ORIGINAL RESEARCH

Development and Validation of a Dynamic Nomogram for Predicting in-Hospital Mortality in Patients with Acute Pancreatitis: A Retrospective Cohort Study in the Intensive Care Unit

Kang Zou[1,2,]*, Shu Huang[3,4,]*, Wensen Ren[1,2,]*, Huan Xu[1,2], Wei Zhang[1,2], Xiaomin Shi[1,2], Lei Shi[1,2], Xiaolin Zhong[1,2], Yan Peng[1,2], Muhan Lü[1,2], Xiaowei Tang 🔴[1,2]

[1]Department of Gastroenterology, the Affiliated Hospital of Southwest Medical University, Luzhou, People's Republic of China; [2]Nuclear Medicine and Molecular Imaging Key Laboratory of Sichuan Province, Luzhou, People's Republic of China; [3]Department of Gastroenterology, Lianshui County People' Hospital, Huaian, People's Republic of China; [4]Department of Gastroenterology, Lianshui People' Hospital of Kangda College Affiliated to Nanjing Medical University, Huaian, People's Republic of China

*These authors contributed equally to this work

Correspondence: Xiaowei Tang; Muhan Lü, Department of Gastroenterology, the Affiliated Hospital of Southwest Medical University, Street Taiping No. 25, Region Jiangyang, Luzhou, Sichuan Province, 646099, People's Republic of China, Tel +86 8303165200, Fax +86 83061641541, Email solitude5834@hotmail.com; lvmuhan@swmu.edu.cn

Purpose: The aim of this study is to develop and validate a predictive model for the prediction of in-hospital mortality in patients with acute pancreatitis (AP) based on the intensive care database.

Patients and Methods: We analyzed the data of patients with AP in the Medical Information Mart for Intensive Care-IV (MIMIC-IV) database and Electronic Intensive Care Unit Collaborative Research Database (eICU-CRD). Then, patients from MIMIC-IV were divided into a development group and a validation group according to the ratio of 8:2, and eICU-CRD was assigned as an external validation group. Univariate logistic regression and least absolute shrinkage and selection operator regression were used for screening the best predictors, and multivariate logistic regression was used to establish a dynamic nomogram. We evaluated the discrimination, calibration, and clinical efficacy of the nomogram, and compared the performance of the nomogram with Acute Physiology and Chronic Health Evaluation II (APACHE-II) score and Bedside Index of Severity in AP (BISAP) score.

Results: A total of 1030 and 514 patients with AP in MIMIC-IV database and eICU-CRD were included in the study. After stepwise analysis, 8 out of a total of 37 variables were selected to construct the nomogram. The dynamic nomogram can be obtained by visiting https://model.sci-inn.com/KangZou/. The area under receiver operating characteristic curve (AUC) of the nomogram was 0.859, 0.871, and 0.847 in the development, internal, and external validation set respectively. The nomogram had a similar performance with APACHE-II (AUC = 0.841, p = 0.537) but performed better than BISAP (AUC = 0.690, p = 0.001) score in the validation group. Moreover, the calibration curve presented a satisfactory predictive accuracy, and the decision curve analysis suggested great clinical application value of the nomogram.

Conclusion: Based on the results of internal and external validation, the nomogram showed favorable discrimination, calibration, and clinical practicability in predicting the in-hospital mortality of patients with AP.

Keywords: acute pancreatitis, medical information mart for intensive care-IV, MIMIC-IV, electronic intensive care unit collaborative research database, eICU-CRD, dynamic nomogram

Introduction

As one of the most common gastrointestinal disorders, acute pancreatitis (AP) is an unpredictable and potentially fatal disease. The annual incidence of AP is about 34 per 100000 person-years.[1] The majority of AP patients usually have a mild course which recovers within one week without any sequelae. However, approximately one-fifth of patients

develop into moderate or severe AP due to organ failure, with a mortality rate of 20–40%.[2] A number of studies have found that severe patients might benefit from early adequate fluid resuscitation,[3] sufficient analgesia,[4] early enteral feeding,[5] use of endoscopy.[6] Therefore, it is highly significant to identify patients at high risk during the early stages of AP. This identification not only aids clinicians in making optimal clinical decisions and improving the prognosis but also facilitates patient and family counseling/explanation as well as resource allocation.[7]

