

Original Paper

Institution-Specific Machine Learning Models for Prehospital Assessment to Predict Hospital Admission: Prediction Model Development Study

Toru Shirakawa^{1,2}, MD; Tomohiro Sonoo^{2,3}, MD; Kentaro Ogura^{2,4}; Ryo Fujimori^{2,4}; Konan Hara^{2,5}, MD, PhD; Tadahiro Goto^{2,6}, MD, MPH; Hideki Hashimoto³, MD; Yuji Takahashi³, MD; Hiromu Naraba³, MD; Kensuke Nakamura^{3,7}, MD, PhD

¹Department of Public Health, Graduate School of Medicine, Osaka University, Suita, Japan

²TXP Medical Co, Ltd, Chuo-ku, Japan

³Department of Emergency Medicine, Hitachi General Hospital, Hitachi, Japan

⁴Faculty of Medicine, The University of Tokyo, Bunkyo-ku, Japan

⁵Department of Public Health, The University of Tokyo, Bunkyo-ku, Japan

⁶Department of Clinical Epidemiology and Health Economics, School of Public Health, The University of Tokyo, Bunkyo-ku, Japan

⁷Department of Emergency Medicine, The University of Tokyo, Bunkyo-ku, Japan

Corresponding Author:

Tomohiro Sonoo, MD

Department of Emergency Medicine

Hitachi General Hospital

Jounan-cho 2-1-1

Hitachi, 317-0077

Japan

Phone: 81 294 23 1111

Email: tomohiro.sonoo@txpmedical.com

Abstract

Background: Although multiple prediction models have been developed to predict hospital admission to emergency departments (EDs) to address overcrowding and patient safety, only a few studies have examined prediction models for prehospital use. Development of institution-specific prediction models is feasible in this age of data science, provided that predictor-related information is readily collectable.

Objective: We aimed to develop a hospital admission prediction model based on patient information that is commonly available during ambulance transport before hospitalization.

Methods: Patients transported by ambulance to our ED from April 2018 through March 2019 were enrolled. Candidate predictors were age, sex, chief complaint, vital signs, and patient medical history, all of which were recorded by emergency medical teams during ambulance transport. Patients were divided into two cohorts for derivation (3601/5145, 70.0%) and validation (1544/5145, 30.0%). For statistical models, logistic regression, logistic lasso, random forest, and gradient boosting machine were used. Prediction models were developed in the derivation cohort. Model performance was assessed by area under the receiver operating characteristic curve (AUROC) and association measures in the validation cohort.

Results: Of 5145 patients transported by ambulance, including deaths in the ED and hospital transfers, 2699 (52.5%) required hospital admission. Prediction performance was higher with the addition of predictive factors, attaining the best performance with an AUROC of 0.818 (95% CI 0.792-0.839) with a machine learning model and predictive factors of age, sex, chief complaint, and vital signs. Sensitivity and specificity of this model were 0.744 (95% CI 0.716-0.773) and 0.745 (95% CI 0.709-0.776), respectively.

Conclusions: For patients transferred to EDs, we developed a well-performing hospital admission prediction model based on routinely collected prehospital information including chief complaints.

(*JMIR Med Inform* 2020;8(10):e20324) doi: [10.2196/20324](https://doi.org/10.2196/20324)

KEYWORDS

prehospital; prediction; hospital admission; emergency medicine; machine learning; data science

Introduction

For patients being transported to an emergency department (ED), predicting hospital admission is important for providing high-quality care. Choosing the appropriate destination hospital with available beds can enhance efficient resource utilization in the context of integrated community health care [1]. Furthermore, accurate risk stratification during transportation can be expected to curb the risk of ED overcrowding and reduce ambulance turnaround times when implemented at hospitals [2].

Although multiple prediction models have been developed to predict hospital admission for ED use [3-11] to address overcrowding and patient safety [12-15], few studies have examined prediction models for prehospital use. Previously reported prehospital prediction models have been limited to patients with a specific disease or to models predicting critically ill conditions or mortality [16-23]. Several studies in the United States and United Kingdom have demonstrated the predictive performance of ED disposition, including hospitalization for general patients transferred by ambulance [24-26]. Nevertheless, these studies were not based on statistical models but on subjective prediction by ambulance staff. Therefore, they have limited generalizability across emergency medical systems and countries. Another study, conducted in Sweden, assessed a prehospital prediction model of hospital admission [27]. However, its predictors included more than 1000 distinct question and answer combinations recorded in a clinical decision support system used at a dispatch center. Therefore, its scalability might not be readily achievable.

Given this context, we aimed to develop prehospital prediction models of hospital admission using machine learning techniques and conventional logistic regression, based on replicable measurements such as chief complaints, vital signs, and past medical histories, which can all be collected routinely in an ambulance in any country. Our goal was to develop an institution-specific model based on readily collectable data with sufficient predictive performance, not a universal model that has broad generalizability.

Methods

Study Design and Setting

This prognostic study used data obtained at a tertiary care hospital in Japan from April 2018 to March 2019. The hospital covers approximately 3 million local residents. Annually, the hospital has about 20,000-25,000 visits, including 5500-6500 ambulance visits. The study protocol was approved by the Ethics Committee of the hospital. They waived informed consent because of the characteristics of the retrospective study design.

Study Participants

We enrolled patients who had been transported to our ED by ambulance. We excluded children aged 6 years or younger because of the difficulty in taking chief complaints and

measuring vital signs such as blood pressure. Patients with cardiopulmonary arrest were not excluded from analyses, thereby facilitating comparison with earlier studies that included patients with cardiopulmonary arrest and examined the predictive performance of ambulance staff [24-26].

