Development and Validation of a Risk Score for Predicting 1-Year Mortality Risk of STEMI Based on a Nationwide Cohort in China

Si Chen
Peking University https://orcid.org/0000-0002-2865-1636
Qianzi Che
Department of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
Yan Zhang
Departments of Cardiology, Peking University First Hospital, Beijing, China, No.8 Xishiku St, Xicheng District, Beijing, China
Jia Jia
Departments of Cardiology, Peking University First Hospital, Beijing, China, No.8 Xishiku St, Xicheng District, Beijing, China
Yiqun Wu
Department of Epidemiology and Biostatistics, School of Public Health, Peking University, No.38, Xueyuan Road, Beijing 100191, China.
Yong Huo
Departments of Cardiology, Peking University First Hospital, Beijing, China, No.8 Xishiku St, Xicheng District, Beijing, China.
Dafang Chen (dafangchen@bjmu.edu.cn)
Department of Epidemiology and Biostatistics, School of Public Health, Peking University, No.38, Xueyuan Road, Beijing 100191, China.

Research

Keywords: ST elevation myocardial infarction, Real-world study, 1-year mortality risk, Risk assessment

Posted Date: February 24th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-236809/v1

Abstract

Background

A risk assessment for identifying long-term risk of post-discharge mortality in Chinese STEMI patients remains a concern. The aim of this study is to establish a bedside available risk scoring system for predicting 1-year mortality risk among Chinese STEMI patients.

Methods

STEMI patients(n=12611) were enrolled from the China STEMI Care Project Phase 2(CSCAP-2) collected between 2015 and 2016. Confounding bias was controlled using propensity score matching. Epidemiological, clinical, laboratory, and imaging variables, treatment strategy and medicine records were screened using extreme gradient boosting and nomogram according to the hazard ratio of Cox regression analysis to construct a predictive score. A validation cohort included 7342 patients collected in 2017 from CSCAP-2 was analyzed using receiver ROC and expectation (E)/observation (O) ratio to validate the risk scoring system.

Results

From 39 potential predictors, 8 variables were independent predictive factor and were included in the risk score: Killip class, early reperfusion strategy, Non-PCI intraoperative anticoagulants, heart rate, gender, age, anterior-wall myocardial infarction (AWMI) and inferior-wall myocardial infarction (IWMI). The new model demonstrated an excellent discrimination and calibration. The c-statistic and E/O ratio were 0.87(95%CI, 0.80-0.93) and 1.14(95%CI, 0.93-1.39) in the train set, 0.88 (95%CI, 0.78-0.96) and 1.15(95%CI, 0.85-1.56) in the test set, meanwhile, 0.89(95%CI, 0.82-0.95) and 1.00(95%CI, 0.81-1.23) in the validation cohort. The score has better sensitivity than the GRACE score and can recognize risk stratification among STEMI patients (P<0.001).

Conclusions

We developed a risk scoring system for predicting 1-year mortality risk of STEMI in a large Chinese population. The new score is easy-to-use and demonstrating a good discriminatory accuracy in predicting both short-term and long-term mortality risk in Chinese patients with STEMI.

Background

Acute coronary syndromes (ACS) represent up to 50% of the global cardiovascular disease burden[1], and are a leading cause of mortality and morbidity especially in China[2]. According to the recently report[3], the mortality of ACS has been on the rise since 2012 in China. In 2016, it was 113.46/100-thousand in city and 118.74/100-thousand in urban. ACS has a poor prognosis. The mortality in in-hospital and by 12 months' post-discharge remains high, particularly in ST-segment elevation myocardial infarction (STEMI)

patients, which is the serious type of ACS. Study reported that the cumulative mortality by 2 years post-discharge of ACS was 5.2%, meanwhile STEMI patients had a mortality of 7.7% in China[4].

Multiple factors, including a situation compounded by multifarious demographic factors, as well as type and severity of ACS, timely diagnosis and treatment, and evidence based secondary and tertiary prevention strategies, have influences on prognosis of ACS patients[5, 6]. Considering combined influences of multiple risk factors, risk assessment is a useful tool for assessing prognosis and guiding management of patients with ACS. The Global Registry of Acute Coronary Events (GRACE) risk model[7] is the most common stratification tool with highest validity for risk stratification among ACS patients. It was developed by population mostly from North America, South America and Europe [8]. Recently, the EPICOR (long-tErm follow uP of antithrombotic management patterns In acute CORonary syndrome patients) and EPICOR Asia studies used a national wide population including 35.5% Chinese patients to establish a risk-scoring model for 2-year post discharge survival in ACS patients[9, 10]. Although prior studies have constructed several risk scoring system for ACS, those risk scoring systems are lack of validation among Chinese STEMI patients. Besides, due to different economic and medical development would lead to different major impact factors of prognosis in STEMI patients, a risk scoring system for Chinese STEMI patients with a comparative performance of long-term outcomes remained to study in detail.

In this study, we aimed to explore relationships between patient demographics, medical history and management, thereby, facilitating the development of a reliable and user-friendly risk-scoring system for individual Chinese STEMI patient one-year survival in a large Chinese population. The subjects analyzed in our study were collected from China STEMI Care Project Phase 2 (CSCAP-2) which has involved more than 20,000 patients hospitalized and discharged following a STEMI event [11].

Methods

Study Population

This study was based on a national healthcare improvement project named China STEMI Care Project Phase 2 (CSCAP-2) [11]. It is a large national prospective cohort study of STEMI patients initiated in 2015. We collected patients between 2015 and 2016 from 236 hospitals among 23 districts in China. More detail about the project has been described previously. 13018 patients with STEMI were eligible for the study if they were aged 18 years or over; where multiple admissions for acute myocardial infarction (AMI) were recorded per person, only the first admission was included; admission within 30 days of onset. 407 patients were excluded for missing outcome value. Finally, 12611 subjects were included in the analysis as the development cohort (Fig. 1).

Data collection procedures

All enrolled subjects would answer a structured baseline questionnaire which was administered by a trained interviewer. The questionnaire collected information of age, gender, adverse life habits, demographic characteristics, medical history, clinical symptoms and auxiliary examination results, diagnosis and treatment process, and secondary preventive drugs. History of smoking was identified positive as smoked within 1 year. Diagnosis of diabetes, hypertension, cerebrovascular disease and chronic renal failure were based on clinical judgmental. Physical examination, including height, weight, heart rate (bmp) and blood pressure (Systolic blood pressure, SBP; Diastolic blood pressure, DBP) were measured according to a standard protocol. Body mass index (BMI) was calculated as weight (Kg)/ [height (m)] ^{2}. Information of treatment strategy, and secondary preventive drugs were collected from the medical records.

Outcomes

Patients were followed up from the admission to 1 year after discharge from the hospital. The outcomes during the hospitalization period were collected at the time of discharge from hospital according to the inpatient medical record. After discharge from the hospital, patients were followed up for mortality status in 1 month, 3 months, 6 months and 1 year by telephone. Follow-up outcome was the incidence of all-cause death. All-cause death analyzed in this study was the mortality of in-hospital, 6-month follow-up and 1-year follow-up.

