Supplementary materials

eTable 1. TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Paragraph
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	Title
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	Abstract
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	Intro Para 1-2
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	Intro Para 3
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	Meth Para 3-4
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	Meth Para 3-4
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centers.	Meth Para 4
	5b	D;V	Describe eligibility criteria for participants.	Meth Para 5 App. A
	5c	D;V	Give details of treatments received, if relevant.	NA
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	Meth Para 8
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Meth Para 7 App. B
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V	Explain how the study size was arrived at.	NA
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Meth Para 7 App. C
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Meth Para 7,9,10
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Meth Para 10,11 App. D
	10c	V	For validation, describe how the predictions were calculated.	Meth Para 7
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Meth Para 11

	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	App. E
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	Meth Para 12
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Meth Para 4
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Res Para 1; App. Fig 1
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Res Para 1; Tab 1; App. C
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Res Para 1; Tab 1
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Res Para 2, 4; Tab 2; App. Fig 2
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Fig 1
	15b	D	Explain how to the use the prediction model.	Res Para 6; Tab 3; Fig 3; App. Tab 3
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Res Para 3,5; Tab 2; Fig 2; App. Tab 2; App. Fig 3-5
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	App. E
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as non-representative sample, few events per predictor, missing data).	Disc Para 6
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Disc Para 2-4
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	Disc Para 1
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	Disc Para 5
Other information				

Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	NA
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	Acknowledgement

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

Appendix A: Sampling method of China PEACE Retrospective AMI Study

The study identified cases for study inclusion using a stratified 2-stage cluster sampling design. In the first stage, representative hospitals were identified using a simple random sampling procedure within each of the 5 study strata: Eastern-rural, Central-rural, Western-rural, Eastern-urban, and Central/Western-urban regions, to reflect current practices. In the second stage, cases were drawn based on the local hospital database for patients with AMI at each sampled hospital using systematic random sampling procedures. In each of the 5 study strata, the study determined the sample size required to achieve a 2% precision for describing the primary outcome (in-hospital mortality). In order to achieve the precision with an α of 0.05, 1150 medical records needed to sampled among hospitals with an average cluster size of 40 in rural areas, and 1750 medical records among hospitals with an average cluster size of 60 in urban areas. These cluster sizes in rural and urban settings appeared reasonable based upon our previous survey of treatment for acute coronary syndromes at more than 1000 hospitals in 2010. The study doubled cluster sizes for 2011 to improve precision in the description of hospital-level treatment patterns and outcomes.

Appendix B: Definition of candidate predictors

Candidate predictors	Definition
Demographics	
Age	Age when enrolled
Sex	Physical sex
Medical history	
Hypertension	Indicate if the patient has been diagnosed previously with hypertension by any one of the following: 1. History of hypertension diagnosed and treated with medication, diet and/or exercise 2. Prior documentation of blood pressure greater than 140 mm Hg systolic and/or 90 mm Hg diastolic for patients without diabetes or chronic kidney disease, or prior documentation of blood pressure greater than 130 mm Hg systolic or 80 mm Hg diastolic on at least two occasions for patients with diabetes or chronic kidney disease. 3. Currently on pharmacological therapy for the treatment of hypertension.
Diabetes Mellitus	Indicate if the patient has a history of diabetes mellitus, regardless of duration of disease or need for antidiabetic agents. Diabetes mellitus is diagnosed by a physician or can be defined as a fasting blood sugar greater than 7 mmol/l or 126 mg/dL. It does not include gestational diabetes. Diabetes mellitus can also be identified by history of pharmacologic treatment for condition.
Myocardial Infarction	Indicate if the patient has had at least one documented previous myocardial infarction
PCI	Indicate whether the patient had a percutaneous coronary intervention (PCI)
Ischemic Stroke	Indicate if the patient has had an ischemic stroke. A prior stroke is defined as any confirmed neurological deficit of abrupt onset caused by a disturbance in cerebral blood supply that did not resolve within 24 hours.
Chronic kidney disease	Indicate if the patient has a history of chronic renal failure. It can be coded for any of the following: 1. A documented history of renal failure, and/or 2. A history of creatinine > 2.0 mg/dL, and/or 3. A documented history of chronic renal disease Prior renal transplant patients are not included unless creatinine has been >2.0 mg/dL since transplantation
In-hospital diagnoses or test at admission	
STEMI	Indicate if the patient was diagnosis as STEMI after this symptom onset
AMI Position	Indicate the site of myocardial injury
The Killip class	Indicate the physician documentation of Killip classification at the patient's arrival to the facility I - No rales over the lung fields and no S3. II - Rales 50% or less over the lung fields or presence of an S3. Includes patients documented as having bibasilar rales. III - Rales more than 50% of the lung fields/frank pulmonary edema. Includes patients documented as having rales throughout. IV - Cardiogenic Shock
Pneumonia	Indicate if there is physician documentation or report of pneumonia on presentation to this facility. Pneumonia is defined as physician

