POLISH HEART JOURNAL

Kardiologia Polska
The Official Peer-reviewed Journal of the Polish Cardiac Society since 1957

Online first
This is a provisional PDF only. Copyedited and fully formatted version will be made available soon

ISSN 0022–9032 e-ISSN 1897–4279

Established machine learning models to predict readmission for elderly patients with ischemic heart disease

Authors: Xuewu Song, Feng Xian, Changyu Zhu, Yi Luo, Yilong Liu, Qing Wen, Rongsheng Tong

Article type: Original article

Received: December 28, 2023

Accepted: June 24, 2024

Early publication date: June 28, 2024

Established machine learning models to predict readmission for elderly patients with ischemic heart disease

Short title: Machine learning predicts readmission for ischemic heart disease

Xuewu Song[1], Feng Xian[2], Changyu Zhu[1], Yi Luo[1], Yilong Liu[1], Qing Wen[3], Rongsheng Tong[1]

[1]Department of Pharmacy, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China
[2]Department of Oncology, Nanchong Central Hospital, the second Clinical Medical College, North Sichuan Medical College, Nanchong, China
[3]Department of Cardiology, Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu, China

Correspondence to:
Rongsheng Tong, PhD,
Department of Pharmacy,
Sichuan Provincial People's Hospital,
University of Electronic Science and Technology of China,
32 West Second Section, First Ring Rd, Qingyang District,
Chengdu 610072, Sichuan, China,
phone: +86 028 8739 3485,
e-mail: 318004031@qq.com

WHAT'S NEW?

Few studies evaluated the readmissions for elderly patients with ischemic heart disease. Therefore, we developed a predictive model based on machine learning. The results showed

that the categorical boosting model had better predictive performance and good calibration in identifying 30-day and 1-year readmissions for elderly patients with ischemic heart disease. Meanwhile, we found that the age-adjusted Charlson comorbidity index, brain natriuretic peptide, heart failure, cholesterol, free thyroxine, thymidine kinase 1, osmotic pressure and red blood cell distribution width-standard deviation were common risk factors for readmission at different time.

ABSTRACT

Background: The contribution of clinical features associated with 30-day or 1-year readmission in elderly patients with ischemic heart disease (IHD) and whether these features can be used to predict the readmission risk of patients have not been studied.

Aims: The study aimed to develop 30-day and 1-year readmission prediction models for elderly IHD patients using machine learning combined features routinely collected at the time of hospital discharge, and to investigate the contribution of features to these predictions.

Methods: Eight machine learning algorithms were used to develop prediction models. Area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve (AUPRC) were used to assess discrimination. Shapley additive explanations (SHAP) analysis was used to explain the contribution of features.

Results: A total of 6687 patients were enrolled. For 30-day readmission, categorical boosting (CB) model had the best predictive performance with the highest AUROC (0.72), and the Brier score was 0.23. For 1-year readmission, CB model had the best predictive performance with the highest AUROC (0.66), and the Brier score was 0.14. The age-adjusted Charlson comorbidity index, brain natriuretic peptide, heart failure, cholesterol, free thyroxine, thymidine kinase 1, osmotic pressure and red blood cell distribution width-standard deviation were the common important features to predict 30-day and 1-year readmission of elderly IHD patients.

Conclusions: Elderly IHD patients with high risk of 30-day or 1-year readmission can be identified using machine learning and features collected at the time of discharge.

Key words: elderly, explainable, ischemic heart disease, machine learning, readmission

INTRODUCTION

Ischemic heart disease (IHD), a common chronic non-communicable cardiovascular disease (CVD), is one of the major contributors to CVD-related disease burden [1]. Due to poor disease control, many IHD patients returned to the hospital after discharge. For elderly IHD patients and their families, frequent readmissions may increase the financial burden, difficult experience and reduce the quality of life [2]. Therefore, identifying high-risk readmission patients and providing individualized therapeutic regimens is the key to reducing the readmission rate of elderly IHD patients.

Many features such as age [3], sex [4], length of stay (LOS) [5] and complication [6] associated with patient readmission have been illustrated. Some routine laboratory tests were also associated with rehospitalization [7, 8]. However, it has not been possible to determine the contribution of these features to patients' readmission or to assess the individualized effect of these features on patient clinical outcomes. Moreover, it is not clear whether the elderly IHD patients at high risk of readmission can be accurately identified using these clinically relevant variables available at the time of hospital discharge. In recent years, some screening tools to identify heart failure or acute myocardial infarction patients who may be readmitted to the hospital have been established [9, 10], which showed that using machine learning to develop a complicated and reliable classification tool is possible. However, to our knowledge, there is no practical readmission evaluation tool for elderly IHD patients.

In this study, we produced 30-day and 1-year readmission prediction models for elderly IHD patients combining 8 machine learning algorithms and features routinely collected at the time of hospital discharge. Additionally, we analyzed the contribution of features to these predictions.

METHODS

Participants and outcome

The subjects were elderly IHD patients who admitted to Sichuan Provincial People's Hospital from August 2018 to April 2020. Inclusion criteria include: I) age ≥60 years; II) diagnosis with IHD [11]. The exclusion criteria were as follows: I) LOS <2 days; II) transferred to other hospitals; III) died in hospital; IV) follow-up time <1 year. The primary outcomes were 30-day and 1-year all-cause readmission. This study was approved by the Ethics Committee of Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital (approval number: 2023-85). Due to the retrospective nature of the study, informed consent was waived.

Data collection and preprocessing

The data were collected from the Hospital Information System, which included basic patient information, medications, comorbidities and laboratory tests. For multiple laboratory test results, we selected the last results of patients before discharge. For multiple readmissions, the first admission record was included.

