data would contribute to the enhanced performance of the clinical prediction model in prehospital settings.

Author Contribution Statement

YHC and JH Park designed and developed the study, analysed and interpreted the data, and drafted the initial manuscript. KJH, YSR, KJS and SDS were involved in the acquisition of data, the development of the research question and assisted with analysis and interpretation of data. All authors revised the drafts for intellectual content and edited the manuscript. All authors reviewed and approved the final draft.

Funding

This study was supported by grant No. '04-2019-0680' from the Seoul National University Hospital Research Fund.

Competing Interests

There are no conflicts of interest for all authors in this study.

Patients consent

Not required

Data availability statement

No data are available. We do not have ethics approval to share data.

Ethical statements

This study complied with the Declaration of Helsinki, and its protocol was approved by the Institutional Review Board of the Seoul National University Hospital with a waiver of informed consent (IRB No: E-2006-004-1128).

References

1. Hsia RY, Markowitz AJ, Lin F, et al. Ten-year trends in traumatic brain injury: a retrospective cohort study of California emergency department and hospital revisits and readmissions. *BMJ Open* 2018;8(12):e022297. doi: 10.1136/bmjopen-2018-022297 [published Online First: 2018/12/16]

2. Finfer SR, Cohen J. Severe traumatic brain injury. *Resuscitation* 2001;48(1):77-90. doi: 10.1016/s0300-9572(00)00321-x [published Online First: 2001/02/13]

3. Spaite DW, Bobrow BJ, Keim SM, et al. Association of Statewide Implementation of the Prehospital Traumatic Brain Injury Treatment Guidelines With Patient Survival Following Traumatic Brain Injury: The Excellence in Prehospital Injury Care (EPIC) Study. *JAMA Surg* 2019;154(7):e191152. doi: 10.1001/jamasurg.2019.1152 [published Online First: 2019/05/09]

4. McHugh GS, Engel DC, Butcher I, et al. Prognostic value of secondary insults in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;24(2):287-93. doi: 10.1089/neu.2006.0031 [published Online First: 2007/03/23]

5. Pelieu I, Kull C, Walder B. Prehospital and Emergency Care in Adult Patients with Acute Traumatic Brain Injury. *Med Sci (Basel)* 2019;7(1) doi: 10.3390/medsci7010012 [published Online First: 2019/01/24]

6. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients:

recommendations of the National Expert Panel on Field Triage, 2011. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 2012;61(1):1-20.

7. Ter Avest E, Taylor S, Wilson M, et al. Prehospital clinical signs are a poor predictor of raised intracranial pressure following traumatic brain injury. *Emerg Med J* 2021;38(1):21-26. doi: 10.1136/emermed-2020-209635 [published Online First: 2020/09/20]

8. Collaborators MCT, Perel P, Arango M, et al. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;336(7641):425-9. doi: 10.1136/bmj.39461.643438.25 [published Online First: 2008/02/14]

9. Steyerberg EW, Mushkudiani N, Perel P, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med* 2008;5(8):e165; discussion e65. doi: 10.1371/journal.pmed.0050165 [published Online First: 2008/08/08]

10. Gozt AK, Hellewell SC, Thorne J, et al. Predicting outcome following mild traumatic brain injury: protocol for the longitudinal, prospective, observational Concussion Recovery (CREST) cohort study. *BMJ Open* 2021;11(5):e046460. doi: 10.1136/bmjopen-2020-046460 [published Online First: 2021/05/15]

11. Huth SF, Slater A, Waak M, et al. Predicting Neurological Recovery after Traumatic Brain Injury in Children: A Systematic Review of Prognostic Models. *J Neurotrauma* 2020;37(20):2141-49. doi: 10.1089/neu.2020.7158 [published Online First: 2020/05/29]

12. Perel P, Edwards P, Wentz R, et al. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* 2006;6:38. doi: 10.1186/1472-6947-6-38 [published Online First: 2006/11/16]

13. Miller PR, Chang MC, Hoth JJ, et al. Predicting Mortality and Independence at Discharge in the Aging Traumatic Brain Injury Population Using Data Available at Admission. *J Am Coll Surg* 2017;224(4):680-85. doi: 10.1016/j.jamcollsurg.2016.12.053 [published Online First: 2017/03/07]

14. Abujaber A, Fadlalla A, Gammoh D, et al. Prediction of in-hospital mortality in patients with post traumatic brain injury using National Trauma Registry and Machine Learning Approach. *Scand J Trauma Resusc Emerg Med* 2020;28(1):44. doi: 10.1186/s13049-020-00738-5 [published Online First: 2020/05/29]

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. doi: 10.1016/j.jclinepi.2020.03.005 [published Online First: 2020/03/24]

16. Roozenbeek B, Lingsma HF, Lecky FE, et al. Prediction of outcome after moderate and severe traumatic brain injury: external validation of the International Mission on Prognosis and Analysis of Clinical Trials (IMPACT) and Corticoid Randomisation After Significant Head injury (CRASH) prognostic models. *Crit Care Med* 2012;40(5):1609-17. doi: 10.1097/CCM.0b013e31824519ce [published Online First: 2012/04/19]

17. Wasserman EB, Shah MN, Jones CM, et al. Identification of a neurologic scale that optimizes EMS detection of older adult traumatic brain injury patients who require transport to a trauma center. *Prehosp Emerg Care* 2015;19(2):202-12. doi: 10.3109/10903127.2014.959225 [published Online First: 2014/10/08]

18. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Ann Intern Med*

2015;162(10):735-6. doi: 10.7326/L15-5093-2 [published Online First: 2015/05/20]

19. Sasser SM, Hunt RC, Sullivent EE, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage. 2009

20. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *MMWR Recomm Rep* 2012;61(RR-1):1-20. [published Online First: 2012/01/13]

21. Andelic N, Anke A, Skandsen T, et al. Incidence of hospital-admitted severe traumatic brain injury and in-hospital fatality in Norway: a national cohort study. *Neuroepidemiology* 2012;38(4):259-67. doi: 10.1159/000338032 [published Online First: 2012/06/09]

22. Ro YS, Shin SD, Holmes JF, et al. Comparison of clinical performance of cranial computed tomography rules in patients with minor head injury: a multicenter prospective study. *Acad Emerg Med* 2011;18(6):597-604. doi: 10.1111/j.1553-2712.2011.01094.x [published Online First: 2011/06/17]

23. Chan V, Thurairajah P, Colantonio A. Defining pediatric traumatic brain injury using International Classification of Diseases Version 10 Codes: a systematic review. *BMC Neurol* 2015;15:7. doi: 10.1186/s12883-015-0259-7 [published Online First: 2015/02/05]

24. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67(2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x

25. Hearst MA, Dumais ST, Osuna E, et al. Support vector machines. *IEEE Intelligent Systems and their Applications* 1998;13(4):18-28. doi: 10.1109/5254.708428

26. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: Association for Computing Machinery, 2016:785–94.

27. Menard S. Applied logistic regression analysis: Sage 2002. pp. 162.

28. Oostema JA, Konen J, Chassee T, et al. Clinical predictors of accurate prehospital stroke recognition. *Stroke* 2015;46(6):1513-7. doi: 10.1161/STROKEAHA.115.008650 [published Online First: 2015/04/30]

29. Swan PY, Nighswonger B, Boswell GL, et al. Factors associated with false-positive emergency medical services triage for percutaneous coronary intervention. *West J Emerg Med* 2009;10(4):208-12. [published Online First: 2010/01/05]

30. Park JH, Shin SD, Song KJ, et al. Prediction of good neurological recovery after out-of-hospital cardiac arrest: A machine learning analysis. *Resuscitation* 2019;142:127-35. doi: 10.1016/j.resuscitation.2019.07.020 [published Online First: 2019/07/31]

31. Esposito TJ, Offner PJ, Jurkovich GJ, et al. Do prehospital trauma center triage criteria identify major trauma victims? *Arch Surg* 1995;130(2):171-6. doi: 10.1001/archsurg.1995.01430020061010 [published Online First: 1995/02/01]

32. Ocak G, Sturms LM, Hoogeveen JM, et al. Prehospital identification of major trauma patients. *Langenbecks Arch Surg* 2009;394(2):285-92. doi: 10.1007/s00423-008-0340-4 [published Online First: 2008/06/27]

33. Fries GR, McCalla G, Levitt MA, et al. A prospective comparison of paramedic judgment and the trauma triage rule in the prehospital setting. *Ann Emerg Med* 1994;24(5):885-9. doi: 10.1016/s0196-0644(94)70207-1 [published Online First: 1994/11/01]

34. Branas CC, MacKenzie EJ, Williams JC, et al. Access to trauma centers in the United States. *JAMA* 2005;293(21):2626-33. doi: 10.1001/jama.293.21.2626 [published Online First: 2005/06/02]

35. Whiting MD, Dengler BA, Rodriguez CL, et al. Prehospital Detection of Life-Threatening Intracranial Pathology: An Unmet Need for Severe TBI in Austere, Rural, and Remote Areas. *Front Neurol* 2020;11:599268. doi: 10.3389/fneur.2020.599268 [published Online First: 2020/11/17]

36. Wojcik SM. Predicting mild traumatic brain injury patients at risk of persistent symptoms in the Emergency Department. *Brain Inj* 2014;28(4):422-30. doi: 10.3109/02699052.2014.884241 [published Online First: 2014/02/26]

37. Management of Concussion/m TBIWG. VA/DoD Clinical Practice Guideline for Management of Concussion/Mild Traumatic Brain Injury. *J Rehabil Res Dev* 2009;46(6):CP1-68. [published Online First: 2010/01/30]

38. Grote S, Bocker W, Mutschler W, et al. Diagnostic value of the Glasgow Coma Scale for traumatic brain injury in 18,002 patients with severe multiple injuries. *J Neurotrauma* 2011;28(4):527-34. doi: 10.1089/neu.2010.1433 [published Online First: 2011/01/27]

39. Badjatia N, Carney N, Crocco TJ, et al. Guidelines for prehospital management of traumatic brain injury 2nd edition. *Prehosp Emerg Care* 2008;12 Suppl 1:S1-52. doi: 10.1080/10903120701732052 [published Online First: 2008/09/06]

40. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019;110:12-22. doi: 10.1016/j.jclinepi.2019.02.004 [published Online First: 2019/02/15]

41. Nakada TA, Masunaga N, Nakao S, et al. Development of a prehospital vital signs chart sharing system. *Am J Emerg Med* 2016;34(1):88-92. doi: 10.1016/j.ajem.2015.09.048 [published Online First: 2015/10/29]

42. Kim Y, Groombridge C, Romero L, et al. Decision Support Capabilities of Telemedicine

in Emergency Prehospital Care: Systematic Review. *J Med Internet Res* 2020;22(12):e18959. doi: 10.2196/18959 [published Online First: 2020/12/09]

43. Reisner AT, Khitrov MY, Chen L, et al. Development and validation of a portable platform for deploying decision-support algorithms in prehospital settings. *Appl Clin Inform* 2013;4(3):392-402. doi: 10.4338/ACI-2013-04-RA-0023 [published Online First: 2013/10/25]

44. Maslove DM, Podchiyska T, Lowe HJ. Discretization of continuous features in clinical datasets. *J Am Med Inform Assoc* 2013;20(3):544-53. doi: 10.1136/amiajnl-2012-000929 [published Online First: 2012/10/13]

Figure legends

Figure 1. Population flow. EMS, emergency medical service; OHCA, out-of-hospital cardiac arrest; TBI, traumatic brain injury.

597 Table 1. Key characteristics of the development and test cohorts.

	n (%) or Median (IQR)			
	Total	Development cohort	Test cohort	P
Total	N = 1169	n = 867	n = 302	
Demographics				
Age, years	53 (36–66)	52 (35–66)	56 (40–69)	< 0.01
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Job, unemployed	299 (25.6)	197 (22.7)	102 (33.8)	< 0.01
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	< 0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	< 0.01
Circumstances of trauma				
Location, road/highway	444 (38.0)	326 (37.6)	118 (39.1)	0.65
Season, summer	336 (28.7)	253 (29.2)	83 (27.5)	0.57
Weekday, weekend	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time, 6 p.m. to midnight	361 (30.9)	265 (30.6)	96 (31.8)	0.69
Mechanism of injury, TA	500 (42.8)	375 (43.3)	125 (41.4)	0.57
Chief complaint				
Fracture/abrasion/laceration	302 (25.8)	204 (23.5)	98 (32.5)	< 0.01
EMS vital sign assessment				
SBP, mmHg	130 (109–150)	130 (104–146)	131 (115–150)	< 0.01
DBP, mmHg	80 (70–91)	80 (69–90)	80 (70–92)	0.21
RR, /min	18 (16–20)	18 (16–20)	18 (16–20)	0.33
HR, /min	86 (75–99)	86 (74–99)	86 (76–100)	0.40
SpO2, %	98 (95–99)	98 (95–99)	98 (96–99)	0.67
AVPU scale, Alert	714 (61.1)	504 (58.1)	210 (69.5)	< 0.01
EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	0.95

598 IQR, interquartile range; TA, traffic accident; SBP, systolic blood pressure; DBP, diastolic
599 blood pressure; RR, respiratory rate; AVPU, mental status in alert, verbal, pain, and
600 unresponsive scale; ED, emergency department; TBI, traumatic brain injury.

Table 2. Discrimination and reclassification of prediction models for outcomes on test cohort.

Outcome	Model	AUROC (95% CI)	p[a]	NRI (95% CI)	p[b]	AUPRC
TBI						
	LR	0.770 (0.698, 0.841)	NA	NA	NA	0.492
	XGB	0.809 (0.743, 0.876)	0.04	0.689 (0.427, 0.951)	< 0.01	0.552
	SVM	0.776 (0.708, 0.844)	0.77	0.339 (0.072, 0.607)	0.01	0.479
	RF	0.800 (0.735, 0.865)	0.13	0.308 (0.047, 0.569)	0.02	0.532
	EN	0.799 (0.732, 0.867)	0.06	0.698 (0.441, 0.954)	< 0.01	0.564
TBI-I						
	LR	0.820 (0.751, 0.890)	NA	NA	NA	0.551
	XGB	0.838 (0.775, 0.901)	0.28	0.539 (0.258, 0.821)	< 0.01	0.554
	SVM	0.812 (0.748, 0.875)	0.66	0.729 (0.464, 0.994)	< 0.01	0.469
	RF	0.836 (0.772, 0.899)	0.38	0.333 (0.058, 0.607)	0.02	0.552
	EN	0.844 (0.779, 0.910)	0.15	1.093 (0.845, 1.342)	< 0.01	0.606
TBI-ND						
	LR	0.767 (0.690, 0.844)	NA	NA	NA	0.482
	XGB	0.800 (0.727, 0.873)	0.07	0.605 (0.326, 0.884)	< 0.01	0.496
	SVM	0.778 (0.704, 0.852)	0.56	0.285 (-0.001, 0.572)	0.05	0.477
	RF	0.809 (0.739, 0.880)	0.03	0.194 (-0.059, 0.448)	0.13	0.535
	EN	0.811 (0.741, 0.882)	0.02	0.768 (0.496, 1.039)	< 0.01	0.551
TBI-D						
	LR	0.664 (0.490, 0.838)	NA	NA	NA	0.138
	XGB	0.714 (0.512, 0.917)	0.64	-0.026 (-0.605, 0.553)	0.93	0.094
	SVM	0.814 (0.718, 0.910)	0.09	0.209 (-0.325, 0.742)	0.44	0.140
	RF	0.889 (0.801, 0.976)	< 0.01	-0.204 (-0.742, 0.334)	0.46	0.196
	EN	0.871 (0.764, 0.978)	0.01	0.119 (-0.415, 0.654)	0.66	0.293

[a]Comparing the AUROC and the logistic regression model.

[b]Comparing the NRI and the logistic regression model.

