



Development and assessment of scoring model for ICU stay and mortality prediction after emergency admissions in ischemic heart disease: a retrospective study of MIMIC-IV databases

Tingting Shu¹ · Jian Huang² · Jiewen Deng³ · Huaqiao Chen⁴ · Yang Zhang^{5,6} · Minjie Duan^{5,6} · Yanqing Wang⁷ · Xiaofei Hu⁸ · Xiaozhu Liu⁹

Received: 11 November 2022 / Accepted: 9 January 2023

© The Author(s), under exclusive licence to Società Italiana di Medicina Interna (SIMI) 2023

Abstract

Ischemic heart disease (IHD) is the leading cause of death and emergency department (ED) admission. We aimed to develop more accurate and straightforward scoring models to optimize the triaging of IHD patients in ED. This was a retrospective study based on the MIMIC-IV database. Scoring models were established by AutoScore formwork based on machine learning algorithm. The predictive power was measured by the area under the curve in the receiver operating characteristic analysis, with the prediction of intensive care unit (ICU) stay, 3d-death, 7d-death, and 30d-death after emergency admission. A total of 8381 IHD patients were included (median patient age, 71 years, 95% CI 62–81; 3035 [36%] female), in which 5867 episodes were randomly assigned to the training set, 838 to validation set, and 1676 to testing set. In total cohort, there were 2551 (30%) patients transferred into ICU; the mortality rates were 1% at 3 days, 3% at 7 days, and 7% at 30 days. In the testing cohort, the areas under the curve of scoring models for shorter and longer term outcomes prediction were 0.7551 (95% CI 0.7297–0.7805) for ICU stay, 0.7856 (95% CI 0.7166–0.8545) for 3d-death, 0.7371 (95% CI 0.6665–0.8077) for 7d-death, and 0.7407 (95% CI 0.6972–0.7842) for 30d-death. This newly accurate and parsimonious scoring models present good discriminative performance for predicting the possibility of transferring to ICU, 3d-death, 7d-death, and 30d-death in IHD patients visiting ED.

Keywords Ischemic heart disease · Emergency department · Scoring model · MIMIC-IV · Machine learning

Tingting Shu, Jian Huang, and Jiewen Deng have contributed equally to this work and share the first authorship.

Xiaofei Hu and Xiaozhu Liu have contributed equally to this work and share the corresponding author ship.

✉ Xiaofei Hu
harryzonetmmu@163.com

✉ Xiaozhu Liu
xiaozhuliu2021@163.com;
xiaozhuliu@hospital.cqmu.edu.cn

¹ Third Military Medical University (Army Medical University), Chongqing 400038, China

² Graduate School, Guangxi University of Chinese Medicine, Nanning, China

³ Department of Neurosurgery, The Second Affiliated Hospital of Chongqing Medical University, Chongqing 400010, China

⁴ Department of Cardiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

⁵ College of Medical Informatics, Chongqing Medical University, Chongqing, China

⁶ Medical Data Science Academy, Chongqing Medical University, Chongqing, China

⁷ The First College of Clinical Medicine, Chongqing Medical University, Chongqing, China

⁸ Department of Radiology, Southwest Hospital, Third Military Medical University (Army Medical University), No. 30, Gaotan Yanzheng Street, Shapingba District, Chongqing 400038, China

⁹ Department of Cardiology, The Second Affiliated Hospital of Chongqing Medical University, No. 288, Tiantian Avenue, Nan'an District, Chongqing 400010, China

Introduction

Ischemic heart disease (IHD) is the leading cause of death and emergency department (ED) admission, which exacerbates the burden of society and individuals [1, 2]. Since a large proportion of ED visits are avoidable [3, 4], identifying those who require immediate attention among the large number of patients who do not require urgent care is critical to ensuring patient safety, especially in overcrowded environment, which might help improve the appropriate use of ED visits and control health expenditures [5, 6].

Triaging as the first layer of emergency care plays an essential role in assessing and stratifying the risks of patients [7]. It is essential to identify the risk of adverse events in emergency patients with specific tool, such as early warning scores [8, 9]. One such tool is the HEART score, which estimates the possibility of major adverse events in patients presenting to the ED due to symptoms of acute coronary syndrome within 30 days [10]. The components are based on medical history, past medical risk factors, age, ECG interpretation, and serum troponin results. However, some variables are not instantaneous and shorter term outcomes are lacking.

To address the need for more accurate and straightforward predictive tools, we aimed to develop scoring models, based on machine learning algorithm, to improve the prioritization of IHD patients visiting the ED.

Methods

Study population

This study was a retrospective study in patients with IHD from the Medical Information Mart for Intensive Care IV (MIMIC-IV) database [11]. The data of MIMIC-IV are sourced from two inpatient database systems: the customized digital electronic health record system (EHR) in hospitals and the intensive care unit (ICU) specific clinical information system. The medical records of patients who admitted to an ICU or ED between 2008 and 2019 were contained in this database and have been reorganized and de-identified [12]. The Institutional Review Board at the Beth Israel Deaconess Medical Center reviewed the research resource, granted a waiver of informed consent, and approved the data sharing initiative.

Patients diagnosed with IHD were extracted according to the 2016 edited International Classification of Diseases, Ninth Revision (ICD-9). IHD related codes from chapter 7 were selected, including code 410 of acute myocardial infarction, code 411 of other acute and subacute

forms of IHD, code 412 of old myocardial infarction, and code 413 of angina pectoris. All patients who visited ED were screened. Patients who died within 24 hours after ED admission or who were under the age of 18 years old were excluded. All the included patients were randomly divided into the training set with the percentage of 70%, the validation set with 10%, and the testing set with 20%.

Outcome

The primary outcomes used to develop and test the tool were ICU stay and 3d-death. ICU stay was defined as transferring to ICU for advanced therapy after emergency admission. The 3d-death was defined as deaths within 72 hours after the time of admission. The secondary outcomes were deaths within 7 and 30 days after emergency admission. Both hospital and post-hospital deaths were recorded in MIMIC-IV database.