The data preprocessing was as follows. Firstly, according to the comorbidities at the time of discharge, we calculated the age-adjusted Charlson comorbidity index (ACCI) of all patients [12]. The calculation method of ACCI is shown in Supplementary material, *Table S1*. Moreover, we counted the number of medications (NOM) taken by each patient at the time of discharge. For the laboratory test results of the patients, the feature with missing data >90% was deleted. Then, we used the random forest (RF) to replace the missing data and the Lasso to feature selection. Before the modeling, the continuous variables were standardized by Z-score.

Model development

We used 8 representative machine learning algorithms including logistic regression (LR), k-nearest neighbor (KNN), support vector machine (SVM), adaptive boosting (AdaBoost), gradient boosting decision tree (GBDT), RF, categorical boosting (CB), and extreme gradient boosting (XGB) to develop prediction models.

In the modeling process, the patients were randomly divided into a training set and a test

set according to 7:3. The training set was used to develop models, and 10-fold cross-validation on the training set was applied for parameter optimization. The test set was used to evaluate the predictive performance of the models. We developed the 30-day and 1-year readmission prediction models, respectively. Notably, compared to non-readmission, fewer patients were readmitted in 30 days. To improve the predictive performance of the models, we randomly matched the non-readmission patients to the 30-day readmission patients by a ratio of 5:1. The matched data were used to establish the 30-day readmission prediction model.

Model assessment

Model performance was mainly assessed using area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve (AUPRC). Meanwhile, we calculated the accuracy, precision, positive predictive value (PPV) and negative predictive value (NPV) to evaluate the predictive performance of the machine learning models [13]. Calibration of model was assessed by brier score, which is an index to evaluate both discrimination and calibration performance [14]. And the model was considered to have favorable calibration when the Brier score ≤ 0.25 [15]. Shapley additive explanations (SHAP) algorithm was used to measure the contribution of each feature to the best model.

Statistical analysis

All statistical analyses were performed using SPSS software version 25. Models building was implemented in Python (Version 3.7.0). The Shapiro-Wilk test was used to assess whether continuous variables followed a normal distribution. Normally distributed variables and skewed distributional data are described as the mean and standard deviation (SD) and median and interquartile range, respectively. The t-test or Mann–Whitney test was used to analyze significant differences, respectively. The χ^2 test was chosen for the analysis of categorical variables, which were expressed as counts and percentages. A two-sided *P*-value of <0.05 was considered statistically significant.

RESULTS

Baseline characteristics

The patient data selection process is shown in Figure 1. This study included 6687 patients, of whom 904 (13.52%) underwent 1-year readmission and 325 (4.86%) underwent 30-day readmission. There were 3673 (54.93%) male patients and 3014 (45.07%) female patients. The average age was 73.5 (8.1) years. All patients were used to develop machine learning models to predict 1-year readmission, and 1950 patients were used to predict 30-day readmission. Table 1 shows the general information of the patients. The characteristics of 30-day readmission and non-readmission patients before matching are shown in Supplementary material, *Table S2*.

Features selection

A total of 174 features were collected, 85 features with missing data >90% were deleted (Supplementary material, *Table S3*). Then, Lasso regression analysis was performed on the remaining 89 features as independent variables with 30-day and 1-year readmissions as dependent variables, respectively (Figure 2). The results showed that when lambda with minimum mean square error, 10 features were selected with 30-day readmission (Supplementary material, *Table S4*). Similarly, for 1-year readmission, the 89 features were reduced to 36 when the lambda with minimum mean square error (Supplementary material, *Table S4*).

Model performance and calibration

Eight machine learning algorithms combined with 10 features selected by Lasso were used to develop prediction models for 30-day readmission. On the training set, RF model had the highest AUROC [0.95 (0.01), standard error = 0.003] (Figure 3A). On the test set, CB model had the highest AUROC (0.72) (Figure 3B). GBDT and XGB model had the highest AUPRC (0.31) on the test set (Figure 3C). And the AUPRC of CB model was 0.30 on the test set (Figure 3C). The accuracy, precision, PPV and NPV of CB model to 30-day readmission on the test set were 0.77, 0.28, 0.28 and 0.87 (Table 2), respectively.

The 8 machine learning algorithms combined with 36 features selected by Lasso were used to develop prediction models for 1-year readmission. On the training set, RF model had the highest AUROC [0.98 (0.01), standard error = 0.002] (Figure 3D). On the test set, CB model had the highest AUROC (0.66) (Figure 3E). Moreover, LR, CB and XGB model had the highest AUPRC (0.20) on the test set (Figure 3F). The accuracy, precision, PPV and NPV of CB model to 30-day readmission on the test set were 0.86, 0.17, 0.17 and 0.88 (Table 2), respectively.

We used the Brier score to assess the calibration of the machine learning models. For 30-day readmission, the XGB model had the optimal brier score (0.19). The Brier score of the CB model was 0.23 (Table 2). For 1-year readmission, the Brier score of the CB model was 0.14 (Table 2).

Features contribution

The contributions of features to predictions of best-performing models by different periods of readmission are presented in Figure 4. For 30-day readmission, the most important features were ACCI, B-type natriuretic peptide (BNP), heart failure, cholesterol and thymidine kinase 1 (TK1) (Figure 4A). For 1-year readmission, most important features were sex, hypertension, nephropathy, NOM, BNP (Figure 4B).

The relationship between SHAP value and feature value is illustrated in more detail for the features of patient BNP and NOM in Figure 4C (30-day readmission) and Figure 4D (1-year readmission), respectively. The results showed that if the patients with high BNP, they may have a high risk of 30-day readmission. Similarly, the patients with ≥8 NOM were more likely to be readmitted in 1 year after discharge.