Data validation

The validation cohort was enrolled from CSCAP-2 collected between 2017 and 2018 from 236 hospitals among 23 districts. After excluding 33 participants for missing outcome value, a total of 7342 patients were included in the validation analysis (Fig. 1). We compared the discrimination and calibration between new model and Fox model (one of the GRACE scores) to identify the validation of our scoring system. The GRACE score provides an estimate of the probability of death within 1 year of hospital discharge in patients with ACS. It was calculated according to the information documented in inpatient medical records.

Statistics analysis

All data represented as mean ± standard deviation (SD) for continuous variables and counts (percentages) for categorical variables. Missing value was imputed by multiple imputation.

As our study was based on real-world data, propensity score matching(PSM) was used to remove confounding bias before variable selection. All of 12611 patients were included for PSM. We performed one-to-ten matching by propensity score calculated using cox regression with a 0.02 caliper width. The PSM results represented 1-year mortality risk to each individual given a set of these covariates (gender, age and comorbidities including atrial fibrillation, heart failure, COPD, cardiac valvular surgery, peripheral artery disease, Hypertension, lipid metabolism disorders, Diabetes, chronic kidney disease and cerebrovascular disease). We considered it reasonable when standardized mean differences in percent balance improvement less than 0.1.

After PSM, 5676 participants (non-death: death of 5043:633) were matched successful and ready for variables selection through Extreme gradient boosting (XGboost) model. The parameters of XGBoost model was brought into GridSearchCV for traversal tuning. The optimal model was defined by the following metrics: accuracy(ACC) and area under curve (AUC) of the receiver operating characteristic curve (ROC), value of precision, recall, and F1-Measure. Interpretation of the XGBoost is carried out as a separate process after the model is trained. The parameter that measures the extent to which features contribute to the predictive ability of the model is called feature importance. Our study mainly applied two kinds of feature importance: gain metric of substitution importance (Permutance Importance) and SHapley Additive explanation (SHAP) value in XGBoost feature importance. The number of variables for risk scoring system was chosen by traversal procedure, according to ACC. Finally, the top 8 variables from the ranking were chosen to develop the Cox regression model and the risk scoring system.

Participants were divided into a training set (70%, 8827 patients) and a test set (30%, 3784 patients) by random sampling. In order to construct the risk scoring system, we developed the Cox regression model with firth penalized maximum likelihood estimation, which was particularly suitable for large number of events and the number of covariates [10]. The performance of the model was defined by and C-statistics and ROC curve. In ROC analysis, Youden index was the value presenting the best sensitivity and specificity for prediction of a given threshold. If the predicting ability is considered reasonable when the AUC is higher than 0.7 and strong when the AUC exceeds 0.8. Expectation (E)/Observation (O) ratios were used to assess model calibration. Usually, E/O ratio between 0.8–1.2 means a good calibration. According to the final Cox regression model and the critical mortality rate, we accessed the risk scoring system by constructing a nomogram to show the correspondence table between the total score and the risk prediction probability.

7342 patients were included in the validation analysis. ROC curve, C-index and E/O ratio was examined in the validation cohort for the scoring system external validation. We further compared the discrimination and calibration between new model and Fox model (one of the GRACE scores). We applied the STEMI risk score for risk stratification as 3 level according to mortality risk for < 1%, 1–3% and ≥ 3%. A Kaplan-Meier survival curve of 1-year all-cause death stratified by the risk level was explored to validate the practicability of the STEMI risk score.

Data analyses were performed using Python (version 3.7) with the scientific libraries “eli5”, “XGBoost” and SAS (version 9.4).

Results

Baseline Characteristics

A total of 12611 STEMI patients collected between 2015 and 2016 based on CSCAP-2 were included in the analysis. The total mortality was 5.19%.

Overall, the mean age(± SD) of participants were 61.3(± 12.6), and 79.11% were male. The diagnosis of hypertension, diabetes and cerebrovascular disease were the highest prevalence comorbidity as 48.5%,20.66%、10.49%, respectively. Chest pain was the most common symptom (95.85%). Most patients have low Killip class (76.88% of level 1). More details were showed in Table 1. All variables listed in Table 1 were used for variable selection procedure in XGBoost model analysis.

Table 1
Baseline characteristics in STEMI patients of the development cohort

Characteristic		Global (n = 12611)		All-cause Death				P
				No(n = 11964)		Yes(N = 647)		
		N	Per(%)	N	Per(%)	N	Per(%)	
Age		12611	61.3(± 12.6)	11964	60.8(± 12.5)	647	70.8 (± 11.8)	< 0.001
Gender	Male	9977	79.11	9577	75.94	400	61.82	< 0.001
	Female	2634	20.89	2387	18.93	247	38.18	
BMI(kg/m2)	18.5–24	4071	32.28	3854	30.56	217	33.54	0.027
	< 18.5	697	5.53	646	5.12	51	7.88	
	24–28	5314	42.14	5064	40.16	250	38.64	
	≥ 28	2529	20.05	2400	19.03	129	19.94	
Current Smoking	No	6420	50.91	5971	47.35	449	69.40	< 0.001
	Yes	6191	49.09	5993	47.52	198	30.60	
Comorbodity								
MI	No	11863	94.07	11277	89.42	586	90.57	< 0.001
	Yes	748	5.93	687	5.45	61	9.43	
PCI	No	11974	94.95	11358	90.06	616	95.21	0.757
	Yes	637	5.05	606	4.81	31	4.79	
CABG	No	12577	99.73	11934	94.63	643	99.38	0.079
	Yes	34	0.27	30	0.24	4	0.62	
Family history of coronary heart disease	No	12082	95.81	11457	90.85	625	96.60	0.301
	Yes	529	4.19	507	4.02	22	3.40	
atrial fibrillation	No	12410	98.41	11789	93.48	621	95.98	< 0.001
	Yes	201	1.59	175	1.39	26	4.02	

Characteristic		Global (n = 12611)		All-cause Death				P
				No(n = 11964)		Yes(N = 647)		
		N	Per(%)	N	Per(%)	N	Per(%)	
heart failure	No	12499	99.11	11875	94.16	624	96.45	< 0.001
	Yes	112	0.89	89	0.71	23	3.55	
COPD	No	12459	98.79	11831	93.81	628	97.06	< 0.001
	Yes	152	1.21	133	1.05	19	2.94	
Cardiac valvular surgery	No	12588	99.82	11944	94.71	644	99.54	0.085
	Yes	23	0.18	20	0.16	3	0.46	
Peripheral artery disease	No	12498	99.1	11862	94.06	636	98.30	0.026
	Yes	113	0.9	102	0.81	11	1.70	
Hypertension	No	6438	51.05	6180	49	258	39.88	< 0.001
	Yes	6173	48.95	5784	45.86	389	60.12	
Lipid metabolism disorders	No	11776	93.38	11163	88.52	613	94.74	0.151
	Yes	835	6.62	801	6.35	34	5.26	
Diabetes	No	10006	79.34	9556	75.78	450	69.55	< 0.001
	Yes	2605	20.66	2408	19.09	197	30.45	
Chronic kidney disease	No	12424	98.52	11805	93.61	619	95.67	< 0.001
	Yes	187	1.48	159	1.26	28	4.33	
Cerebrovascular disease	No	11300	89.6	10783	85.5	517	79.91	
	Yes	1311	10.4	1181	9.36	130	20.09	
Symptom								
Chest pain	No	527	4.18	500	3.96	27	4.17	< 0.001