Patient Information in the Prehospital Setting

Vital signs were measured at the scene when the patient was placed in the ambulance. After emergency medical service (EMS) staff members recorded patient information and conditions during transportation, they transmitted the information via telephone to ED staff members at the destination hospital. This information was input into an ED database through the Next Stage ER system (TXP Medical Co, Ltd), which structures information related to the chief complaint and past medical history with flexible input templates and a minor natural language processing algorithm [28]. The recorded chief complaint was translated automatically into 231 chief complaint categories based on the Japan Triage and Acuity Scale (JTAS) [29], which was developed based on the Canadian Triage and Acuity System [30]. Past medical histories were encoded corresponding to the International Statistical Classification of Diseases, 10th Revision (ICD-10) codes [31].

Candidate Predictors

Candidate predictors were age, sex, chief complaints, prehospital vital signs, and past medical histories. Although chief complaints were grouped into 231 categories based on JTAS, 75 complaints were not observed (ie, none of the included patients presented with these complaints). Therefore, 156 complaints were used. Vital signs include the level of consciousness, systolic blood pressure, diastolic blood pressure, pulse rate, respiratory rate, body temperature, and oxygen saturation with oxygen administration during transportation. The level of consciousness was assessed according to the Japan Coma Scale, which can be summarized briefly into four categories of alert, possible eye opening but not lucid, possible eye opening upon stimulation, and no eye opening and coma [32]. Past medical histories were grouped using the first 3 characters (1 alphabet letter and 2 digits) of the ICD-10 code. The 156 chief complaints and 505 past medical histories observed in our study were encoded to dummy variables. In all, 832 predictors were identified as candidate predictors.

Outcome Measures

The primary outcome was the composite of hospitalization, transfer to other care facilities, and death at the ED. These outcomes were recorded at the time patients left the ED. Sensitivity analysis was performed by excluding mortality from the hospitalization outcomes.

Data Analysis

Model Development

To predict hospital admission, we developed four models using candidate predictors as explained above: (1) logistic regression, (2) logistic regression with lasso penalization (logistic lasso),

(3) random forest [33], and (4) gradient boosting machine (GBM) [34]. For the GBM model, we used the extreme gradient boosting (XGBoost) implementation [35]. For each model, to evaluate the incremental benefit of adding each predictor, we further developed four models according to the predictors. Model 1 consists of age and sex only. Model 2 further includes 156 chief complaints. Model 3 further includes vital signs. Model 4 further includes 505 past medical histories. These modalities were designed according to the typical temporal order of information collection processing: call by a patient or bystander, arrival of an emergency medical team, and examination in the ambulance.

Feature Processing

To account for potential nonlinear relations between continuous features and the risk of hospital admission, we categorized the values of age and vital signs into deciles for logistic regression and logistic lasso. Since random forest and GBM can accommodate the nonlinear relations, we used continuous age and vital signs in those models.

Study Cohorts and Missing Values

We used 70.0% (3601/5145) of the available data for the derivation cohort. The remaining 30.0% (1544/5145) of data were used for the validation cohort. We divided patients into the two groups by random allocation. Hyperparameters for machine learning models were determined using a grid search with 5-fold cross-validation in the derivation cohort. Among the 5145 patients, frequencies (proportions) of missing values were 25 (0.5%) for sex, 552 (10.7%) for orientation, 593 (11.5%) for systolic blood pressure, 647 (12.6%) for diastolic blood pressure, 511 (9.9%) for pulse rate, 1152 (22.4%) for respiratory rate, 1040 (20.2%) for oxygen saturation, and 1086 (21.1%) for body temperature. The number of patients with at least one missing vital sign was 2174 (42.3%). To address the missing data, we used a missing indicator for logistic regression and lasso, assigned 0 for random forest, and left missing data in GBM, for which XGBoost can accommodate missing values.

Model Validation

In the validation cohort, we examined the prediction ability of the models by calculating the area under the receiver operating

characteristic curve (AUROC) and the area under the precision-recall curve (AUPRC). Calibration of the models was depicted by plotting predicted probabilities and the observed admission rates according to deciles of the predicted probabilities. Sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were estimated with predictors in the most accurate model at the threshold probability that maximizes the Youden indices [36].

For comparison with earlier studies of hospital admission prediction in the ED including walk-in patients and those transported by ambulance [3-11], we evaluated the prediction performance of the model described above including walk-in patients.

All analyses were conducted using Python 3.7 with scikit-learn [37], XGBoost [35], and tableone [38] packages. We used 200 bootstrap samples to calculate 95% confidence intervals for performance measures. Two-tailed *P* values of <.05 were inferred as statistically significant.

Results

During the study period, 5530 patients were transported to our ED by ambulance. From these, we excluded 385 visits by patients aged 6 years or younger. In all, 5145 visits were included in the analyses. Among the 5145 visits with ambulance transport, 2507 visits (48.7%) led to hospital admission, 96 visits (1.9%) led to death in the ED, and 96 visits (1.9%) required hospital transfer. The number of patients who required hospital admission, died in the ED, or required hospital transfer was 1889 of 3601 patients (52.5%) in the derivation cohort and 810 of 1544 patients (52.5%) in the validation cohort. Compared to patients who were not admitted to the hospital, patients who were admitted to the hospital (including those who died or were transferred) had worse vital signs (eg, lower level of consciousness, lower blood pressure). Moreover, they were older, were likely to have altered mental status or fever, and were likely to have a history of circulatory and respiratory system symptoms (Table 1).