Data/variables extraction

All data and variables were evaluated and recorded when patients visited ED. Age and gender were included as demographic characteristics [13–15]. Vital signs were included due to the availability and immediacy responding to critical states, including temperature, heart rate, respiratory rate, oxygen saturation, systolic blood pressure (SBP), and diastolic blood pressure (DBP) [16]. History of smoke was included as a common risk factor for cardiovascular disease [17]. Major comorbidities were included as following: hypertension, atrial fibrillation, heart failure, chronic pulmonary disease, diabetes without chronic complication, diabetes with complication, renal disease, and malignant cancer. We also included variables that reflect the severity or urgency of IHD, including arrival transport, pain_max score, and intubated state. The arrival transport was included because it could reflect the urgency of ED admission [18, 19]. Pain_max score is based on the pain level when the patient enters ED [20], from no pain to total intolerability corresponding to a score from 0 to 10. Patient with intubated state suggests a severe condition [21, 22], such as respiratory failure or loss of consciousness.

Statistical analysis

In the description of the baseline characteristics, categorical variables were reported as numbers with percentages. The continuous variables in this study were represented as the median with quartile (Q1, Q3) with nonnormal distributions. Values of vital signs that were beyond the range of clinical and physiological cognition were considered outliers. We used K-Nearest Neighbor to interpolate the missing values.

AutoScore, a machine learning-based automatic clinical score generator, was applied to derive the scoring models in this study [23]. Its application to mortality prediction using EHR data increases the implementation and validation in the clinic. The training set was used to generate a preliminary scoring model using the AutoScore framework. The validation set was used to evaluate parameter tuning and model selection for candidate scoring models. Four scoring models were finally derived to predict the possibility of ICU stay, 3d-, 7d-, and 30d-death, respectively.

The performance metrics of the final scoring models were analyzed in the testing cohort. In addition, all the four scoring models were applied to predict the possibility of mortality within 3 days after emergency admission in the testing set and the performance were evaluated. The predictive power of scoring models was measured with areas under the curve (AUC) in the receiver operating characteristic (ROC) analysis. Sensitivity, specificity, positive predictive value, and negative predictive value were calculated under the optimal threshold and were reported with 95% confidence interval (95% CI). The data were analyzed using R software (version 3.5.3). $P < 0.05$ is considered statistically significant.

Results

Baseline characteristics

A total of 8381 IHD patients from ED were included in the present study (median patient age, 71 years, 95% CI 62–81; 3035 [36%] female), in which 5867 episodes were randomly assigned to the training set (median patient age, 72 years, 95% CI 62–81; 2142 [37%] female), 838 episodes to the validation set (median patient age, 71 years, 95% CI 62–81; 311 [37%] female), and 1676 to the testing set (median patient age, 71 years, 95% CI 62–80; 582 [35%] female). The flow diagram of patient selection and study cohort formation is presented in Fig. 1. In total, there were 2551 (30%) patients transferred into ICU after visiting ED. The mortality rates observed in the total cohort were 1% with 123 cases at 3 days, 3% with 245 cases at 7 days, and 7% with 622 cases at 30 days. The baseline characteristics in the training, validation, and testing cohorts were similar in terms of age, gender, vital signs, comorbidities, and other characteristics (Table 1).

Selected variables and scoring models

We performed AutoScore to select the most discriminative variables from all 20 baseline variables. Parsimony plots based on the validation set were applied for determining the choice of variables (Supplemental Figs. 1–4). With good balance in the parsimony plot, we chose 7 variables

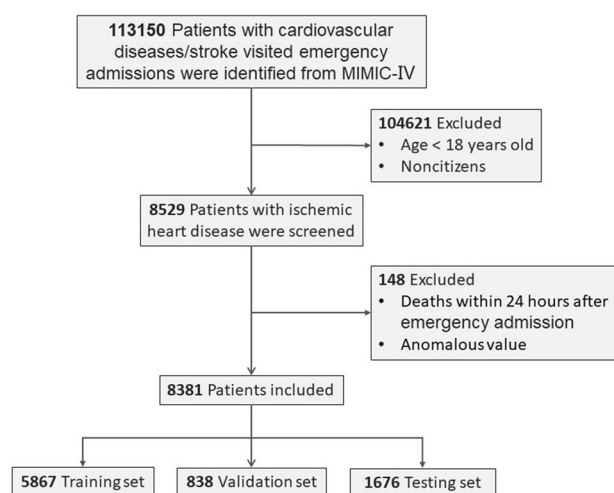


Fig. 1 Flow of patient selection and study cohort formation

for the prediction of ICU stay, including age, SBP, DBP, temperature, heart rate, respiratory rate, and arrival transport. Nine variables were chosen as the parsimonious choice for 3d-death and 7d-death predicting with the addition of oxygen saturation and pain_max score. On this basis, 10 variables with the addition of atrial fibrillation were chosen for 30d-death estimation achieving good performance in the scoring model. When more variables were added to the scoring model for the outcomes, their performance were not improved significantly.

In the ICU-stay scoring model, the scores are mainly contributed by the arrival transport of helicopter, SBP < 90 mmHg, respiratory rate ≥ 24 beats per minute (bpm), heart rate ≥ 115 bpm, temperature < 35.8°C, DBP < 50 mmHg, and age < 45 or ≥ 65 years old (Table 2). In the 3d-death scoring model, the scores are mainly contributed by the arrival transport of helicopter, temperature < 35.8 °C, SBP < 90 mmHg, age ≥ 83 years old, respiratory rate ≥ 30 bpm, heart rate ≥ 96 bpm, oxygen saturation < 93%, DBP ≥ 95 mmHg, and pain_max score ≥ 7 (Table 2). In the scoring models of 7d-death and 30d-death, the trends of contribution to scores were similar with 3d-death, including more urgent admission transport, older age, lower SBP, lower temperature and oxygen saturation, faster respiratory rates and heart rates, higher DBP, and with painful symptoms. Patients with the comorbidity of malignant cancer were added extra scores in the prediction of 30d-death.