Furthermore, SHAP can provide the contribution of every feature to the predicted outcome of each patient. An example of 30-day readmission is illustrated in Figure 4E. Total triiodothyronine (TT3), BNP, osmotic pressure (OSM), free thyroxine (FT4), TK1 and red blood cell distribution width-standard deviation (RDW-SD) provided a positive contribution, while heart failure, ACCI and cholesterol provided a negative contribution. Another example of 1-year readmission is illustrated in Figure 4F. High-sensitivity cardiac troponin I, LOS, total

bilirubin, TK1 and procalcitonin provided a positive contribution, while sex, nephropathy, hypertension and heart failure provided a negative contribution.

DISCUSSION

In this study, we reported the development and validation of machine learning models to identify elderly IHD patients at high risk for 30-day or 1-year readmission. We used 8 representative machine learning algorithms including LR, KNN, SVM, AdaBoost, GBDT, RF, CB, and XGB, combined with the features selected by Lasso to develop prediction models. For 30-day or 1-year readmission, CB model had the best predictive performance.

Several studies have assessed 30-day or 1-year all-cause readmissions after acute myocardial infarction, but few are specific to IHD, especially for elderly patients. Dodson et al [16] established a risk model by backward selection and Bayesian to predict all-cause readmission at 30 days for 3006 elderly acute myocardial infarction patients (≥75 years). The C statistic of the model was 0.63 in validation cohort. Another study established RF was the best predictive model with an AUROC of 0.66 [17]. The model developed by Dreyer et al [18] to predict 1-year readmission among younger acute myocardial infarction adults had good calibration, and the C statistic of the model was 0.69. Compared to other studies, the best model in predicting 30-day or 1-year readmissions in our study had a consistent predictive performance. Meanwhile, the Brier score of the CB model was 0.23 on 30-day readmission and 0.14 on 1-year readmission, which showed the model had a good calibration ability.

An important challenge in optimizing the clinical utility of machine learning models resides in balancing the PPV and NPV in specific clinical settings [13]. As the proportion of patients at high risk of readmission is low, benefit from machine learning models should ensure those patients with high risk of readmission can be identified before discharge. In our study, all machine learning models had good NPV, which suggested that our model can accurately identify those patients at low risk for readmissions. However, the PPV values of all models were unsatisfactory, and we need to optimize the models to improve the identification of high-risk readmission patients in further studies.

Many of the top predictors contributing to identified readmission of elderly IHD patients in this study have been previously reported, such as BNP, LOS, sex and ACCI [19–23]. Heart failure is one of the significant risk factors for 30-day readmission in cardiovascular patients [24]. Our results also clarify that elderly IHD patients with heart failure were more likely to be readmitted within 30 days or 1 year. Poly-comorbidity often means polypharmacy. Care for the hospitalized elderly patients is made more complex by polypharmacy, which may increase the risks of adverse drug events [25]. Our study was unable to trace the cause of readmission. And this is one of the limitations of the study. However, those polypharmacy elderly IHD patients should receive more attention to reduce readmissions. Furthermore, a new finding is that OSM is also a common important feature for the prediction of 30-day or 1-year readmission of elderly IHD patients. Severe reduction of OSM predicts adverse outcomes in patients with cardiopulmonary failure. Taniguchi and colleagues recommended using OSM measurement to monitor patients with unstable angina pectoris [26]. One interesting finding was that TK1 was a common important feature in predicting 30-day and 1-year readmission of elderly IHD patients. However, TK1 is not routinely tested in patients with cardiovascular disease, which limits the clinical utility of the model.

Although our results provided overall risk estimates for the population, they are of limited clinical assistance. Because significant individual differences existed in readmission risk features. The same features may have different clinical significance for different patients. A major advantage of the model developed in our study is that it provides individual risk estimates, which may assist clinicians in providing optimal care and appropriate interventions for patients.

Limitations

There are several limitations to this study. First, although the complications of the elderly IHD patients were enrolled, relevant information, such as the duration of the complications and disease severity were not included in our study. Second, this was a single-center study, and it would need to be tested in other medical institutions to assess the predictive performance of the model. Third, the inclusion of some unconventional laboratory test features such as TK1 in the

final model may limit its clinical utility. Finally, the inclusion of additional features, such as health literacy, frailty and socioeconomic status, could improve the readmission risk assessment and should be validated in further studies.

CONCLUSIONS

In this study, we used clinically relevant features available at the time of hospital discharge combined with 8 machine learning algorithms to predict 30-day and 1-year readmission in elderly IHD patients, and investigated features contributing to these predictions. The CB model had better predictive performance and good calibration which showed that machine learning models had potential clinical application value.

Supplementary material

Supplementary material is available at https://journals.viamedica.pl/polish_heart_journal.

Article information

Conflict of interest: None declared.

Funding: This research was funded by the National Key Research and Development Program of China (2020YFC2005506).

REFERENCES

1. Li Y, Zhang J. Disease burden and risk factors of ischemic heart disease in China during 1990-2019 based on the Global Burden of Disease 2019 report: A systematic analysis.

Front Public Health. 2022; 10: 973317, doi: 10.3389/fpubh.2022.973317, indexed in Pubmed: 36408039.