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; NRI, net reclassification index; AUPRC, area under precision-recall curve; TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support vector machine; RF, random forest; EN, elastic net

614 Table 3. Test characteristics of prediction models for outcomes on test cohort.

Outcome	Model	Specificity (95% CI)	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)	Cutoff
TBI						
	LR	47.5 (40.9, 54.0)	80.3 (68.7, 89.1)	29.9 (23.3, 37.3)	89.6 (82.9, 94.3)	0.136
	XGB	72.5 (66.3, 78.1)	80.3 (68.7, 89.1)	44.9 (35.7, 54.3)	92.9 (88.2, 96.2)	0.268
	SVM	64.8 (58.4, 70.9)	80.3 (68.7, 89.1)	39.0 (30.7, 47.7)	92.2 (87.0, 95.8)	0.191
	RF	68.2 (61.9, 74.1)	80.3 (68.7, 89.1)	41.4 (32.8, 50.4)	92.5 (87.6, 96.0)	0.185
	EN	61.0 (54.5, 67.3)	80.3 (68.7, 89.1)	36.6 (28.7, 44.9)	91.7 (86.3, 95.5)	0.205
TBI-I						
	LR	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.164
	XGB	74.0 (68.0, 79.4)	80.4 (67.6, 89.8)	41.3 (31.9, 51.1)	94.3 (90.0, 97.1)	0.143
	SVM	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.172
	RF	76.0 (70.2, 81.2)	80.4 (67.6, 89.8)	43.3 (33.6, 53.3)	94.4 (90.3, 97.2)	0.205
	EN	81.3 (75.9, 86.0)	80.4 (67.6, 89.8)	49.5 (38.8, 60.1)	94.8 (90.9, 97.4)	0.204
TBI-ND						
	LR	46.1 (39.8, 52.6)	80.7 (68.1, 90.0)	25.8 (19.6, 32.9)	91.1 (84.7, 95.5)	0.090
	XGB	66.5 (60.2, 72.4)	80.7 (68.1, 90.0)	35.9 (27.7, 44.9)	93.7 (89.0, 96.8)	0.242
	SVM	59.2 (52.7, 65.4)	80.7 (68.1, 90.0)	31.5 (24.1, 39.7)	92.9 (87.7, 96.4)	0.147
	RF	60.4 (54.0, 66.6)	80.7 (68.1, 90.0)	32.2 (24.6, 40.5)	93.1 (88.0, 96.5)	0.138
	EN	74.3 (68.3, 79.6)	80.7 (68.1, 90.0)	42.2 (32.8, 52.0)	94.3 (90.0, 97.1)	0.201
TBI-D						
	LR	42.6 (36.9, 48.5)	81.8 (48.2, 97.7)	5.1 (2.4, 9.5)	98.4 (94.4, 99.8)	0.005
	XGB	57.7 (51.8, 63.5)	81.8 (48.2, 97.7)	6.8 (3.2, 12.5)	98.8 (95.8, 99.9)	0.002
	SVM	74.2 (68.8, 79.2)	81.8 (48.2, 97.7)	10.7 (5.0, 19.4)	99.1 (96.7, 99.9)	0.039
	RF	74.9 (69.5, 79.8)	81.8 (48.2, 97.7)	11.0 (5.1, 19.8)	99.1 (96.8, 99.9)	0.005
	EN	79.0 (73.9, 83.6)	81.8 (48.2, 97.7)	12.9 (6.1, 23.0)	99.1 (96.9, 99.9)	0.033

615 TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-
616 ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with
617 death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support
618 vector machine; RF, random forest; EN, elastic net.
619

Table 4. Top 5 important variables for outcomes in descending order using model specific metrics

Outcome	Rank	LR	XGB	RF	EN
TBI					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Motor, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	4	Light reflex	Other mechanism	Light reflex	GCS, Eye, 1
	5	GCS, Motor, 1	GCS, Verbal, 2	GCS, Motor, 1	GCS, Verbal, 1
TBI-I					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	GCS, Eye, 1
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	Loss of consciousness
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	Light reflex	GCS, Verbal, 1
	5	GCS, Motor, 1	Other mechanism	GCS, Motor, 1	Light reflex
TBI-ND					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	GCS, Verbal, 2	GCS, Verbal, 1
	5	GCS, Motor, 1	GCS, Motor, 1	GCS, Motor, 4	Light reflex
TBI-D					
	1	Loss of consciousness	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	2	GCS, Verbal, 1	Oxygen saturation<96%	Light reflex	GCS, Verbal, 1
	3	GCS, Eye, 1	Fall mechanism	Loss of consciousness	Loss of consciousness
	4	Light reflex	Afternoon	GCS, Eye, 1	Age over 80
	5	GCS, Motor, 1	Light reflex	GCS, Motor, 1	HR 87-99

TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression; XGB, extreme gradient boosting; RF, random forest; EN, elastic net; GCS, Glasgow coma scale; HR, heart rate.

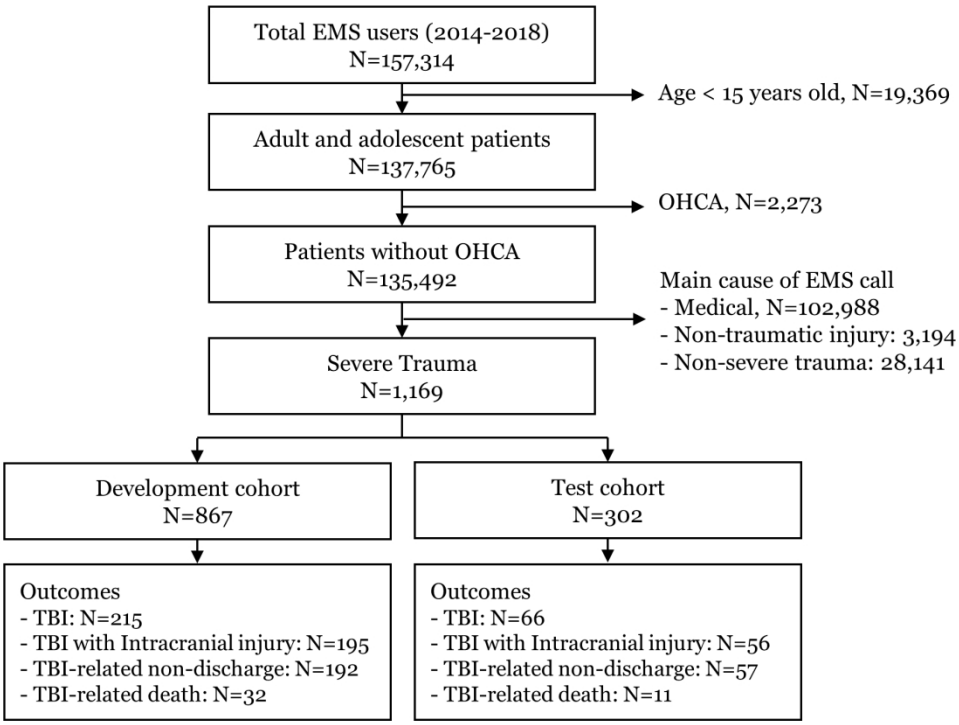


Figure 1

165x119mm (300 x 300 DPI)

Supplementary Table 1. List of analyzed variables.

Variables	Descriptions	Type of raw data	Category	Preprocessing
Gender	Sex of the patients	Binary	Male, Female	
Age	Age of patients	Continuous	15-39 years, 40-59 years, 60-79 years, and 80- years	Discretization and one hot encoding
Job	Job of patients	Categorical	Unemployed, Student/Housewife; Office/Commercial/Service workers; Industrial/Agricultural/Fishery/Miner worker; Others	One hot encoding Missing data were classified into others
Diabetes	History of diabetes mellitus	Binary	Yes, No	Missing data were classified into no
Hypertension	History of hypertension	Binary	Yes, No	Missing data were classified into no
Location of injury	Location of injury	Categorical	home/residential area/medical facility/school/gym; Road/highway; Off-road traffic area; Others	One hot encoding Missing data were classified into others
Season	Season when injury occurred	Categorical	Spring, Summer, Fall, Winter	One hot encoding
Weekend	Whether Injury occurred on weekday or weekend	Binary	Weekday, Weekend	
Daytime	When injury was occurred	Categorical	Night (Midnight to 5AM), Morning (6AM to 11AM), Afternoon (Midday to 5PM), Evening (6PM to 11PM)	One hot encoding Missing time were imputed using EMS call time
Mechanism of injury	Mechanism of injury	Categorical	Slip down, Fall down, Traffic accident, Other	One hot encoding Missing data were classified into others
Glasgow coma scale eye	Eye element of Glasgow coma scale	Categorical	1;2;3;4;Unknown	One hot encoding
Glasgow coma scale Verbal	Verbal element of Glasgow coma scale	Categorical	1;2;3;4;5;Unknown	One hot encoding
Glasgow coma scale Motor	Motor element of Glasgow coma scale	Categorical	1;2;3;4;5;6;Unknown	One hot encoding
Light Reflex any Abnormal	Any abnormality of light reflex on any side	Categorical	No, Yes, Unknown	One hot encoding Missing data were classified into unknown

Systolic blood pressure	Systolic blood pressure	Continuous	-107 mmHg, 108-130 mmHg, 131-145 mmHg, 146- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Diastolic blood pressure	Diastolic blood pressure	Continuous	-69 mmHg, 70-80 mmHg, 81-91 mmHg, 92- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Heart rate	Heart rate	Continuous	-74/min, 75-86/min, 87-99/min, 100-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Respiratory rate	Respiratory rate	Continuous	-16/min, 17-18/min, 19-20/min, 21-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Oxygen saturation	Oxygen saturation	Continuous	-95%, 96-98%, 99%, 100%, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Body temperature	Body temperature	Continuous	-36℃, 36.1-36.3℃, 36.4-36.8℃, 36.9-℃, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Chest pain or abdominal pain	Symptom of chest pain or abdominal pain	Binary	Yes, No	

Fracture, abrasion, or laceration	Symptom of fracture, abrasion, or laceration	Binary	Yes, No	
Loss of consciousness	Symptom of loss of consciousness (whether patients had loss of consciousness between injury and EMS provider's assessment)	Binary	Yes, No	
Dyspnea	Symptom of dyspnea	Binary	Yes, No	
Nose bleeding	Symptom of nose bleeding	Binary	Yes, No	
Nausea or vomiting	Symptom of nausea or vomiting	Binary	Yes, No	
Headache, paralysis or dizziness	Symptom of headache, paralysis or dizziness	Binary	Yes, No	

Supplementary Table 2. Demographic characteristics of development and test cohorts

	N (%) or Median (IQR)			
Characteristics	Total	Development	Test	P-value
Total	1169	867	302	
Demographics				
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Age, years	53 (36-66)	52 (35-66)	56 (40-69)	<0.01
Job of patients				<0.01
Unemployed	299 (25.6)	197 (22.7)	102 (33.8)	
Student/Housewife	161 (13.8)	129 (14.9)	32 (10.6)	
Office/Commercial/Service worker	283 (24.2)	176 (20.3)	107 (35.4)	
Industrial/Agricultural/Fishery/Minery worker	36 (3.1)	25 (2.9)	11 (3.6)	
Others	390 (33.4)	340 (39.2)	50 (16.6)	
Past medical history				
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	<0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	<0.01
Circumstances of Trauma				
Location of trauma				0.52
Residential/Nursing/Education/Exercise facility	303 (25.9)	218 (25.1)	85 (28.1)	
Road/Highway	444 (38.0)	326 (37.6)	118 (39.1)	
Off-road traffic area	181 (15.5)	140 (16.1)	41 (13.6)	
Others	241 (20.6)	183 (21.1)	58 (19.2)	
Season of trauma				<0.01
Spring	249 (21.3)	167 (19.3)	82 (27.2)	
Summer	336 (28.7)	253 (29.2)	83 (27.5)	
Fall	304 (26.0)	242 (27.9)	62 (20.5)	
Winter	280 (24.0)	205 (23.6)	75 (24.8)	
Weekday	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time of trauma				0.83
6A-MD	281 (24.0)	206 (23.8)	75 (24.8)	
MD-6P	266 (22.8)	203 (23.4)	63 (20.9)	
6P-MN	361 (30.9)	265 (30.6)	96 (31.8)	
MN-6A	261 (22.3)	193 (22.3)	68 (22.5)	
Mechanism of Trauma				0.60
Traffic accident	500 (42.8)	375 (43.3)	125 (41.4)	
Slip down	325 (27.8)	232 (26.8)	93 (30.8)	
Fall down	171 (14.6)	129 (14.9)	42 (13.9)	
Others	173 (14.8)	131 (15.1)	42 (13.9)	
Chief complaint				
Altered mentality	279 (23.9)	223 (25.7)	56 (18.5)	0.01
Facture/Abrasion/Laceration	302 (25.8)	204 (23.5)	98 (32.5)	<0.01
Chest/Abdominal pain	47 (4.0)	31 (3.6)	16 (5.3)	0.19
Dyspnea	25 (2.1)	20 (2.3)	5 (1.7)	0.50

Epistaxis	44 (3.8)	30 (3.5)	14 (4.6)	0.36
Headache/Paralysis/Dizziness/Vertigo	95 (8.1)	64 (7.4)	31 (10.3)	0.11
Nausea/Vomiting	32 (2.7)	20 (2.3)	12 (4.0)	0.13
EMS Vital sign assessment				
SBP, mmHg	130 (109-150)	130 (104-146)	131 (115-150)	<0.01
Missing	65 (5.6)	56 (6.5)	9 (3.0)	0.02
DBP, mmHg	80 (70-91)	80 (69-90)	80 (70-92)	<0.01
Missing	75 (6.4)	65 (7.5)	10 (3.3)	0.01
HR, /min	86 (75-99)	86 (74-99)	86 (76-100)	<0.01
Missing	31 (2.7)	28 (3.2)	3 (1.0)	0.04
RR, /min	18 (16-20)	18 (16-20)	18 (16-20)	<0.01
Missing	36 (3.1)	33 (3.8)	3 (1.0)	0.01
SpO2, %	98 (95-99)	98 (95-99)	98 (96-99)	<0.01
Missing	38 (3.3)	33 (3.8)	5 (1.7)	0.07
Temperature, ℃	36.5 (36-36.8)	36.5 (36-36.8)	36.5 (36-36.7)	<0.01
Missing	94 (8.0)	65 (7.5)	29 (9.6)	0.25
AVPU scale				<0.01
Alert	714 (61.1)	504 (58.1)	210 (69.5)	
Verbal	168 (14.4)	136 (15.7)	32 (10.6)	
Pain	199 (17.0)	158 (18.2)	41 (13.6)	
Unresponsive	88 (7.5)	69 (8.0)	19 (6.3)	
Abnormal light reflex	165 (14.1)	132 (15.2)	33 (10.9)	<0.01
Missing	66 (5.6)	57 (6.6)	9 (3.0)	
GCS scale component				
Glasgow coma scale eye				<0.01
4	558 (47.7)	380 (43.8)	178 (58.9)	
3	128 (10.9)	109 (12.6)	19 (6.3)	
2	110 (9.4)	82 (9.5)	28 (9.3)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	
Glasgow coma scale Verbal				0.01
5	520 (44.5)	359 (41.4)	161 (53.3)	
4	118 (10.1)	88 (10.1)	30 (9.9)	
3	25 (2.1)	19 (2.2)	6 (2.0)	
2	132 (11.3)	105 (12.1)	27 (8.9)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	200 (17.1)	155 (17.9)	45 (14.9)	
Glasgow coma scale Motor				<0.01
6	499 (42.7)	333 (38.4)	166 (55.0)	
5	124 (10.6)	103 (11.9)	21 (7.0)	
4	158 (13.5)	123 (14.2)	35 (11.6)	
3	47 (4.0)	39 (4.5)	8 (2.6)	
2	17 (1.5)	15 (1.7)	2 (0.7)	
1	125 (10.7)	99 (11.4)	26 (8.6)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	

EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Advanced airway management	4 (0.3)	2 (0.2)	2 (0.7)	0.28
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
Field triage decision scheme criteria*				
Physiological criteria				
SBP<90 mmHg	58 (5.0)	42 (4.8)	16 (5.3)	0.75
RR<10 or >29 /min	11 (0.9)	11 (1.3)	0 (0.0)	0.08
Non-Alert	429 (36.7)	343 (39.6)	86 (28.5)	<0.01
Anatomic criteria				
All penetrating injuries to head, neck, torso and extremities proximal to elbow or knee	34 (2.9)	23 (2.7)	11 (3.6)	0.38
Chest wall instability or deformity	4 (0.3)	4 (0.5)	0 (0.0)	0.58
Two or more proximal long bone fractures	19 (1.6)	13 (1.5)	6 (2.0)	0.60
Crush, degloved, mangled or pulseless extremity	15 (1.3)	13 (1.5)	2 (0.7)	0.38
Amputation proximal to wrist or ankle	9 (0.8)	9 (1.0)	0 (0.0)	0.12
Pelvic fractures	8 (0.7)	6 (0.7)	2 (0.7)	>0.95
Open or depressed skull fracture	17 (1.5)	9 (1.0)	8 (2.6)	0.05
Paralysis	21 (1.8)	11 (1.3)	10 (3.3)	0.02
Mechanism of injury criteria				
Fall > 6 meter	113 (9.7)	84 (9.7)	29 (9.6)	>0.95
High-risk auto crash	96 (8.2)	73 (8.4)	23 (7.6)	0.66
Auto vs pedestrian/bicyclist thrown, run over, or with significant (>30km/h) impact	119 (10.2)	83 (9.6)	36 (11.9)	0.25
Motorcycle crash > 30 km/hour	105 (9.0)	70 (8.1)	35 (11.6)	0.07
ED disposition				0.11
Discharge	320 (27.4)	241 (27.8)	79 (26.2)	
Transfer	444 (38.0)	316 (36.4)	128 (42.4)	
Admitted	366 (31.3)	276 (31.8)	90 (29.8)	
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	>0.95

*EMS providers check specific criteria orderly from physiologic, anatomical, and mechanism of injury. If the preceding criteria are satisfied, the information of the latter criteria is not collected.