Table 1. Baseline characteristics of patients according to hospital admission status.

Characteristic	Admission		P value
	No (n=2446)	Yes (n=2699)	
Age (years), mean (SD)	63.0 (23.7)	73.4 (16.2)	<.001
Male sex, n (%)	1308 (53.5)	1591 (58.9)	.25
Selected chief complaint, n (%)^a			
Altered mental status	220 (9.0)	373 (13.8)	<.001
Dyspnea	176 (7.2)	403 (14.9)	.75
Chest pain	188 (7.7)	200 (7.4)	.54
Abdominal pain	169 (6.9)	174 (6.4)	<.001
Fever	97 (4.0)	175 (6.5)	<.001
Vital signs			
Level of consciousness, n (%)			
Alert	1619 (66.2)	1338 (49.6)	<.001
Possible eye opening, not lucid	427 (17.5)	642 (23.8)	<.001
Possible eye opening upon stimulation	85 (3.5)	228 (8.4)	<.001
No eye opening and coma	36 (1.5)	218 (8.1)	<.001
Systolic blood pressure (mm Hg), mean (SD)	148.0 (33.2)	140.7 (43.8)	<.001
Diastolic blood pressure (mm Hg), mean (SD)	83.5 (21.2)	80.8 (27.5)	<.001
Pulse rate (bpm), mean (SD)	88.7 (21.3)	92.7 (24.7)	<.001
Respiratory rate (bpm), mean (SD)	21.5 (5.3)	22.8 (6.2)	<.001
Body temperature (°C), mean (SD)	36.8 (7.6)	36.9 (2.7)	.76
Oxygen saturation (%), mean (SD)	97.0 (3.1)	93.8 (7.7)	<.001
Oxygen administration during transportation, n (%)	287 (11.7)	947 (35.1)	<.001
Selected past medical history, n (%)^a			
R09 Other symptoms and signs involving the circulatory and respiratory systems	624 (25.5)	805 (29.8)	<.001
E11 Type 2 diabetes mellitus	419 (17.1)	582 (21.6)	<.001
I63 Cerebral infarction	202 (8.3)	303 (11.2)	.001
E78 Disorders of lipoprotein metabolism and other lipidemias	151 (6.2)	212 (7.9)	.02
I51 Complications and ill-defined descriptions of heart disease	159 (6.5)	180 (6.7)	.85

^aThe five most frequent chief complaints and past medical history items are shown.

Overall, the GBM model achieved the highest AUROCs and AUPRCs in models 3 and 4 (Tables 2 and 3). The most accurate model was GBM in model 3, with AUROC of 0.818 (95% CI 0.792-0.839), AUPRC of 0.831 (95% CI 0.804-0.855), sensitivity of 0.744 (95% CI 0.716-0.773), and specificity of 0.745 (95% CI 0.709-0.776) (Tables 2-4). The highest AUROC of logistic regression was 0.805 (95% CI 0.782-0.827) in model 3. It was lower in model 4: 0.750 (95% CI 0.720-0.774) (Figure 1). In models 2-4, precision-recall curve analysis showed

superior performance of machine learning models compared to that of logistic regression among patients with higher risk of hospital admission (Figure 2). The lasso and GBM showed good calibration in all models (Figure 3). Hyperparameters of machine learning models are shown in Table S1 in Multimedia Appendix 1. The exclusion of mortality at the ED showed slightly lower predictive performance, with AUROC of 0.803 (95% CI 0.775-0.823) for GBM in model 3 (Tables S2-S4 in Multimedia Appendix 1).

Table 2. Areas under the receiver operating characteristic curve and 95% confidence intervals of hospital admission prediction models according to machine learning methods and prediction models.

Model type	Model 1 ^a	Model 2 ^b	Model 3 ^c	Model 4 ^d
Logistic regression	0.631 (0.602-0.657)	0.750 (0.723-0.774)	0.805 (0.782-0.827)	0.750 (0.720-0.774)
Lasso	0.631 (0.602-0.657)	0.755 (0.730-0.779)	0.817 (0.793-0.839)	0.811 (0.787-0.832)
Random forest	0.594 (0.567-0.619)	0.735 (0.710-0.763)	0.813 (0.786-0.834)	0.814 (0.786-0.833)
Gradient boosting machine	0.624 (0.598-0.652)	0.758 (0.734-0.783)	0.818 (0.792-0.839)	0.815 (0.788-0.833)

^aModel 1: Age and sex.^bModel 2: Age, sex, and chief complaint.^cModel 3: Age, sex, chief complaint, and vital signs.^dModel 4: Age, sex, chief complaint, vital signs, and past medical history.**Table 3.** Areas under the precision-recall curve and 95% confidence intervals of hospital admission prediction models according to machine learning models and predictor modalities.

Model type	Model 1 ^a	Model 2 ^b	Model 3 ^c	Model 4 ^d
Logistic regression	0.614 (0.578-0.653)	0.729 (0.700-0.766)	0.794 (0.764-0.827)	0.709 (0.667-0.744)
Lasso	0.614 (0.578-0.654)	0.766 (0.737-0.795)	0.829 (0.805-0.853)	0.820 (0.793-0.845)
Random forest	0.580 (0.550-0.615)	0.734 (0.703-0.770)	0.828 (0.802-0.853)	0.828 (0.801-0.851)
Gradient boosting machine	0.609 (0.580-0.647)	0.766 (0.734-0.799)	0.831 (0.804-0.855)	0.828 (0.803-0.852)

^aModel 1: Age and sex.^bModel 2: Age, sex, and chief complaint.^cModel 3: Age, sex, chief complaint, and vital signs.^dModel 4: Age, sex, chief complaint, vital signs, and past medical history.**Table 4.** Measures of predictive performance and 95% confidence intervals for prediction model 3 at optimal thresholds^a.