The scores of these scoring models based on the analysis of the variables included were all around 0–69. The possibility of predicted risk, patient ratios, and the evaluation of performance in each scoring models for different scoring intervals based on the testing cohort are shown in Table 3. For ICU stay and 3d-death scoring models, most patients

Table 1 Baseline characteristics of the study cohort

Variables	Total (<i>n</i> = 8381)	Training set (<i>n</i> = 5867)	Validation set (<i>n</i> = 838)	Testing set (<i>n</i> = 1676)	<i>P</i> [#]
Age [†] , years old	71 (62, 81)	72 (62, 81)	71 (62, 81)	71 (62, 80)	0.565
Gender, <i>n</i> (%)					0.346
Female	3035 (36)	2142 (37)	311 (37)	582 (35)	
Male	5346 (64)	3725 (63)	527 (63)	1094 (65)	
History of smoke, <i>n</i> (%)					0.767
No	5293 (63)	3691 (63)	536 (64)	1066 (64)	
Yes	3088 (37)	2176 (37)	302 (36)	610 (36)	
Arrival transport, <i>n</i> (%)					0.882
Helicopter	73 (1)	50 (1)	8 (1)	15 (1)	
Ambulance	5168 (62)	3615 (62)	506 (60)	1047 (62)	
Walk in	2315 (28)	1631 (28)	242 (29)	442 (26)	
Others	825 (10)	571 (10)	82 (10)	172 (10)	
Temperature [†] , °C	36.67 (36.39, 36.89)	36.67 (36.39, 36.89)	36.61 (36.39, 36.89)	36.67 (36.39, 36.89)	0.969
Heart rate [†] , bpm	80 (69, 93)	80 (69, 92)	80 (69, 92.75)	80 (69, 93.25)	0.557
Respiratory rate [†] , bpm	18 (16, 19)	18 (16, 19)	18 (16, 19)	18 (16, 18)	0.458
Oxygen saturation [†] , %	98 (96, 99)	98 (96, 99)	98 (96, 99)	98 (96, 99)	0.027
SBP [†] , mmHg	135 (118, 153)	135 (118, 153)	135 (117, 154)	135 (119, 153)	0.932
DBP [†] , mmHg	74 (64, 84.94)	74 (64, 84)	75 (64, 86)	74 (64, 84)	0.745
Pain_max score, <i>n</i> (%)					0.798
0	4246 (51)	2957 (50)	435 (52)	854 (51)	
1	169 (2)	124 (2)	14 (2)	31 (2)	
2	406 (5)	287 (5)	40 (5)	79 (5)	
3	370 (4)	255 (4)	33 (4)	82 (5)	
4	437 (5)	320 (5)	35 (4)	82 (5)	
5	559 (7)	398 (7)	45 (5)	116 (7)	
6	406 (5)	273 (5)	50 (6)	83 (5)	
7	355 (4)	245 (4)	33 (4)	77 (5)	
8	577 (7)	409 (7)	55 (7)	113 (7)	
9	205 (2)	140 (2)	25 (3)	40 (2)	
10	651 (8)	459 (8)	73 (9)	119 (7)	
Intubated, <i>n</i> (%)					0.655
No	8349 (100)	5845 (100)	836 (100)	1668 (100)	
Yes	32 (0)	22 (0)	2 (0)	8 (0)	
Hypertension, <i>n</i> (%)					0.790
No	1746 (21)	1214 (21)	182 (22)	350 (21)	
Yes	6635 (79)	4653 (79)	656 (78)	1326 (79)	
Atrial fibrillation, <i>n</i> (%)					0.752
No	6002 (72)	4206 (72)	591 (71)	1205 (72)	
Yes	2379 (28)	1661 (28)	247 (29)	471 (28)	
Heart failure, <i>n</i> (%)					0.546
No	5483 (65)	3833 (65)	538 (64)	1112 (66)	
Yes	2898 (35)	2034 (35)	300 (36)	564 (34)	
Chronic pulmonary disease, <i>n</i> (%)					0.407
No	6464 (77)	4520 (77)	661 (79)	1283 (77)	
Yes	1917 (23)	1347 (23)	177 (21)	393 (23)	
Diabetes without chronic complication, <i>n</i> (%)					0.636
No	6230 (74)	4349 (74)	620 (74)	1261 (75)	
Yes	2151 (26)	1518 (26)	218 (26)	415 (25)	
Diabetes with chronic complication, <i>n</i> (%)					0.408

Table 1 (continued)

Variables	Total (<i>n</i> =8381)	Training set (<i>n</i> =5867)	Validation set (<i>n</i> =838)	Testing set (<i>n</i> =1676)	<i>P</i> [#]
No	6891 (82)	4844 (83)	678 (81)	1369 (82)	0.199
Yes	1490 (18)	1023 (17)	160 (19)	307 (18)	
Renal disease, <i>n</i> (%)					0.199
No	6156 (73)	4306 (73)	598 (71)	1252 (75)	
Yes	2225 (27)	1561 (27)	240 (29)	424 (25)	0.916
Malignant cancer, <i>n</i> (%)					
No	7548 (90)	5280 (90)	754 (90)	1514 (90)	0.215
Yes	833 (10)	587 (10)	84 (10)	162 (10)	
ICU stay, <i>n</i> (%)					0.215
No	5830 (70)	4084 (70)	601 (72)	1145 (68)	
Yes	2551 (30)	1783 (30)	237 (28)	531 (32)	0.563
3d-death, <i>n</i> (%)					
Survival	8258 (99)	5786 (99)	825 (98)	1647 (98)	0.251
Death	123 (1)	81 (1)	13 (2)	29 (2)	
7d-death, <i>n</i> (%)					0.357
Survival	8136 (97)	5707 (97)	811 (97)	1618 (97)	
Death	245 (3)	160 (3)	27 (3)	58 (3)	0.357
30d-death, <i>n</i> (%)					
Survival	7759 (93)	5445 (93)	776 (93)	1538 (92)	0.357
Death	622 (7)	422 (7)	62 (7)	138 (8)	

n number, *bpm* beats per minute, *SBP* systolic blood pressure, *DBP* diastolic blood pressure, *ICU* intensive care unit, *3d* 3 days, *7d* 7 days, *30d* 30 days

[†]Data are presented as median with quartile (Q1, Q3);

[#]The *p* value reflects the difference between baseline characteristics in the training, validation, and test sets. *P*<0.05 is considered statistically significant

had a risk score between 10 and 34, and few patients had scores under 5 or above 45. For patients who died within 7 days and 30 days of emergency admission, the risk score appeared to increase to the range of 26–48, with few patients scoring below 23 or above 53.