Currently, there were some score systems such as Ranson, Glasgow, sequential organ failure assessment (SOFA), Acute Physiology and Chronic Health Evaluation II (APACHE II), and Bedside Index of Severity in AP (BISAP) score for risk stratification in AP. Ranson[8] and Glasgow[9] scores were developed for predicting the severity of AP, BISAP[10,11] and SOFA[12] scores were for the mortality of AP, APACHE II[13,14] score was for both severity and mortality. However, the wide clinical application of these score systems was restricted to their complex calculation, lagging assessment, and suboptimal accuracy.

Nomogram is a visual statistical model which was widely used to evaluate the prognosis of related diseases in medical research. In the era of "big data", medical databases are increasingly used for scientific research. Although several studies have developed nomograms to predict in-hospital mortality of AP based on critical care databases,[15–17] small sample size and lacking further external validation were their main limitation. Moreover, the clinical applicability of traditional nomogram may be restricted to time-consuming and lacking an exact calculation. Therefore, the aim of the study was as follows: First, we developed a dynamic nomogram based on the data from the Medical Information Mart for Intensive Care-IV (MIMIC-IV) database to predict in-hospital mortality of AP patients admitted to the Intensive Care Unit (ICU) and compared the performance of the nomogram and traditional score systems. Furthermore, the data from the Electronic Intensive Care Unit Collaborative Research Database (eICU-CRD) was used for external validation to improve the level of evidence.

Materials and Methods
Data Source and Outcome
This retrospective cohort study extracted critical care data from the MIMIC-IV (v.2.0) database. It contains over 70,000 ICU patients admitted to the Beth Israel Deaconess Medical Center from 2008 to 2019.[18] The MIMIC-IV, a large, single-center, and de-identified patient database, includes plenty of patient information such as demographics, bedside monitors, International Classification of Diseases codes version 9 and version 10 (ICD-9 and ICD-10) revision diagnoses, laboratory data, therapeutic measures, radiology reports, and clinical outcome.[19] In addition, the eICU-CRD (version 2.0), a multicenter database composed of more than 200,000 ICU admissions in the United States, was served as an independent external validation set.[20] After an in-depth study of a training course "Protecting Human Research Participants", an author (Kang Zou) obtained permission to access the two databases for research purposes (certification number: 11739201). The ethics committee of the Affiliated Hospital of Southwest Medical University, China, approved the study and waived the ethics approval as this retrospective study was based on public databases and personal information provided in the databases was de-identified. The primary outcome of this study was the all-cause in-hospital mortality of patients with AP admitted to ICU.

Study Patients
All patients in the MIMIC-IV and eICU-CRD who were diagnosed as AP were recruited (only including the ICU admission). The patients were identified by a manual search of ICD-9 (577.0) and ICD10 (K85) codes. The exclusion criteria were as follows: (1) Age below 18, (2) Repeat ICU admission, (3) Primary outcome missing. The flow chart of patient selection is shown in Figure 1.