Model type	Sensitivity	Specificity	PPV ^b	NPV ^c	Accuracy
Logistic regression	0.760 (0.724-0.786)	0.731 (0.698-0.766)	0.752 (0.728-0.783)	0.737 (0.700-0.771)	0.746 (0.723-0.770)
Lasso	0.724 (0.684-0.751)	0.774 (0.742-0.800)	0.776 (0.748-0.803)	0.722 (0.683-0.752)	0.749 (0.723-0.767)
Random forest	0.720 (0.687-0.749)	0.777 (0.742-0.804)	0.776 (0.745-0.807)	0.718 (0.685-0.748)	0.745 (0.720-0.768)
Gradient boosting machine	0.736 (0.696-0.768)	0.743 (0.712-0.772)	0.757 (0.726-0.785)	0.721 (0.680-0.756)	0.739 (0.713-0.765)

^aPredictors were age, sex, chief complaint, and vital signs.^bPPV: positive predictive value.^cNPV: negative predictive value.

Figure 1. ROC curves of hospital admission prediction models. ROC curves for the three machine learning models are similar to those of logistic regression in models 1, 2, and 3, but superior to those of logistic regression in model 4. ROC: receiver operating characteristic.

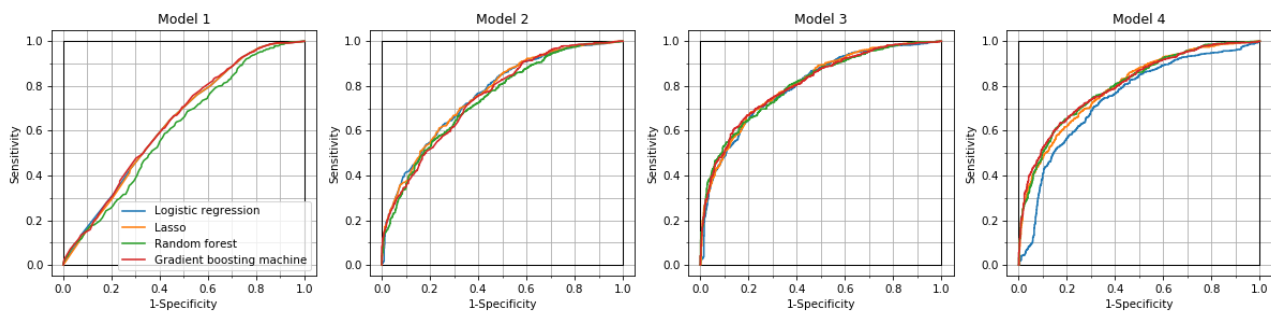


Figure 2. Precision-recall curves of hospital admission prediction models. Precision-recall curves of the three machine learning models are similar. Logistic regression model showed inferior performance for patients with higher predicted probabilities (left side on the horizontal axis) in models 2, 3, and 4.

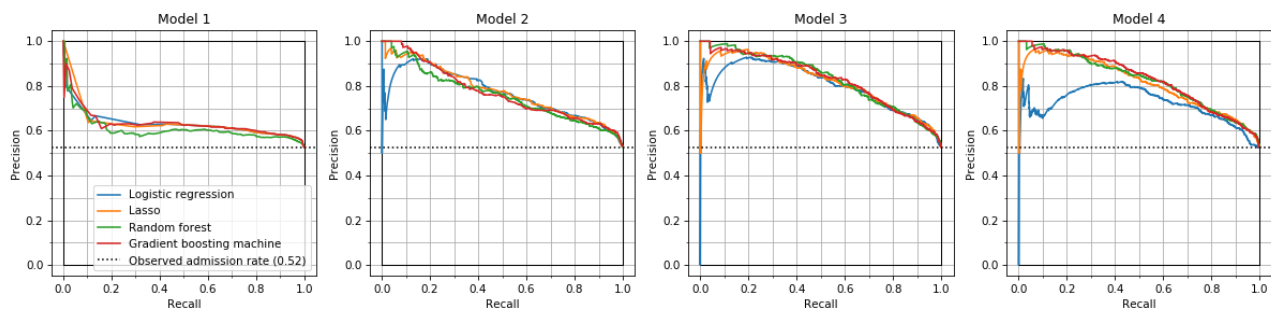
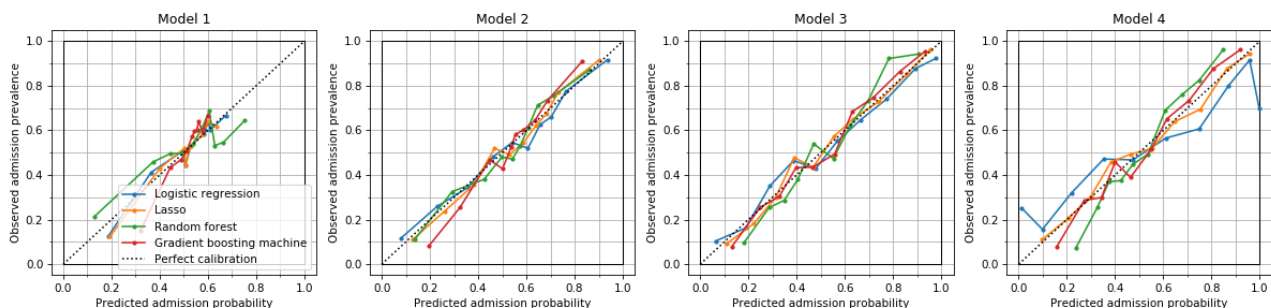


Figure 3. Calibration curves of hospital admission prediction models. Lasso and gradient boosting machine showed good calibration in all models. Logistic regression was ill-calibrated for patients with the lowest and the highest deciles of predicted probability in model 4.