	documentation of pneumonia plus the presence of a demonstrable infiltrate by chest radiograph or other imaging.
LVEF	Indicate whether the EF was documented in the echocardiogram report.
Heart rate	Indicate the first measurement or earliest record of pulse rate (in beats per minute).
GLU	Indicate the first measurement or earliest record of blood glucose (mmol/L).
SBP	Indicate the first measurement or earliest record of systolic blood pressure (mm Hg).
WBC	Indicate the first measurement or earliest record of white blood cell count ($\times 10^9$/L).
Serum creatinine	Indicate the first measurement or earliest record of serum creatinine (μmol/L).
Troponin	Indicate the first measurement or earliest record of troponin.

* PCI: Percutaneous Coronary Intervention, HF: Heart Failure, STEMI: ST Elevation Myocardial Infarction, LVEF: Left Ventricular Ejection Fraction, GLU: Blood glucose, SBP: Systolic Blood Pressure, WBC: White blood cell count

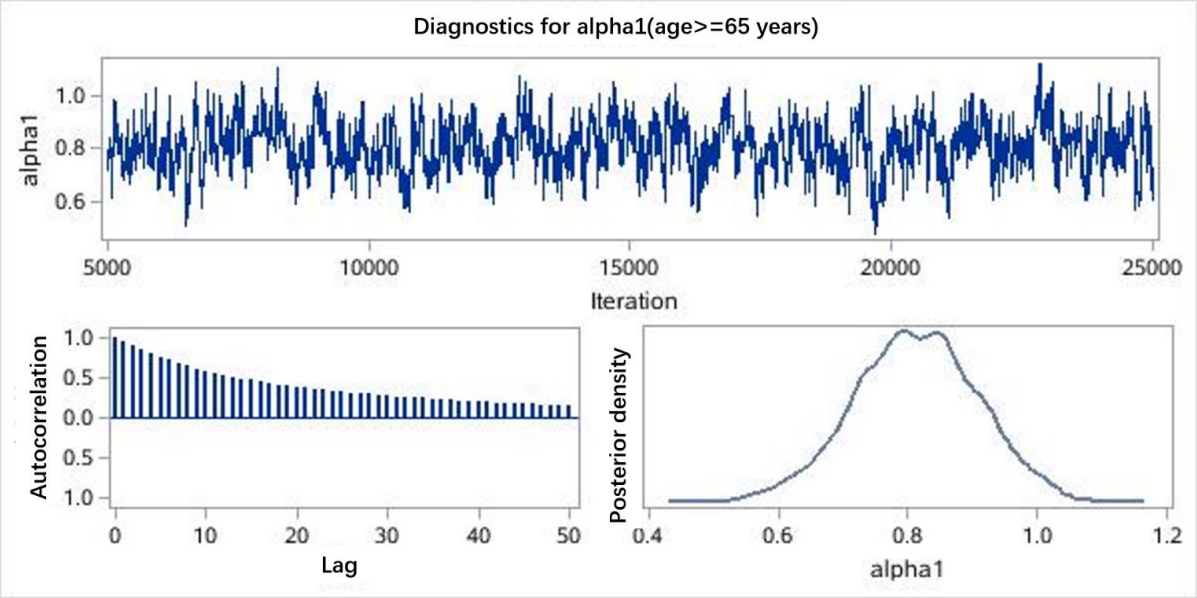
Appendix C: Multiple imputation for predictors