IQR, interquartile range; SBP, systolic blood pressure; RR, respiratory rate; ED, emergency department; TBI, traumatic brain injury.

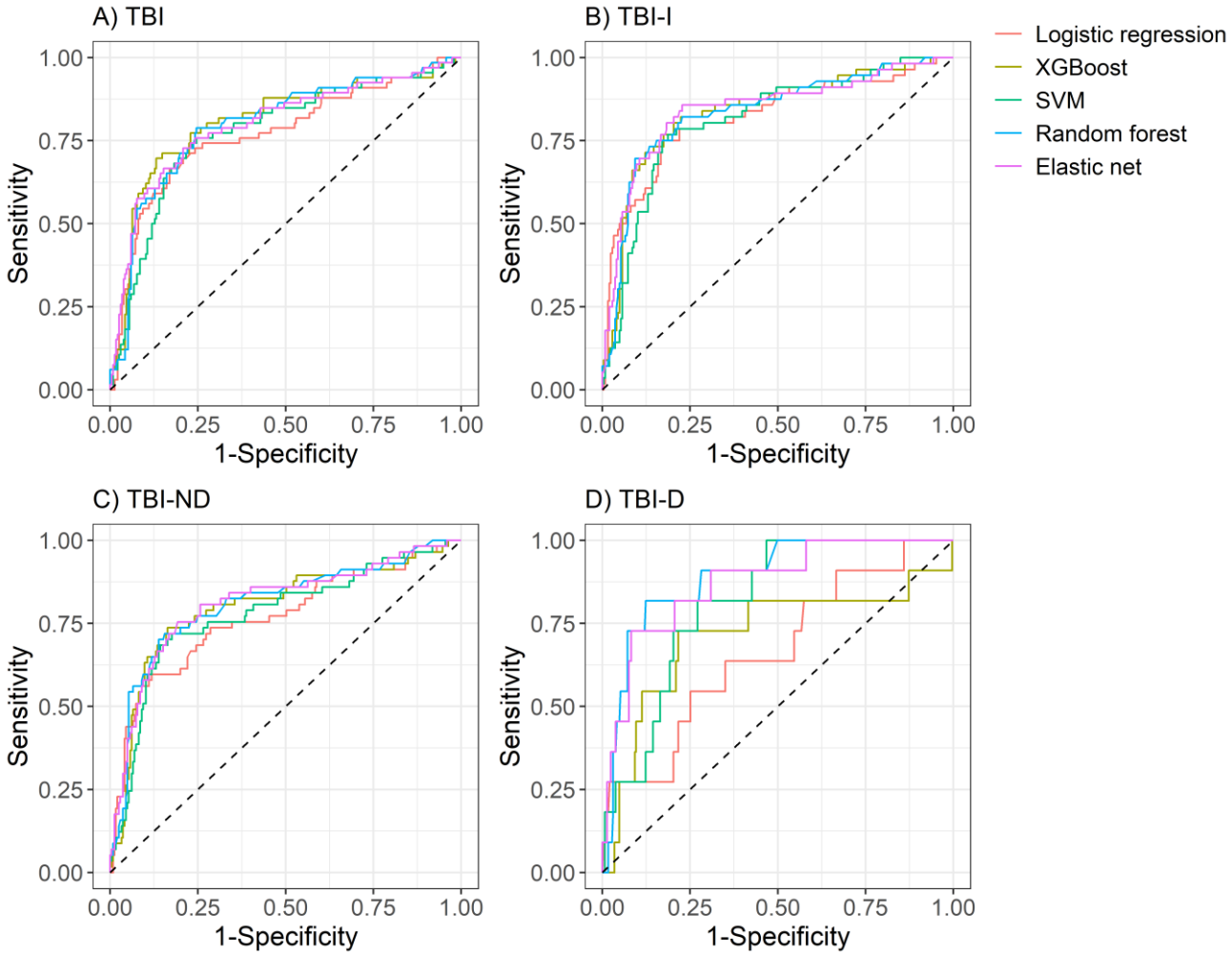
Supplementary Table 3. Hyperparameters of the final prediction models*

Model	Outcome	Hyperparameters
Elastic net	TBI	alpha: 0.325, lambda: 0.07506346
	TBI-I	alpha: 0.325, lambda: 0.07506346
	TBI-ND	alpha: 0.325, lambda: 0.07017153
	TBI-D	alpha: 0.325, lambda: 0.01565599
Random forest	TBI	ntree:500, mtry: 18
	TBI-I	ntree:500, mtry: 18
	TBI-ND	ntree:500, mtry: 18
	TBI-D	ntree:500, mtry: 15
Support vector machine	TBI	sigma: 0.008047; C: 4
	TBI-I	sigma: 0.008047; C: 4
	TBI-ND	sigma: 0.008047; C: 4
	TBI-D	sigma: 0.008047; C: 4
Extreme gradient boosting	TBI	nrounds: 299; max_depth: 1; eta: 0.4807096; gamma: 2.336623; colsample_bytree: 0.3657893; min_child_weight: 8; subsample: 0.8182623
	TBI-I	nrounds: 299; max_depth: 1; eta: 0.4807096; gamma: 2.336623; colsample_bytree: 0.3657893; min_child_weight: 8; subsample: 0.8182623
	TBI-ND	nrounds: 301; max_depth: 1; eta: 0.02154674; gamma: 4.696105; colsample_bytree: 0.590754; min_child_weight: 1; subsample: 0.5070866
	TBI-D	nrounds: 50; max_depth: 0.3; eta: 0.3; gamma: 0; colsample_bytree: 0.8; min_child_weight: 1; subsample: 0.5510204

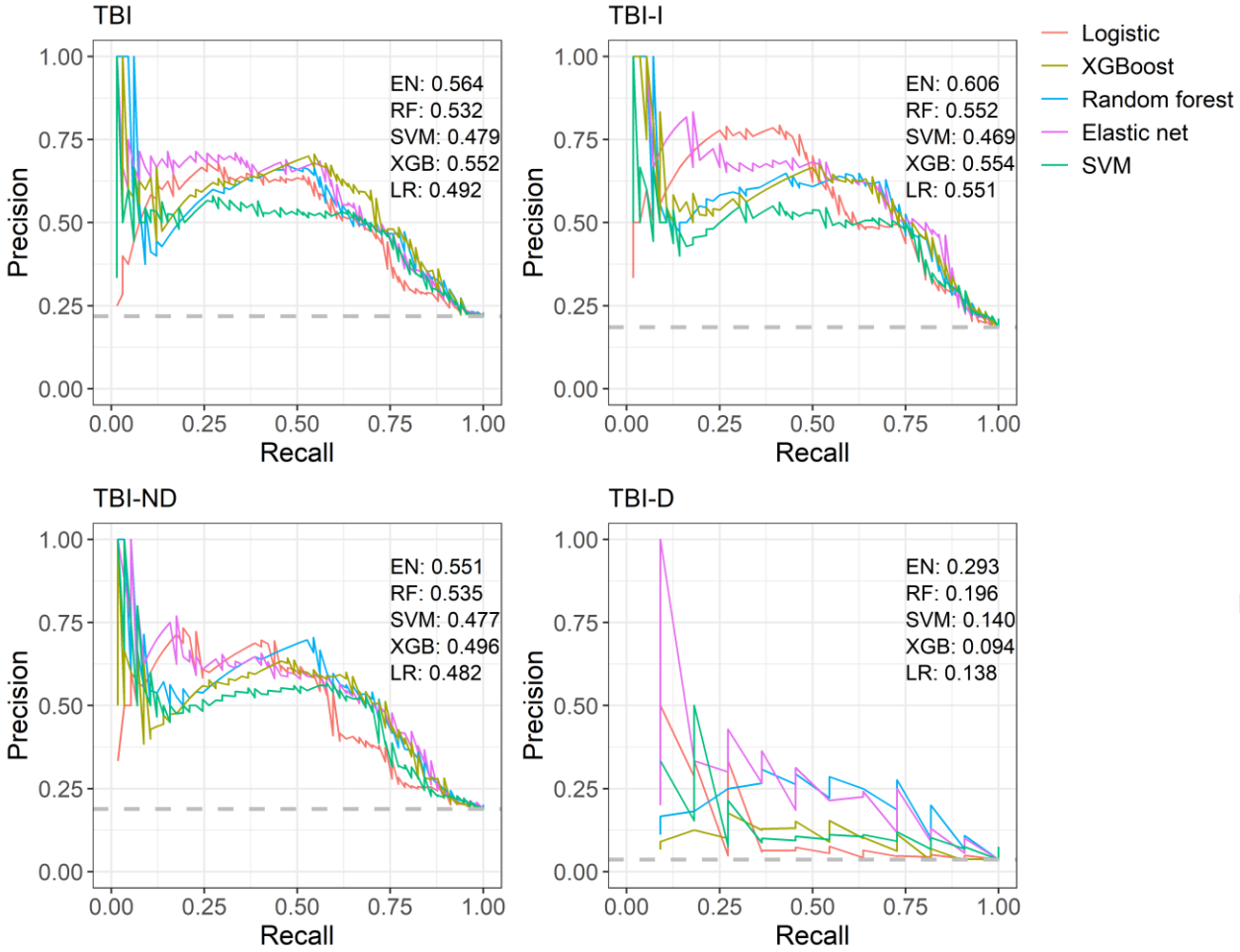
*Aside from the hyperparameters mentioned, all other hyperparameters are used as the default value.

TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death.

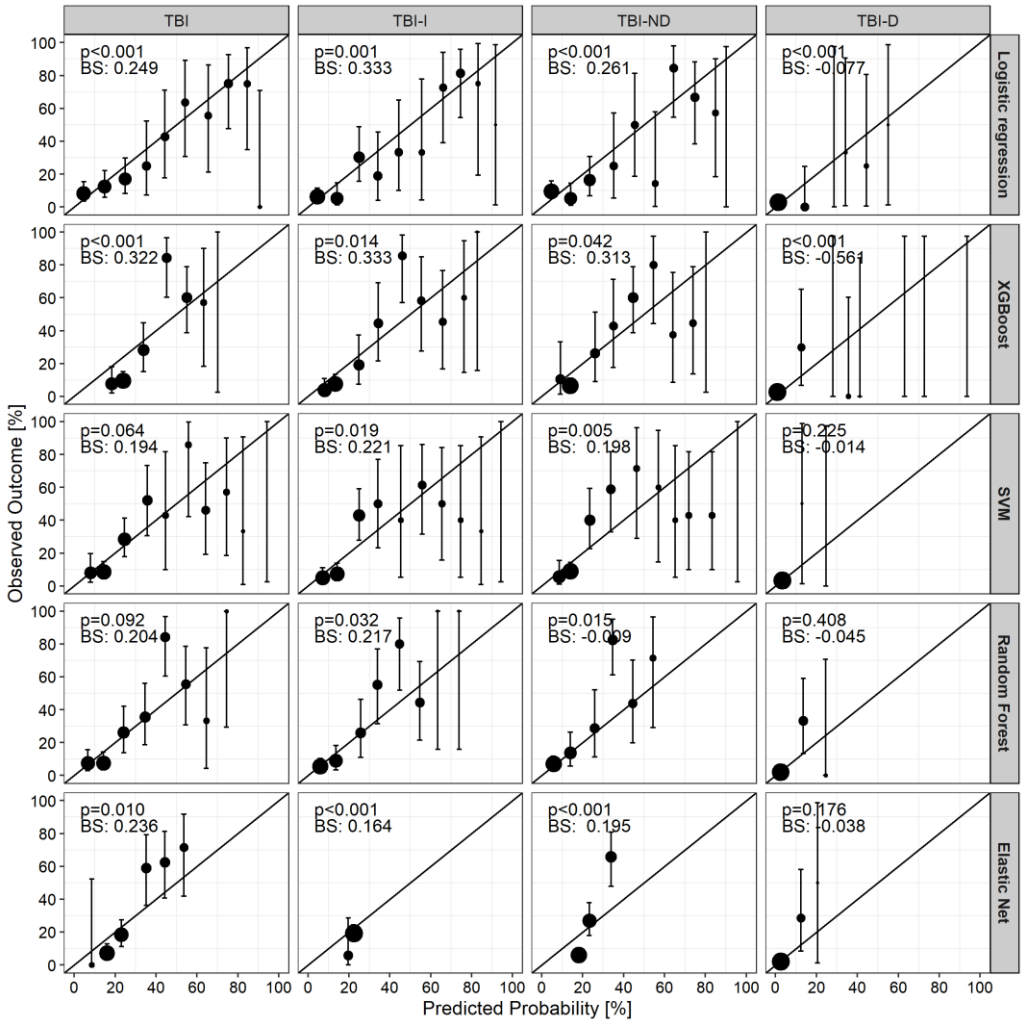
Supplementary Figure 1. Receiver operating characteristics of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death.



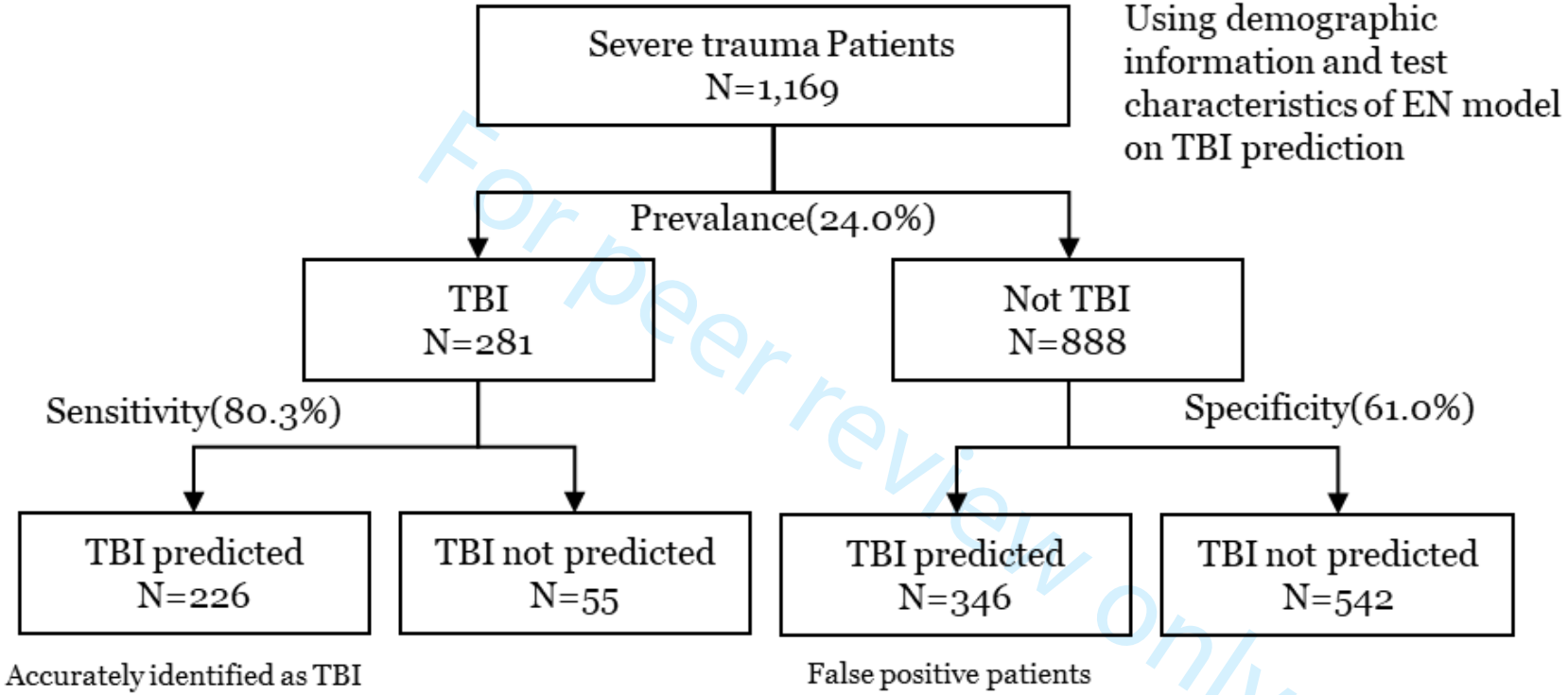
Supplementary Figure 2. Precision-recall curve of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; LR, logistic regression analysis; XGB, extreme gradient boosting; RF, random forest, EN, elastic net.