A GBM model with data of both walk-in and ambulance visits to our ED during the study period ($n=16,857$) demonstrated higher performance than that for patients transported by ambulance, with AUROC of 0.873 (95% CI 0.860-0.883), sensitivity of 0.830 (95% CI 0.807-0.850), and specificity of 0.743 (95% CI 0.712-0.772) in the validation set.

Discussion

Principal Findings

To our knowledge, this report is the first of a study developing and validating prediction models for hospital admission based on common prehospital information for patients transported to EDs by ambulance. Information used for this study was collected in prehospital settings within a routine clinical practice. Therefore, the method of the prediction model development is readily applicable to other facilities that support clinical decision making by EMSs.

Our results are comparable to those presented in earlier reports describing the performance of subjective prediction by ambulance staff for patients they transported. A prospective study in the United Kingdom revealed a response rate of 99.7% (396/397). Analyses of 396 cases demonstrated sensitivity of 0.717 (95% CI 0.65-0.78) and specificity of 0.770 (95% CI 0.71-0.81) [24]. Another prospective study conducted in the United States found a response rate of 24.6% (101/411) from the cases analyzed [25]. Sensitivity of prediction by EMS staff members was 0.733 (95% CI 0.658-0.798), and the specificity was 0.850 (95% CI 0.798-0.891). Another study in the United States examined 932 transports to a hospital and reported the performance of EMS staff prediction of hospitalization as 0.62 (95% CI 0.54-0.68) for sensitivity and 0.89 (95% CI 0.86-0.91) for specificity [26]. However, prediction by EMS staff in this study was done at the time they left the ED. The results might be affected by incorporation bias because of observation or direct discussion with physicians and nurses in the ED. Therefore, the true performance might be lower. These studies

are based on the impressions of paramedics. Therefore, their performance in other emergency medical systems remains unknown. However, our method relies on common prehospital measurements, which present the benefit of applicability to other standard emergency medical systems.

The AUROC achieved using the proposed model was lower than those reported from earlier studies for patients after arrival at the ED, reporting values of 0.80-0.87 [3-11]. However, these earlier models included both walk-in and ambulance patients. Because our prediction model was restricted to patients transferred by ambulance, the target population might be more severely affected by health issues than walk-in patients, making it difficult to discriminate patients who need inpatient care and patients who do not. Indeed, prediction performance including both walk-in and ambulance visits to our ED demonstrated comparable performance to that of an AUROC of 0.873.

The logistic regression model demonstrated comparable performance to that obtained with other machine learning models, with <0.02 difference in AUROCs in models 1-3 and lower performance in model 4. Two recent reports have described similar predictive performance in logistic regression and machine learning models for predicting hospital admission after ED visits [39,40]. However, the ratios of the number of variables to the number of patients were smaller in those studies than in this study: previous studies reported 972 variables to 560,486 patients [39], and 111 variables to 1,721,294 patients [40], whereas this study reported 832 variables to 5145 patients. The lower predictive performance of logistic regression can be attributed to overfitting. By selecting important predictors by lasso or other methods, a logistic regression model might be built with comparable performance to those of other machine learning models, as suggested by our result obtained for lasso, which virtually reduces the number of variables in logistic regression.

Acknowledgments

The authors are grateful to Dr Kei Taneishi at RIKEN Advanced Institute for Computational Science for valuable comments and discussion.

Conflicts of Interest

T Shirakawa, T Sonoo, KO, RF, TG, and KH are owners or employees of TXP Medical Co, Ltd.

Multimedia Appendix 1

Supplementary tables.

[DOCX File, 17 KB-Multimedia Appendix 1]

References

1. EMS Agenda 2050. U.S. Department of Transportation NHTSA. 2019. URL: <https://www.ems.gov/projects/ems-agenda-2050.html> [accessed 2020-10-15]
2. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo CA. A conceptual model of emergency department crowding. *Ann Emerg Med* 2003 Aug;42(2):173-180. [doi: [10.1067/mem.2003.302](https://doi.org/10.1067/mem.2003.302)] [Medline: [12883504](https://pubmed.ncbi.nlm.nih.gov/12883504/)]
3. Parker CA, Liu N, Wu SX, Shen Y, Lam SSW, Ong MEH. Predicting hospital admission at the emergency department triage: A novel prediction model. *Am J Emerg Med* 2019 Aug;37(8):1498-1504 [FREE Full text] [doi: [10.1016/j.ajem.2018.10.060](https://doi.org/10.1016/j.ajem.2018.10.060)] [Medline: [30413365](https://pubmed.ncbi.nlm.nih.gov/30413365/)]