Data Collection and Variable Extraction
For this study, we collected information from the electronic medical record including the general demographic, vital signs, laboratory test results, advanced life support strategy, accompanied diseases, and score systems. The process of data collection was following Cosgriff[21] with a slight modification. The following variables were excluded: (1) Variables
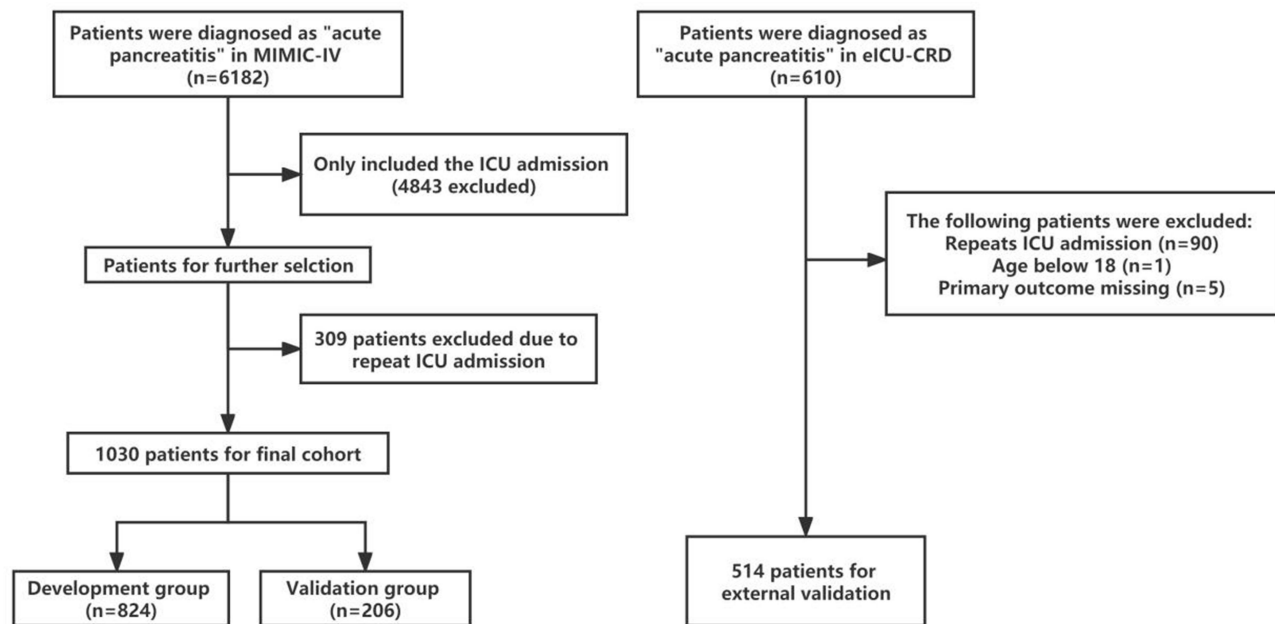
Figure 1 Flow chart of patients selection.

missing over 30%, (2) Variables that were not shared by the two databases. Finally, the following demographic data collected admission were extracted: age, gender, race, and hospital expire flag (in-hospital death). The mean over all values of the vital sign in the first 24 hours was selected: heart rate, respiratory rate, systolic blood pressure (SBP), diastolic blood pressure (DBP), oxyhemoglobin saturation (Spo2), mean blood pressure (MBP), and temperature. The maximum value within 24 hours of admission was used for the following variables: red cell distribution width (RDW), creatinine, blood urea nitrogen (BUN), total bilirubin (TBIL), aminotransferase alanine (ALT), aminotransferase aspartate (AST), and aniongap. The minimum value within the first 24 hours was extracted for bicarbonate, albumin, platelets, hemoglobin, hematocrit, red blood cell (RBC), calcium, and chloride. Both the minimum and maximum value within 24 hours after admission was selected for the following variables: the maximum of white blood cells (WBC_max), the minimum of white blood cells (WBC_min), the minimum of sodium (sodium_min), the maximum of sodium (sodium_-max), the minimum of potassium (potassium_min), the maximum of potassium (potassium_max). In addition, advanced life support (mechanical ventilation and renal replacement therapy), comorbidities (hypertension and diabetes), and score systems (BISAP and APACHE II score) were also included. For missing values, a K-nearest neighbor classification algorithm was applied to minimize the impact of missing data on results.

Statistical Analysis

The continuous variables with a normal distribution were presented as the mean ± standard deviation ($\bar{x} \pm s$), and Student's *t*-test was used for comparisons between groups, while those having an abnormal distribution were expressed as the median and interquartile ranges [M (Q1, Q3)] and compared with the Mann–Whitney U rank-sum test. Categorical variables were expressed by counts and percentages [n (%)], and χ2 test or Fisher's exact test was used for comparison between groups. All patients from MIMIC-IV were randomly assigned into the development group and validation group according to a ratio of 8:2. The univariate logistic regression and least absolute shrinkage and selection operator (LASSO) regression method was used for screening the best predictors in the development group. On the basis of the screened variables, multivariate logistic regression was used to build the nomogram in the development group and then validated in the validation group. Furthermore, the data from eICU-CRD was used for external validation. As the nomogram was developed and validated in ICU patients, the results cannot be generalized to patients in a general ward well patient with mild AP. Various indicators, including the area under the receiver operating characteristic curve (AUC),

sensitivity, specificity, and Yoden index were used to assess the discrimination of this predictive model. Delong test was utilized to compare the predictive performance of the nomogram and score systems. Moreover, the calibration of nomogram was evaluated by the calibration curve, and the clinical value of the predictive model was also evaluated through decision curve analysis (DCA). All statistical analyses were carried out using R software (v.3.6.3), and statistical significance was set at $p < 0.05$.