Characteristic		Global (n = 12611)		All-cause Death				P
				No(n = 11964)		Yes(N = 647)		
		N	Per(%)	N	Per(%)	N	Per(%)	
	Yes	12084	95.82	11464	90.9	620	95.83	
SBP (mmHg)	≤ 140	9135	72.44	8613	68.3	522	80.68	0.994
	> 140	3476	27.56	3351	26.57	125	19.32	
DBP (mmHg)	≤ 90	10124	80.28	9567	75.86	557	86.09	< 0.001
	> 90	2487	19.72	2397	19.01	90	13.91	
Heart rate(per minute)	≤ 100	11526	91.4	11031	87.47	495	76.51	< 0.001
	> 100	1085	8.6	933	7.4	152	23.49	
Killip class	1	9747	77.29	9469	75.09	278	42.97	< 0.001
	2	1739	13.79	1601	12.7	138	21.33	
	3	481	3.81	430	3.41	51	7.88	
	4	644	5.11	464	3.68	180	27.82	
AWMI	No	6914	54.83	6628	52.56	286	44.20	< 0.001
	Yes	5697	45.17	5336	42.31	361	55.80	
IWMI	No	6644	52.68	6249	49.55	395	61.05	< 0.001
	Yes	5967	47.32	5715	45.32	252	38.95	
Left bundle branch block	No	12552	99.53	11916	94.49	636	98.30	< 0.001
	Yes	59	0.47	48	0.38	11	1.70	
Pathological Q wave	No	11306	89.65	10740	85.16	566	87.48	< 0.001
	Yes	1305	10.35	1224	9.71	81	12.52	
Time from onset to consultation ≤ 12 hours	No	2626	20.82	2452	19.44	174	26.89	0.063
	Yes	9985	79.18	9512	75.43	473	73.11	

Characteristic		Global (n = 12611)		All-cause Death				P
				No(n = 11964)		Yes(N = 647)		
		N	Per(%)	N	Per(%)	N	Per(%)	
Transfer from other hospitals	No	8219	65.17	7783	61.72	436	67.39	< 0.001
	Yes	4392	34.83	4181	33.15	211	32.61	
Medication								
Aspirin	No	1764	13.99	1618	12.83	146	22.57	0.225
	Yes	10847	86.01	10346	82.04	501	77.43	
Clopidogrel	No	5392	42.76	5088	40.35	304	46.99	< 0.001
	Yes	7219	57.24	6876	54.52	343	53.01	
Ticagrelor	No	8671	68.76	8208	65.09	463	71.56	0.026
	Yes	3940	31.24	3756	29.78	184	28.44	
Loading statin	No	8026	63.64	7585	60.15	441	68.16	0.114
	Yes	4585	36.36	4379	34.72	206	31.84	
Beta blockers applied within 24 hours of onset	No	7691	60.99	7219	57.24	472	72.95	0.014
	Yes	4920	39.01	4745	37.63	175	27.05	
Non-PCI intraoperative anticoagulants	No	2482	19.68	2247	17.82	235	36.32	< 0.001
	Yes	10129	80.32	9717	77.05	412	63.68	
Early reperfusion strategy	Ppci	8239	65.33	7961	63.13	278	42.97	< 0.001
	PCI after Trombolysis	197	1.56	190	1.51	7	1.08	
	Thrombolysis	645	5.11	606	4.81	39	6.03	
	No or others	3530	27.99	3207	25.43	323	49.92	
Secondary prevention medication								
DAPT	No	1270	10.07	749	5.94	521	80.53	< 0.001
	Yes	11341	89.93	11215	88.93	126	19.47	

Characteristic		Global (n = 12611)		All-cause Death				P
				No(n = 11964)		Yes(N = 647)		
		N	Per(%)	N	Per(%)	N	Per(%)	
ACEI/ARB	No	6648	52.72	6045	47.93	603	93.20	< 0.001
	Yes	5963	47.28	5919	46.94	44	6.80	
β-blockers	No	5264	41.74	4675	37.07	589	91.04	< 0.001
	Yes	7347	58.26	7289	57.8	58	8.96	
Statins	No	2575	20.42	2015	15.98	560	86.55	< 0.001
	Yes	10036	79.58	9949	78.89	87	13.45	
Hospital characteristics								
Hospital degree	2nd	757	6	725	5.75	32	4.95	< 0.001
	3rd	11854	94	11239	89.12	615	95.05	
Hospital category	General	12052	95.57	11431	90.64	621	95.98	0.245
	Specialist	559	4.43	533	4.23	26	4.02	
Regional distribution	Eastern	8878	70.4	8443	66.95	435	67.23	0.599
	Central	2165	17.17	2059	16.33	106	16.38	
	Western	1568	12.43	1462	11.59	106	16.38	
Chest Pain Center Verification	No	7706	61.11	7311	57.97	395	61.05	0.977
	Verified	4905	38.89	4653	36.9	252	38.95	
Onset season	Spring	5549	44	5278	41.85	271	41.89	0.193
	Summer	1220	9.67	1167	9.25	53	8.19	
	Autumn	3086	24.47	2909	23.07	177	27.36	
	Winter	2756	21.85	2610	20.7	146	22.57	

Variable selection

PSM was applied among 12611 STEMI patients to control the confounding. The performed max standardized mean differences in percent balance improvement was 0.054. After matching gender, age and comorbidities including atrial fibrillation, heart failure, COPD, cardiac valvular surgery, peripheral artery disease, Hypertension, lipid metabolism disorders, Diabetes, chronic kidney disease, cerebrovascular disease, a total of 5657 patients (non-death: death of 5043:633) were included in the XGBoost model construction.

39 variables mentioned at Table 1 were analyzed in the XGboost model. The dataset was divided into 70% and 30% randomly for training set and test set. ACC and AUC were 0.91 and 0.88 in the training set, while 0.89 and 0.70 in the test set. The precision, F1-meansure and recall value of the model is 0.84, 0.85 and 0.89, respectively. Concerning the accuracy and simplicity, the optimal model included 8 variables. Therefore, according to the importance ranking of XGBoost model and SHAP value (Fig. 2), 8 variables were selected to develop the COX regression model and risk score, including Killip class, Early reperfusion strategy, Non-PCI intraoperative anticoagulants, heart rate, gender, age, anterior-wall myocardial infarction(AMWI) and inferior-wall myocardial infarction(IWMI).

Risk scoring system for 1-year mortality on STEMI patients

The development cohort was divided into training set (70%) and test set (30%). The results of Cox regression analysis using training set are presented in Table 2. Finally, 8827 participates were analyzed in Cox regression model as the training set.