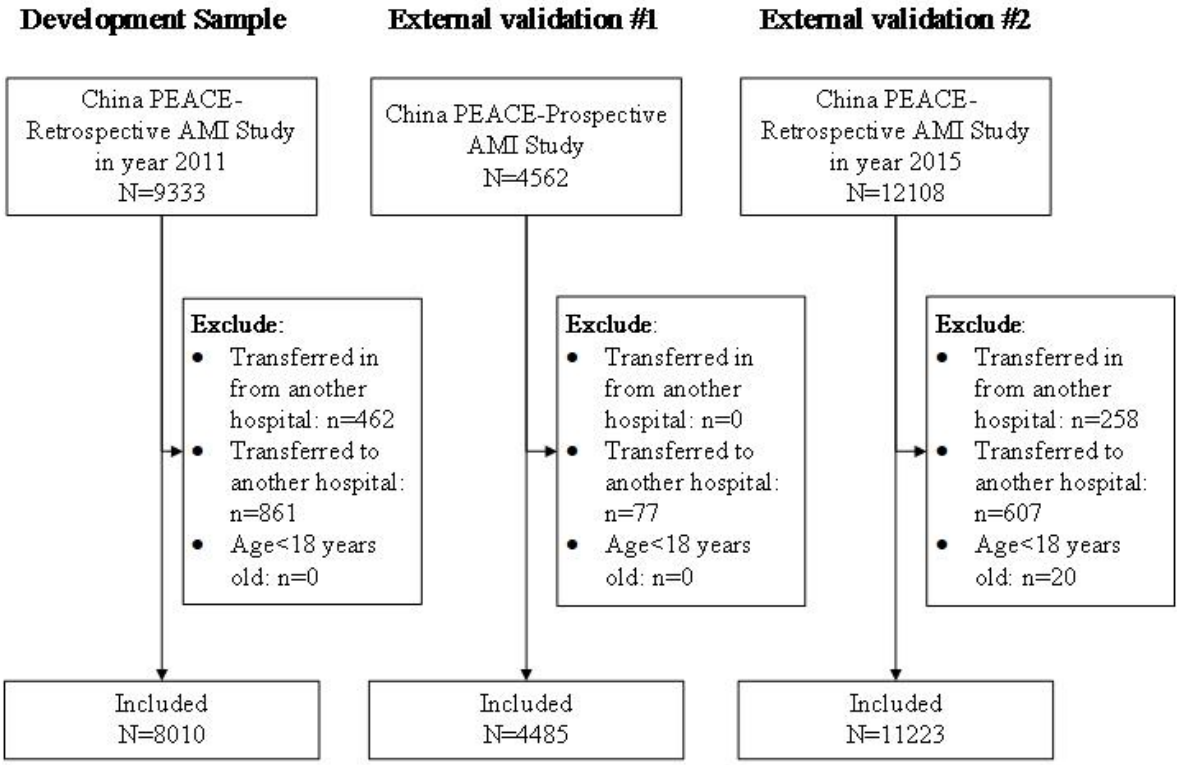
For predictors, including age, SBP, heart rate, white blood cell count, blood glucose and serum creatinine, we applied a multiple imputation method base on Markov Chain Monte Carlo (MCMC) by PROC MI procedure in SAS to impute the missing value. In detail, we built separate multiple variable model to estimate the value of missing cases for each variable, with the patients' characteristics as independent variables (including sex, history of hypertension, diabetes, CHD, MI, PCI, stroke, whether STEMI, Killip class, subtypes of AMI, whether pneumonia at admission, LEVF and chronic kidney disease). We repeated the estimated 10 times for each variable, and the average of them was used as the imputation results. The missing rates of each variable, and the distribution of original values and values after imputation were shown in the below table.

Variables	Missing (N, %)	Original			After imputation		
		Mean ±SD	Median (Q1, Q3)	Min, max	Mean ±SD	Median (Q1, Q3)	Min, max
Age	6,0.07%	65.63±12.72	67(57,76)	20,94	65.64±12.72	67(57,76)	20,94
SBP	59,0.74%	130.35±27.27	130(110,149)	10,267	130.23±27.23	130(110,149)	10,267
Heart rate	27,0.34%	78.59±20.86	78(66,90)	5,223	78.61±20.84	78(66,90)	5,223
White blood cell count	511,6.38%	9.93±4.01	9.16(7.1,11.9)	1,65	9.92±3.91	9.24(7.26,11.79)	1,65
Blood glucose	695,8.68%	8.18±4.08	6.97(5.63,9.2)	0.4,34.75	8.18±3.96	7(5.71,9.2)	0.4,34.75
Serum creatinine	569,7.1%	92.83±61.88	80(65,100)	2,1415	91.6±60.11	79(65,99)	2,1415