Supplementary Figure 3. Calibration plot of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; p, p-value of Hosmer-Lemeshow test; BS, scaled Brier score.



Supplementary Table 4. Example of calculating false-positive patients for accurately identified patients. TBI, traumatic brain injury; EN, elastic net.



False-positive patients for every 10 patients that are accurately identified as TBI :
346/226 x 10 = 15.3, rounded up to 16 patients

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	4
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	7
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	8
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8-9
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	9
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8-9
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	N/A
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	10-11
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	11
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	8	D;V	Explain how the study size was arrived at.	14
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	11
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	11
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	11-12
	10c	V	For validation, describe how the predictions were calculated.	12-13
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	12-13
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	12
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	14
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	14
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	14
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
	15b	D	Explain how to the use the prediction model.	14-15
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	14-15
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	19-20
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	16-17
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	16
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	18-19
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Suppl
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	20

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

BMJ Open

Development and validation of a prehospital-stage prediction tool for traumatic brain injury: a multicentre retrospective cohort study in Korea

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-055918.R2
Article Type:	Original research
Date Submitted by the Author:	16-Dec-2021
Complete List of Authors:	Choi, Yeong Ho; Seoul National University Hospital, Emergency Department; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Park, Jeong Ho; Seoul National University Hospital, Emergency Department; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Hong, Ki Jeong; Seoul National University Hospital, Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Ro, Young Sun; Seoul National University Hospital, Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Song, Kyoung Jun; Seoul Metropolitan Boramae Hospital, Department of Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services Shin, Sang Do; Seoul National University Hospital, Department of Emergency Medicine; Seoul National University Hospital Biomedical Research Institute, Laboratory of Emergency Medical Services
Primary Subject Heading:	Emergency medicine
Secondary Subject Heading:	Emergency medicine, Health informatics
Keywords:	ACCIDENT & EMERGENCY MEDICINE, Neurological injury < NEUROLOGY, Trauma management < ORTHOPAEDIC & TRAUMA SURGERY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title page

1. Title

Development and validation of a prehospital-stage prediction tool for traumatic brain injury: a multicentre retrospective cohort study in Korea

2. Authors

Yeong Ho Choi, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: d2uk87@gmail.com

Jeong Ho Park, MD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital Biomedical Research Institute, Seoul, Korea

E-mail: timthe@gmail.com

Ki Jeong Hong, MD, PhD

Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea

Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: emkjhong@gmail.com

Young Sun Ro, MD, DrPH
Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: ro.youngsun@gmail.com

Kyoung Jun Song, MD, PhD
Department of Emergency Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: skciva@gmail.com

Sang Do Shin, MD, PhD
Department of Emergency Medicine, Seoul National University College of Medicine and Hospital, Seoul, Republic of Korea
Laboratory of Emergency Medical Services, Seoul National University Hospital
Biomedical Research Institute, Seoul, Korea
E-mail: shinsangdo@gmail.com

3. Address correspondence and requests for reprints: Jeong Ho Park, MD

Address: Seoul National University Hospital, 101 Daehak-Ro, Jongno-Gu, Seoul 03080, Korea

Phone: +82-2-2072-1800

FAX: +82-2-741-7855

E-mail: timthe@gmail.com

Abstract

Objectives: Predicting diagnosis and prognosis of traumatic brain injury (TBI) at the prehospital stage is challenging; however, using comprehensive prehospital information and machine learning may improve the performance of the predictive model. We developed and tested predictive models for TBI that use machine learning algorithms using information that can be obtained in the prehospital stage.

Design: This was a multi-center retrospective study.

Setting and participants: This study was conducted at three tertiary academic emergency departments (EDs) located in an urban area of South Korea. The data from adult patients with severe trauma who were assessed by emergency medical service (EMS) providers and transported to three participating hospitals between 2014 to 2018 were analyzed.

Results: We developed and tested five machine learning algorithms—logistic regression analyses, extreme gradient boosting, support vector machine, random forest, and elastic net (EN)—to predict TBI, TBI with intracranial hemorrhage or injury (TBI-I), TBI with emergency department or admission result of admission or transferred (TBI-ND), and TBI with emergency department or admission result of death (TBI-D). A total of 1,169 patients were included in the final analysis, and the proportions of TBI, TBI-I, TBI-ND, and TBI-D were 24.0%, 21.5%, 21.3%, and 3.7%, respectively. The EN model yielded an AUROC of 0.799 for TBI, 0.844 for TBI-I, 0.811 for TBI-ND, and 0.871 for TBI-D. The EN model also yielded the highest specificity, and significant reclassification improvement. Variables related to loss of consciousness, Glasgow Coma Scale, and light reflex were the three most important variables to predict all outcomes.

Conclusion: Our results inform the diagnosis and prognosis of TBI. Machine learning models resulted in significant performance improvement over that with logistic regression

analyses, and the best performing model was EN.

Keywords: brain injuries; traumatic; outcome; prognosis; machine learning.

Strengths and limitations of this study

- This study presented prehospital factors that could predict traumatic brain injury in trauma patients chosen by model-specific metrics.
- We treated the missing variables as a different category, reflecting prehospital field uncertainties and increasing data utilization.
- The retrospective observational study design could lead to certain types of bias (eg, selection bias, confounding bias).
- External validation for other areas should be conducted to generalize the developed prediction model.

Introduction

Traumatic brain injury (TBI) is a significant health burden worldwide.[1] It is the leading cause of mortality and disability among young individuals.[2] Patients with TBI are vulnerable to hypoxia and hypotension in the early period of their course and these insults are associated with poor outcomes.[3 4] Prehospital assessment and management of patients with TBI is important,[5] as early prediction of TBI and correcting hypoxia and hypotension during the prehospital stage could be beneficial.[3] However, the identification of TBI can often be challenging in the prehospital area.[5] Vulnerable patients, including the elderly or patients who take medications like anti-platelet or anticoagulant drugs, often have TBI owing to low energy insults.[6] Prehospital clinical signs are also reported to have poor sensitivity for raised intracranial pressure following TBI.[7]

Several prediction models to target patients with TBI have been reported.[8-12] However, most incorporated information that is available only in the hospital, such as laboratory results or image findings.[8 9 13] In addition, most previous prediction models focused on the outcomes of patients with TBI,[14-16] not the identification of TBI. Previously, predictors of older adult patients with TBI who required transport to a trauma center were identified. However, this was consensus-based; therefore, there is a lack of clinical data.[17] Accurate prehospital prediction of TBI and its severity could prevent delays to definite care for patients with TBI. Most emergency medical service (EMS) providers collect various information including demographics, past medical history, circumstances of the trauma, and clinical signs including vital signs; but those variables have not been evaluated together as predictors of TBI and its severity. Using a variety of prehospital information, and adapting newly emerging machine learning algorithms for predicting diagnosis, disposition, and outcome of TBI, might improve the accuracy of identification of TBI and its severity.

The aim of this study was to develop and test prediction models for the diagnosis and prognosis of TBI using prehospital information and machine learning algorithms among patients with severe trauma. We hypothesized that incorporating prehospital information could achieve acceptable performance in predicting TBI, and machine learning algorithms could contribute to performance improvement.

Materials and Methods

Study design and settings

This was a multi-center retrospective study conducted at three tertiary academic emergency departments (EDs) located in an urban area (Seoul and Bundang) of South Korea. These EDs received 50,000–90,000 visits annually and are not designated trauma centers. We adhered to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement on reporting predictive models.[18]

The EMS system in South Korea is operated by the National Fire Agency. The EMS level is considered intermediate, as EMS providers can perform bleeding control, spinal motion restriction, immobilization and splintage, advanced airway management, and administer fluid intravenously. As only physicians can declare death in South Korea, EMS providers cannot stop resuscitation and must transport all patients including those in cardiac arrest to the ED. For all EMS transport, EMS providers record an ambulance run-sheet by law. Since 2012, the National Fire Agency adapted the United States Centers for Disease Control and Prevention of the United States field triage decision scheme to evaluate patients with trauma,[19] and they developed an EMS severe trauma in-depth registry. For said patients, EMS providers evaluate whether patients met trauma center transport criteria in the field triage decision scheme. If they did, the in-depth registry should be recorded, and EMS

transport protocol recommends that patients are transferred to a near regional trauma center; but it is not mandatory.

The Ministry of Health and Welfare designated three ED levels according to the resources and functional requirements; level 1 (n = 36) and level 2 (n = 118) EDs have more resources and better facilities for emergency care and must be staffed by emergency physicians 24 hours a day/365 days a year; whereas level 3 EDs (n = 248) can be staffed by general physicians. In accordance with the EMS Act, all EDs participated annually in a nationwide functional performance evaluation program, which was administered by the Ministry of Health and Welfare. The three participating hospitals in this study were all level 1 EDs that can perform acute trauma care for patients with TBI 24 hours a day/365 days a year—including emergency neurosurgical operation and angiographic interventions. The Ministry of Health and Welfare also designated trauma centers in Korea. Total 16 trauma centers were designated as trauma centers in 2018. Among them, 15 were Level I EDs.

Data source

We used an EMS ambulance run-sheet, EMS trauma in-depth registry, and ED administrative database. The EMS database information, including ambulance run-sheet and trauma in-depth registry, was collected electronically by EMS providers using tablets. The EMS record review for each severe trauma has been performed by EMS medical directors of each fire department since 2012. The ED administrative database contains patients' demographic characteristics, route of visit, time of visit, and diagnosis and disposition. We merged the EMS database with the ED administrative database based on patients' arrival time, age, and sex.

Study population

We included adult (age ≥ 15) EMS users who were transported to participating hospitals with severe trauma from January 1, 2014 to December 31, 2018. Severe trauma was assessed by EMS providers and defined as patients who fulfilled trauma center transport criteria (physiologic criteria, anatomic criteria, mechanism of injury criteria, or special patients or system consideration criteria) in the field triage decision scheme.[20] Patients were excluded if they had out-of-hospital cardiac arrest or their main cause of EMS call was medical or nontraumatic injury including choking, drowning, fire, flame, heat, cold, poisoning, chemical, sexual assault, weather, or natural disaster. Patients with an unknown outcome were also excluded.

Outcome measure

The primary outcome measure was the diagnosis of TBI. TBI diagnosis was defined as patients whose diagnostic code, according to the International Statistical Classification of Diseases and Related Health Problems (ICD-10), was between S06.0 and S06.9.[21 22] Although S06.7 is codes for the duration of unconscious, we included S06.7 in our study outcome according to the previous studies. [21-23] However, no patients only have S06.7 code for TBI diagnosis in our study. The ED administrative database has two types of primary diagnostic codes: the final diagnostic codes at ED discharge and at hospital discharge. We extracted up to 20 codes for each. We defined the diagnostic code as positive for TBI if a confirmative diagnostic code was found in any level of the discharge record. Because ICD 10 code is not directly linked to the severity of TBI, we further included a variety of additional outcome measures to perform analysis that take into account severity. A secondary outcome measure was TBI diagnosis with intracranial hemorrhage or injury (TBI-I), defined as TBI

patients excluding concussion (ICD 10 code with S06.0). A tertiary outcome was TBI with non-discharge (TBI-ND), defined as TBI patients excluding ED discharged patients. Because TBI-ND patients needed further management by hospitalization or transfer, we thought that this group of patients had clinically significant severity. A quaternary outcome measure was TBI with death (TBI-D), defined as TBI patients who died in ED or hospital. Because TBI-D patients are most severe group, TBI-D patients were also included in TBI-ND.

Variables and preprocessing

We collected patients' demographic data, circumstances of trauma, chief complaints, EMS vital sign assessment, EMS management and hospital outcomes. The detailed descriptions of each variable are described in Supplementary Table 1. Categorical variables were preprocessed with the one-hot encoding (dummy variable encoding) method. Continuous variables were divided into four quantiles and unknown or missing values were categorized as a fifth category. One-hot encoding was also applied to discretized continuous variables. Preprocessing measures including discretization results of continuous variables are presented in Supplementary Table 1.

Model development

We developed prediction models for outcomes by using five machine learning algorithms: traditional logistic regression analyses (LR), extreme gradient boost (XGB), random forest (RF), support vector machine (SVM), and elastic net (EN). The LR algorithm was chosen as baseline comparison algorithm because it is widely used in the medical field and has been used for previous prediction model development in TBI studies.[12] Backward stepwise LR was selected for feature selection, and we used the default parameter of stepAIC function from MASS package (version 7.3-53.1) in R for the selection. The other four algorithms were

selected based on their ability to model nonlinear associations, their relative ease of implementation, and their general acceptance in the machine learning community.[24-26] All algorithms have a method to calculate the probability of the outcome occurring and algorithms other than LR need hyperparameter tuning for proper training and prediction.

The study population was split into training cohorts that included development, validation, and test cohorts. The development cohort included a training cohort from which each of the machine learning prediction models were derived and a validation cohort in which the prediction models were applied to adjust the hyperparameters of the algorithm. The test cohort was used for the final evaluation of the performance of the prediction models. Chronological split was used for data split. Patients enrolled from January 1, 2014 to December 31, 2016 were used as the training cohort; patients from January 1, 2017 to December 31, 2017 were used as the validation cohort; and patients from January 1, 2018 to December 31, 2018 were used as the test cohort. Hyperparameter tuning using validation data was conducted by, first, a random search within 10,000 randomly generated hyperparameters; then, grid search hyperparameters chosen around from random search with five candidates per each hyperparameter. Finally, hyperparameter with best area under receiver-operation curve (AUROC) in validation cohorts were selected. Test data were separated during training and tuning processes and used to measure algorithm performance.

Statistical analysis

The demographic findings and outcomes of the study population were described in this study. Additionally, the baseline characteristics of the training cohort and the validation cohort were compared. The continuous variables were compared by using Student's T-test or the Wilcoxon rank sum test, and the categorical variables were compared by using the chi-squared test or the Fisher exact test, as appropriate.

We assessed discrimination performance by comparing the AUROC for each model in the test cohort. We considered an AUROC of 0.5 as no discrimination, 0.7 to 0.8 as acceptable, 0.8 to 0.9 as excellent, and more than 0.9 is considered outstanding.[27] Area under the precision-recall curve (AUPRC) was assessed for each model in the test cohort. We assessed the calibration power by using the Hosmer–Lemeshow test, the scaled Brier score, and a calibration plot in the test cohort. For the delineation of test characteristics, the sensitivity, specificity, and positive and negative predictive values with 95% CIs were determined using a cutoff probability at a sensitivity of 80%. Given that poor sensitivity of clinical predictors for TBI in previous studies,[7] and almost 75% sensitivity level for other severe disease prediction in prehospital settings,[28 29] we thought that 80% sensitivity was an appropriate target for our prediction model. We calculated false positive rate as 1 – specificity. The added prognostic power of each prediction model compared to the LR model was also evaluated by continuous net reclassification index (NRI). NRI is a statistical method to quantify how well a new model correctly reclassifies the study population with the other models. Details of NRI are described elsewhere.[30]

By using a model-specific metric, the variable importance of each model was assessed, except for the SVM algorithm. The variable importance was determined by the coefficient effect sizes for the LR model. The XGB and RF models were ranked by variable importance on the selection frequency of the variable as a decision node. The absolute value of the coefficients corresponding to the tuned model were used for the measurement of variable importance in the EN algorithm. To compare the variable importance of each prediction models efficiently, top 5 variables of each model was presented.