Table 2
Cox regression model for 1-year mortality of STEMI patients

Parameter		β	HR	95%CI		P
age		0.05	1.05	1.04	1.06	< 0.001
Gender	Male	Ref				< 0.001
	Female	0.32	1.38	1.13	1.69	
Killip class	1	Ref				< 0.001
	2	0.61	1.85	1.44	2.38	
	3	0.74	2.10	1.45	3.02	
	4	1.95	7.06	5.54	9.00	
Heart rate(per minute)	≤ 100	Ref				< 0.001
	> 100	0.47	1.61	1.27	2.04	
AWMI	No	Ref				0.009
	Yes	0.36	1.44	1.08	1.92	
IWMI	No	Ref				0.486
	Yes	-0.11	0.90	0.67	1.21	
Early reperfusion strategy	Ppci	Ref				< 0.001
	PCI after thrombolysis	0.68	1.97	0.95	4.10	
	Thrombolysis	0.66	1.94	1.30	2.91	
	Others	0.53	1.70	1.39	2.07	
Non-PCI intraoperative anticoagulants	No	Ref				< 0.001
	Yes	-0.81	0.45	0.37	0.54	

On Cox regression analysis, with results reported as hazard ratio (95% CI), elder (HR,1.05 per year; 95%CI,1.04–1.06), female(HR,1.38; 95%CI, 1.13–1.69]), KILLIP class (for level2 vs level1: HR, 1.85, 95%CI; 1.44–2.38; level3 vs level1: HR,2.10; 95%CI,1.45–3.02; level4 vs level1:HR, 7.06; 95%CI, 5.54-9.00), Heart rate over 100 per minute (HR,1.61; 95%CI,1.27–2.04), symptom of AMWI (HR,1.44, 95%CI,1.08–1.92),

early reperfusion strategy (for Thrombolysis vs Ppci: HR,1.93, 95%CI,1.30–2.91; Othrs vs pPCI: HR,1.70; 95%CI,1.39–2.07) and non-PCI intraoperative anticoagulants(HR,0.45; 95%CI,0.37–0.54) were independently associated with 1-year mortality (Table 2).

The 8 risk factors selected by XGBoost selection procedure were used to form mortality risk estimation nomogram (Fig. 3). The risk scoring system was internally validated using the test set. The nomogram demonstrated good accuracy in estimating the 1-year mortality, with a C statistic of 0.87(95%CI, 0.80–0.93) in training set and 0.88 (95%CI, 0.78–0.96) in test set. The E/O ratio was 1.14(95%CI, 0.93–1.39) in trainning set and 1.15(95%CI, 0.85, 1.56) in test set.

Performance of the STEMI risk scoring system

In the training set of development cohort, for 8827 patients with STEMI, discharge, 6-month, and 1-year survival rates were 96.5%, 95.2%, and 94.9%, respectively. In the test set of development cohort, for 3784 patients with STEMI, the corresponding rates for overall survival were 96.2%, 95.0% and 94.7%. By the outcome of 1-year mortality, the AUC(95%CI), sensitivity and specificity of risk scoring system for patients in training set were 0.84(95%CI ,0.82–0.88), 80.3% and 72.5%. The results in the test set were similar, which had an AUC (95%CI) of 0.85(95%CI, 0.82–0.87), sensitivity of 81.6% and specificity of 72.8%. For different follow-up duration of outcomes, fitting the model with the outcomes of in-hospital and 6-month death risk, the performance of the model was similar with 1-year developed model. (Table 3)

Table 3
Performance of the STEMI risk score in STEMI patients

Outcome duration		Mortality	AUC	95%CI		Sensitivity	Specificity
Discharged	Developed model	0.035	0.84	0.82	0.87	80.3%	72.5%
	Internal verification	0.038	0.85	0.82	0.88	81.6%	72.8%
	External verification	0.039	0.85	0.82	0.87	81.6%	72.8%
6 month	Developed model	0.048	0.84	0.82	0.86	81.6%	71.5%
	Internal verification	0.502	0.85	0.83	0.88	77.4%	76.8%
	External verification	0.049	0.84	0.82	0.86	79.5%	74.5%
1 year	Developed model	0.051	0.83	0.81	0.85	80.3%	72.5%
	Internal verification	0.053	0.85	0.82	0.87	81.6%	72.8%
	External verification	0.050	0.84	0.82	0.86	83.0%	71.8%

Validation of the STEMI risk scoring system

A total of 7342 patients with similar distribution with the development cohort were included in the external validation analysis regarded as the validation cohort (Supplement Table 1). Overall, the mortality for 1- year after discharged in the validation cohort is 95.03%. The mean age(± SD) of participants were 61.6(± 12.7), and 77.67% were male. The risk scoring system demonstrated good accuracy in estimating the 1-year mortality, with a c-statistics of 0.89(95%CI, 0.82–0.95) and an E/O ratio of 1.00(95%CI, 0.81–1.23). The performance of this risk score predicting 1-year mortality was satisfactory with accuracy based on AUC of 0.84(95%CI, 0.82–0.86), meanwhile, the sensitivity and specificity were 83.0% and 71.8%.

To further validate the scoring system, we compared the performance of the STEMI risk scoring system with the GRACE score. We developed the GRACE scoring model in the validation cohort. The performance of GRACE score predicting 1-year mortality have a similar performance with the STEMI risk score with an AUC (95%CI) of 0.85(95%CI, 0.85–0.86), meanwhile, the sensitivity and specificity of the GRACE risk score were 83.8% and 69.9 % (Table 4).

Table 4
Performance of the GRACE score for STEMI patients in the validation cohort

Outcome duration	Mortality	AUC	95%CI		Sensitivity	Specificity
Discharged	0.039	0.85	0.85	0.86	76.0%	80.4%
6-month	0.049	0.85	0.83	0.87	73.0%	80.9%
1-year	0.050	0.85	0.83	0.87	83.8%	69.9%

Risk stratification by STEMI risk scoring system

We divided patients into 3 groups for low risk, median risk and high risk according to the mortality risk calculated by the risk scoring system for < 1%(score between 0–43), 1–3%(scored between 44–70) and ≥ 3% (scored over 71), respectively. We explored a Kaplan-Meier survival curve of 1-year all-cause death according to the stratification mentioned above using the validation cohort. The risk stratification was defined to have significant differences (< 0.001) (Fig. 4).

Discussion

This study presents a risk scoring system which was constructed by nomogram of COX regression including 8 variables selected by XGBoost that predicts 1-year mortality risk of STEMI patients from a nationwide cohort in China, including 244 hospitals. The risk scoring system was comprised of 8 variables: level of Killip class, early reperfusion strategy, use of non-PCI intraoperative anticoagulants, heart rate, gender, age, diagnosis of AWMI and IWMI. It was derived from patients hospitalized from 2015 to 2016 and was validated in patients hospitalized in 2017. The performance of this risk score was

satisfactory with discriminatory accuracy based on c statistics in the training set, test set and validation cohort of 0.87, 0.88, and 0.89, respectively.

Previous studies have established several risk scoring systems for estimating the prognosis of STEMI, such as the GRACE score and thrombolysis in myocardial infarction (TIMI) risk score. Those risk scores are derived from studies mainly enrolling patients from Western countries, and both were developed initially only to assess short-term prognoses. Few risk score was developed by Chinese population with STEMI. Although several studies validated the present risk scores in Chinese STEMI patients, those studies were limited by single-center study[11]. Recently, the EPICOR and EPICOR Asia studies used a national wide population to establish a risk-scoring model for 2-year post discharge survival in ACS patients[9, 10]. But the risk score presented a limited clinical value with a c statistics of 0.712 (95% CI, 0.650–0.772) among STEMI patients in China [12]. A risk assessment tool for STEMI patients with a comparative performance of long-term outcomes in Chinese populations still remained to study in detail. Therefore, in our study, we constructed a scoring system based on a national wide population in China to predict 1-year mortality risk for STEMI.