Performance evaluation

We used scoring models of ICU stay, 3d-death, 7d-death, and 30d-death to predict the possibility of transferring to ICU, death within 3 days, death within 7 days, and death within 30 days in patients after emergency admissions, respectively (Table 4). Among the testing cohorts, the scoring model achieved good performance for shorter term and longer term mortality prediction, with an AUC of 0.7551 (95% CI 0.7297–0.7805) for ICU stay, an AUC of 0.7856 (95% CI 0.7166–0.8545) for 3d-death, an AUC of 0.7371 (95% CI 0.6665–0.8077) for 7d-death, and an AUC of 0.7407 (95% CI 0.6972–0.7842) for 30d-death. Among the three cohorts in each outcome, the scoring model achieved the best performance for mortality with 3 days prediction, with an AUC of 0.8243 (95% CI 0.7771–0.8716) in training set, an AUC of 0.7877 (95% CI 0.6593–0.9161) in validation

set, and an AUC of 0.7856 (95% CI 0.7166–0.8545) in testing set. We also applied the four scoring models to predict the possibility of mortality within 3 days after emergency admission in testing cohort and the performance are reported in Fig. 2. The 30d-death scoring model achieved the best performance for 3d-death prediction with an AUC of 0.790.

Discussion

In this retrospective study with a population of IHD patients presenting to the ED, newly developed scoring models were able to identify patients with high-urgency of ICU stay, 3d-death, 7d-death, and 30d-death. To the best of our knowledge, this is the first validated point-based clinical tool based on machine learning algorithm to be developed specifically for identifying patients who transferred into ICU or died during shorter or longer term care in the ED. Scoring models showing good discriminative performance in which 7–10 variables were selected to calculate a score as following: age, arrival transport, temperature, SBP, DBP, respiratory

Table 2 Scoring models derived from the primary and secondary outcomes

Variable	ICU stay		3d-death score		7d-death score		30d-death score	
	Interval	Point	Interval	Point	Interval	Point	Interval	Point
SBP, mmHg	<90	19	<90	15	<90	11	<90	20
	[90, 115)	7	[90, 120)	5	[90, 140)	6	[90, 130)	17
	≥ 115	0	[120, 180)	0	[140, 160)	0	[130, 160)	14
			≥ 180	4	≥ 160	3	[160, 200)	11
							≥ 200	0
Temperature, °C	<35.8	8	<35.8	19	<35.8	14	<35.8	11
	[35.8, 36.3)	0	[35.8, 36.3)	0	[35.8, 36.3)	0	[35.8, 36.3)	0
	[36.3, 37.3)	2	≥ 36.3	4	[36.3, 37.1)	2	[36.3, 37)	2
	[37.3, 38)	0			≥ 37.1	1	[37, 38)	2
	≥ 38	6					≥ 38	5
DBP, mmHg	<50	7	<80	0	<70	1	<70	3
	[50, 60)	1	[80, 95)	2	[70, 100)	0	[70, 90)	2
	[60, 90)	0	≥ 95	5	≥ 100	6	[90, 110)	0
	≥ 90	3					≥ 110	8
Heart rate, bpm	<56	2	<66	0	<50	2	<65	0
	[56, 90)	0	[66, 96)	4	[50, 90)	0	[65, 95)	3
	[90, 115)	6	≥ 96	7	[90, 120)	5	[95, 115)	7
	≥ 115	9			≥ 120	6	≥ 115	3
Age, years old	<45	2	<48	0	<45	0	<48	0
	[45, 65)	0	[48, 60)	7	[45, 65)	6	[48, 60)	9
	≥ 65	2	[60, 83)	4	[65, 85)	9	[60, 83)	12
			≥ 83	12	≥ 85	15	≥ 83	18
Respiratory rate, bpm	<20	0	<22	0	<15	0	<20	0
	[20, 24)	5	[22, 30)	9	[15, 20)	7	[20, 24)	6
	≥ 24	14	≥ 30	11	[20, 24)	12	≥ 24	8
					≥ 24	13		
Arrival transport	Helicopter	41	Helicopter	23	Helicopter	28	Helicopter	16
	Ambulance	13	Ambulance	11	Ambulance	15	Ambulance	6
	Walk in	0	Walk in	0	Walk in	0	Walk in	0
	Others	32	Others	19	Others	22	Others	10
Oxygen saturation, %	–		<93	7	<93	7	<91	7
			[93, 96)	0	≥ 93	0	≥ 91	0
			≥ 96	1				
Pain_max score	–		<7	0	<8	0	<7	0
			≥ 7	2	≥ 8	0	≥ 7	1
Malignant cancer	–		–		–		No	0
							Yes	3

bpm beats per minute, SBP systolic blood pressure, DBP diastolic blood pressure, ICU intensive care unit, 3d 3 days, 7d 7 days, 30d 30 days

Table 3 Varying cutoffs of predicted risk based on the scoring models of ICU stay, 3d-death, 7d-death, and 30d-death, the proportion of patients stratified for the primary and secondary outcomes, and the corresponding sensitivity, specificity, positive and negative predictive values on the testing cohort