All statistical analyses were performed with R Statistical Software (version 4.0.1; R Foundation for Statistical Computing, Vienna, Austria). Packages included caret, e1071, xgboost, randomForest, and glmnet for the analysis of the machine learning algorithms.

Patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

Result

Demographic findings

Among the 157,134 EMS users transported to three hospitals from 2014 to 2018, 1,169 patients were included in the final analysis (Figure 1). Patients were split into 2 datasets: data from 2014 to 2017, consisting of 867 patients (74.2%) in the development cohort; and the remaining data from 2018 consisting of 302 patients (25.8%) in the test cohort (Figure 1). Among the development cohort, data from 2014 to 2016—consisting of 661 patients—were used as the training cohort, and 2017 data—consisting of 206 patients—were used as the validation cohort in the model.

Table 1 shows key demographic findings of the development and test cohorts. Median (IQR) age was 52 years (35–66) in the development cohort and 56 years (40–69) in the test cohort. Traffic accident was most common mechanism of trauma (43.3% for the development cohort and 41.4% for the test cohort). The proportion of patients with alert mental status was

58.1% for the development cohort and 69.5% in the test cohort. Overall, TBI, TBI-I, TBI-ND, TBI-D occurred in 215 (24.8%), 195 (22.5%), 192 (22.1%), and 32 (3.7%) in the development cohort; and 66 (21.9%), 56 (18.5%), 57 (18.9%), and 11 (3.6%) in the test cohort. All demographic characteristics of the development and test cohorts are described in Supplementary Table 2.

Main analysis

The final hyperparameters of prediction models are described in Supplementary Table 3. The discrimination and NRI of the prediction models on the test cohort are presented in Table 2. The AUROC for outcomes were 0.770–0.806 for TBI, 0.820–0.844 for TBI-I, 0.767–0.811 for TBI-ND, and 0.664–0.889 for TBI-D (Table 2 and Supplementary Figure 1). Compared to LR, XGB performed significantly well in predicting TBI, and RF and EN performed well in predicting TBI-ND and TBI-D. EN model generally performed well on all outcomes. The AUROC of the EN model for outcomes were 0.799 (95% CI: 0.732–0.867), 0.844 (95% CI: 0.779–0.910), 0.811 (95% CI: 0.741–0.882), and 0.871 (95% CI: 0.764–0.978) for TBI, TBI-I, TBI-ND, and TBI-D, respectively. Machine learning models generally resulted in significant reclassification improvement compared to LR for TBI, TBI-I, and TBI-ND. For prediction TBI-D, AUROC difference, and reclassification improvement compared to LR was non-significant in all machine learning models. The precision-recall curve is shown in Supplementary Figure 2. AUPRC were 0.479–0.564 for TBI, 0.469–0.606 for TBI-I, 0.477–0.551 for TBI-ND and 0.094–0.140 for TBI-D. EN model showed highest AUPRC among all prediction models. Supplementary Figure 3 shows the calibration plot of prediction models according to outcomes. All prediction models generally showed poor calibration. Given the high AUROC and AUPRC among prediction models, and reclassification improvement

compared to LR, we determined EN as a best-performing prediction model in our analysis.

Using cutoff of 80% sensitivity, specificity was 47.5–68.2% for TBI, 71.1–81.3% for TBI-I, 46.1–74.3% for TBI-ND, and 42.6-–.0 for TBI-D. EN showed the highest specificity and PPV among all outcomes. False positive rate (1 – specificity) was almost 19.7–39.0% according to outcomes in the EN model. The 95% CI of specificity of the EN model was not overlapped with LR in TBI, TBI-ND, and TBI-D predictions. NPV was almost 89–99% for all outcomes in the prediction models (Table 3).

Table 4 shows the top 5 variable importance of prediction models according to outcomes. Variables related to patients' symptom of loss of consciousness, Glasgow Coma Scale component, and light reflex were the three most important variables to predict all outcomes. Compared to other outcomes, the difference between variable importance for TBI-D was prominent, and the mechanism of injury, heart rate, and age showed the highest importance for predicting TBI-D.

Discussion

By using prehospital data from EMS users visiting three teaching hospitals, we developed and validated prediction models for the diagnosis and prognosis of TBI using machine learning algorithms among patients with severe trauma, identified by EMS providers in South Korea. We found that 24% of patients were diagnosed with TBI, 22% showed intracranial injury, 21% could not be discharged from the ED with a TBI diagnosis, and 4% showed TBI-related death. Machine learning models showed acceptable-to-excellent discrimination performance (AUROCs were 0.799–0.871 according to outcomes in the best-performing EN model). When identifying 80% of target patients with TBI, the false positive rate was almost 19.7–39.0%. Consciousness status related variables ranging from patients' symptom to EMS

providers' assessment showed the highest importance for predicting all outcomes. This study adds considerably to the understanding of prehospital prediction performance of TBI among patients with severe trauma. Use of comprehensive prehospital information and certain machine learning approaches led to increased performance with a diminished false positive rate compared to those of the traditional statistical model.

Several studies reported that EMS providers' assessment using prehospital information is effective for the identification of patients with severe trauma who require direct transport to a trauma center.[31-33] Because TBI accounts for a significant portion of patients with severe trauma,[32] and the majority of patients have poor access to trauma centers,[34] identification of TBI among patients with severe trauma by EMS providers could contribute to proper prehospital management and destination hospital decisions.[3] However, prehospital identification of TBI is challenging.[35] Prehospital clinical signs showed poor predictive performance for differentiating patients with TBI. [7], and previous prediction models related to TBI mostly focused on TBI outcomes.[8 9 13] One study reported the predictors for mild TBI with persistent symptoms; but a single-center case-control study design and ED-based model development lacks applicability to prehospital settings.[36] In this study, we developed and tested TBI prediction models that used prehospital information, and we found acceptable discrimination power for the prediction of diagnosis and prognosis of TBI. Uniquely, we incorporated various demographic variables, trauma circumstances, patients' complaints, and EMS assessment information in the prediction models, and we adapted the machine learning algorithms.

When using a cutoff for 80% sensitivity for TBI detection, the false positive rate was 19.7–39.0% (Table 2). Those false positive rate levels are plausible for detecting severe diseases in EMS settings. A previous study reported a 26% of false positive rate of EMS

triage for myocardial infarction with a sensitivity of 74% and 50% of false positive rate of EMS recognition of stroke in sensitivity of 74%.[28 29] Considering the prevalence of outcomes (24% in TBI, 22% in TBI-I, 21% in TBI-ND, and 4% in TBI-D; Table 1), there would be 16, 9, 12, and 67 false-positive patients for every 10 patients that are accurately identified as TBI, TBI-I, TBI-ND, and TBI-D, respectively (Supplementary Table 4). Because of the low prevalence of TBI-D, a similar specificity of the prediction model for outcomes resulted in a very low positive predictive value and a high proportion of false positive cases, which suggested the limited applicability of prediction models for TBI-D in prehospital settings.

Consciousness-status-related variables ranging from patients' complaints to EMS assessment showed the highest importance regardless of models and outcomes in our study. Consciousness status is closely associated with head trauma. Head trauma can result in structural brain injury or physiological disruption of brain function, which could result in altered mental status.[37] Mental status is also associated with TBI severity, [38] and its association with TBI outcomes have been reported.[8 9 13] History taking and physical examination for altered mental status is key to early diagnosis and proper management of TBI in prehospital settings.[39]

We adapted machine learning algorithms for the prediction of TBI-related outcomes and found an improvement in discrimination and an increase in specificity with the same sensitivity thresholds. However, the LR model also showed acceptable or similar performance compared to machine learning models, according to the outcomes. In clinical prediction models, a previous systematic review reported no performance benefit of the machine learning model over LR.[40] The previous study stated that machine learning models tend to show high performance with a strong signal-to-noise ratio problem like gaming, image recognition. However, clinical prediction problems often result in a poor signal-to-

noise ratio.[40] If we could use unstructured data, which has strong signal-to-noise ratio like continuous vital sign monitoring data or audiovisual data for patients' appearance, machine learning models might perform better than LR models. In addition, if we analyzed more patient data, the performance improvement of machine models might be elucidated.

Precise assessment in prehospital field could contribute to improved patient-related outcomes. High demand of EMS call and response, disparity in accessibility to definitive care capable hospitals according to regions,[34] and the importance of timely management in acute disease care are the chief reasons behind the necessity for the accurate assessment of EMS providers. Although information acquisition and processing is quite difficult in prehospital areas, various instruments and information systems could attribute to diminish those problems. Complex data acquisition like mobile CT or other unstructured data[41], information sharing through telemedicine,[42] and decision support tools in prehospital environments[43] could contribute to the accurate assessment of EMS providers. More information acquisition and real-time processing of those data could improve the clinical prediction models in prehospital areas, which could lead to the improvement of patients' safety and outcomes.

Our study had several limitations. First, our data were collected at three teaching hospitals in urban areas of South Korea. Therefore, external validation for other areas should be conducted to generalize the developed prediction model. Second, we used retrospective analysis of electronically collected prehospital and hospital data. There might be various information loss and missing data. We treated missing status as a separate category for our analysis;[44] however, there could be different reasons for missing data. Third, there is a possibility that the prediction model was overfitted or underfitted. The use of large number of predictors also can contribute to overfitting. To minimize this issue, we rigorously searched hyperparameters and carefully chose hyperparameters according to the performance in

independent validation cohorts. Fourth, we selected our study population using trauma center transport criteria for EMS providers in Korea. Although those criteria are based on the field triage decision scheme which is the most widely used prehospital trauma triage protocol,[6] extrapolation to another EMS setting or general trauma patients would be limited. Fifth, Abbreviated Injury Scale (AIS) codes were not used to identify our study outcome because of a lack of information. To compensate for this limitation, we further identified TBI-I, TBI-ND, and TBI-D patients to consider severity. However, different definitions of clinical severity, including ICU admission or emergency operation, might be possible. Lastly, this study was performed in an intermediate-service-level EMS system. The generalization of our study findings to different EMS settings should be made with caution.

In conclusion, we presented data on TBI among patients with severe trauma assessed by EMS providers, and our results inform the development of prediction models for the diagnosis and prognosis of TBI in our population. We used various information that can be obtained in prehospital settings and showed acceptable outcome performance. The consistent importance of consciousness-status-related variables emphasizes the importance of assessment and monitoring of consciousness status in prehospital areas. Although prospective, and implementation studies are needed for TBI prediction in prehospital areas, our study outlined a novel method for the precise assessment of EMS providers using a machine-learning-based prediction model. Further collection of various types of patient-related data would contribute to the enhanced performance of the clinical prediction model in prehospital settings.

Author Contribution Statement

YHC and JH Park designed and developed the study, analysed and interpreted the data, and drafted the initial manuscript. KJH, YSR, KJS and SDS were involved in the acquisition of data, the development of the research question and assisted with analysis and interpretation of data. All authors revised the drafts for intellectual content and edited the manuscript. All authors reviewed and approved the final draft.

Funding

This study was supported by grant No. '04-2019-0680' from the Seoul National University Hospital Research Fund.

Competing Interests

There are no conflicts of interest for all authors in this study.

Patients consent

Not required

Data availability statement

No data are available. We do not have ethics approval to share data.

Ethical statements

This study complied with the Declaration of Helsinki, and its protocol was approved by the Institutional Review Board of the Seoul National University Hospital with a waiver of informed consent (IRB No: E-2006-004-1128).

References

1. Hsia RY, Markowitz AJ, Lin F, et al. Ten-year trends in traumatic brain injury: a retrospective cohort study of California emergency department and hospital revisits and readmissions. *BMJ Open* 2018;8(12):e022297. doi: 10.1136/bmjopen-2018-022297 [published Online First: 2018/12/16]

2. Finfer SR, Cohen J. Severe traumatic brain injury. *Resuscitation* 2001;48(1):77-90. doi: 10.1016/s0300-9572(00)00321-x [published Online First: 2001/02/13]

3. Spaite DW, Bobrow BJ, Keim SM, et al. Association of Statewide Implementation of the Prehospital Traumatic Brain Injury Treatment Guidelines With Patient Survival Following Traumatic Brain Injury: The Excellence in Prehospital Injury Care (EPIC) Study. *JAMA Surg* 2019;154(7):e191152. doi: 10.1001/jamasurg.2019.1152 [published Online First: 2019/05/09]

4. McHugh GS, Engel DC, Butcher I, et al. Prognostic value of secondary insults in traumatic brain injury: results from the IMPACT study. *J Neurotrauma* 2007;24(2):287-93. doi: 10.1089/neu.2006.0031 [published Online First: 2007/03/23]

5. Pelieu I, Kull C, Walder B. Prehospital and Emergency Care in Adult Patients with Acute Traumatic Brain Injury. *Med Sci (Basel)* 2019;7(1) doi: 10.3390/medsci7010012 [published Online First: 2019/01/24]

6. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients:

recommendations of the National Expert Panel on Field Triage, 2011. *Morbidity and Mortality Weekly Report: Recommendations and Reports* 2012;61(1):1-20.

7. Ter Avest E, Taylor S, Wilson M, et al. Prehospital clinical signs are a poor predictor of raised intracranial pressure following traumatic brain injury. *Emerg Med J* 2021;38(1):21-26. doi: 10.1136/emermed-2020-209635 [published Online First: 2020/09/20]

8. Collaborators MCT, Perel P, Arango M, et al. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ* 2008;336(7641):425-9. doi: 10.1136/bmj.39461.643438.25 [published Online First: 2008/02/14]

9. Steyerberg EW, Mushkudiani N, Perel P, et al. Predicting outcome after traumatic brain injury: development and international validation of prognostic scores based on admission characteristics. *PLoS Med* 2008;5(8):e165; discussion e65. doi: 10.1371/journal.pmed.0050165 [published Online First: 2008/08/08]

10. Gozt AK, Hellewell SC, Thorne J, et al. Predicting outcome following mild traumatic brain injury: protocol for the longitudinal, prospective, observational Concussion Recovery (CREST) cohort study. *BMJ Open* 2021;11(5):e046460. doi: 10.1136/bmjopen-2020-046460 [published Online First: 2021/05/15]

11. Huth SF, Slater A, Waak M, et al. Predicting Neurological Recovery after Traumatic Brain Injury in Children: A Systematic Review of Prognostic Models. *J Neurotrauma* 2020;37(20):2141-49. doi: 10.1089/neu.2020.7158 [published Online First: 2020/05/29]

12. Perel P, Edwards P, Wentz R, et al. Systematic review of prognostic models in traumatic brain injury. *BMC Med Inform Decis Mak* 2006;6:38. doi: 10.1186/1472-6947-6-38 [published Online First: 2006/11/16]

13. Miller PR, Chang MC, Hoth JJ, et al. Predicting Mortality and Independence at Discharge in the Aging Traumatic Brain Injury Population Using Data Available at Admission. *J Am Coll Surg* 2017;224(4):680-85. doi: 10.1016/j.jamcollsurg.2016.12.053 [published Online First: 2017/03/07]

14. Abujaber A, Fadlalla A, Gammoh D, et al. Prediction of in-hospital mortality in patients with post traumatic brain injury using National Trauma Registry and Machine Learning Approach. *Scand J Trauma Resusc Emerg Med* 2020;28(1):44. doi: 10.1186/s13049-020-00738-5 [published Online First: 2020/05/29]

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. doi: 10.1016/j.jclinepi.2020.03.005 [published Online First: 2020/03/24]

16. Roozenbeek B, Lingsma HF, Lecky FE, et al. Prediction of outcome after moderate and severe traumatic brain injury: external validation of the International Mission on Prognosis and Analysis of Clinical Trials (IMPACT) and Corticoid Randomisation After Significant Head injury (CRASH) prognostic models. *Crit Care Med* 2012;40(5):1609-17. doi: 10.1097/CCM.0b013e31824519ce [published Online First: 2012/04/19]

17. Wasserman EB, Shah MN, Jones CM, et al. Identification of a neurologic scale that optimizes EMS detection of older adult traumatic brain injury patients who require transport to a trauma center. *Prehosp Emerg Care* 2015;19(2):202-12. doi: 10.3109/10903127.2014.959225 [published Online First: 2014/10/08]

18. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD). *Ann Intern Med*

2015;162(10):735-6. doi: 10.7326/L15-5093-2 [published Online First: 2015/05/20]

19. Sasser SM, Hunt RC, Sullivent EE, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage. 2009

20. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *MMWR Recomm Rep* 2012;61(RR-1):1-20. [published Online First: 2012/01/13]

21. Andelic N, Anke A, Skandsen T, et al. Incidence of hospital-admitted severe traumatic brain injury and in-hospital fatality in Norway: a national cohort study. *Neuroepidemiology* 2012;38(4):259-67. doi: 10.1159/000338032 [published Online First: 2012/06/09]

22. Ro YS, Shin SD, Holmes JF, et al. Comparison of clinical performance of cranial computed tomography rules in patients with minor head injury: a multicenter prospective study. *Acad Emerg Med* 2011;18(6):597-604. doi: 10.1111/j.1553-2712.2011.01094.x [published Online First: 2011/06/17]

23. Chan V, Thurairajah P, Colantonio A. Defining pediatric traumatic brain injury using International Classification of Diseases Version 10 Codes: a systematic review. *BMC Neurol* 2015;15:7. doi: 10.1186/s12883-015-0259-7 [published Online First: 2015/02/05]

24. Zou H, Hastie T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67(2):301-20. doi: 10.1111/j.1467-9868.2005.00503.x

25. Hearst MA, Dumais ST, Osuna E, et al. Support vector machines. *IEEE Intelligent Systems and their Applications* 1998;13(4):18-28. doi: 10.1109/5254.708428

26. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, California, USA: Association for Computing Machinery, 2016:785–94.