We compared our risk scoring system with the GRACE score among follow-up outcomes at discharged and 6-month, 1-year long follow-up duration (Table 4.). The GRACE score is the most notable risk score for STEMI with highest vitality [13, 14]. It consists of 8 indicators: age, heart rate, SBP, creatinine, level of Killip class and symptoms of cardiac arrest on admission, ECG ST-segment changes and elevated myocardial necrosis markers. Limited study has validated the GRACE score among Chinses patients. Only few single-center study[11] estimated that the accuracy of the GRACE score in predicting all-cause mortality which was ranged 0.766 to 0.789 for c-statistics. In our study, the results showed that the GRACE score have a similar discriminatory accuracy and lower sensitivity than the STEMI risk score. More high-risk STEMI patients would be recognized by the STEMI risk score. It is helpful for health management among Chinese STEMI patients by accessing preventive intervention or more frequently follow-up to decrease death risk at the early stage.

The features selected by XGBoost are mostly coherent with previous reports [15]. Older age, male sex and worse Killip class have been constantly emphasized as major risk factors for adverse outcomes in patients with STEMI. A prior study showed that every additional year of age of the patient with AMI will lead to an increasing risk of 9.3% higher on death[16]. Another study using large population-based cohort observed that female has an higher in-hospital mortality risk than male among patients with STEMI (odds ratio, 1.42; 95% CI, 1.24–1.64)[17]. As reported in previous studies, admission HR are predictors of in-hospital and long-term mortality. Every increase of 5 bpm. in heart rate was associated with a 29% increased risk of cardiovascular mortality at 1-year follow-up [18]. AWMI and IWMI were pointed to have different prognosis. With AWMI, significant reduction in left ventricular function is responsible for the hemodynamic compromise. The myocardial mass supplied by the right coronary artery or left circumflex is usually smaller than that of the left atrial dimension, but with proximal right coronary artery associated IWMI, there may be additional hemodynamic compromise due to right ventricular infarction[19–21]. Since

the myocardium at risk is greater in patients presenting with AWMI than IWMI, the reported incidence of cardiogenic shock is higher in patients with AWMI compared with those with IWMI[21].And in-hospital mortality was modestly lower in patients with cardiac shock complicating IWMI vs. AWMI (30.3% vs. 31.9%; OR, 0.80; 95% CI, 0.75–0.86)[19].

In addition to previous reports, we take an added concern of treatment indicators which have more value in the risk stratification for long-term prognosis prediction, especially in China[22, 23]. Poverty, limited healthcare infrastructure for PCI, and poor accessibility to acute emergency medical services are most important system-level limitations for STEMI care in middle-income countries [24, 25]. In 2013, only 36% of AMI patients could receive reperfusion therapy and visited at hospitals equipped with catheterization labs within 12 hours after onset in China. The proportion would be lower among primary hospitals for lacking medical resources [26]. It has been proved that PCI can decrease the 1-year mortality in patients with STEMI [27]. When PCI could not access within 90 minutes, STEMI patients are suggested to have anticoagulants [28]. Under this condition, the factors of early reperfusion strategy and non-PCI intraoperative anticoagulants have import predictive value in China.

Our risk scoring system based on simple patient demographics and subjective symptoms using machine learning and nomogram can be useful for the health management of STEMI patients after discharged. First, our risk scoring system is easy-to-use. It uses information that can be easily provided, such as age, sex, and pre-hospitalize diagnosis. Previous scoring systems [4, 11, 15], including a recently reported deep learning model [29], require laboratory or radiographic findings as the main variables. Although such models can be helpful in fully equipped medical facilities, they initially consume a certain amount of medical resources and time. Besides, we used nomogram to present the risk scoring system which can calculate the risk directly using graphs, so that the results are more readable and have higher practical value than conventional models. Second, because the variables were selected by a data-driven method, it can recognize predictive factors of the risk scoring system with sufficient accuracy and discrimination. XGBoost is a newly integrated machine learning algorithm based on decision tree. It is more effective than previous methods by using clever penalization of trees and implementation on single, distributed systems and out-of-core computation. With regularization to prevent overfitting and extra parameter defining loss function, XGBoost is expected to be useful in the field of real-world data based study. Previous study[30] demonstrated that XGBoost algorithm showed an improvement in the prediction ability compared with other machine classifiers and conventional risk scores. As our study was drawn from a real-world observation cohort, XGBoost is optimal for our study to identify the independent risk factors of 1-year mortality. The variables included in the risk scoring system were selected by traversal procedure of XGBoost model, according to ACC. This selection method was a data-driven method and can select variables from a large scale of potential factors and can powerfully reduce variance and reduce bias caused by subjective experiences. Then the nomogram was developed based on the selected variables for practical value in clinical application. However, small number of outcomes were observed in the analysis which may cause over matched during the PSM procedure. Extending follow-up time of the

cohort or expanding the population scale may improve the results by increasing the observed number of outcomes. We may further validate the risk scoring system in larger scale of population in the future.

Conclusion

We developed a risk scoring system for predicting 1-year mortality risk of STEMI in a large Chinese population. The new score is easy-to-use and demonstrating a good discriminatory accuracy in predicting both short-term and long-term mortality risk in Chinese patients with STEMI. And it has a better sensitivity than the GRACE score.

Abbreviate

Anterior-wall Myocardial Infarction, AWMI

Acute coronary syndromes, ACS

Confidence interval, CI

China STEMI Care Project Phase 2, CSCAP-2

The Global Registry of Acute Coronary Events, GRACE

Hazard Ratio, HR

long-tErm follow uP of antithrombotic management patterns In acute CORonary syndrome patients, EPICOR

Inferior-wall Myocardial Infarction, IWMI

Receiver Operating Characteristic curve, ROC curve

Primary Percutaneous Coronary, PCI

ST-segment Elevation Myocardial Infarction, STEMI

Declarations

Ethics approval and consent to participate

This study was approved by the ethics committee of Peking University First Hospital, and each participant provided written informed consent. We adhered to the principles of the Declaration of Helsinki. The procedures followed were in accordance with institutional guidelines.

Available of data and materials

The data underlying this article cannot be shared publicly for the privacy of individuals that participated in the study. The data will be shared on reasonable request to the corresponding author.

Competing of interests

The authors report no conflict of interest.

Funding

The study supported by a grant from the National Natural Science Foundation of China (81172768 81872692) and Beijing Municipal Natural Science Foundation (7182085). We appreciate the staffs in CSCAP-2 and research coordinators for the assistance to data acquisition.

Authors' contribution

Chen Si contributed to data analysis and drafting the article. Qianzi Che helped to complete the analysis and supervised for methology. Jia jia and Yan Zhang contributed to data collection.Yiqun Wu helped to contributions to conception and design. Yong Huo contribute to the acquisition of data. Dafang Chen was the supervisor of the article and takes charge of the fnal approval of the version to be published.

Acknowledgements

We appreciate the staffs in CSCAP-2 and research coordinators for the assistance to data acquisition. All authors already know the publication of the study, and none of them have any potential conflicts of interest.

Conflict of interest

The authors report no conflict of interest.