Score cut-off	Predicted risk	Percentage of patients (%)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
ICU stay scoring model							
≥ 0	≥ 1%	100	31.7% (31.7–31.7%)	100% (100–100%)	0% (0–0%)	31.7% (31.7–31.7%)	NA% (NA–NA%)
≥ 5	≥ 9%	90	39.4% (38–40.8%)	97% (95.5–98.3%)	12.7% (10.8–14.6%)	34% (33.4–34.6%)	90.2% (85.4–94.2%)
≥ 7	≥ 10%	86	42.8% (41.2–44.5%)	96% (94.4–97.6%)	18.1% (15.9–20.4%)	35.2% (34.5–36%)	90.8% (86.9–94.2%)
≥ 10	≥ 13.3%	80	47.6% (45.8–49.4%)	93.4% (91.3–95.5%)	26.3% (23.8–28.7%)	37% (36.1–38%)	89.6% (86.4–92.7%)
≥ 16	≥ 20%	69	54.7% (52.6–56.9%)	87.8% (84.9–90.4%)	39.4% (36.5–42.3%)	40.1% (38.9–41.6%)	87.4% (84.7–90%)
≥ 25	≥ 35%	30	73.2% (71.2–75.1%)	55.2% (50.8–59.5%)	81.5% (79.3–83.6%)	58.1% (54.6–61.4%)	79.7% (78.1–81.3%)
≥ 32	≥ 50%	18	76.1% (74.3–77.7%)	41.4% (37.5–45.6%)	92.1% (90.5–93.7%)	71% (66.3–75.5%)	77.2% (76–78.5%)
≥ 35	≥ 56.7%	14	75.6% (74–77.2%)	33.7% (29.6–37.9%)	95% (93.8–96.2%)	75.9% (70.7–81%)	75.6% (74.4–76.7%)
≥ 45	≥ 75.5%	5	72.1% (71.1–73.1%)	13.9% (11.1–16.9%)	99% (98.4–99.6%)	87.3% (79.7–93.8%)	71.3% (70.6–72%)
3d-death scoring model							
≥ 4	≥ 4.6%	100	31.5% (31.3–31.7%)	99.4% (98.7–100%)	0% (0–0%)	31.6% (31.4–31.7%)	0% (0–0%)
≥ 5	≥ 5%	100	31.6% (31.3–31.7%)	99.4% (98.7–100%)	0.1% (0–0.3%)	31.6% (31.4–31.7%)	20% (0–100%)
≥ 14	≥ 10%	92	38.4% (37.1–39.7%)	97.6% (96.2–98.9%)	10.9% (9.2–12.8%)	33.7% (33.2–34.3%)	90.7% (85.8–95.2%)
≥ 23	≥ 20%	72	52.5% (50.5–54.6%)	89.1% (86.3–91.7%)	35.5% (32.8–38.3%)	39.1% (37.8–40.3%)	87.6% (84.7–90.3%)
≥ 24	≥ 22.4%	62	57.5% (55.2–59.8%)	81.4% (78.2–84.6%)	46.6% (43.4–49.4%)	41.4% (39.7–43%)	84.3% (82–86.8%)
≥ 29	≥ 30%	46	65.9% (63.5–68%)	68.5% (64.6–72.5%)	64.6% (61.8–67.4%)	47.3% (44.9–49.7%)	81.6% (79.6–83.5%)
≥ 34	≥ 41.2%	27	71.2% (69.2–73.1%)	47.1% (42.9–51.4%)	82.4% (80.2–84.6%)	55.4% (51.5–59.1%)	77.1% (75.6–78.6%)
≥ 38	≥ 50%	16	73.1% (71.5–74.8%)	32.4% (28.2–36.5%)	92.1% (90.5–93.6%)	65.3% (60.2–70.6%)	74.6% (73.4–75.8%)
≥ 44	≥ 63.1%	6	71.5% (70.4–72.6%)	14.1% (11.1–17.1%)	98.2% (97.4–99%)	78.2% (69.4–86.2%)	71.1% (70.4–71.9%)
≥ 54	≥ 80.6%	2	69.6% (68.9–70.2%)	4.9% (3.2–6.8%)	99.6% (99.1–99.9%)	84.6% (69.6–96%)	69.3% (68.9–69.7%)
7d-death scoring model							
≥ 23	≥ 0.8%	85	17.9% (16.2–19.6%)	94.8% (87.9–100%)	15.1% (13.3–16.9%)	3.9% (3.6–4.1%)	98.8% (97.3–100%)
≥ 26	≥ 1%	81	22.4% (20.6–24.4%)	94.8% (87.9–100%)	19.8% (17.9–21.8%)	4.1% (3.8–4.3%)	99.1% (97.9–100%)
≥ 33	≥ 1.8%	65	37.3% (35–39.6%)	87.9% (79.3–94.8%)	35.5% (33.1–37.8%)	4.7% (4.2–5.1%)	98.8% (97.9–99.5%)
≥ 43	≥ 4%	28	72.3% (70.2–74.3%)	60.3% (46.6–72.4%)	72.7% (70.5–74.8%)	7.4% (5.8–8.7%)	98.1% (97.5–98.6%)
≥ 46	≥ 5%	20	80.4% (78.5–82.2%)	48.3% (36.2–60.3%)	81.5% (79.6–83.4%)	8.6% (6.4–10.7%)	97.8% (97.3–98.3%)

Table 3 (continued)