4. Kraaijvanger N, Rijpsma D, Roovers L, van Leeuwen H, Kaasjager K, van den Brand L, et al. Development and validation of an admission prediction tool for emergency departments in the Netherlands. *Emerg Med J* 2018 Aug;35(8):464-470. [doi: [10.1136/emmermed-2017-206673](https://doi.org/10.1136/emmermed-2017-206673)] [Medline: [29627769](https://pubmed.ncbi.nlm.nih.gov/29627769/)]
5. Barak-Corren Y, Israelit SH, Reis BY. Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. *Emerg Med J* 2017 May;34(5):308-314. [doi: [10.1136/emmermed-2014-203819](https://doi.org/10.1136/emmermed-2014-203819)] [Medline: [28188202](https://pubmed.ncbi.nlm.nih.gov/28188202/)]
6. Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. *Acad Emerg Med* 2011 Aug;18(8):844-850. [doi: [10.1111/j.1553-2712.2011.01125.x](https://doi.org/10.1111/j.1553-2712.2011.01125.x)] [Medline: [21843220](https://pubmed.ncbi.nlm.nih.gov/21843220/)]
7. Lucke JA, de Gelder J, Clarijs F, Heringhaus C, de Craen AJM, Fogteloo AJ, et al. Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years. *Emerg Med J* 2018 Jan;35(1):18-27. [doi: [10.1136/emmermed-2016-205846](https://doi.org/10.1136/emmermed-2016-205846)] [Medline: [28814479](https://pubmed.ncbi.nlm.nih.gov/28814479/)]
8. Rahimian F, Salimi-Khorshidi G, Payberah AH, Tran J, Ayala Solares R, Raimondi F, et al. Predicting the risk of emergency admission with machine learning: Development and validation using linked electronic health records. *PLoS Med* 2018 Nov;15(11):e1002695 [FREE Full text] [doi: [10.1371/journal.pmed.1002695](https://doi.org/10.1371/journal.pmed.1002695)] [Medline: [30458006](https://pubmed.ncbi.nlm.nih.gov/30458006/)]
9. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. *Emerg Med J* 2015 Mar;32(3):174-179. [doi: [10.1136/emmermed-2013-203200](https://doi.org/10.1136/emmermed-2013-203200)] [Medline: [24421344](https://pubmed.ncbi.nlm.nih.gov/24421344/)]
10. Peck JS, Gaehde SA, Nightingale DJ, Gelman DY, Huckins DS, Lemons MF, et al. Generalizability of a simple approach for predicting hospital admission from an emergency department. *Acad Emerg Med* 2013 Nov;20(11):1156-1163 [FREE Full text] [doi: [10.1111/acem.12244](https://doi.org/10.1111/acem.12244)] [Medline: [24238319](https://pubmed.ncbi.nlm.nih.gov/24238319/)]
11. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care* 2019 Feb 22;23(1):64 [FREE Full text] [doi: [10.1186/s13054-019-2351-7](https://doi.org/10.1186/s13054-019-2351-7)] [Medline: [30795786](https://pubmed.ncbi.nlm.nih.gov/30795786/)]
12. Bernstein SL, Aronsky D, Duseja R, Epstein S, Handel D, Hwang U, Society for Academic Emergency Medicine, Emergency Department Crowding Task Force. The effect of emergency department crowding on clinically oriented outcomes. *Acad Emerg Med* 2009 Jan;16(1):1-10 [FREE Full text] [doi: [10.1111/j.1553-2712.2008.00295.x](https://doi.org/10.1111/j.1553-2712.2008.00295.x)] [Medline: [19007346](https://pubmed.ncbi.nlm.nih.gov/19007346/)]
13. Fatovich DM, Nagree Y, Sprivilis P. Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia. *Emerg Med J* 2005 May;22(5):351-354. [doi: [10.1136/emj.2004.018002](https://doi.org/10.1136/emj.2004.018002)] [Medline: [15843704](https://pubmed.ncbi.nlm.nih.gov/15843704/)]
14. McCarthy ML, Zeger SL, Ding R, Levin SR, Desmond JS, Lee J, et al. Crowding delays treatment and lengthens emergency department length of stay, even among high-acuity patients. *Ann Emerg Med* 2009 Oct;54(4):492-503.e4. [doi: [10.1016/j.annemergmed.2009.03.006](https://doi.org/10.1016/j.annemergmed.2009.03.006)] [Medline: [19423188](https://pubmed.ncbi.nlm.nih.gov/19423188/)]
15. Schull MJ, Kiss A, Szalai J. The effect of low-complexity patients on emergency department waiting times. *Ann Emerg Med* 2007 Mar;49(3):257-64, 264.e1. [doi: [10.1016/j.annemergmed.2006.06.027](https://doi.org/10.1016/j.annemergmed.2006.06.027)] [Medline: [17049408](https://pubmed.ncbi.nlm.nih.gov/17049408/)]
16. Uchida K, Yoshimura S, Hiyama N, Oki Y, Matsumoto T, Tokuda R, et al. Clinical Prediction Rules to Classify Types of Stroke at Prehospital Stage. *Stroke* 2018 Aug;49(8):1820-1827 [FREE Full text] [doi: [10.1161/STROKEAHA.118.021794](https://doi.org/10.1161/STROKEAHA.118.021794)] [Medline: [30002147](https://pubmed.ncbi.nlm.nih.gov/30002147/)]
17. Peltan ID, Rowhani-Rahbar A, Vande Vusse LK, Caldwell E, Rea TD, Maier RV, et al. Development and validation of a prehospital prediction model for acute traumatic coagulopathy. *Crit Care* 2016 Nov 16;20(1):371 [FREE Full text] [doi: [10.1186/s13054-016-1541-9](https://doi.org/10.1186/s13054-016-1541-9)] [Medline: [27846895](https://pubmed.ncbi.nlm.nih.gov/27846895/)]
18. Pineskoski J, Kuisma M, Olkkola KT, Nurmi J. Prehospital National Early Warning Score predicts early mortality. *Acta Anaesthesiol Scand* 2019 May;63(5):676-683. [doi: [10.1111/aas.13310](https://doi.org/10.1111/aas.13310)] [Medline: [30623422](https://pubmed.ncbi.nlm.nih.gov/30623422/)]
19. Koyama S, Yamaguchi Y, Gibo K, Nakayama I, Ueda S. Use of prehospital qSOFA in predicting in-hospital mortality in patients with suspected infection: A retrospective cohort study. *PLoS One* 2019;14(5):e0216560 [FREE Full text] [doi: [10.1371/journal.pone.0216560](https://doi.org/10.1371/journal.pone.0216560)] [Medline: [31063494](https://pubmed.ncbi.nlm.nih.gov/31063494/)]
20. Williams TA, Tohira H, Finn J, Perkins GD, Ho KM. The ability of early warning scores (EWS) to detect critical illness in the prehospital setting: A systematic review. *Resuscitation* 2016 May;102:35-43. [doi: [10.1016/j.resuscitation.2016.02.011](https://doi.org/10.1016/j.resuscitation.2016.02.011)] [Medline: [26905389](https://pubmed.ncbi.nlm.nih.gov/26905389/)]
21. Kievlan DR, Martin-Gill C, Kahn JM, Callaway CW, Yealy DM, Angus DC, et al. External validation of a prehospital risk score for critical illness. *Crit Care* 2016 Aug 11;20(1):255 [FREE Full text] [doi: [10.1186/s13054-016-1408-0](https://doi.org/10.1186/s13054-016-1408-0)] [Medline: [27515164](https://pubmed.ncbi.nlm.nih.gov/27515164/)]
22. Jouffroy R, Saade A, Ellouze S, Carpentier A, Michaloux M, Carli P, et al. Prehospital triage of septic patients at the SAMU regulation: Comparison of qSOFA, MRST, MEWS and PRESEP scores. *Am J Emerg Med* 2018 May;36(5):820-824. [doi: [10.1016/j.ajem.2017.10.030](https://doi.org/10.1016/j.ajem.2017.10.030)] [Medline: [29056391](https://pubmed.ncbi.nlm.nih.gov/29056391/)]
23. van Rein EAJ, van der Sluijs R, Voskens FJ, Lansink KWW, Houwert RM, Lichtveld RA, et al. Development and Validation of a Prediction Model for Prehospital Triage of Trauma Patients. *JAMA Surg* 2019 May 01;154(5):421-429 [FREE Full text] [doi: [10.1001/jamasurg.2018.4752](https://doi.org/10.1001/jamasurg.2018.4752)] [Medline: [30725101](https://pubmed.ncbi.nlm.nih.gov/30725101/)]
24. Price TG, Hooker EA, Neubauer J. Prehospital provider prediction of emergency department disposition: implications for selective diversion. *Prehosp Emerg Care* 2005;9(3):322-325. [doi: [10.1080/10903120590962012](https://doi.org/10.1080/10903120590962012)] [Medline: [16147483](https://pubmed.ncbi.nlm.nih.gov/16147483/)]