Data Availability Statement

The data underlying this article cannot be shared publicly for the privacy of individuals that participated in the study. The data will be shared on reasonable request to the corresponding author.

Support

The study supported by a grant from the National Natural Science Foundation of China (81172768 81872692) and Beijing Municipal Natural Science Foundation (7182085). We appreciate the staffs in

CSCAP-2 and research coordinators for the assistance to data acquisition.

References

1. Huo Y, Thompson P, Buddhari W, et al. Challenges and solutions in medically managed ACS in the Asia-Pacific region: expert recommendations from the Asia-Pacific ACS Medical Management Working Group[J]. Int J Cardiol. 2015, 183: 63-75.
2. Ohira T, Iso H. Cardiovascular disease epidemiology in Asia: an overview[J]. Circ J. 2013, 77(7): 1646-1652.
3. Hu Shengshou, Gao Runlin, Liu Lisheng, et al. Summary of "Chinese Cardiovascular Disease Report 2018"[J]. Chinese Journal of Circulation. 2019, 34(03): 209-220.
4. Huo Y, Lee S W, Sawhney J, et al. Two-year outcomes post-discharge in Asian patients with acute coronary syndrome: Findings from the EPICOR Asia study[J]. Int J Cardiol. 2020, 315: 1-8.
5. Chan M Y, Du X, Eccleston D, et al. Acute coronary syndrome in the Asia-Pacific region[J]. Int J Cardiol. 2016, 202: 861-869.
6. Cheng C I, Chen C P, Kuan P L, et al. The causes and outcomes of inadequate implementation of existing guidelines for antiplatelet treatment in patients with acute coronary syndrome: the experience from Taiwan Acute Coronary Syndrome Descriptive Registry (T-ACCORD Registry)[J]. Clin Cardiol. 2010, 33(6): E40-E48.
7. Granger C B, Goldberg R J, Dabbous O, et al. Predictors of hospital mortality in the global registry of acute coronary events[J]. Arch Intern Med. 2003, 163(19): 2345-2353.
8. Chu C Y, Lin T H, Lai W T. The Management and Prognostic Factors of Acute Coronary Syndrome: Evidence from the Taiwan Acute Coronary Syndrome Full Spectrum Registry[J]. Acta Cardiol Sin. 2017, 33(4): 329-338.
9. Pocock S J, Huo Y, Van de Werf F, et al. Predicting two-year mortality from discharge after acute coronary syndrome: An internationally-based risk score[J]. Eur Heart J Acute Cardiovasc Care. 2019, 8(8): 727-737.
10. Huo Y, Lee S W, Sawhney J P, et al. Rationale, Design, and Baseline Characteristics of the EPICOR Asia Study (Long-tErm follow-uP of antithrombotic management patterns In Acute CORonary Syndrome patients in Asia)[J]. Clin Cardiol. 2015, 38(9): 511-519.
11. Chen Y H, Huang S S, Lin S J. TIMI and GRACE Risk Scores Predict Both Short-Term and Long-Term Outcomes in Chinese Patients with Acute Myocardial Infarction[J]. Acta Cardiol Sin. 2018, 34(1): 4-12.
12. Liu Y, Yao Y, Tang X F, et al. Evaluation of a novel score for predicting 2-year outcomes in patients with acute coronary syndrome after percutaneous coronary intervention[J]. J Chin Med Assoc. 2019, 82(8): 616-622.
13. Komiyama K, Nakamura M, Tanabe K, et al. In-hospital mortality analysis of Japanese patients with acute coronary syndrome using the Tokyo CCU Network database: Applicability of the GRACE risk

score[J]. J Cardiol. 2018, 71(3): 251-258.

14. Abu-Assi E, Ferreira-Gonzalez I, Ribera A, et al. "Do GRACE (Global Registry of Acute Coronary events) risk scores still maintain their performance for predicting mortality in the era of contemporary management of acute coronary syndromes?"[J]. Am Heart J. 2010, 160(5): 826-834.
15. Fox K A, Dabbous O H, Goldberg R J, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: prospective multinational observational study (GRACE)[J]. BMJ. 2006, 333(7578): 1091.
16. Zhang Chunyan. Logistic regression analysis of influencing factors of mortality in hospitalized patients with acute myocardial infarction[J]. Chinese Medical Record. 2015, 16(10): 54-55.
17. Ezekowitz J A, Savu A, Welsh R C, et al. Is There a Sex Gap in Surviving an Acute Coronary Syndrome or Subsequent Development of Heart Failure?[J]. Circulation. 2020, 142(23): 2231-2239.
18. Antoni M L, Boden H, Delgado V, et al. Relationship between discharge heart rate and mortality in patients after acute myocardial infarction treated with primary percutaneous coronary intervention[J]. Eur Heart J. 2012, 33(1): 96-102.
19. Gupta T, Weinreich M, Kolte D, et al. Comparison of Incidence and Outcomes of Cardiogenic Shock Complicating Posterior (Inferior) Versus Anterior ST-Elevation Myocardial Infarction[J]. Am J Cardiol. 2020, 125(7): 1013-1019.
20. Tewelde S Z, Liu S S, Winters M E. Cardiogenic Shock[J]. Cardiol Clin. 2018, 36(1): 53-61.
21. Niccoli G, Scalone G, Lerman A, et al. Coronary microvascular obstruction in acute myocardial infarction[J]. Eur Heart J. 2016, 37(13): 1024-1033.
22. Scholz K H, Maier S, Maier L S, et al. Impact of treatment delay on mortality in ST-segment elevation myocardial infarction (STEMI) patients presenting with and without haemodynamic instability: results from the German prospective, multicentre FITT-STEMI trial[J]. Eur Heart J. 2018, 39(13): 1065-1074.
23. Koul S, Andell P, Martinsson A, et al. Delay from first medical contact to primary PCI and all-cause mortality: a nationwide study of patients with ST-elevation myocardial infarction[J]. J Am Heart Assoc. 2014, 3(2): e486.
24. Baliga R R, Bahl V K, Alexander T, et al. Management of STEMI in low- and middle-income countries[J]. Glob Heart. 2014, 9(4): 469-510.
25. Nallamothu B K. A race for the base: ST-segment-elevation myocardial infarction systems of care in low- and middle-income countries[J]. Circ Cardiovasc Qual Outcomes. 2013, 6(1): 5-6.
26. Centers H O C C. The best reperfusion treatment strategy for acute myocardial infarction in primary hospitals– on-site thrombolysis or referral to PCI[Z].
27. Alexander T, Mullasari A S, Joseph G, et al. A System of Care for Patients With ST-Segment Elevation Myocardial Infarction in India: The Tamil Nadu-ST-Segment Elevation Myocardial Infarction Program[J]. JAMA Cardiol. 2017, 2(5): 498-505.

28. Jung C, Elsasser A. [Update ESC Guideline 2017 - Acute Myocardial Infarction (STEMI)][J]. Dtsch Med Wochenschr. 2018, 143(11): 797-801.
29. Nakanishi R, Slomka P J, Rios R, et al. Machine Learning Adds to Clinical and CAC Assessments in Predicting 10-Year CHD and CVD Deaths[J]. JACC Cardiovasc Imaging. 2020.
30. Li Y M, Jiang L C, He J J, et al. Machine Learning to Predict the 1-Year Mortality Rate After Acute Anterior Myocardial Infarction in Chinese Patients[J]. Ther Clin Risk Manag. 2020, 16: 1-6.