2. Huang J, Zeng Q, Pan X, et al. Projections of the effects of global warming on the disease burden of ischemic heart disease in the elderly in Tianjin, China. BMC Public Health. 2019; 19(1): 1465, doi: 10.1186/s12889-019-7678-0, indexed in Pubmed: 31694683.
3. Kwok CS, Rao SV, Gilchrist I, et al. Relation Between Age and Unplanned Readmissions After Percutaneous Coronary Intervention (Findings from the Nationwide Readmission Database). Am J Cardiol. 2018; 122(2): 220–228, doi: 10.1016/j.amjcard.2018.03.367, indexed in Pubmed: 29861049.
4. Steitieh DA, Lu DY, Kalil RK, et al. Sex-Based Differences in Revascularization and 30-Day Readmission After ST-Segment-Elevation Myocardial Infarction in the United States. Cardiovasc Revasc Med. 2021; 31: 41–47, doi: 10.1016/j.carrev.2020.12.016, indexed in Pubmed: 33358184.
5. Reynolds K, Butler MG, Kimes TM, et al. Relation of Acute Heart Failure Hospital Length of Stay to Subsequent Readmission and All-Cause Mortality. Am J Cardiol. 2015; 116(3): 400–405, doi: 10.1016/j.amjcard.2015.04.052, indexed in Pubmed: 26037295.
6. Bachli C, Wagner U, Müller B, et al. Impact of comorbidities and patient age on clinical outcomes in patients with myocardial infarction-a nationwide cohort study. 39th International Symposium on Intensive Care and Emergency Medicine. Critical Care. 2019; 23(S2), doi: 10.1186/s13054-019-2358-0.
7. Giustino G, Brener S, Maehara A, et al. TCT-265 Impact of platelet count on myocardial infarct size and adverse events in anterior ST-Elevation myocardial infarction: results from the INFUSE-AMI trial. JACC. 2015; 66(15_Supplement): B104–B105, doi: 10.1016/j.jacc.2015.08.948.
8. Nakahashi H, Kosuge M, Kiyokuni M, et al. Impact of change in red cell distribution width on long-term outcomes in patients with acute myocardial infarction undergoing primary percutaneous coronary intervention. Circulation. 2016; 134(1_suppl).

9. Chen S, Hu W, Yang Y, et al. Predicting Six-Month Re-Admission Risk in Heart Failure Patients Using Multiple Machine Learning Methods: A Study Based on the Chinese Heart Failure Population Database. J Clin Med. 2023; 12(3), doi: 10.3390/jcm12030870, indexed in Pubmed: 36769515.
10. Matheny ME, Ricket I, Goodrich CA, et al. Development of Electronic Health Record-Based Prediction Models for 30-Day Readmission Risk Among Patients Hospitalized for Acute Myocardial Infarction. JAMA Netw Open. 2021; 4(1): e2035782, doi: 10.1001/jamanetworkopen.2020.35782, indexed in Pubmed: 33512518.
11. Safiri S, Karamzad N, Singh K, et al. Burden of ischemic heart disease and its attributable risk factors in 204 countries and territories, 1990-2019. Eur J Prev Cardiol. 2022; 29(2): 420–431, doi: 10.1093/eurjpc/zwab213, indexed in Pubmed: 34922374.
12. Koppie TM, Serio AM, Vickers AJ, et al. Age-adjusted Charlson comorbidity score is associated with treatment decisions and clinical outcomes for patients undergoing radical cystectomy for bladder cancer. Cancer. 2008; 112(11): 2384–2392, doi: 10.1002/cncr.23462, indexed in Pubmed: 18404699.
13. Castela Forte J, Yeshmagambetova G, van der Grinten ML, et al. Comparison of Machine Learning Models Including Preoperative, Intraoperative, and Postoperative Data and Mortality After Cardiac Surgery. JAMA Netw Open. 2022; 5(10): e2237970, doi: 10.1001/jamanetworkopen.2022.37970, indexed in Pubmed: 36287565.
14. Sun LY, Chung AM, Farkouh ME, et al. Defining an Intraoperative Hypotension Threshold in Association with Stroke in Cardiac Surgery. Anesthesiology. 2018; 129(3): 440–447, doi: 10.1097/ALN.0000000000002298, indexed in Pubmed: 29889106.
15. Lin J, Yin M, Liu Lu, et al. The Development of a Prediction Model Based on Random Survival Forest for the Postoperative Prognosis of Pancreatic Cancer: A SEER-Based Study. Cancers (Basel). 2022; 14(19), doi: 10.3390/cancers14194667, indexed in Pubmed: 36230593.
16. Dodson JA, Hajduk AM, Murphy TE, et al. Thirty-Day Readmission Risk Model for Older Adults Hospitalized With Acute Myocardial Infarction. Circ Cardiovasc Qual

Outcomes. 2019; 12(5): e005320, doi: 10.1161/CIRCOUTCOMES.118.005320, indexed in Pubmed: 31010300.

17. Garg N, Ray S, Mathur A. Abstract 10932: Prediction of 30-Day Hospital Readmission in High-Risk Atherosclerotic Cardiovascular Disease Patients Using Machine Learning Methods on Electronic Health Record Data from Medical Information Mart for Intensive Care-3 Database. Circulation. 2021; 144(Suppl_1), doi: 10.1161/circ.144.suppl_1.10932.
18. Dreyer RP, Murphy TE, Raparelli V, et al. Development of a Risk Prediction Model for 1-year Readmission Among Younger Adults Hospitalized for Acute Myocardial Infarction. Circulation. 2020; 142(suppl_3), doi: 10.1161/circ.142.suppl_3.14508.
19. Almeida I, Miranda H, Santos H, et al. Prognostic value of BNP in STEMI patients. European Heart Journal. Acute Cardiovascular Care. 2021; 10(Supplement_1), doi: 10.1093/ehjacc/zuab020.093.
20. Bierman AS, Guan J, Turial D, et al. Multimorbidity and causes of hospital readmission at 30 days and one year. Journal of the American Geriatrics Society. 2015; 63(S1): S209, doi: 10.1111/jgs.13439.
21. Arakaki AJ, Raparelli V, Murphy T, et al. Gender-Based Risk Prediction for 1-year Readmission Among Younger Women Hospitalized for Acute Myocardial Infarction. Circulation. 2021; 144(Suppl_1), doi: 10.1161/circ.144.suppl_1.13264.
22. Bierman AS, Guan J, Stukel T. The relationship between multi morbidity and concordant and discordant causes of hospital readmission at 30-days and one year. In: Abstracts from the 38th Annual Meeting of the Society of General Internal Medicine. J Gen Intern Med. 2015; 30(Suppl_2): S280.
23. Miñana G, Bosch MJ, Núñez E, et al. Length of stay and risk of very early readmission in acute heart failure. Eur J Intern Med. 2017; 42: 61–66, doi: 10.1016/j.ejim.2017.04.003, indexed in Pubmed: 28400077.