Score cut-off	Predicted risk	Percentage of patients (%)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
≥ 53	≥ 8.7%	5	93.2% (92.1–94.2%)	24.1% (13.8–36.2%)	95.7% (94.6–96.6%)	16.7% (9.9–24.4%)	97.2% (96.9–97.7%)
≥ 55	≥ 10%	5	93.5% (92.5–94.5%)	24.1% (13.8–36.2%)	96% (95–96.9%)	17.6% (10.6–25.4%)	97.2% (96.9–97.7%)
≥ 63	≥ 17.8%	1	96.2% (95.7–96.7%)	12.1% (5.2–20.7%)	99.3% (98.8–99.6%)	36.8% (15.8–58.8%)	96.9% (96.7–97.2%)
≥ 65	≥ 20%	1	96.4% (96.1–96.8%)	6.9% (1.7–13.8%)	99.6% (99.3–99.9%)	40% (12.5–72.7%)	96.8% (96.6–97%)
30d-death scoring model							
≥ 24	≥ 1%	98	10.3% (9.7–11%)	100% (100–100%)	2.3% (1.6–3.1%)	8.4% (8.4–8.5%)	100% (100–100%)
≥ 29	≥ 1.8%	94	14.3% (13.2–15.5%)	98.6% (96.4–100%)	6.8% (5.5–8.1%)	8.7% (8.4–8.8%)	98.2% (95.1–100%)
≥ 39	≥ 4.7%	62	44.3% (42–46.6%)	87% (81.2–92%)	40.5% (38–42.9%)	11.6% (10.8–12.4%)	97.2% (96–98.3%)
≥ 40	≥ 5%	62	44.3% (42.1–46.5%)	87% (81.9–92%)	40.5% (38.1–42.9%)	11.6% (10.8–12.4%)	97.2% (96.1–98.3%)
≥ 48	≥ 10%	24	76.6% (74.5–78.5%)	55.1% (46.4–63.8%)	78.5% (76.4–80.4%)	18.7% (16–21.5%)	95.1% (94.2–96%)
≥ 49	≥ 11.8%	22	78.5% (76.5–80.4%)	51.4% (43.5–60.1%)	80.8% (78.8–82.8%)	19.5% (16.5–22.5%)	94.9% (94–95.7%)
≥ 56	≥ 20%	6	89% (87.9–90.1%)	20.3% (13.8–26.8%)	95.2% (94.1–96.2%)	27.5% (19.4–35.4%)	93% (92.5–93.6%)
≥ 59	≥ 26.9%	3	90.7% (89.9–91.5%)	14.5% (8.7–20.3%)	97.5% (96.7–98.3%)	34.4% (22.7–46.3%)	92.7% (92.2–93.2%)
≥ 69	≥ 50%	0	91.5% (91.2–91.8%)	1.4% (0–3.6%)	99.6% (99.3–99.9%)	25% (0–62.5%)	91.8% (91.7–92%)

PPV positive predictive value, NPV negative predictive value, ICU intensive care unit, 3d 3 days, 7d 7 days, 30d 30 days

rate, heart rate, oxygen saturation, pain_max score, and comorbidity of malignant cancer.

IHD remains a major threat to public health, and overall burden is increasing globally [1]. The total number of prevalent cases, deaths, and DALYs due to IHD increased steadily from 1990, reaching 197 million prevalent cases, 9.14 million deaths, and 182 million DALYs due to IHD in 2019 [1]. In the present study, we found 30% IHD patients transferred into ICU after visiting ED and the mortality within 3 days, 7 days, and 30 days were 1%, 3%, and 7%, respectively. Although our study excluded patients who died within 24 hours in ED to ensure the reliability of the analysis results, a mortality of 0.98% was observed in the process of screening. Therefore, the disease burden caused by IHD is practically heavy. More accurate and parsimonious scoring models in this study help stratifies patients with high-risk in ED, which is conducive to patient safety and rational allocation of medical resources.

In the crowded and rapidly changing ED, doctors need to make the fastest judgment on the patient's condition and provide the timeliest intervention [24]. Physicians must

undergo dual specialist and general internal medicine training, as a high proportion of inpatients have multiple comorbidities [25]. Combined with machine learning and logistic regression, AutoScore automatically develops a minimalistic sparse score risk model for predefined outcomes, enabling users to quickly and seamlessly build interpretable clinical scores that can be easily implemented and validated in clinical practice [23]. One study has revealed that objective criteria may be useful to physicians who are evaluating the appropriateness of admitting patients presenting to the ED to a short stay unit and to predict the LOS, which can both shorten hospital stays and reduce the cost of hospitalization [26]. In addition, the reported scoring systems have complex requirements for variables and lack prediction of adverse events in the very short term of ED [27–30]. In this study, both the candidate variables and those were ultimately selected in the scoring models are readily available in ED, making this clinical tool more implementable. Demographics, vital signs, and arrival transport are all indicators that can be obtained in ED in the first time, while more complex inspection

Table 4 Comparison of AUC values achieved by different triage scores on the training, validation, and testing cohort

	AUC	Sensitivity	Specificity	PPV	NPV
ICU stay scoring model					
Training set	0.7305 (0.7162–0.7447)	0.5294 (0.5048–0.5519)	0.8102 (0.7982–0.8225)	0.5492 (0.5294–0.569)	0.7977 (0.789–0.8059)
Validation set	0.7521 (0.7153–0.7888)	0.5593 (0.4999–0.6186)	0.8083 (0.7750–0.8433)	0.5341 (0.4855–0.5894)	0.8233 (0.8023–0.8447)
Testing set	0.7551 (0.7297–0.7805)	0.6573 (0.6158–0.6949)	0.7197 (0.6917–0.7459)	0.5205 (0.4919–0.5477)	0.8193 (0.8004–0.8365)
3d-death scoring model					
Training set	0.8243 (0.7771–0.8716)	0.7531 (0.6543–0.8395)	0.7475 (0.7366–0.7589)	0.0401 (0.0351–0.0451)	0.9954 (0.9936–0.9970)
Validation set	0.7877 (0.6593–0.9161)	0.7692 (0.5385–1.0000)	0.7351 (0.7035–0.7643)	0.0439 (0.0295–0.0565)	0.9951 (0.9900–1.0000)
Testing set	0.7856 (0.7166–0.8545)	0.8276 (0.6897–0.9310)	0.6284 (0.6047–0.6527)	0.0378 (0.0310–0.0439)	0.9952 (0.9913–0.9981)
7d-death scoring model					
Training set	0.8012 (0.7701–0.8324)	0.7495 (0.6312–0.7688)	0.7599 (0.7491–0.7705)	0.0758 (0.0678–0.0835)	0.9891 (0.9865–0.9916)
Validation set	0.7813 (0.6888–0.8738)	0.7407 (0.5556–0.8889)	0.7196 (0.6871–0.7497)	0.0820 (0.0620–0.1000)	0.9880 (0.9798–0.9948)
Testing set	0.7371 (0.6665–0.8077)	0.7414 (0.6207–0.8621)	0.6335 (0.6100–0.6557)	0.0675 (0.0571–0.0786)	0.9855 (0.9791–0.9920)
30d-death scoring model					
Training set	0.7672 (0.7455–0.7890)	0.7488 (0.7062–0.7891)	0.6621 (0.649–0.6753)	0.1467 (0.1381–0.1551)	0.9715 (0.9666–0.976)
Validation set	0.7603 (0.6921–0.8285)	0.6129 (0.4839–0.7258)	0.826 (0.8003–0.8531)	0.2186 (0.1782–0.2635)	0.9639 (0.9528–0.9747)
Testing set	0.7407 (0.6972–0.7842)	0.7029 (0.6232–0.7754)	0.7061 (0.684–0.7302)	0.1769 (0.1566–0.1973)	0.9637 (0.9544–0.9725)