Figures

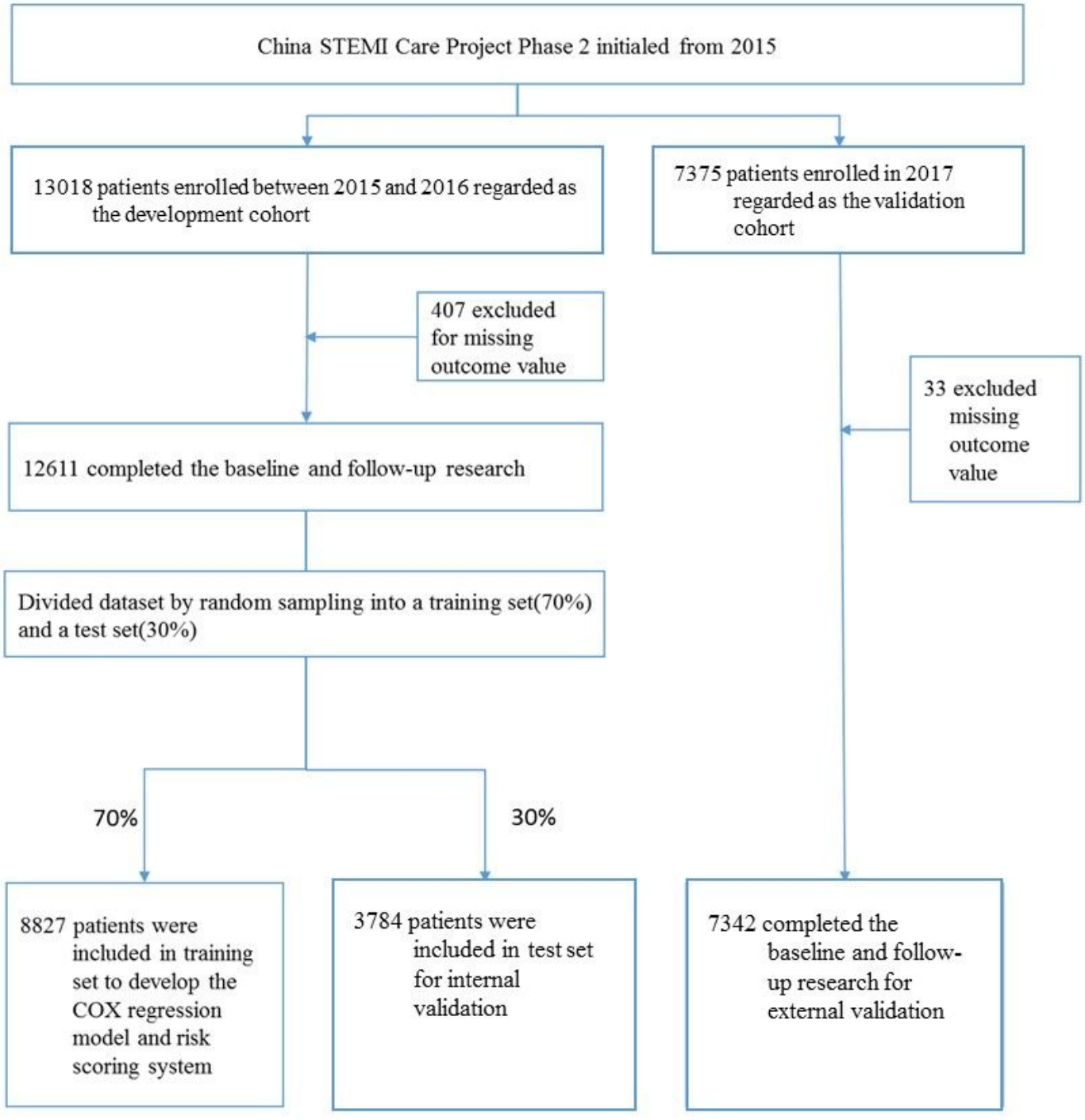

Figure 1

Flow chart for research analysis

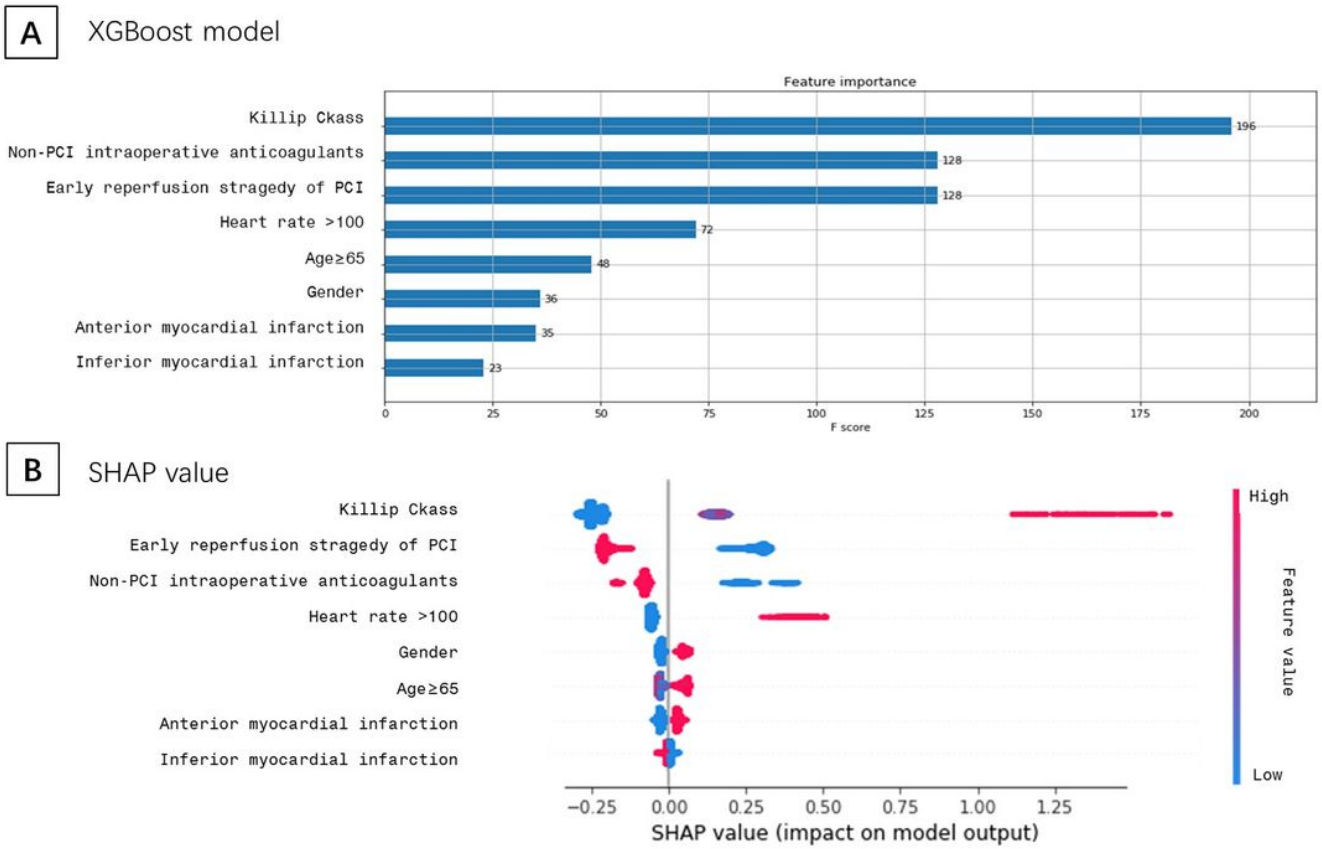

Figure 2

Variable importance ranking according to XGBoost and SHAP value

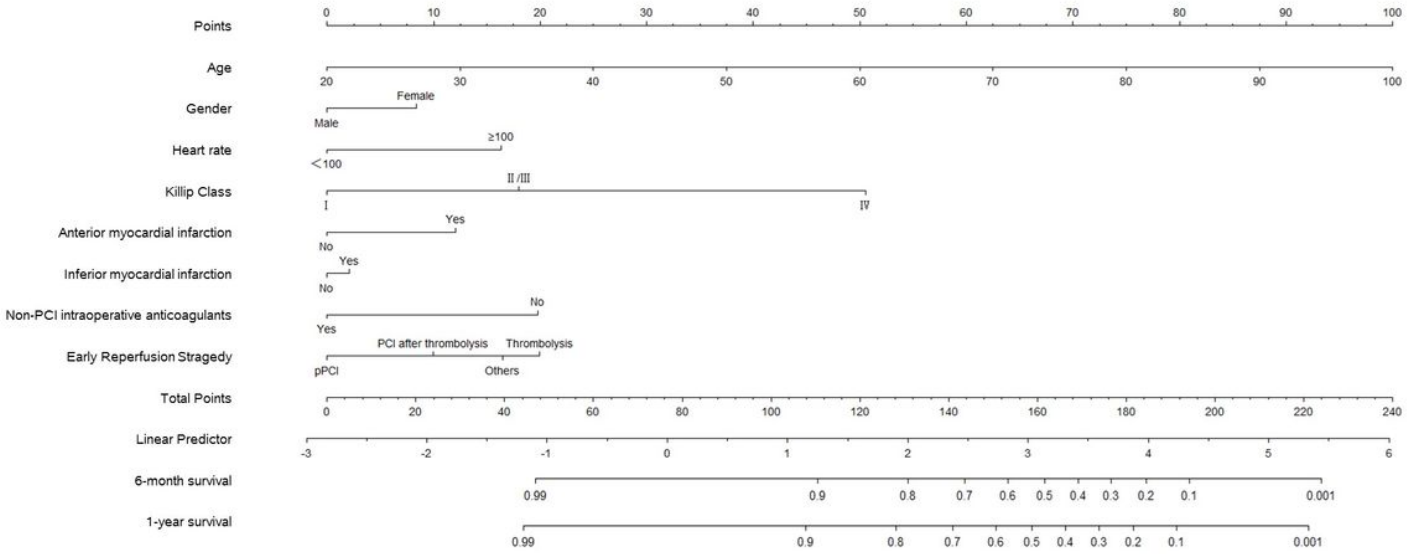

Figure 3