25. Levine SD, Colwell CB, Pons PT, Gravitz C, Haukoos JS, McVaney KE. How well do paramedics predict admission to the hospital? A prospective study. *J Emerg Med* 2006 Jul;31(1):1-5. [doi: [10.1016/j.jemermed.2005.08.007](https://doi.org/10.1016/j.jemermed.2005.08.007)] [Medline: [16798145](https://pubmed.ncbi.nlm.nih.gov/16798145/)]
26. Clesham K, Mason S, Gray J, Walters S, Cooke V. Can emergency medical service staff predict the disposition of patients they are transporting? *Emerg Med J* 2008 Oct;25(10):691-694. [doi: [10.1136/emj.2007.054924](https://doi.org/10.1136/emj.2007.054924)] [Medline: [18843076](https://pubmed.ncbi.nlm.nih.gov/18843076/)]
27. Spangler D, Hermansson T, Smekal D, Blomberg H. A validation of machine learning-based risk scores in the prehospital setting. *PLoS One* 2019;14(12):e0226518 [FREE Full text] [doi: [10.1371/journal.pone.0226518](https://doi.org/10.1371/journal.pone.0226518)] [Medline: [31834920](https://pubmed.ncbi.nlm.nih.gov/31834920/)]
28. Sonoo T, Naraba H, Hashimoto H, Nakamura K, Morimura N. Development and evaluation of computer system that enables emergency department efficiency improvement and collection of coded data during normal clinical workflow. *J Japanese Assoc Acute Med* 2018;29(2):45-55. [doi: [10.1002/jja2.12276](https://doi.org/10.1002/jja2.12276)]
29. Kuriyama A, Ikegami T, Kaihara T, Fukuoka T, Nakayama T. Validity of the Japan Acuity and Triage Scale in adults: a cohort study. *Emerg Med J* 2018 Jun;35(6):384-388. [doi: [10.1136/emj.2017.207214](https://doi.org/10.1136/emj.2017.207214)] [Medline: [29535086](https://pubmed.ncbi.nlm.nih.gov/29535086/)]
30. Bullard MJ, Musgrave E, Warren D, Unger B, Skeldon T, Grierson R, et al. Revisions to the Canadian Emergency Department Triage and Acuity Scale (CTAS) Guidelines 2016. *CJEM* 2017 Jul;19(S2):S18-S27. [doi: [10.1017/cem.2017.365](https://doi.org/10.1017/cem.2017.365)] [Medline: [28756800](https://pubmed.ncbi.nlm.nih.gov/28756800/)]
31. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. World Health Organization. 1992. URL: <http://www.who.int/classifications/icd/en/bluebook.pdf> [accessed 2020-10-15]
32. Shigematsu K, Nakano H, Watanabe Y. The eye response test alone is sufficient to predict stroke outcome--reintroduction of Japan Coma Scale: a cohort study. *BMJ Open* 2013;3(4). [doi: [10.1136/bmjopen-2013-002736](https://doi.org/10.1136/bmjopen-2013-002736)] [Medline: [23633419](https://pubmed.ncbi.nlm.nih.gov/23633419/)]
33. Breiman L. Random forests. *Mach Learn* 2001;45(1):5-32. [doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
34. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat* 2001 Oct;29(5):1189-1232. [doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)]
35. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, New York, USA: Association for Computing Machinery; 2016 Presented at: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, California, USA p. 785-794. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
36. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950 Jan;3(1):32-35. [Medline: [15405679](https://pubmed.ncbi.nlm.nih.gov/15405679/)]
37. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 2011;12:2825-2830 [FREE Full text]
38. Pollard TJ, Johnson AEW, Raffa JD, Mark RG. tableone: An open source Python package for producing summary statistics for research papers. *JAMIA Open* 2018 Jul;1(1):26-31 [FREE Full text] [doi: [10.1093/jamiaopen/ooy012](https://doi.org/10.1093/jamiaopen/ooy012)] [Medline: [31984317](https://pubmed.ncbi.nlm.nih.gov/31984317/)]
39. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. *PLoS One* 2018;13(7):e0201016 [FREE Full text] [doi: [10.1371/journal.pone.0201016](https://doi.org/10.1371/journal.pone.0201016)] [Medline: [30028888](https://pubmed.ncbi.nlm.nih.gov/30028888/)]
40. Rendell K, Koprinska I, Kyme A, Ebker-White AA, Dinh MM. The Sydney Triage to Admission Risk Tool (START2) using machine learning techniques to support disposition decision-making. *Emerg Med Australas* 2019 Jun;31(3):429-435. [doi: [10.1111/1742-6723.13199](https://doi.org/10.1111/1742-6723.13199)] [Medline: [30469164](https://pubmed.ncbi.nlm.nih.gov/30469164/)]
41. Landi F, Onder G, Cesari M, Barillaro C, Lattanzio F, Carbonin PU, et al. Comorbidity and social factors predicted hospitalization in frail elderly patients. *J Clin Epidemiol* 2004 Aug;57(8):832-836. [doi: [10.1016/j.jclinepi.2004.01.013](https://doi.org/10.1016/j.jclinepi.2004.01.013)] [Medline: [15551473](https://pubmed.ncbi.nlm.nih.gov/15551473/)]
42. Clay OJ, Roth DL, Safford MM, Sawyer PL, Allman RM. Predictors of overnight hospital admission in older African American and Caucasian Medicare beneficiaries. *J Gerontol A Biol Sci Med Sci* 2011 Aug;66(8):910-916 [FREE Full text] [doi: [10.1093/gerona/glr082](https://doi.org/10.1093/gerona/glr082)] [Medline: [21565981](https://pubmed.ncbi.nlm.nih.gov/21565981/)]
43. Amegbor PM, Plumb KB, Rosenberg MW. Determinants of Overnight Stay in Health Centres and Length of Admission: A Study of Canadian Seniors. *Can J Aging* 2020 Feb 24:1-12. [doi: [10.1017/S0714980819000771](https://doi.org/10.1017/S0714980819000771)] [Medline: [32089138](https://pubmed.ncbi.nlm.nih.gov/32089138/)]
44. Friedman CP, Wong AK, Blumenthal D. Achieving a nationwide learning health system. *Sci Transl Med* 2010 Nov 10;2(57):57cm29. [doi: [10.1126/scitranslmed.3001456](https://doi.org/10.1126/scitranslmed.3001456)] [Medline: [21068440](https://pubmed.ncbi.nlm.nih.gov/21068440/)]
45. Konečný J, McMahan H, Ramage D, Richtárik P. Federated Optimization: Distributed Machine Learning for On-Device Intelligence. arXiv. 2016. URL: <http://arxiv.org/abs/1610.02527> [accessed 2020-10-15]