Results

Baseline Data from the MIMIC-IV Database

As shown in Table 1, a total of 1030 patients extracted from the MIMIC-IV database were included in this study, in which, 133 patients were death and 897 survival during hospitalization, respectively. Among all the characteristics, age ($p < 0.001$), respiratory rate ($p < 0.001$), temperature ($p < 0.001$), SBP ($p < 0.001$), DBP ($p < 0.001$), MBP ($p < 0.001$),

Table 1 Summary of Demographic and Clinical Features of MIMIC-IV Patients with Acute Pancreatitis

	Survival	**Death**	**P-value**
Number (sample size), n (%)	897 (87.09)	133 (12.91)	
Baseline characteristics			
Age, year	58.27 (46.23, 71.90)	70.34 (57.45, 80.84)	<0.001
Height, cm	168.76 (162.86, 176.53)	168.82 (160.02, 173.00)	0.054
Weight, kg	81.10 (70.00, 98.20)	81.20 (68.00, 96.00)	0.602
Gender, n (%)			
Female	373 (41.58)	61 (45.87)	0.351
Male	524 (58.42)	72 (54.14)	
Race, n (%)			
White	567 (63.21)	82 (61.65)	0.729
Others/Unknown	330 (36.79)	51 (38.35)	
Vital signs			
Heart rate, /min	93.08 (79.40, 106.75)	93.04 (85.31, 109.94)	0.172
Respiratory rate, /min	19.90 (17.12, 23.16)	22.22 (19.72, 25.13)	<0.001
Temperature, °C	36.95 (36.68, 37.34)	36.75 (36.49, 37.13)	<0.001
SBP, mmHg	121.82 (109.75, 136.81)	109.00 (100.76, 117.787)	<0.001
DBP, mmHg	68.00 (59.68, 77.63)	58.60 (53.22, 66.90)	<0.001
MBP, mmHg	81.76 (73.67, 91.68)	72.60 (67.42, 81.37)	<0.001
Spo2, %	96.32 (94.94, 97.78)	95.77 (94.00, 97.32)	<0.001
Laboratory parameters			
WBC_max, 10^9/L	13.50 (9.70, 19.20)	16.30 (11.40, 22.30)	<0.001
WBC_min, 10^9/L	10.40 (7.00, 15.00)	12.40 (7.50, 17.20)	0.023
Hematocrit, %	31.88 ± 6.50	29.13 ± 6.72	<0.001
RDW, %	14.80 (13.80, 16.00)	16.20 (14.70, 18.60)	<0.001
RBC, 10^9/L	3.46 (3.01, 3.98)	3.06 (2.54, 3.70)	<0.001
Hemoglobin, g/dL	10.62 ± 2.23	9.47 ± 2.28	<0.001
Platelets, 10^9/L	171.00 (119.00, 241.00)	143.00 (76.00, 221.00)	0.002
ALT, Units/L	58.00 (27.00, 144.00)	58.00 (30.00, 204.00)	0.249

(Continued)

Table 1 (Continued).