AUC area under the curve, PPV positive predictive value, NPV negative predictive value, ICU intensive care unit, 3d 3 days, 7d 7 days, 30d 30 days

indicators and imaging results are not screened because they are not available in time. The tendency of the selected variable values to be assigned points through cut-off segments is correspond to common clinical knowledge [24, 31]. For such example, in our scoring models, patients transported by helicopter to ED with shock blood pressure, breathless, and rapid heart rates are at high risk for transferring to ICU for advanced therapy. On this basis, patients with hypoxemia, elevated DBP, and painful symptoms have a higher risk of death within 3 days. Furthermore, the scoring models in our study for risk prediction show satisfactory performance through sensitivity and accuracy, and the prediction results obtained are highly reliable.

Transferring patients to ICU for advanced therapy and assessing outcomes in short term are important parts for emergency physician to triage patients from ED, and the correct decision to transfer to advanced care or not is both responsible for patient safety and rational use of medical resources [32–34]. Our study found that predictions of 3d-death showed

the best performance among the scoring models of 30d-death. This may be contributed by more variables were included in this model, pain_max score, oxygen saturation, and the comorbidity of malignant cancer, which increases the identification of high-urgency, resulting in greater reliability and accuracy in 3d-death prediction.

There are several limitations in the present study. Firstly, the variables included in the scoring models in this study are based on clinical routine collection, which may cause some parameters to appear similar between high-risk patients and other patients during an ED visit, making the prediction models less accurate. Then, we did not compare the performance of scoring models in this study with other scoring systems due to the limited types of variables in the MIMIC-IV database [11]. In addition, the scoring models based on the database in the United States have not been tested in other countries and regions, which may have present prediction performance due to differences in economic and medical conditions [35].

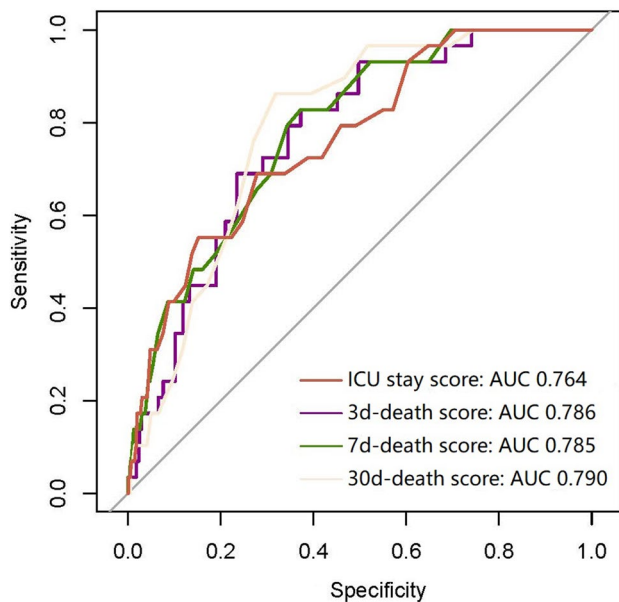


Fig. 2 Receiver operating characteristic curves of scoring models for predicting mortality within 3 days after emergency admission. *ACU* area under the curve, *ICU* intensive care unit, *3d* 3 days, *7d* 7 days, *30d* 30 days

Conclusion

The point-based clinical tool that we have successfully implemented based on machine learning algorithm offers an innovative approach to optimize the triaging of IHD patients in ED. This newly accurate and parsimonious scoring models present good discriminative performance for predicting the possibility of transferring to ICU, 3d-death, 7d-death, and 30d-death in IHD patients visiting ED.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11739-023-03199-7>.

Author contributions TS was responsible for the drafting and reviewing of the manuscript; JH was responsible for the statistical analysis; JD participated was responsible for data acquiring; HC and YZ participated in material support; YW participated in the discussion part; XL and XH provided the study concept and design. All authors have read and approved the final version of the manuscript.

Funding None.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The collection of patient information and creation of the research resource were reviewed by the Institutional Review Board

at the Beth Israel Deaconess Medical Center, who granted a waiver of informed consent and approved the data sharing initiative.