27. Menard S. Applied logistic regression analysis: Sage 2002. pp. 162.

28. Oostema JA, Konen J, Chassee T, et al. Clinical predictors of accurate prehospital stroke recognition. *Stroke* 2015;46(6):1513-7. doi: 10.1161/STROKEAHA.115.008650 [published Online First: 2015/04/30]

29. Swan PY, Nighswonger B, Boswell GL, et al. Factors associated with false-positive emergency medical services triage for percutaneous coronary intervention. *West J Emerg Med* 2009;10(4):208-12. [published Online First: 2010/01/05]

30. Park JH, Shin SD, Song KJ, et al. Prediction of good neurological recovery after out-of-hospital cardiac arrest: A machine learning analysis. *Resuscitation* 2019;142:127-35. doi: 10.1016/j.resuscitation.2019.07.020 [published Online First: 2019/07/31]

31. Esposito TJ, Offner PJ, Jurkovich GJ, et al. Do prehospital trauma center triage criteria identify major trauma victims? *Arch Surg* 1995;130(2):171-6. doi: 10.1001/archsurg.1995.01430020061010 [published Online First: 1995/02/01]

32. Ocak G, Sturms LM, Hoogeveen JM, et al. Prehospital identification of major trauma patients. *Langenbecks Arch Surg* 2009;394(2):285-92. doi: 10.1007/s00423-008-0340-4 [published Online First: 2008/06/27]

33. Fries GR, McCalla G, Levitt MA, et al. A prospective comparison of paramedic judgment and the trauma triage rule in the prehospital setting. *Ann Emerg Med* 1994;24(5):885-9. doi: 10.1016/s0196-0644(94)70207-1 [published Online First: 1994/11/01]

34. Branas CC, MacKenzie EJ, Williams JC, et al. Access to trauma centers in the United States. *JAMA* 2005;293(21):2626-33. doi: 10.1001/jama.293.21.2626 [published Online First: 2005/06/02]

35. Whiting MD, Dengler BA, Rodriguez CL, et al. Prehospital Detection of Life-Threatening Intracranial Pathology: An Unmet Need for Severe TBI in Austere, Rural, and Remote Areas. *Front Neurol* 2020;11:599268. doi: 10.3389/fneur.2020.599268 [published Online First: 2020/11/17]

36. Wojcik SM. Predicting mild traumatic brain injury patients at risk of persistent symptoms in the Emergency Department. *Brain Inj* 2014;28(4):422-30. doi: 10.3109/02699052.2014.884241 [published Online First: 2014/02/26]

37. Management of Concussion/m TBIWG. VA/DoD Clinical Practice Guideline for Management of Concussion/Mild Traumatic Brain Injury. *J Rehabil Res Dev* 2009;46(6):CP1-68. [published Online First: 2010/01/30]

38. Grote S, Bocker W, Mutschler W, et al. Diagnostic value of the Glasgow Coma Scale for traumatic brain injury in 18,002 patients with severe multiple injuries. *J Neurotrauma* 2011;28(4):527-34. doi: 10.1089/neu.2010.1433 [published Online First: 2011/01/27]

39. Badjatia N, Carney N, Crocco TJ, et al. Guidelines for prehospital management of traumatic brain injury 2nd edition. *Prehosp Emerg Care* 2008;12 Suppl 1:S1-52. doi: 10.1080/10903120701732052 [published Online First: 2008/09/06]

40. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol* 2019;110:12-22. doi: 10.1016/j.jclinepi.2019.02.004 [published Online First: 2019/02/15]

41. Nakada TA, Masunaga N, Nakao S, et al. Development of a prehospital vital signs chart sharing system. *Am J Emerg Med* 2016;34(1):88-92. doi: 10.1016/j.ajem.2015.09.048 [published Online First: 2015/10/29]

42. Kim Y, Groombridge C, Romero L, et al. Decision Support Capabilities of Telemedicine

in Emergency Prehospital Care: Systematic Review. *J Med Internet Res* 2020;22(12):e18959. doi: 10.2196/18959 [published Online First: 2020/12/09]

43. Reisner AT, Khitrov MY, Chen L, et al. Development and validation of a portable platform for deploying decision-support algorithms in prehospital settings. *Appl Clin Inform* 2013;4(3):392-402. doi: 10.4338/ACI-2013-04-RA-0023 [published Online First: 2013/10/25]

44. Maslove DM, Podchiyska T, Lowe HJ. Discretization of continuous features in clinical datasets. *J Am Med Inform Assoc* 2013;20(3):544-53. doi: 10.1136/amiajnl-2012-000929 [published Online First: 2012/10/13]

Figure legends

Figure 1. Population flow. EMS, emergency medical service; OHCA, out-of-hospital cardiac arrest; TBI, traumatic brain injury.

Table 1. Key characteristics of the development and test cohorts.

	n (%) or Median (IQR)			
	Total	Development cohort	Test cohort	P
Total	N = 1169	n = 867	n = 302	
Demographics				
Age, years	53 (36–66)	52 (35–66)	56 (40–69)	< 0.01
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Job, unemployed	299 (25.6)	197 (22.7)	102 (33.8)	< 0.01
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	< 0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	< 0.01
Circumstances of trauma				
Location, road/highway	444 (38.0)	326 (37.6)	118 (39.1)	0.65
Season, summer	336 (28.7)	253 (29.2)	83 (27.5)	0.57
Weekday, weekend	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time, 6 p.m. to midnight	361 (30.9)	265 (30.6)	96 (31.8)	0.69
Mechanism of injury, TA	500 (42.8)	375 (43.3)	125 (41.4)	0.57
Chief complaint				
Fracture/abrasion/laceration	302 (25.8)	204 (23.5)	98 (32.5)	< 0.01
EMS vital sign assessment				
SBP, mmHg	130 (109–150)	130 (104–146)	131 (115–150)	< 0.01
DBP, mmHg	80 (70–91)	80 (69–90)	80 (70–92)	0.21
RR, /min	18 (16–20)	18 (16–20)	18 (16–20)	0.33
HR, /min	86 (75–99)	86 (74–99)	86 (76–100)	0.40
SpO2, %	98 (95–99)	98 (95–99)	98 (96–99)	0.67
AVPU scale, Alert	714 (61.1)	504 (58.1)	210 (69.5)	< 0.01
EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	0.95

IQR, interquartile range; TA, traffic accident; SBP, systolic blood pressure; DBP, diastolic blood pressure; RR, respiratory rate; AVPU, mental status in alert, verbal, pain, and unresponsive scale; ED, emergency department; TBI, traumatic brain injury.

Table 2. Discrimination and reclassification of prediction models for outcomes on test cohort.

Outcome	Model	AUROC (95% CI)	p[a]	NRI (95% CI)	p[b]	AUPRC
TBI						
	LR	0.770 (0.698, 0.841)	NA	NA	NA	0.492
	XGB	0.809 (0.743, 0.876)	0.04	0.689 (0.427, 0.951)	< 0.01	0.552
	SVM	0.776 (0.708, 0.844)	0.77	0.339 (0.072, 0.607)	0.01	0.479
	RF	0.800 (0.735, 0.865)	0.13	0.308 (0.047, 0.569)	0.02	0.532
	EN	0.799 (0.732, 0.867)	0.06	0.698 (0.441, 0.954)	< 0.01	0.564
TBI-I						
	LR	0.820 (0.751, 0.890)	NA	NA	NA	0.551
	XGB	0.838 (0.775, 0.901)	0.28	0.539 (0.258, 0.821)	< 0.01	0.554
	SVM	0.812 (0.748, 0.875)	0.66	0.729 (0.464, 0.994)	< 0.01	0.469
	RF	0.836 (0.772, 0.899)	0.38	0.333 (0.058, 0.607)	0.02	0.552
	EN	0.844 (0.779, 0.910)	0.15	1.093 (0.845, 1.342)	< 0.01	0.606
TBI-ND						
	LR	0.767 (0.690, 0.844)	NA	NA	NA	0.482
	XGB	0.800 (0.727, 0.873)	0.07	0.605 (0.326, 0.884)	< 0.01	0.496
	SVM	0.778 (0.704, 0.852)	0.56	0.285 (-0.001, 0.572)	0.05	0.477
	RF	0.809 (0.739, 0.880)	0.03	0.194 (-0.059, 0.448)	0.13	0.535
	EN	0.811 (0.741, 0.882)	0.02	0.768 (0.496, 1.039)	< 0.01	0.551
TBI-D						
	LR	0.664 (0.490, 0.838)	NA	NA	NA	0.138
	XGB	0.714 (0.512, 0.917)	0.64	-0.026 (-0.605, 0.553)	0.93	0.094
	SVM	0.814 (0.718, 0.910)	0.09	0.209 (-0.325, 0.742)	0.44	0.140
	RF	0.889 (0.801, 0.976)	< 0.01	-0.204 (-0.742, 0.334)	0.46	0.196
	EN	0.871 (0.764, 0.978)	0.01	0.119 (-0.415, 0.654)	0.66	0.293

[a]Comparing the AUROC and the logistic regression model.

[b]Comparing the NRI and the logistic regression model.

AUROC, area under the receiver operating characteristic curve; CI, confidence interval; NRI, net reclassification index; AUPRC, area under precision-recall curve; TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support vector machine; RF, random forest; EN, elastic net

Table 3. Test characteristics of prediction models for outcomes on test cohort.

Outcome	Model	Specificity (95% CI)	Sensitivity (95% CI)	PPV (95% CI)	NPV (95% CI)	Cutoff
TBI						
	LR	47.5 (40.9, 54.0)	80.3 (68.7, 89.1)	29.9 (23.3, 37.3)	89.6 (82.9, 94.3)	0.136
	XGB	72.5 (66.3, 78.1)	80.3 (68.7, 89.1)	44.9 (35.7, 54.3)	92.9 (88.2, 96.2)	0.268
	SVM	64.8 (58.4, 70.9)	80.3 (68.7, 89.1)	39.0 (30.7, 47.7)	92.2 (87.0, 95.8)	0.191
	RF	68.2 (61.9, 74.1)	80.3 (68.7, 89.1)	41.4 (32.8, 50.4)	92.5 (87.6, 96.0)	0.185
	EN	61.0 (54.5, 67.3)	80.3 (68.7, 89.1)	36.6 (28.7, 44.9)	91.7 (86.3, 95.5)	0.205
TBI-I						
	LR	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.164
	XGB	74.0 (68.0, 79.4)	80.4 (67.6, 89.8)	41.3 (31.9, 51.1)	94.3 (90.0, 97.1)	0.143
	SVM	71.1 (65.0, 76.7)	80.4 (67.6, 89.8)	38.8 (29.9, 48.3)	94.1 (89.7, 97.0)	0.172
	RF	76.0 (70.2, 81.2)	80.4 (67.6, 89.8)	43.3 (33.6, 53.3)	94.4 (90.3, 97.2)	0.205
	EN	81.3 (75.9, 86.0)	80.4 (67.6, 89.8)	49.5 (38.8, 60.1)	94.8 (90.9, 97.4)	0.204
TBI-ND						
	LR	46.1 (39.8, 52.6)	80.7 (68.1, 90.0)	25.8 (19.6, 32.9)	91.1 (84.7, 95.5)	0.090
	XGB	66.5 (60.2, 72.4)	80.7 (68.1, 90.0)	35.9 (27.7, 44.9)	93.7 (89.0, 96.8)	0.242
	SVM	59.2 (52.7, 65.4)	80.7 (68.1, 90.0)	31.5 (24.1, 39.7)	92.9 (87.7, 96.4)	0.147
	RF	60.4 (54.0, 66.6)	80.7 (68.1, 90.0)	32.2 (24.6, 40.5)	93.1 (88.0, 96.5)	0.138
	EN	74.3 (68.3, 79.6)	80.7 (68.1, 90.0)	42.2 (32.8, 52.0)	94.3 (90.0, 97.1)	0.201
TBI-D						
	LR	42.6 (36.9, 48.5)	81.8 (48.2, 97.7)	5.1 (2.4, 9.5)	98.4 (94.4, 99.8)	0.005
	XGB	57.7 (51.8, 63.5)	81.8 (48.2, 97.7)	6.8 (3.2, 12.5)	98.8 (95.8, 99.9)	0.002
	SVM	74.2 (68.8, 79.2)	81.8 (48.2, 97.7)	10.7 (5.0, 19.4)	99.1 (96.7, 99.9)	0.039
	RF	74.9 (69.5, 79.8)	81.8 (48.2, 97.7)	11.0 (5.1, 19.8)	99.1 (96.8, 99.9)	0.005
	EN	79.0 (73.9, 83.6)	81.8 (48.2, 97.7)	12.9 (6.1, 23.0)	99.1 (96.9, 99.9)	0.033

TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression analysis; XGB, extreme gradient boosting; SVM, support vector machine; RF, random forest; EN, elastic net.

Table 4. Top 5 important variables for outcomes in descending order using model specific metrics

Outcome	Rank	LR	XGB	RF	EN
TBI					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Motor, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	4	Light reflex	Other mechanism	Light reflex	GCS, Eye, 1
	5	GCS, Motor, 1	GCS, Verbal, 2	GCS, Motor, 1	GCS, Verbal, 1
TBI-I					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	GCS, Eye, 1
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	Loss of consciousness
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	Light reflex	GCS, Verbal, 1
	5	GCS, Motor, 1	Other mechanism	GCS, Motor, 1	Light reflex
TBI-ND					
	1	Loss of consciousness	Loss of consciousness	Loss of consciousness	Loss of consciousness
	2	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1	GCS, Eye, 1
	3	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 1
	4	Light reflex	GCS, Verbal, 2	GCS, Verbal, 2	GCS, Verbal, 1
	5	GCS, Motor, 1	GCS, Motor, 1	GCS, Motor, 4	Light reflex
TBI-D					
	1	Loss of consciousness	GCS, Verbal, 1	GCS, Verbal, 1	GCS, Motor, 2
	2	GCS, Verbal, 1	Oxygen saturation<96%	Light reflex	GCS, Verbal, 1
	3	GCS, Eye, 1	Fall mechanism	Loss of consciousness	Loss of consciousness
	4	Light reflex	Afternoon	GCS, Eye, 1	Age over 80
	5	GCS, Motor, 1	Light reflex	GCS, Motor, 1	HR 87-99

TBI, traumatic brain injury, TBI-I, traumatic brain injury with intracranial injury; TBI-ND; traumatic brain injury with non-discharge; TBI-D, traumatic brain injury with death; LR, logistic regression; XGB, extreme gradient boosting; RF, random forest; EN, elastic net; GCS, Glasgow coma scale; HR, heart rate.