	Survival	Death	*P*-value
AST, Units/L	79.00 (39.00, 173.00)	127.00 (50.00, 310.00)	<0.001
Albumin, g/dL	3.10 ± 0.61	2.67 ± 0.64	<0.001
TBIL, mg/dL	1.20 (0.60, 2.90)	2.00 (0.90, 7.00)	<0.001
BUN, mg/dL	19.00 (12.00, 34.00)	40.00 (25.00, 62.00)	<0.001
Creatinine, mg/dL	1.10 (0.70, 1.80)	2.00 (1.20, 3.30)	<0.001
Aniongap, mmHg	17.00 (14.00, 20.00)	20.00 (16.00, 24.00)	<0.001
Bicarbonate, mmol/L	21.00 (17.00, 24.00)	17.00 (13.00, 21.00)	<0.001
Calcium, mg/dL	7.80 (7.20, 8.30)	7.50 (6.80, 8.10)	0.002
Chloride, mmol/L	102.00 (98.00, 106.00)	102.00 (97.00, 107.00)	0.895
Sodium_min, mmol/L	137.00 (134.00, 140.00)	137.00 (132.00, 141.00)	0.975
Sodium_max, mmol/L	140.00 (137.00, 143.00)	140.00 (136.00, 143.00)	0.367
Potassium_min, mmol/L	3.70 (3.40, 4.10)	4.00 (3.50, 4.40)	<0.001
Potassium_max, mmol/L	4.30 (3.90, 4.80)	4.60 (4.10, 5.40)	<0.001
Advanced life support			
Mechanical ventilation, n (%)	279 (31.10)	68 (51.13)	<0.001
Renal replacement therapy, n (%)	17 (1.90)	16 (12.03)	<0.001
Accompanied diseases			
Hypertension, n (%)	556 (61.98)	94 (70.68)	0.053
Diabetes, n (%)	279 (31.10)	42 (31.58)	0.912
Score system			
BISAP	2.00 (1.00, 3.00)	3.00 (2.00, 4.00)	<0.001
APACHE II	18.00 (13.00, 24.00)	28.00 (22.00, 34.00)	<0.001

Abbreviations: SBP, systolic blood pressure; DBP, diastolic blood pressure; MBP, mean blood pressure; Spo2, oxyhemoglobin saturation; WBC_max, the maximum of white blood cells; WBC_min, the minimum of white blood cells; RDW, red cell distribution width; RBC, red blood cells; ALT, aminotransferase alanine; AST, aminotransferase aspartate; TBIL, total bilirubin; BUN, blood urea nitrogen; Sodium_min, the minimum of sodium; Sodium_max, the maximum of sodium; Potassium_min, the minimum of potassium; Potassium_max, the maximum of potassium; BISAP, Bedside Index of Severity in Acute Pancreatitis; APACHE II, Acute Physiology and Chronic Health Evaluation II.

Spo2 ($p < 0.001$), WBC_max ($p < 0.001$), WBC_min ($p = 0.023$), hematocrit ($p < 0.001$), RDW ($p < 0.001$), RBC ($p < 0.001$), Hemoglobin ($p < 0.001$), platelets ($p = 0.002$), AST ($p < 0.001$), albumin ($p < 0.001$), TBIL ($p < 0.001$), BUN ($p < 0.001$), creatinine ($p < 0.001$), aniongap ($p < 0.001$), bicarbonate ($p < 0.001$), calcium ($p = 0.002$), potassium_min ($p < 0.001$), potassium_max ($p < 0.001$), mechanical ventilation ($p < 0.001$), renal replacement therapy ($p < 0.001$), BISAP score ($p < 0.001$), and APACHE II score ($p < 0.001$) differed significantly between the two groups.