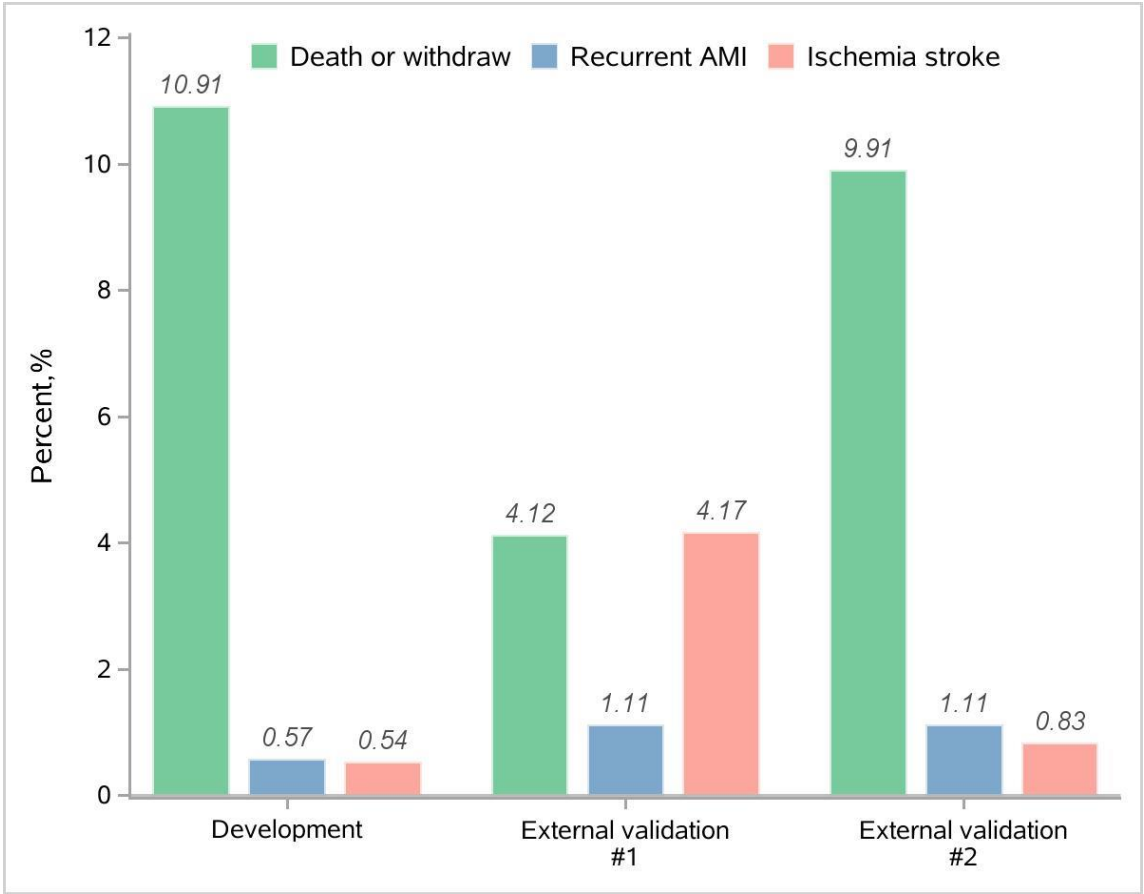
Appendix D: Markov Chain Monte Carlo (MCMC) simulation for selecting predictors

We fitted the model with Markov Chain Monte Carlo (MCMC) simulation method to calculate a posterior probability for each selected prediction (Petros Dellaportas, Jonathan J. Forster, Ioannis Ntzoufras. On Bayesian model and variable selection using MCMC. Statistics and Computing 12:27–36, 2002). In detail, by the PROC MCMC procedure of SAS software, we built a model with the outcome as dependent variable and all potential predictors as independent variables. By this method, we could get the posterior probability density function of specific predictor and the outcome, which measures the strength of association between them. For example, the bellowing figure shows the results of the predictor "age greater than 65 years old". Based on that, we could further calculate the probability of positive association or negative association among all iteration, which ranges between 0 to 100%. In order to select stable factors, only potential predictors with posterior probability of 100% for positive association would be included in the final predictors list.