References

- Roth GA, Mensah GA, Johnson CO et al (2020) Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *J Am Coll Cardiol* 76(25):2982–3021
- Zhou M, Wang H, Zeng X et al (2019) Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* (London, England) 394(10204):1145–1158
- Adams JG (2013) Emergency department overuse: perceptions and solutions. *JAMA* 309(11):1173–1174
- McCusker J, Verdon J (2006) Do geriatric interventions reduce emergency department visits? a systematic review. *J Gerontol A Biol Sci Med Sci* 61(1):53–62
- Fingar KR, Barrett ML, Elixhauser A, Stocks C, Steiner CA (2015) Trends in Potentially Preventable Inpatient Hospital Admissions and Emergency Department Visits. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US)
- McDermott KW, Jiang HJ (2020) Characteristics and Costs of Potentially Preventable Inpatient Stays, 2017. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US)
- Hinson JS, Martinez DA, Cabral S et al (2019) Triage performance in emergency medicine: a systematic review. *Ann Emerg Med* 74(1):140–152
- Churpek MM, Yuen TC, Park SY, Meltzer DO, Hall JB, Edelson DP (2012) Derivation of a cardiac arrest prediction model using ward vital signs*. *Crit Care Med* 40(7):2102–2108
- Fernandes M, Vieira SM, Leite F, Palos C, Finkelstein S, Sousa JMC (2020) Clinical decision support systems for triage in the emergency department using intelligent systems: a review. *Artif Intell Med* 102:101762
- Sharp AL, Wu YL, Shen E et al (2018) The HEART score for suspected acute coronary syndrome in U.S. emergency departments. *J Am Coll Cardiol* 72(15):1875–1877
- Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R (2022) MIMIC-IV (version 2.0). *PhysioNet*
- Goldberger AL, Amaral LA, Glass L et al (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):E215–E220
- Hayashi T, Kawashima S, Nomura H et al (2011) Age, gender, insulin and blood glucose control status alter the risk of ischemic heart disease and stroke among elderly diabetic patients. *Cardiovasc Diabetol* 10:86
- Vaccarino V, Sullivan S, Hammad M et al (2018) Mental stress-induced-myocardial ischemia in young patients with recent myocardial infarction: sex differences and mechanisms. *Circulation* 137(8):794–805
- Senior R, Reynolds HR, Min JK et al (2022) Predictors of left main coronary artery disease in the ISCHEMIA trial. *J Am Coll Cardiol* 79(7):651–661
- Vidal-Petiot E, Ford I, Greenlaw N et al (2016) Cardiovascular event rates and mortality according to achieved systolic and diastolic blood pressure in patients with stable coronary artery disease: an international cohort study. *Lancet* (London, England) 388(10056):2142–2152
- Barry J, Mead K, Nabel EG et al (1989) Effect of smoking on the activity of ischemic heart disease. *JAMA* 261(3):398–402

18. Schull MJ, Morrison LJ, Vermeulen M, Redelmeier DA (2003) Emergency department overcrowding and ambulance transport delays for patients with chest pain. *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne* 168(3):277–283
19. Reiner-Deitemyer V, Teuschl Y, Matz K et al (2011) Helicopter transport of stroke patients and its influence on thrombolysis rates: data from the Austrian Stroke Unit Registry. *Stroke* 42(5):1295–1300
20. Mehta PK, Bess C, Elias-Smale S et al (2019) Gender in cardiovascular medicine: chest pain and coronary artery disease. *Eur Heart J* 40(47):3819–3826
21. Seitz KP, Gandotra S, Casey JD (2022) Emergency tracheal intubation: a procedure in need of better evidence. *Am J Respir Crit Care Med* 206(4):370–371
22. Andersen LW, Granfeldt A, Callaway CW et al (2017) Association between tracheal intubation during adult in-hospital cardiac arrest and survival. *JAMA* 317(5):494–506
23. Xie F, Chakraborty B, Ong MEH, Goldstein BA, Liu N (2020) AutoScore: a machine learning-based automatic clinical score generator and its application to mortality prediction using electronic health records. *JMIR Med Inform* 8(10):e21798
24. Calder LA, Perry J, Yan JW et al (2021) Adverse events among emergency department patients with cardiovascular conditions: a multicenter study. *Ann Emerg Med* 77(6):561–574
25. Smyth HD, Gorey SD, O'Keefe HD et al (2022) Generalist vs specialist acute medical admissions—what is the impact of moving towards acute medical subspecialty admissions on efficacy of care provision? *Eur J Intern Med* 98:47–52
26. Capone F, Molinari L, Noale M et al (2021) Admission criteria for a cardiovascular short stay unit: a retrospective analysis on a pilot unit. *Intern Emerg Med* 16(8):2087–2095
27. Williams JM, Greenslade JH, Chu K, Brown AF, Lipman J (2016) Severity scores in emergency department patients with presumed infection: a prospective validation study. *Crit Care Med* 44(3):539–547
28. Moman RN, Loprinzi Brauer CE, Kelsey KM, Havyer RD, Lohse CM, Belloio MF (2017) PREDICTing mortality in the emergency department: external validation and derivation of a clinical prediction tool. *Acad Emerg Med* 24(7):822–831
29. Cardona M, O'Sullivan M, Lewis ET et al (2019) Prospective validation of a checklist to predict short-term death in older patients after emergency department admission in Australia and Ireland. *Acad Emerg Med* 26(6):610–620
30. Xia Y, Zou L, Li D et al (2020) The ability of an improved qSOFA score to predict acute sepsis severity and prognosis among adult patients. *Medicine* 99(5):e18942
31. Carpenter CR, Shelton E, Fowler S et al (2015) Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis. *Acad Emerg Med* 22(1):1–21
32. Armitage M, Flanagan D (2001) Improving quality measures in the emergency services. *J R Soc Med* 94(Suppl 39):9–12
33. Mahmud E, Dauerman HL, Welt FGP et al (2020) Management of acute myocardial infarction during the COVID-19 pandemic: a position statement from the Society for Cardiovascular Angiography and Interventions (SCAI), the American College of Cardiology (ACC), and the American College of Emergency Physicians (ACEP). *J Am Coll Cardiol* 76(11):1375–1384
34. Pendyal A, Rothenberg C, Scofi JE et al (2020) National trends in emergency department care processes for acute myocardial infarction in the United States, 2005 to 2015. *J Am Heart Assoc* 9(20):e017208
35. Gupta R, Yusuf S (2019) Challenges in management and prevention of ischemic heart disease in low socioeconomic status people in LLMICs. *BMC Med* 17(1):209

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.