A Predictive Nomogram Development

There were 824 and 206 patients in the development and validation group. Supplementary Table 1 displayed the baseline characteristics of the two groups. There were no significant differences between the two data sets in any variable (P > 0.05), which meant that the distribution of clinical features for the two data sets was well balanced. Univariate logistic regression analysis for variable selection in the development group is shown in Supplementary Table 2. LASSO regression was used to further screen variables with non-zero coefficients among the results of univariate logistic regression. It showed that the best predictors of in-hospital mortality in AP were age, oxygen saturation, SBP, temperature, respiratory rate, RDW, Wbc_max, bicarbonate, BNU, albumin, and TBIL (Figure 2).
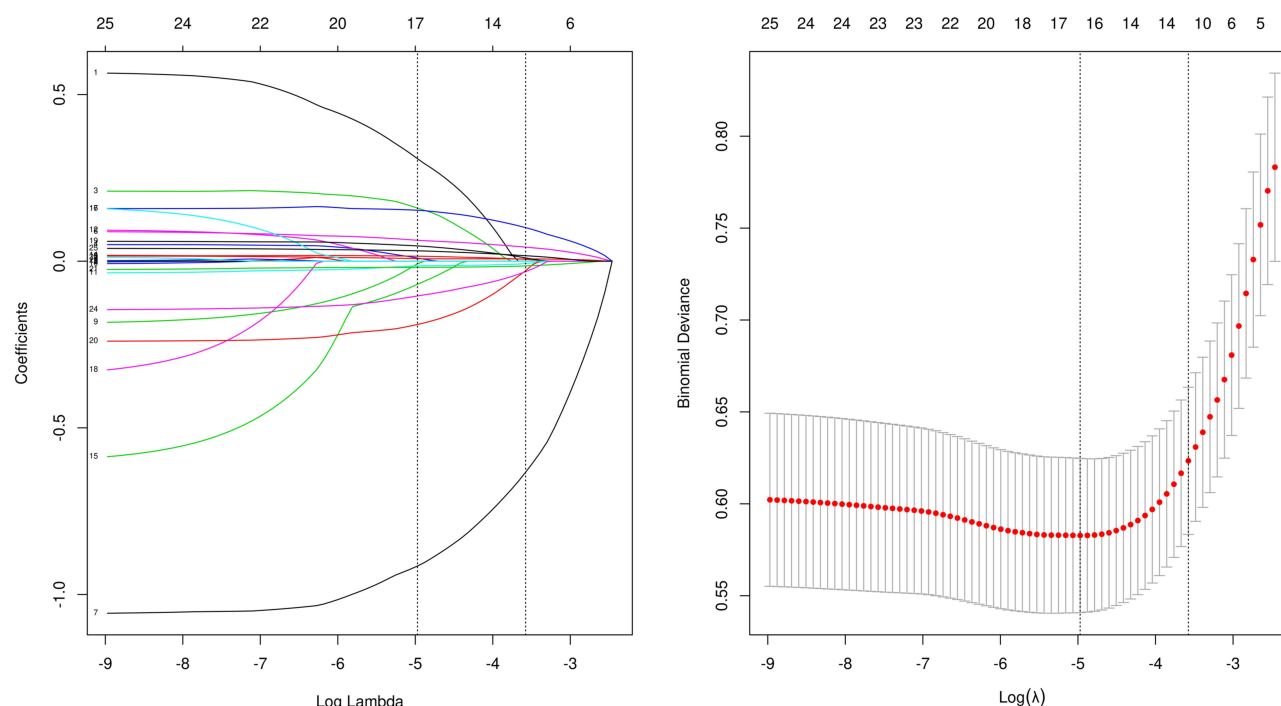
Figure 2 Selection of predictive factors using the least absolute shrinkage and selection operator logistic regression algorithm. Least absolute shrinkage and selection operator coefficient profiles of the 26 candidate variables. The best value was determined by the two dashed vertical lines which were drawn according to the minimum mean-square error criterion (left dashed line) and the standard error criterion (right dashed line). In the present study, 11 predictors were selected according to the minimum standard error criterion (λ=0.028).

Based on the above variables, multivariate logistic regression was used to establish a predictive model represented by a nomogram (Figure 3A). Multivariate logistic regression showed that there were 8 independent predictors of in-hospital mortality in AP, as shown in Table 2.

To obtain the in-hospital mortality risk of an individual patient, we can get an exact variable value according to clinic characteristics of the patient, and then draw a vertical line to the "Points" axis to get points of one variable. Subsequently, we can obtain total points by summing all of the points for 8 variables. Finally, the probability of in-hospital mortality for the patient can be located by drawing a vertical line from the "Total Points" axis to the "Prob of in-hospital mortality" axis. A precise probability of in-hospital mortality can be obtained from the dynamic nomogram by visiting https://model.sci-inn.com/KangZou/. For example, an 83-year-old patient with AP who was admitted in ICU, the mean SBP, respiratory rate, and Spo2 were 96mmHg, 28 breaths/min, and 87%, respectively, in the first 24 hours. The maximum value within 24 hours of RDW, TBIL and BUN were 16%, 3mg/dL, and 39mg/dL, respectively. The minimum of albumin in the first 24 hours was 2g/dL. We can obtain the exact probability and 95% confidence interval of in-hospital mortality of this patient was 0.842 (0.663–0.944), as shown in Figure 3B.