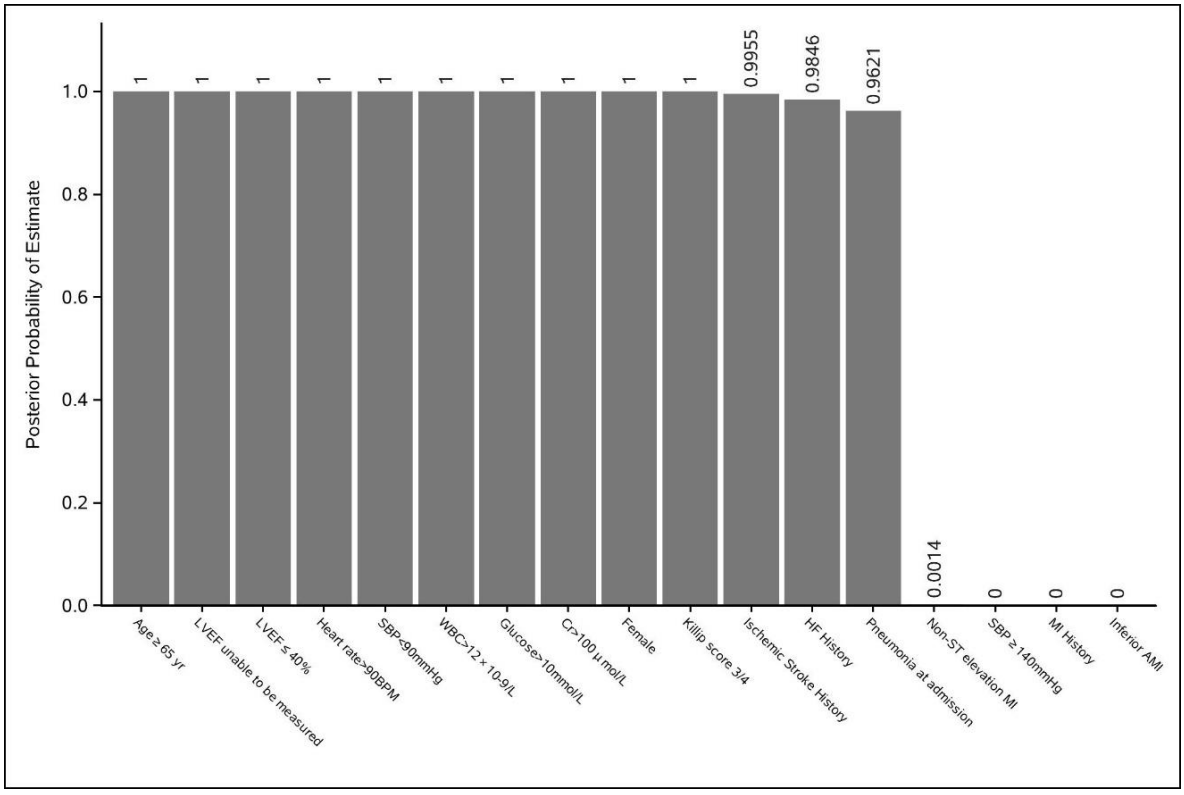




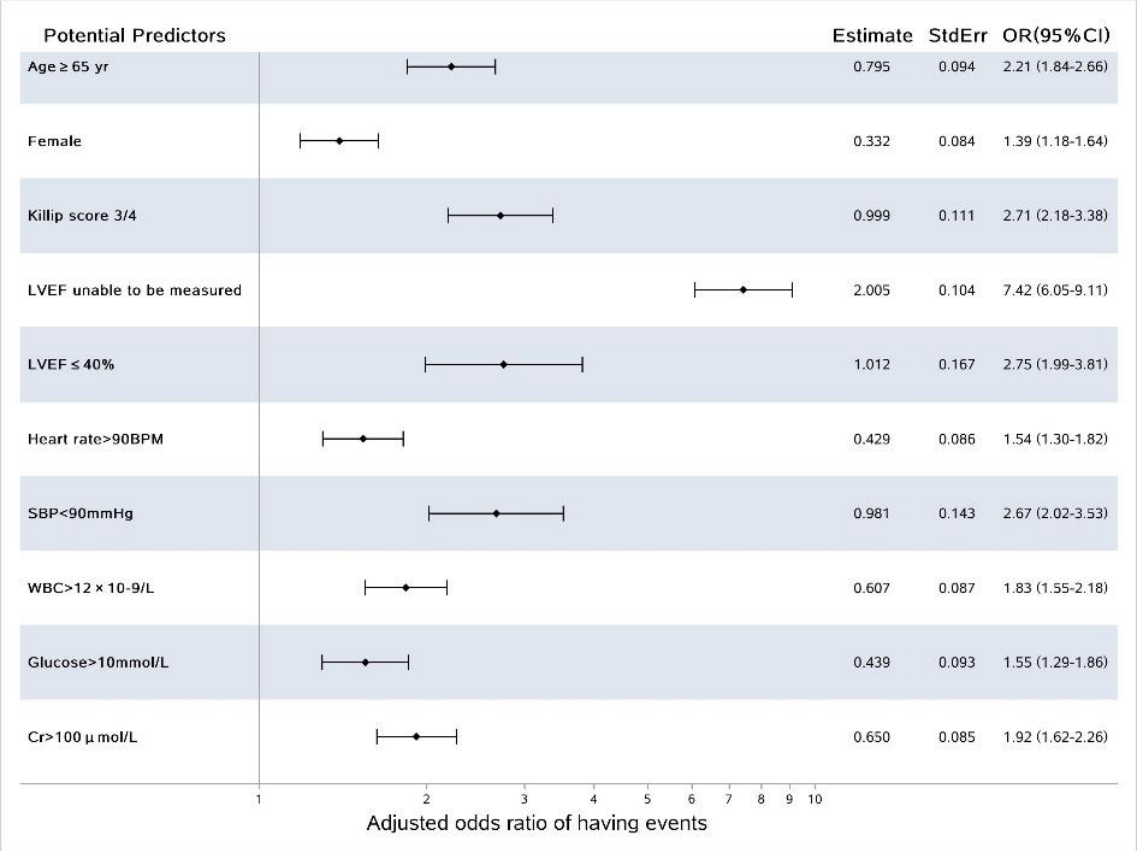
eFigure 1. Flowchart of study participants' selection



eFigure 2. Rates of all-cause death, recurrent AMI and ischemic stroke in hospital among development and external validation samples