Nomogram for 1-year mortality of STEMI patients

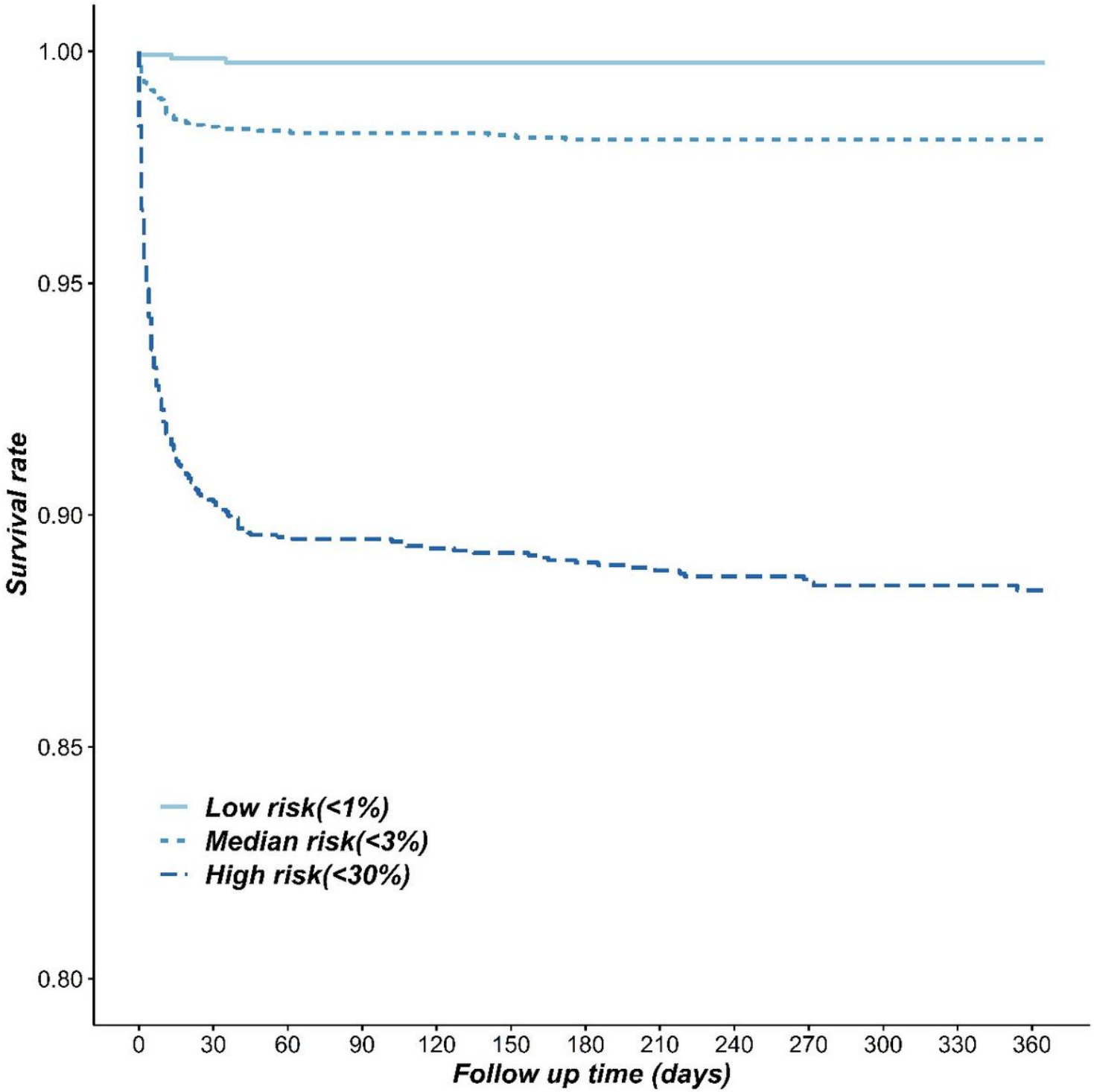

Figure 4

The Kaplan-Meier survival curve of 1-year all-cause death in STEMI patients within different risk stratification

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement1118.docx