Evaluation and Validation of the Nomogram

The receiver operating characteristic (ROC) curve and AUC of the nomogram and score system in both the development and validation groups were calculated and compared (Figure 4A and B). In the development group, the AUC of the nomogram, APACHE II, and BISAP scores were 0.859, 0.781, and 0.700, respectively. The discrimination of our nomogram was significantly higher than APACHE II (p = 0.001) and BISAP (p < 0.001) score. In the validation group, the AUC was 0.871, 0.841, and 0.690, respectively. Compared to BISAP score, our nomogram performed better in the discrimination (p = 0.013). However, there was no significant difference between APACHE II score and the nomogram (p = 0.537) (Table 3).
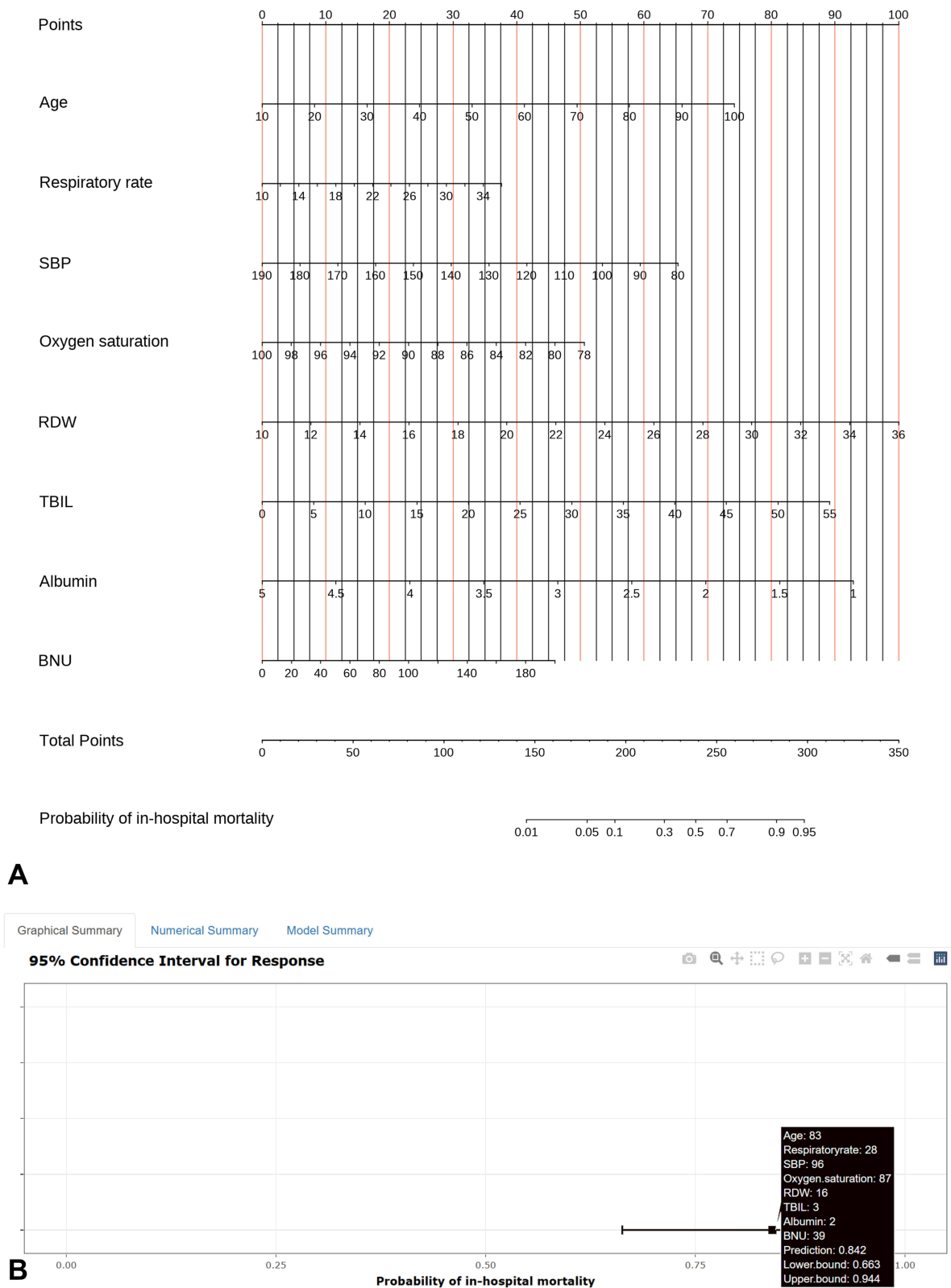
Figure 3 Nomogram to estimate the risk of in-hospital mortality in acute pancreatitis. (**A**) Ordinary nomogram. (**B**) The schematic outline of dynamic nomogram. To obtain the accurate probability of in-hospital mortality, select the value of each variable and then click on the "Predict" button. The probability and 95% confidence interval will receive. For properly function of the site, please click the "Quit" button after finish predicting.
Abbreviations: SBP, systolic blood pressure; RDW, red cell distribution width; TBIL, total bilirubin; BUN, blood urea nitrogen.