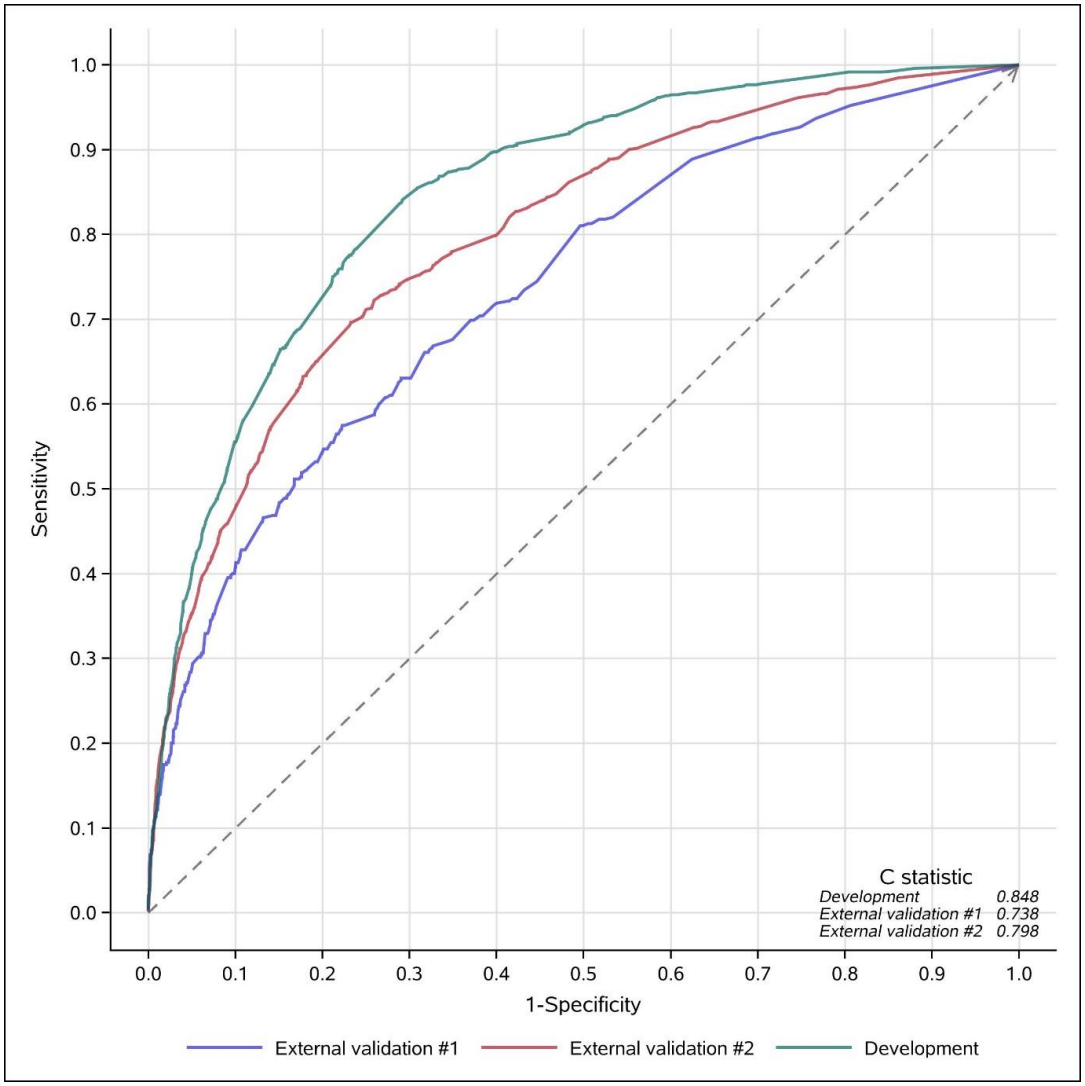


eFigure 3. Posterior probability of potential predictors with MCMC simulation



*StdErr: standard error

eFigure 4. Estimate coefficients and odds ratios of predictors of the prediction model base on development sample without missing data imputation



eFigure 5. ROC curve of prediction model among development and external validation samples

eTable 2. Performance of prediction model among subgroups in the development sample

Subgroup of development sample	Number of patients	Number of events	Events rate (95% CI), %	AUC
Time from symptom to admission				
Symptom to admission>24h	3061	329	10.75(9.67-11.90)	0.837
Symptom to admission<=24h	4949	606	12.24(11.34-13.19)	0.854
AMI type				
Non-STEMI	1429	149	10.43(8.89-12.13)	0.823
STEMI	6581	786	11.94(11.17-12.75)	0.853
Primary PCI				
No	7336	918	12.51(11.76-13.29)	0.843
Yes	674	17	2.52(1.48-4.01)	0.831
Gender				
Male	5425	522	9.62(8.85-10.44)	0.859
Female	2585	413	15.98(14.58-17.45)	0.816
Type of hospital				
Tertiary hospital	5294	563	10.63(9.82-11.50)	0.861
Secondary hospital	2716	372	13.70(12.42-15.05)	0.817