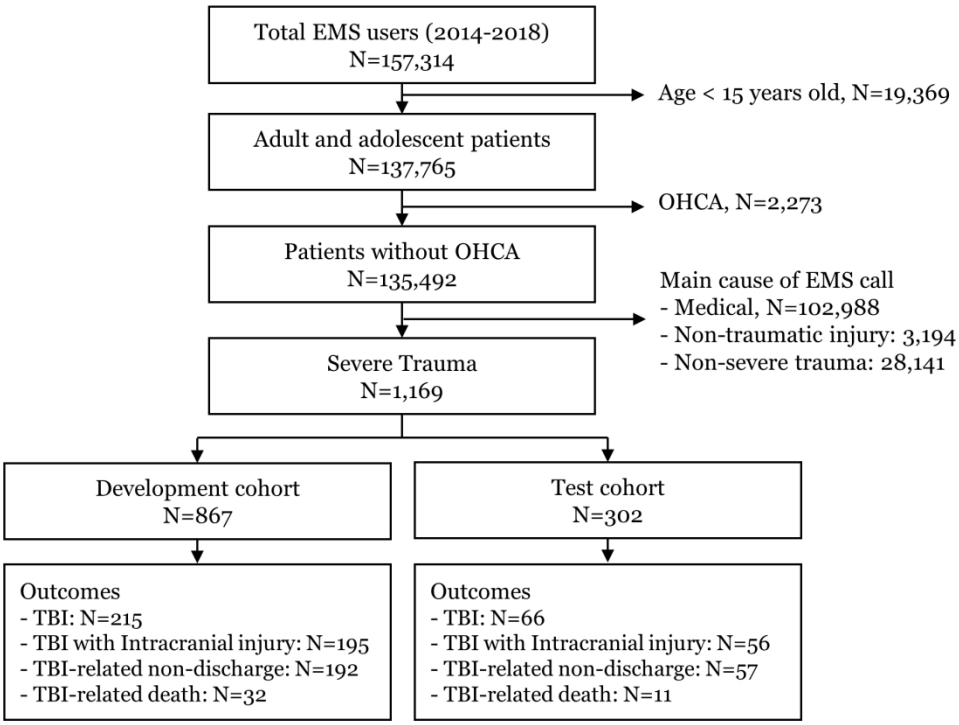


Figure 1. Population flow. EMS, emergency medical service; OHCA, out-of-hospital cardiac arrest; TBI, traumatic brain injury.

165x119mm (300 x 300 DPI)

Supplementary Table 1. List of analyzed variables.

Variables	Descriptions	Type of raw data	Category	Preprocessing
Gender	Sex of the patients	Binary	Male, Female	
Age	Age of patients	Continuous	15-39 years, 40-59 years, 60-79 years, and 80- years	Discretization and one hot encoding
Job	Job of patients	Categorical	Unemployed, Student/Housewife; Office/Commercial/Service workers; Industrial/Agricultural/Fishery/Miner worker; Others	One hot encoding Missing data were classified into others
Diabetes	History of diabetes mellitus	Binary	Yes, No	Missing data were classified into no
Hypertension	History of hypertension	Binary	Yes, No	Missing data were classified into no
Location of injury	Location of injury	Categorical	home/residential area/medical facility/school/gym; Road/highway; Off-road traffic area; Others	One hot encoding Missing data were classified into others
Season	Season when injury occurred	Categorical	Spring, Summer, Fall, Winter	One hot encoding
Weekend	Whether Injury occurred on weekday or weekend	Binary	Weekday, Weekend	
Daytime	When injury was occurred	Categorical	Night (Midnight to 5AM), Morning (6AM to 11AM), Afternoon (Midday to 5PM), Evening (6PM to 11PM)	One hot encoding Missing time were imputed using EMS call time
Mechanism of injury	Mechanism of injury	Categorical	Slip down, Fall down, Traffic accident, Other	One hot encoding Missing data were classified into others
Glasgow coma scale eye	Eye element of Glasgow coma scale	Categorical	1;2;3;4;Unknown	One hot encoding
Glasgow coma scale Verbal	Verbal element of Glasgow coma scale	Categorical	1;2;3;4;5;Unknown	One hot encoding
Glasgow coma scale Motor	Motor element of Glasgow coma scale	Categorical	1;2;3;4;5;6;Unknown	One hot encoding
Light Reflex any Abnormal	Any abnormality of light reflex on any side	Categorical	No, Yes, Unknown	One hot encoding Missing data were classified into unknown

Systolic blood pressure	Systolic blood pressure	Continuous	-107 mmHg, 108-130 mmHg, 131-145 mmHg, 146- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Diastolic blood pressure	Diastolic blood pressure	Continuous	-69 mmHg, 70-80 mmHg, 81-91 mmHg, 92- mmHg, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Heart rate	Heart rate	Continuous	-74/min, 75-86/min, 87-99/min, 100-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Respiratory rate	Respiratory rate	Continuous	-16/min, 17-18/min, 19-20/min, 21-/min, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Oxygen saturation	Oxygen saturation	Continuous	-95%, 96-98%, 99%, 100%, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Body temperature	Body temperature	Continuous	-36℃, 36.1-36.3℃, 36.4-36.8℃, 36.9-℃, Unknown	Discretization and one hot encoding Cutoff values for categories were calculated from median and interquartile range of training cohort Missing data were classified into unknown
Chest pain or abdominal pain	Symptom of chest pain or abdominal pain	Binary	Yes, No	

Fracture, abrasion, or laceration	Symptom of fracture, abrasion, or laceration	Binary	Yes, No	
Loss of consciousness	Symptom of loss of consciousness (whether patients had loss of consciousness between injury and EMS provider's assessment)	Binary	Yes, No	
Dyspnea	Symptom of dyspnea	Binary	Yes, No	
Nose bleeding	Symptom of nose bleeding	Binary	Yes, No	
Nausea or vomiting	Symptom of nausea or vomiting	Binary	Yes, No	
Headache, paralysis or dizziness	Symptom of headache, paralysis or dizziness	Binary	Yes, No	

Supplementary Table 2. Demographic characteristics of development and test cohorts

	N (%) or Median (IQR)			
Characteristics	Total	Development	Test	P-value
Total	1169	867	302	
Demographics				
Male	809 (69.2)	592 (68.3)	217 (71.9)	0.25
Age, years	53 (36-66)	52 (35-66)	56 (40-69)	<0.01
Job of patients				<0.01
Unemployed	299 (25.6)	197 (22.7)	102 (33.8)	
Student/Housewife	161 (13.8)	129 (14.9)	32 (10.6)	
Office/Commercial/Service worker	283 (24.2)	176 (20.3)	107 (35.4)	
Industrial/Agricultural/Fishery/Minery worker	36 (3.1)	25 (2.9)	11 (3.6)	
Others	390 (33.4)	340 (39.2)	50 (16.6)	
Past medical history				
Diabetes	62 (5.3)	35 (4.0)	27 (8.9)	<0.01
Hypertension	105 (9.0)	61 (7.0)	44 (14.6)	<0.01
Circumstances of Trauma				
Location of trauma				0.52
Residential/Nursing/Education/Exercise facility	303 (25.9)	218 (25.1)	85 (28.1)	
Road/Highway	444 (38.0)	326 (37.6)	118 (39.1)	
Off-road traffic area	181 (15.5)	140 (16.1)	41 (13.6)	
Others	241 (20.6)	183 (21.1)	58 (19.2)	
Season of trauma				<0.01
Spring	249 (21.3)	167 (19.3)	82 (27.2)	
Summer	336 (28.7)	253 (29.2)	83 (27.5)	
Fall	304 (26.0)	242 (27.9)	62 (20.5)	
Winter	280 (24.0)	205 (23.6)	75 (24.8)	
Weekday	811 (69.4)	599 (69.1)	212 (70.2)	0.72
Time of trauma				0.83
6A-MD	281 (24.0)	206 (23.8)	75 (24.8)	
MD-6P	266 (22.8)	203 (23.4)	63 (20.9)	
6P-MN	361 (30.9)	265 (30.6)	96 (31.8)	
MN-6A	261 (22.3)	193 (22.3)	68 (22.5)	
Mechanism of Trauma				0.60
Traffic accident	500 (42.8)	375 (43.3)	125 (41.4)	
Slip down	325 (27.8)	232 (26.8)	93 (30.8)	
Fall down	171 (14.6)	129 (14.9)	42 (13.9)	
Others	173 (14.8)	131 (15.1)	42 (13.9)	
Chief complaint				
Altered mentality	279 (23.9)	223 (25.7)	56 (18.5)	0.01
Facture/Abrasion/Laceration	302 (25.8)	204 (23.5)	98 (32.5)	<0.01
Chest/Abdominal pain	47 (4.0)	31 (3.6)	16 (5.3)	0.19
Dyspnea	25 (2.1)	20 (2.3)	5 (1.7)	0.50

Epistaxis	44 (3.8)	30 (3.5)	14 (4.6)	0.36
Headache/Paralysis/Dizziness/Vertigo	95 (8.1)	64 (7.4)	31 (10.3)	0.11
Nausea/Vomiting	32 (2.7)	20 (2.3)	12 (4.0)	0.13
EMS Vital sign assessment				
SBP, mmHg	130 (109-150)	130 (104-146)	131 (115-150)	<0.01
Missing	65 (5.6)	56 (6.5)	9 (3.0)	0.02
DBP, mmHg	80 (70-91)	80 (69-90)	80 (70-92)	<0.01
Missing	75 (6.4)	65 (7.5)	10 (3.3)	0.01
HR, /min	86 (75-99)	86 (74-99)	86 (76-100)	<0.01
Missing	31 (2.7)	28 (3.2)	3 (1.0)	0.04
RR, /min	18 (16-20)	18 (16-20)	18 (16-20)	<0.01
Missing	36 (3.1)	33 (3.8)	3 (1.0)	0.01
SpO2, %	98 (95-99)	98 (95-99)	98 (96-99)	<0.01
Missing	38 (3.3)	33 (3.8)	5 (1.7)	0.07
Temperature, ℃	36.5 (36-36.8)	36.5 (36-36.8)	36.5 (36-36.7)	<0.01
Missing	94 (8.0)	65 (7.5)	29 (9.6)	0.25
AVPU scale				<0.01
Alert	714 (61.1)	504 (58.1)	210 (69.5)	
Verbal	168 (14.4)	136 (15.7)	32 (10.6)	
Pain	199 (17.0)	158 (18.2)	41 (13.6)	
Unresponsive	88 (7.5)	69 (8.0)	19 (6.3)	
Abnormal light reflex	165 (14.1)	132 (15.2)	33 (10.9)	<0.01
Missing	66 (5.6)	57 (6.6)	9 (3.0)	
GCS scale component				
Glasgow coma scale eye				<0.01
4	558 (47.7)	380 (43.8)	178 (58.9)	
3	128 (10.9)	109 (12.6)	19 (6.3)	
2	110 (9.4)	82 (9.5)	28 (9.3)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	
Glasgow coma scale Verbal				0.01
5	520 (44.5)	359 (41.4)	161 (53.3)	
4	118 (10.1)	88 (10.1)	30 (9.9)	
3	25 (2.1)	19 (2.2)	6 (2.0)	
2	132 (11.3)	105 (12.1)	27 (8.9)	
1	174 (14.9)	141 (16.3)	33 (10.9)	
Unknown	200 (17.1)	155 (17.9)	45 (14.9)	
Glasgow coma scale Motor				<0.01
6	499 (42.7)	333 (38.4)	166 (55.0)	
5	124 (10.6)	103 (11.9)	21 (7.0)	
4	158 (13.5)	123 (14.2)	35 (11.6)	
3	47 (4.0)	39 (4.5)	8 (2.6)	
2	17 (1.5)	15 (1.7)	2 (0.7)	
1	125 (10.7)	99 (11.4)	26 (8.6)	
Unknown	199 (17.0)	155 (17.9)	44 (14.6)	

EMS management				
Intravenous route	176 (15.1)	129 (14.9)	47 (15.6)	0.77
Hemorrhage control	586 (50.1)	426 (49.1)	160 (53.0)	0.25
Spinal motion restriction	811 (69.4)	606 (69.9)	205 (67.9)	0.51
Advanced airway management	4 (0.3)	2 (0.2)	2 (0.7)	0.28
Oxygen supply	233 (19.9)	176 (20.3)	57 (18.9)	0.59
Field triage decision scheme criteria*				
Physiological criteria				
SBP<90 mmHg	58 (5.0)	42 (4.8)	16 (5.3)	0.75
RR<10 or >29 /min	11 (0.9)	11 (1.3)	0 (0.0)	0.08
Non-Alert	429 (36.7)	343 (39.6)	86 (28.5)	<0.01
Anatomic criteria				
All penetrating injuries to head, neck, torso and extremities proximal to elbow or knee	34 (2.9)	23 (2.7)	11 (3.6)	0.38
Chest wall instability or deformity	4 (0.3)	4 (0.5)	0 (0.0)	0.58
Two or more proximal long bone fractures	19 (1.6)	13 (1.5)	6 (2.0)	0.60
Crush, degloved, mangled or pulseless extremity	15 (1.3)	13 (1.5)	2 (0.7)	0.38
Amputation proximal to wrist or ankle	9 (0.8)	9 (1.0)	0 (0.0)	0.12
Pelvic fractures	8 (0.7)	6 (0.7)	2 (0.7)	>0.95
Open or depressed skull fracture	17 (1.5)	9 (1.0)	8 (2.6)	0.05
Paralysis	21 (1.8)	11 (1.3)	10 (3.3)	0.02
Mechanism of injury criteria				
Fall > 6 meter	113 (9.7)	84 (9.7)	29 (9.6)	>0.95
High-risk auto crash	96 (8.2)	73 (8.4)	23 (7.6)	0.66
Auto vs pedestrian/bicyclist thrown, run over, or with significant (>30km/h) impact	119 (10.2)	83 (9.6)	36 (11.9)	0.25
Motorcycle crash > 30 km/hour	105 (9.0)	70 (8.1)	35 (11.6)	0.07
ED disposition				0.11
Discharge	320 (27.4)	241 (27.8)	79 (26.2)	
Transfer	444 (38.0)	316 (36.4)	128 (42.4)	
Admitted	366 (31.3)	276 (31.8)	90 (29.8)	
In-hospital mortality	90 (7.7)	74 (8.5)	16 (5.3)	0.07
Outcomes				
TBI	281 (24.0)	215 (24.8)	66 (21.9)	0.30
TBI with intracranial injury	251 (21.5)	195 (22.5)	56 (18.5)	0.15
TBI-related non-discharge	249 (21.3)	192 (22.1)	57 (18.9)	0.23
TBI-related death	43 (3.7)	32 (3.7)	11 (3.6)	>0.95

*EMS providers check specific criteria orderly from physiologic, anatomical, and mechanism of injury. If the preceding criteria are satisfied, the information of the latter criteria is not collected.

IQR, interquartile range; SBP, systolic blood pressure; RR, respiratory rate; ED, emergency department; TBI, traumatic brain injury.