24. Wang H, Zhao T, Wei X, et al. The prevalence of 30-day readmission after acute myocardial infarction: A systematic review and meta-analysis. Clin Cardiol. 2019; 42(10): 889–898, doi: 10.1002/clc.23238, indexed in Pubmed: 31407368.
25. DiConti-Gibbs A, Chen KY, Coffey CE. Polypharmacy in the Hospitalized Older Adult: Considerations for Safe and Effective Treatment. Clin Geriatr Med. 2022; 38(4): 667–684, doi: 10.1016/j.cger.2022.07.011, indexed in Pubmed: 36210083.
26. Taniguchi H, Takayama Y, Sugiura T, et al. Clinical significance of serum colloid osmotic pressure in relation to pulmonary edema and coronary instability in patients with unstable angina. Clin Cardiol. 1992; 15(5): 348–352, doi: 10.1002/clc.4960150508, indexed in Pubmed: 1623655.

Table 1. Patient characteristics

Characteristics		**30–day**			**1–year**		
		Readmission (n = 325)	**Non-readmission (n = 1625)**	***P***	**Readmission (n = 904)**	**Non-readmission (n = 5783)**	***P*-value**
Sex				0.208			<0.001
	Males	185 (56.9)	863 (53.1)		557 (61.6)	3116 (53.9)	
	Females	140 (43.1)	762 (46.9)		347 (38.4)	2667 (46.1)	
Age, years				<0.01			<0.001
	60–69	87 (26.8)	602 (37.0)		267 (29.5)	2150 (37.2)	
	70–79	143 (44.0)	648 (39.9)		371 (41.0)	2296 (39.7)	
	≥80	95 (29.2)	375 (23.1)		266 (29.4)	1337 (23.1)	
ACCI				<0.001			<0.001
	2–5	204 (62.8)	1198 (73.7)		559 (61.8)	4214 (72.9)	
	6–10	115 (35.4)	411 (25.3)		331 (36.6)	1508 (26.1)	
	≥11	6 (1.8)	16 (1.0)		14 (1.5)	61 (1.1)	
NOM				<0.001			<0.001
	1–5	66 (20.3)	335 (20.6)		129 (14.3)	1182 (20.4)	
	6–10	143 (44.0)	900 (55.4)		474 (52.4)	3233 (55.9)	
	≥11	116 (35.7)	390 (24.0)		301 (33.3)	1368 (23.7)	
Total protein, g/l		66.8 (61.0–72.0)	67.8 (62.7–72.8)	<0.01	66.7 (61.0–71.8)	68.0 (62.8–72.7)	<0.001
Total bilirubin,		13.3 (9.3–	13.3 (10.3–	0.333	12.8 (9.2–	13.4 (10.1–	<0.01

μmol/l	18.1)	18.0)		17.3)	17.9)	
Cholesterol, mmol/l	3.78 (2.94–4.43)	4.07 (3.29–4.92)	<0.001	3.84 (3.13–4.63)	4.06 (3.32–4.88)	<0.001
AST, U/l	28 (22–37)	28 (22–36)	0.516	27 (22–35)	28 (22–36)	<0.05
Osmotic pressure, mOsm/l	283 (278–287)	283 (279–287)	<0.05	283 (278–287)	283 (279–287)	<0.05
Urea, mmol/l	6.75 (5.00–8.99)	6.13 (4.95–7.72)	<0.01	6.47 (4.95–8.53)	6.08 (4.86–7.78)	<0.001
Creatine kinase, U/l	83 (51–119)	92 (64–142)	<0.001	88 (59–129)	90 (63–138)	<0.05
Myoglobin, ng/ml	57.2 (39.2–100.8)	47.2 (33.8–75.0)	<0.001	55.2 (37.4–91.1)	48.0 (34.2–76.2)	<0.001
Creatinine, μmol/l	78.4 (64.1–103.6)	72.0 (59.9–89.8)	<0.001	78.7 (63.0–101.0)	72.2 (59.7–90.5)	<0.001
AST/ALT	1.30 (1.00–1.80)	1.28 (0.95–1.71)	0.167	1.28 (0.96–1.70)	1.28 (0.96–1.70)	0.511
eGFR, ml/min	79.0 (56.4–90.0)	84.1 (66.6–93.0)	<0.001	79.7 (57.7–90.5)	83.9 (65.3–93.1)	<0.001
HDL-C, mmol/l	1.14 (0.97–1.39)	1.18 (0.99–1.43)	0.084	1.16 (0.98–1.37)	1.19 (1.00–1.44)	<0.01
Triglyceride, mmol/l	1.24 (0.83–1.90)	1.40 (0.98–2.05)	<0.01	1.32 (0.91–1.96)	1.37 (0.97–1.99)	0.062
LDL-C, mmol/l	1.90 (1.36–2.61)	2.18 (1.58–2.86)	<0.001	1.98 (1.46–2.69)	2.18 (1.61–2.81)	<0.001
ALT, U/l	20 (15–30)	22 (16–34)	<0.01	21 (15–32)	22 (15–33)	0.098
GGT, U/l	28 (19–46)	26 (17–44)	0.088	27 (18–48)	25 (17–44)	<0.01
BNP, pg/ml	193.6 (58.6–	66.6 (26.2–	<0.00	112.5 (45.9–	72.7 (28.3–	<0.00