eTable 3. Correspondence of the obtained score to the estimated risk

Risk score	Num. of patients	Mean of predicted event rate,%	Median of predicted event rate,%	Risk score	Num. of patients	Mean of predicted event rate,%	Median of predicted event rate,%
0	854	0.75	0.75	44	33	22.52	22.28
4	177	1.03	1.03	45	95	23.56	23.46
5	94	1.11	1.11	46	106	25.41	25.53
7	129	1.34	1.34	47	121	26.74	26.5
8	110	1.49	1.5	48	37	28.41	27.96
9	245	1.56	1.57	49	45	30.74	31
10	510	1.68	1.68	50	76	31.7	31.9
11	34	1.83	1.83	51	83	34.04	33.75
12	99	2.01	2	52	37	35.36	34.67
13	48	2.16	2.13	53	44	38.1	38.06
14	416	2.29	2.29	54	61	39.84	39.88
15	111	2.5	2.47	55	52	41.5	41.76
16	58	2.75	2.78	56	41	43.88	43.75
17	117	3.01	2.98	57	11	45.89	45.8
18	182	3.31	3.32	58	48	48.1	48.43
19	141	3.42	3.47	59	36	49.61	49.29
20	36	3.86	3.87	60	39	51.88	51.55
21	113	4.05	4.03	61	14	54.9	55.16
22	148	4.43	4.41	62	27	56.31	56.35
23	89	4.75	4.69	63	36	58.41	58.25
24	464	5.25	5.25	64	7	59.82	59.24
25	60	5.71	5.79	65	16	62.59	62.72
26	130	5.95	5.93	66	15	64.25	64.5
27	75	6.57	6.48	67	22	65.74	65.59
28	153	7.03	7.05	68	21	67.95	68.06
29	117	7.69	7.57	69	5	70.14	69.83
30	93	8.36	8.46	70	14	71.7	72.11
31	102	8.88	8.98	71	21	72.89	72.26
32	88	9.94	10	72	9	74.55	74.45
33	168	10.37	10.44	73	4	76.74	76.77
34	476	11.13	11.12	74	15	77.72	77.5
35	76	11.88	11.96	75	21	79.16	79.24
36	51	12.93	12.79	77	1	81.88	81.88
37	65	14.01	14.1	78	10	83.3	83.46
38	351	14.63	14.61	79	8	83.93	83.93
39	115	15.81	15.59	80	3	85.05	85.05
40	41	16.98	17.2	82	4	87.35	87.35
41	87	18.53	18.31	83	6	88.17	88.17
42	184	20	20.05	87	3	91.07	91.07
43	156	20.45	20.16				

Appendix E: Comparison between our risk score and GRACE, TIMI scores

We compared the AUC and the Net Reclassification Index (NRI) between our score and GRACE score for in-hospital mortality/MI and for ischemic stroke, TIMI score among patients with STEMI and among patients with NSTEMI. In detail, we calculated four GRACE or TIMI risk score for different outcome:

(1) GRACE risk score for in-hospital death or MI: based on 'Nomogram for translating eight Granger risk factors into an in-hospital death/MI risk score (as used in Palm Pilot software available on GRACE website)', which uses age, pulse, SBP, creatinine, killio class, cardiac arrest on presentation, positive initial cardiac enzymes and whether STEMI (*Fred Anderson, Gordon FitzGerald. Methods and formulas used to calculate the GRACE Risk Scores for patients presenting to hospital with an acute coronary syndrome: Coordinating Center for the Global Registry of Acute Coronary Events, Center for Outcomes Research, University of Massachusetts Medical School*).

(2) GRACE risk score for Ischemic Stroke: using bellowing variables: age, weight, AF/atrial flutter, positive biomarkers, SBP>=160mmHg, Killip class II-IV, ST change and no history of smoking (*Park KL, Budaj A, Goldberg RJ, et al. Risk-prediction model for ischemic stroke in patients hospitalized with an acute coronary syndrome (from the global registry of acute coronary events [GRACE]). The American journal of cardiology, 2012*).

(3) TIMI risk score for STEMI: predicting the risk of 30-day mortality at presentation, using age, history of DM/HTN or angina,SBP<100mmHg, heart rate>100 BPM, Killip class II-IV, weight<67kg, anterior STE or LBBB at presentation, and time to treatment>4hours (*Morrow DA, Antman EM, Charlesworth A, et al. TIMI risk score for ST-elevation myocardial infarction: A convenient, bedside, clinical score for risk assessment at presentation: An intravenous nPA for treatment of infarcting myocardium early II trial substudy. Circulation, 2000*).

(4) TIMI risk score for UA/NSTEMI: predicting the risk of developing at least 1 component of all-cause mortality, new or recurrent MI, severe recurrent ischemia requiring urgent revascularization through 14 days, using age, risk factors for CAD (history of CAD, hypertension, hypercholesterolemia, diabetes, current smoker), significant coronary stenosis, ST deviation, severe angina symptoms, use of aspirin in last 7 days, and elevated serum cardiac markers (*Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: A method for prognostication and therapeutic decision making. JAMA, 2000*).