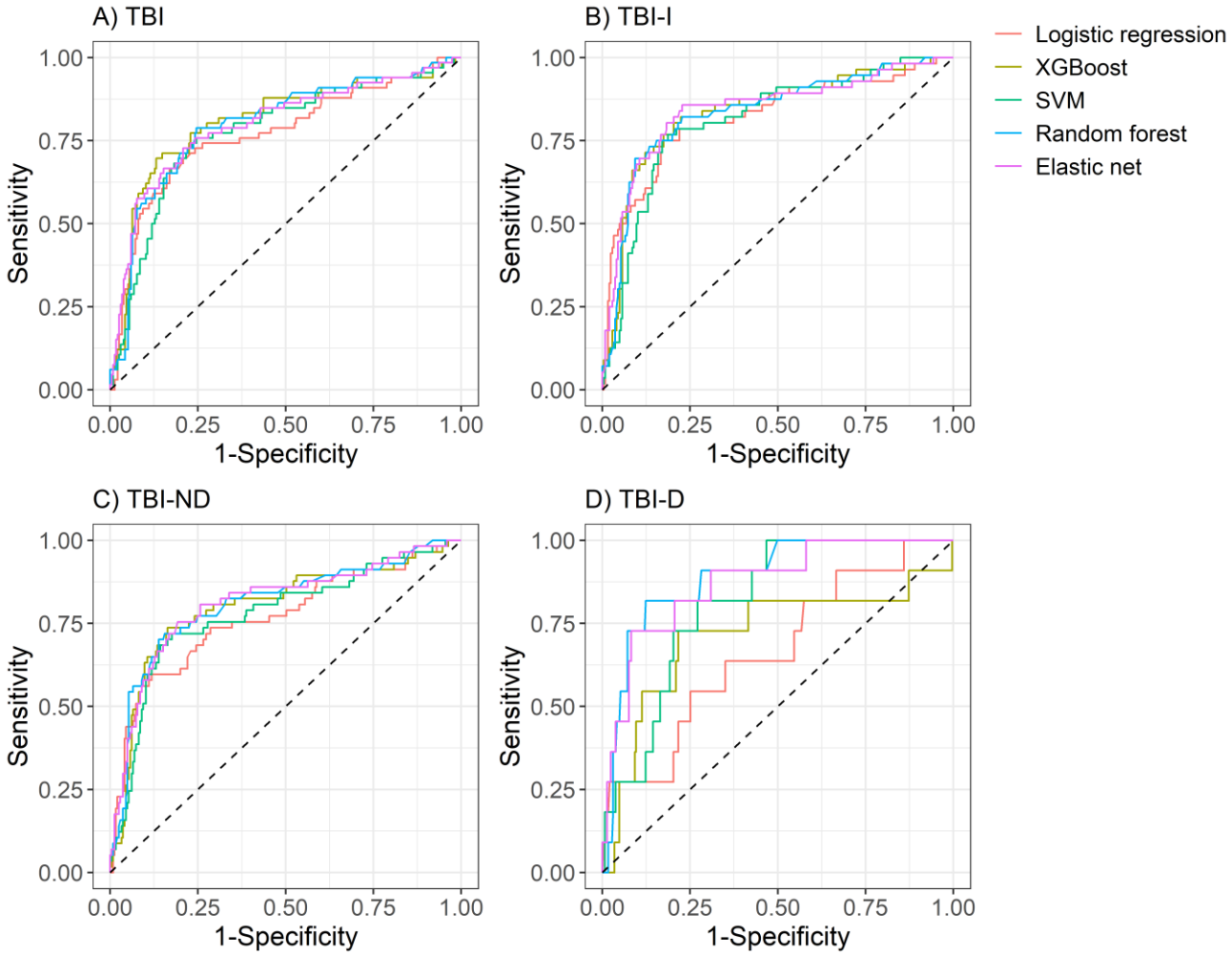
Supplementary Table 3. Hyperparameters of the final prediction models*

Model	Outcome	Hyperparameters
Elastic net	TBI	alpha: 0.325, lambda: 0.07506346
	TBI-I	alpha: 0.325, lambda: 0.07506346
	TBI-ND	alpha: 0.325, lambda: 0.07017153
	TBI-D	alpha: 0.325, lambda: 0.01565599
Random forest	TBI	ntree:500, mtry: 18
	TBI-I	ntree:500, mtry: 18
	TBI-ND	ntree:500, mtry: 18
	TBI-D	ntree:500, mtry: 15
Support vector machine	TBI	sigma: 0.008047; C: 4
	TBI-I	sigma: 0.008047; C: 4
	TBI-ND	sigma: 0.008047; C: 4
	TBI-D	sigma: 0.008047; C: 4
Extreme gradient boosting	TBI	nrounds: 299; max_depth: 1; eta: 0.4807096; gamma: 2.336623; colsample_bytree: 0.3657893; min_child_weight: 8; subsample: 0.8182623
	TBI-I	nrounds: 299; max_depth: 1; eta: 0.4807096; gamma: 2.336623; colsample_bytree: 0.3657893; min_child_weight: 8; subsample: 0.8182623
	TBI-ND	nrounds: 301; max_depth: 1; eta: 0.02154674; gamma: 4.696105; colsample_bytree: 0.590754; min_child_weight: 1; subsample: 0.5070866
	TBI-D	nrounds: 50; max_depth: 0.3; eta: 0.3; gamma: 0; colsample_bytree: 0.8; min_child_weight: 1; subsample: 0.5510204

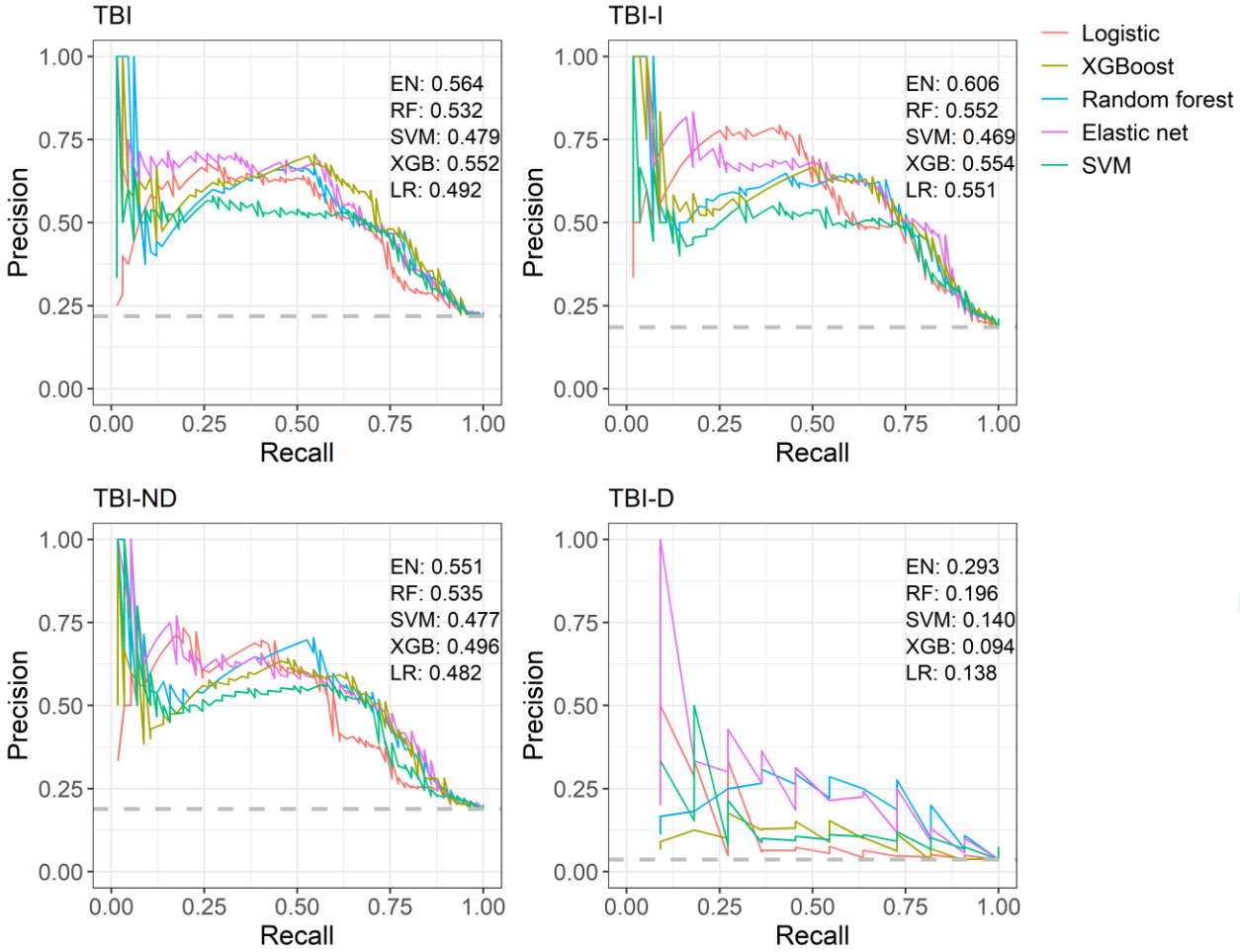
*Aside from the hyperparameters mentioned, all other hyperparameters are used as the default value.

TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death.

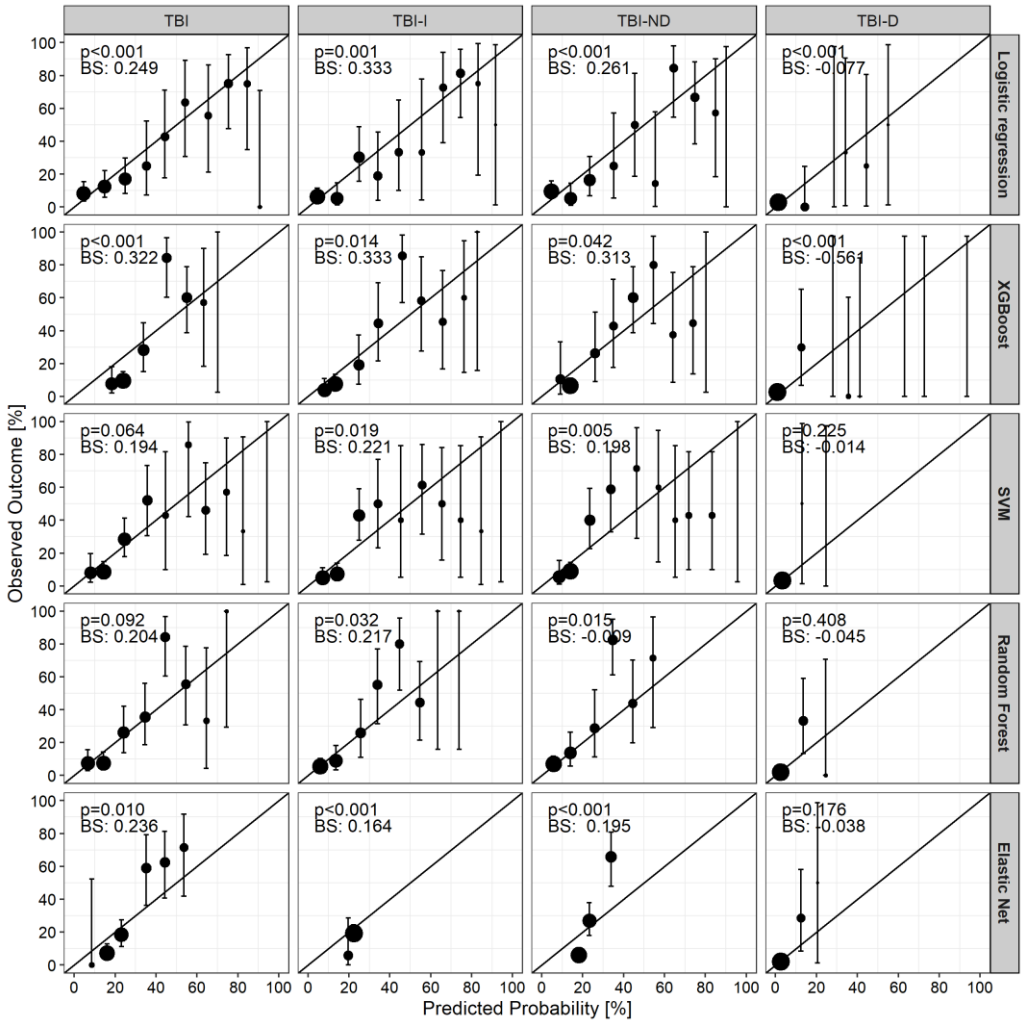
Supplementary Figure 1. Receiver operating characteristics of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death.



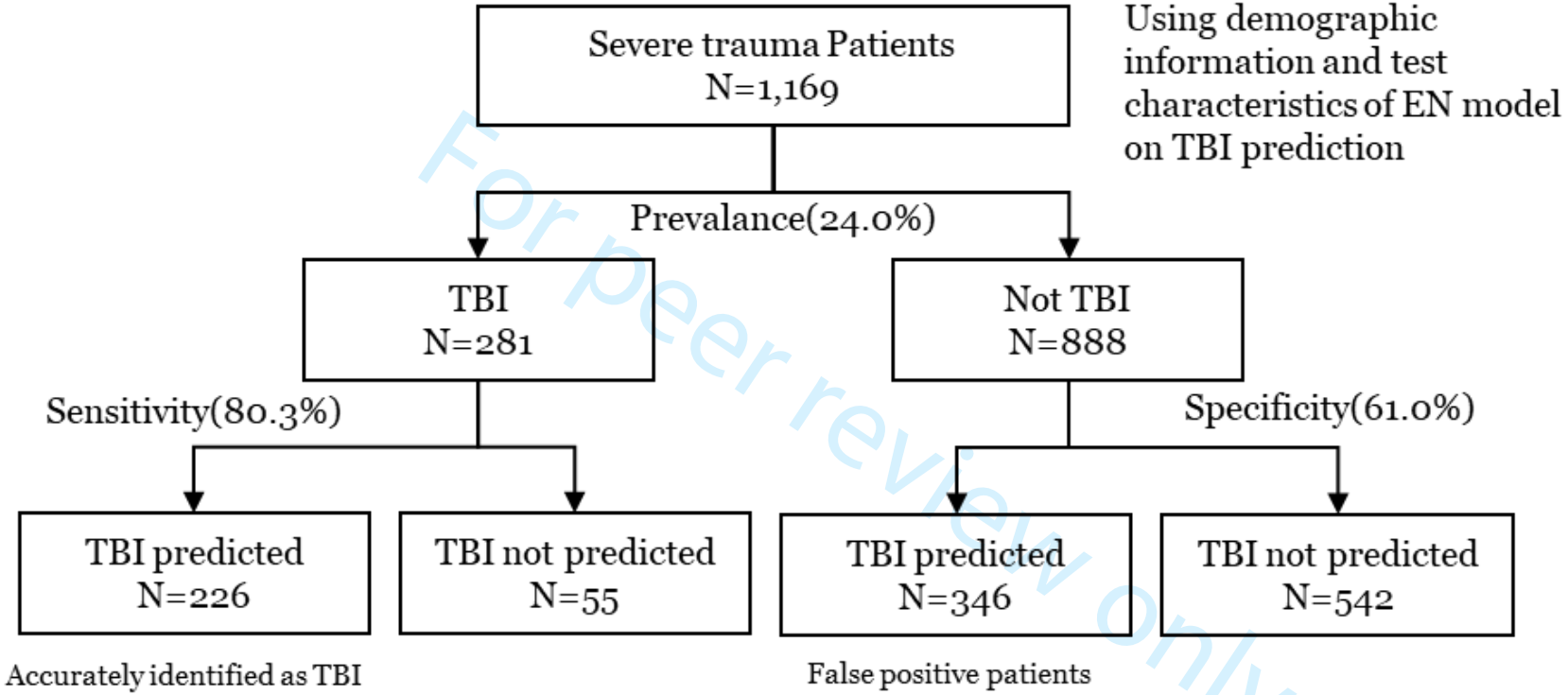
Supplementary Figure 2. Precision-recall curve of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; LR, logistic regression analysis; XGB, extreme gradient boosting; RF, random forest, EN, elastic net.



Supplementary Figure 3. Calibration plot of prediction models according to outcomes. TBI, traumatic brain injury; TBI-I, TBI with intracranial hemorrhage or injury; TBI-ND, TBI non-discharge; TBI-D, TBI with death; p, p-value of Hosmer-Lemeshow test; BS, scaled Brier score.



Supplementary Table 4. Example of calculating false-positive patients for accurately identified patients. TBI, traumatic brain injury; EN, elastic net.



False-positive patients for every 10 patients that are accurately identified as TBI :
346/226 x 10 = 15.3, rounded up to 16 patients

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	4
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	7
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	8
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	8-9
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	9
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	8-9
	5b	D;V	Describe eligibility criteria for participants.	10
	5c	D;V	Give details of treatments received, if relevant.	N/A
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	10-11
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	N/A
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	11
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A
Sample size	8	D;V	Explain how the study size was arrived at.	14
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	11
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	11
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	11-12
	10c	V	For validation, describe how the predictions were calculated.	12-13
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	12-13
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	12
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	14
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	14
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	14
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	14
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	N/A
	15b	D	Explain how to the use the prediction model.	14-15
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	14-15
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	19-20
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	16-17
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	16
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	18-19
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Suppl
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	20

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.