	511)	214.1)	1	349.4)	219.7)	1
D–dimer, mg/l	0.65 (0.35–1.43)	0.46 (0.24–1.08)	<0.001	0.54 (0.28–1.19)	0.47 (0.25–1.13)	<0.01
Hemoglobin, g/l	124 (109–138)	130 (118–142)	<0.001	126 (112–139)	130 (117–141)	<0.001
Hematocrit	0.37 (0.31–0.42)	0.39 (0.33–0.42)	<0.01	0.38 (0.32–0.42)	0.39 (0.32–0.42)	<0.05
RDW–SD, fl	45.3 (43.0–48.1)	44.6 (42.5–47.0)	<0.01	45.1 (42.8–47.9)	44.7 (42.6–47.3)	<0.05
TK1, pmol/l	0.61 (0.30–1.31)	0.54 (0.23–1.24)	0.198	0.61 (0.33–1.32)	0.53 (0.23–1.23)	<0.01

Data presented as number (%) or median (Q1–Q3)

Abbreviations: ACCI, age-adjusted Charlson comorbidity index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BNP, brain natriuretic peptide; eGFR, estimated glomerular filtration rate; GGT, γ-glutamyl transpeptidase; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; NOM, number of medications; RDW-SD, red blood cell distribution width-standard deviation; TK1, thymidine kinase 1

Table 2. Models performance

Prediction	Model	Accuracy	Precision	PPV	NPV	Brier score
30-day readmission	LR	0.63	0.26	0.26	0.92	0.37
	KNN	0.57	0.23	0.23	0.90	0.43
	SVM	0.59	0.25	0.25	0.93	0.41
	AdaBoost	0.71	0.28	0.28	0.89	0.29
	GBDT	0.68	0.26	0.26	0.89	0.31
	RF	0.78	0.30	0.29	0.87	0.22
	CB	0.77	0.28	0.28	0.87	0.23
	XGB	0.81	0.37	0.37	0.88	0.19
1-year readmission	LR	0.62	0.18	0.18	0.91	0.38
	KNN	0.41	0.15	0.15	0.93	0.59
	SVM	0.69	0.19	0.19	0.91	0.31
	AdaBoost	0.83	0.22	0.22	0.88	0.17
	GBDT	0.87	0.21	0.21	0.88	0.13
	RF	0.87	0.30	0.30	0.88	0.13
	CB	0.86	0.17	0.17	0.88	0.14
	XGB	0.86	0.26	0.26	0.88	0.14

Abbreviations: AdaBoost, adaptive boosting; CB, categorical boosting; GBDT, gradient boosting decision tree; KNN, k-nearest neighbor; LR, logistic regression; NPV, negative predictive value; PPV, positive predictive value; RF, random forest; SVM, support vector machine; XGB, extreme gradient boosting