Abbreviations

- AUPRC:** area under the precision-recall curve
- AUROC:** area under the receiver operating characteristic curve
- ED:** emergency department
- EMS:** emergency medical service
- GBM:** gradient boosting machine
- ICD-10:** International Statistical Classification of Diseases, 10th Revision
- JTAS:** Japan Triage and Acuity Scale

XGBoost: extreme gradient boosting

Edited by G Eysenbach; submitted 04.06.20; peer-reviewed by R Lieu, Y Katayama; comments to author 29.06.20; revised version received 24.08.20; accepted 16.09.20; published 27.10.20

Please cite as:

*Shirakawa T, Sonoo T, Ogura K, Fujimori R, Hara K, Goto T, Hashimoto H, Takahashi Y, Naraba H, Nakamura K
Institution-Specific Machine Learning Models for Prehospital Assessment to Predict Hospital Admission: Prediction Model Development Study*

JMIR Med Inform 2020;8(10):e20324

URL: <http://medinform.jmir.org/2020/10/e20324/>

doi: [10.2196/20324](https://doi.org/10.2196/20324)

PMID: [33107830](https://pubmed.ncbi.nlm.nih.gov/33107830/)

©Toru Shirakawa, Tomohiro Sonoo, Kentaro Ogura, Ryo Fujimori, Konan Hara, Tadahiro Goto, Hideki Hashimoto, Yuji Takahashi, Hiromu Naraba, Kensuke Nakamura. Originally published in JMIR Medical Informatics (<http://medinform.jmir.org>), 27.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.