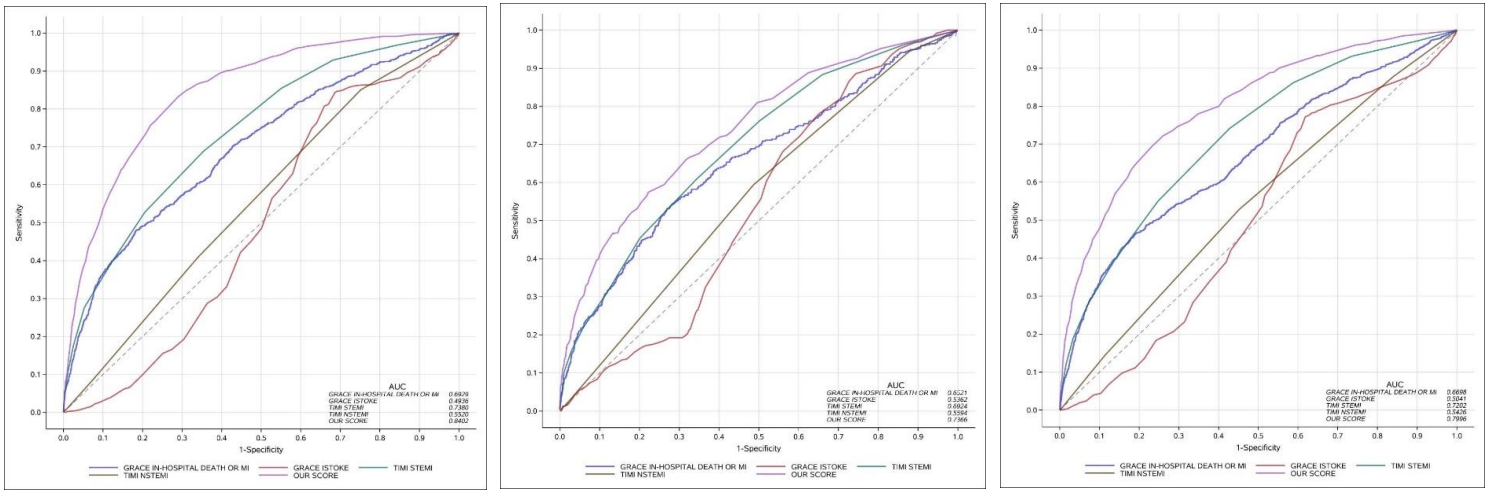
After getting the risk scores, we built logistic models among development and validation cohorts, with in-hospital MACE as dependent variable and each risk score as independent variable. By this method, we got ROC curve and AUC for each risk score. We also compared the statistical difference of the AUC between our risk score and other risk score. To calculate the NRI between our risk score and GRACE, TIMI risk score, we

divided samples into three equal parts based on each score. In this way, each score can be grouped by a relatively uniform standard.

Result showed significant improvement of reclassification ability of our score among all development and validation samples (bellowing table and figure). Considering the difference of outcome definition among these scores, to a certain extent, the positive result still indicated the significance of our risk score in predicting the in-hospital comprehensive outcomes.

AUC Comparison and NRI between Risk Score with GRACE and TIMI Scores

SAMPLE GROUP	SCORE	AUC	$AUC_{OUR\ SORE} - AUC_{OTHER\ SCORE}$		NRI	
			DIFFERENCE(95% CI)	P VALUE	%	P VALUE
Development	Our Score	0.84				
	GRACE In-Hospital Death or MI	0.693	0.147(0.128,0.167)	<0.0001	30.337	<0.0001
	GRACE Ischemic Stroke	0.494	0.347(0.325,0.368)	<0.0001	67.495	<0.0001
	TIMI STEMI	0.738	0.102(0.087,0.117)	<0.0001	17.153	<0.0001
	TIMI NSTEMI	0.552	0.288(0.267,0.31)	<0.0001	45.081	<0.0001
External validation #1	Our Score	0.737				
	GRACE In-Hospital Death or MI	0.652	0.084(0.052,0.117)	<0.0001	18.992	<0.0001
	GRACE Ischemic Stroke	0.536	0.2(0.163,0.238)	<0.0001	41.024	<0.0001
	TIMI STEMI	0.692	0.044(0.018,0.071)	<0.0001	5.277	<0.0001
	TIMI NSTEMI	0.559	0.177(0.141,0.214)	<0.0001	34.183	<0.0001
External validation #2	Our Score	0.8				
	GRACE In-Hospital Death or MI	0.67	0.13(0.113,0.147)	<0.0001	26.521	<0.0001
	GRACE Ischemic Stroke	0.504	0.296(0.276,0.315)	<0.0001	59.622	<0.0001
	TIMI STEMI	0.72	0.079(0.066,0.093)	<0.0001	13.103	<0.0001
	TIMI NSTEMI	0.543	0.257(0.237,0.277)	<0.0001	40.870	<0.0001



(a) **(b)** **(c)**

Figure. ROC curve and AUC of risk scores among development and external validation samples
(a) development sample (b) external validation sample #1 (c) external validation sample #2