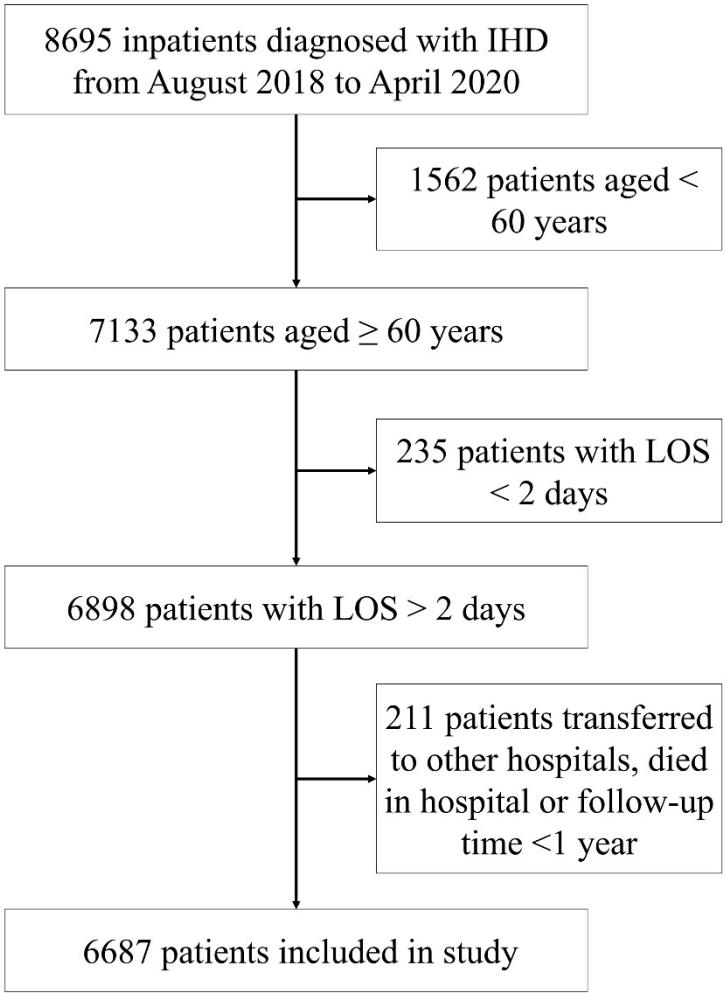

Figure 1. Patient data selection process

Abbreviations: IHD, ischemic heart disease; LOS, length of stay

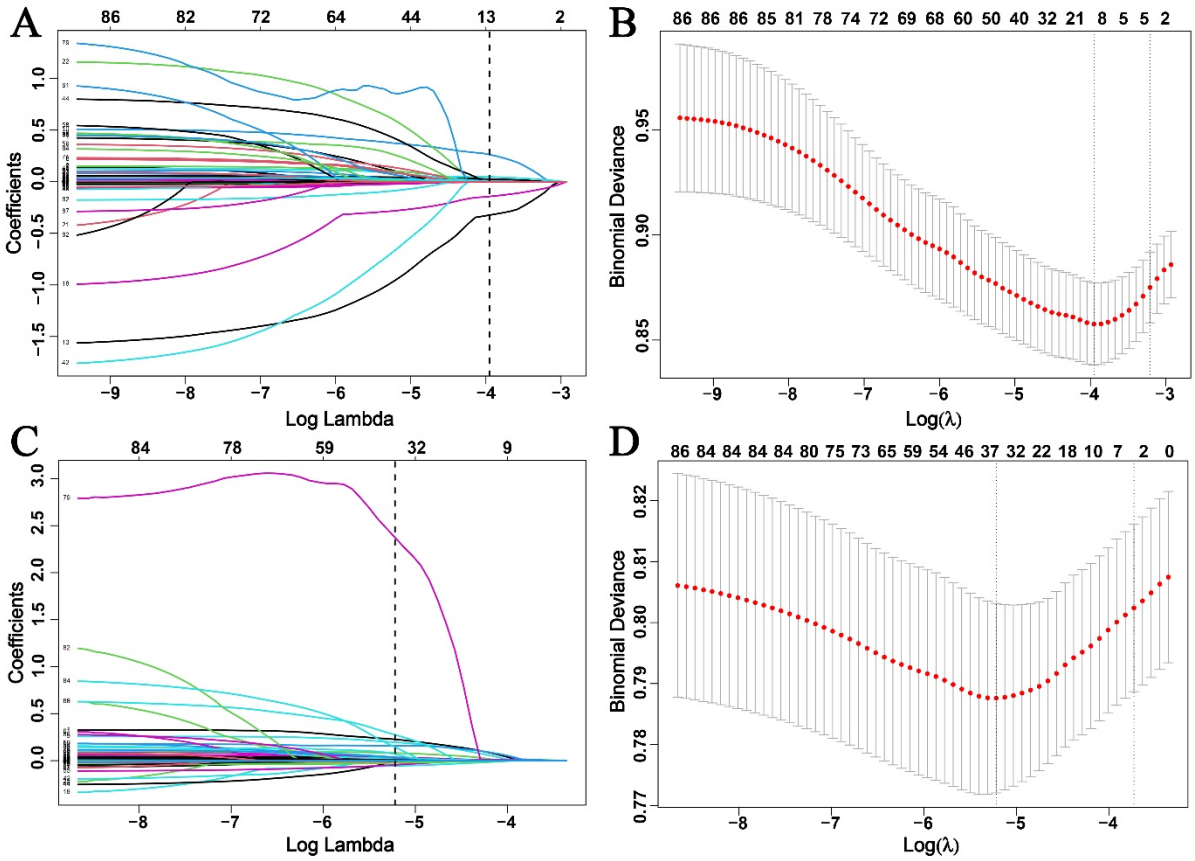

Figure 2. A. and **B.** Lasso regression analysis for 30-day readmission. **C.** and **D.** Lasso regression analysis for 1-year readmission. **A.** and **C.** Vertical lines are drawn at selected values by applying 10-fold cross-validation. **B.** and **D.** In the Lasso model, the coefficient profiles of 89 texture features were draw from the log (λ) sequence. Vertical dotted lines are drawn at the minimum mean square error and the standard error of the minimum distance

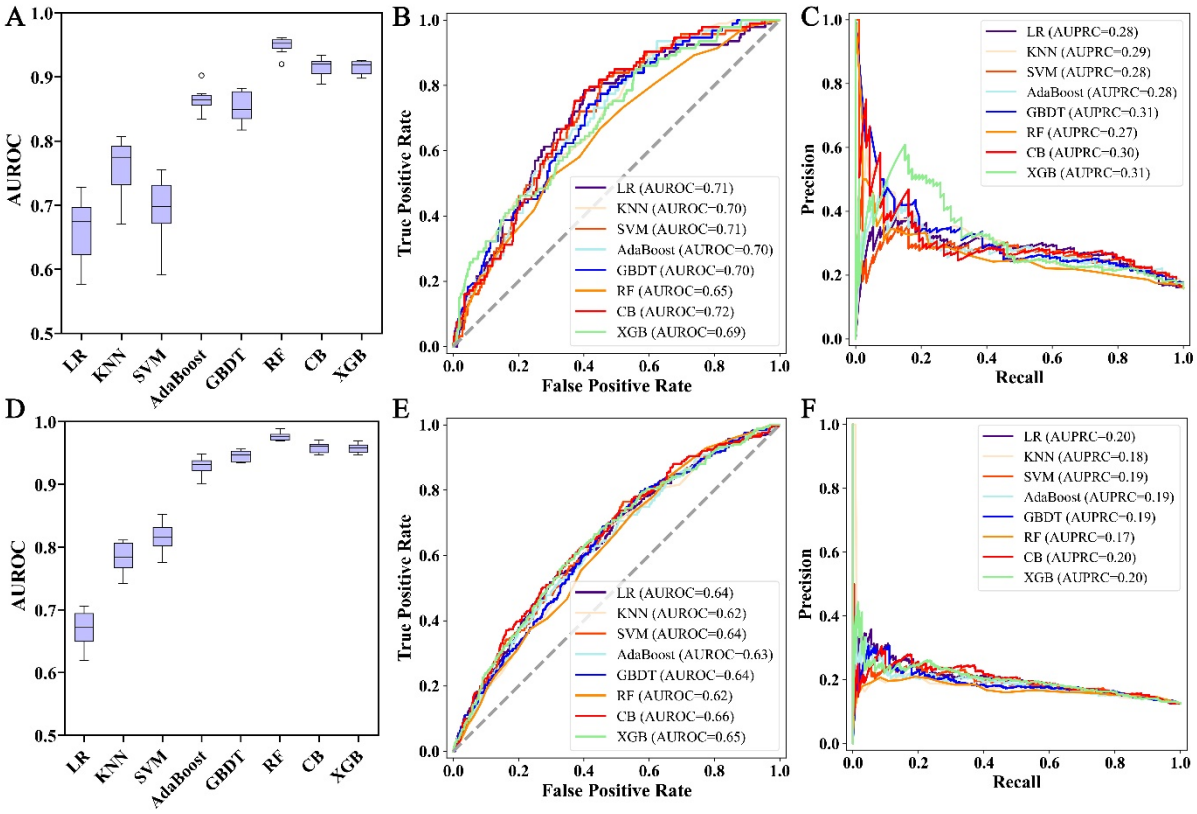

Figure 3. Models predictive performance. **A.**, **B.** and **C.** 30-day readmission. **A.** The AUROC on the training set. **B.** The AUROC on the test set. **C.** The AUPRC on the test set. **D.**, **E.** and **F.** 1-year readmission. **D.** The AUROC on the training set. **E.** The AUROC on the test set. **F** The AUPRC on the test set

Abbreviations: AUPRC, area under the precision recall curve; AUROC, area under the receiver operating characteristic curve; other — see Table 2

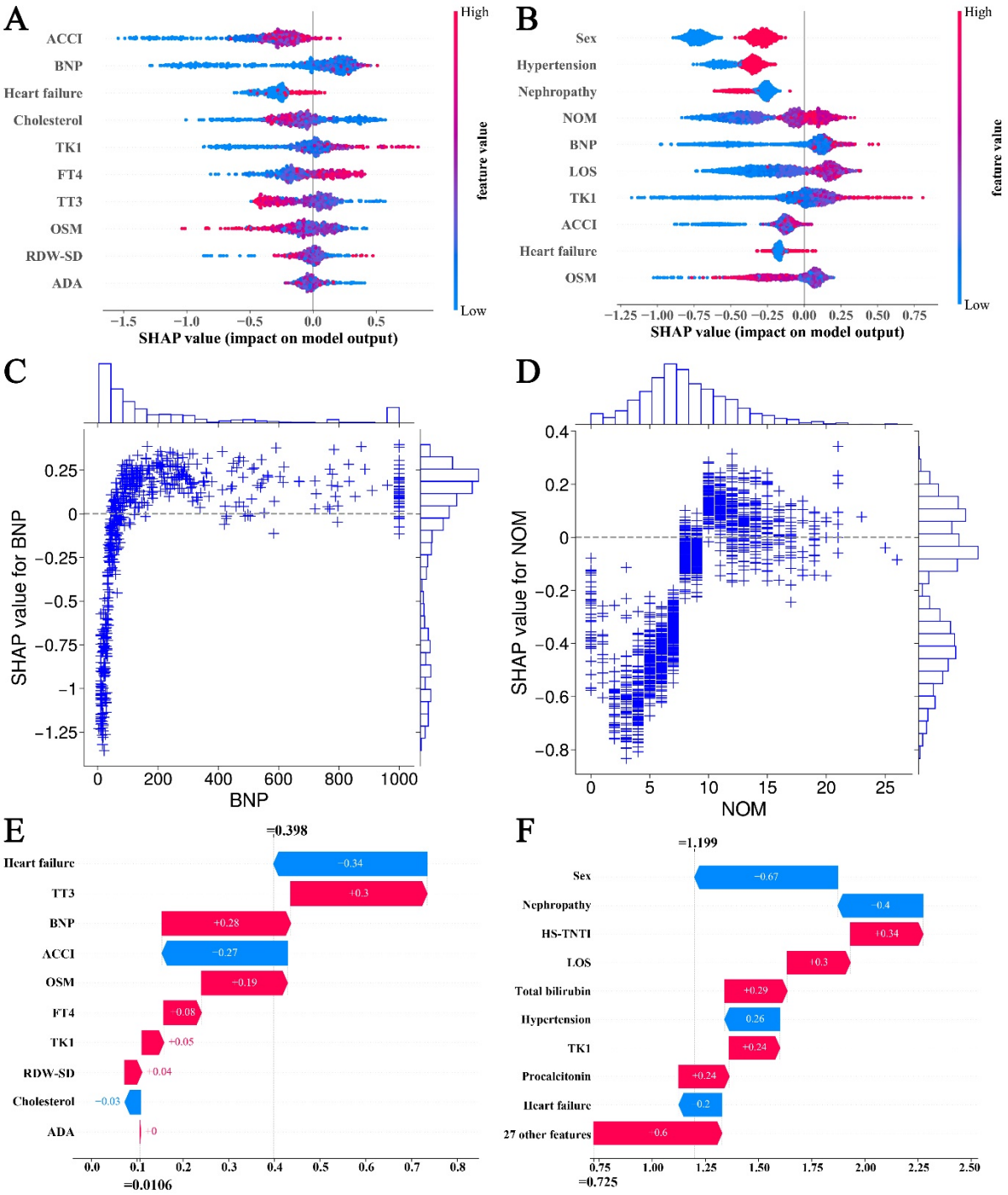

Figure 4. Contributions of input features to readmission predictions. The SHAP summary plot of the 10 most important variables of the CB model for 30-day readmission (**A**) and 1-year readmission (**B**). In the plots, the x-axis indicates the SHAP value; and the y-axis indicates the relationship between features and SHAP values. The color of the dot represents the value of the features. Red represents higher feature values, and blue represents lower feature values (**C**). **D.** Scatter plot shows the relationship between the feature value and SHAP value. **C.** BNP for 30-

day readmission. **D.** NOM for 1-year readmission. In the plots, the distributions of the SHAP value and feature values are shown as histograms on the right and top of the scatter graph. **E.** and **F**. Contribution of every feature to the predicted outcome of one sample. Red represents positive contribution, and blue represents negative contribution. **E.** 30-day readmission. **F.** 1-year readmission

Abbreviations: ADA, adenosine deaminase; FT4, free thyroxine; HS-TNTI, high-sensitivity cardiac troponin I; OSM, osmotic pressure; SHAP, Shapley additive explanations; TT3, total triiodothyronine; other — see Table 1 and Figure 1