Table 2 Risk Factors Selected After Multivariate Logistic Regression Analysis

Variable	OR with CI	*P*-value
Age	1.042 (1.025–1.059)	<0.001
SBP	0.971 (0.956–0.986)	<0.001
Respiratory rate	1.074 (1.011–1.741)	0.020
Oxygen saturation	0.893 (0.802–0.989)	0.033
RDW	1.209 (1.106–1.324)	<0.001
BUN	1.011 (1.003–1.019)	0.005
Albumin	0.318 (0.205–0.483)	<0.001
TBIL	1.083 (1.043–1.129)	<0.001
(Intercept)	2605.968	0.152

Abbreviations: OR, odds ratio; CI, confidence interval; SBP, systolic blood pressure; RDW, red cell distribution width; BUN, blood urea nitrogen; TBIL, total bilirubin.

The calibration curves constructed through 1000 bootstrap replicates is shown in Figure 4C and D. The Brier score was 0.08 in validation group. The calibration curves showed good consistency between the predicted probability and the actual probability in both the development and validation groups.

The DCA in Figure 4E and F indicated that the nomogram, APACHE II, and BISAP score showed more benefit than the treat-all-patients or treat-none scheme. Moreover, compared with the APACHE II and BISAP score, the net benefit of our nomogram was significantly higher, suggesting that our nomogram was more clinically useful.

To improve the generalization of the model, we used the data from eICU-CRD for external validation. There were 514 patients extracted from eICU-CRD, of which 39 patients were death and 475 survival. In the external validation set, age (p = 0.003), respiratory rate (p = 0.010), SBP (p < 0.001), DBP (p < 0.001), MBP (p < 0.001), WBC_max (p = 0.006), hematocrit (p = 0.001), RDW (p < 0.001), RBC (p < 0.001), hemoglobin (p < 0.001), albumin (p < 0.001), TBIL (p = 0.009), BUN (p < 0.001), creatinine (p < 0.001), aniongap (p = 0.042), potassium_max (p = 0.004), mechanical ventilation (p < 0.001), renal replacement therapy (p < 0.001) were significantly different between the survival and death patients (Supplementary Table 3). The ROC of the external cohort is shown in Figure 5A. The AUC of the nomogram in the external cohort was 0.847. The calibration curve is shown in Figure 5B. The Brier score in the external cohort was 0.077, which indicated that the predicted probability of nomogram was also highly in agreement with the actual probability in the external cohort.

Discussion

AP is a heterogeneous disease with a mild to severe course. Most severe AP patients require intensive care and advanced life support. Determining the prognosis of AP is the base of both patient-centered care and shared decision-making, such as choice of treatment strategies and informed consent.[19] This present study constructed a dynamic nomogram with the data from MIMIC-IV database. Based on multivariate regression analyses, 8 predictors including age, oxygen saturation, SBP, respiratory rate, RDW, albumin, TBIL, and BUN were the independent prognostic factors for in-hospital mortality of AP. The nomogram showed great performance in predicting in-hospital mortality of AP and external validation data from eICU-CRD received equally satisfactory results. In addition, compared to the traditional score system, our nomogram showed better discrimination and clinical practicability.

There were several prognostic scores systems such as APACHE II, BISAP, and Ranson score performed well in predicting the mortality of AP.[22–24] However, these score systems were mainly applied to scientific research due to various limitations, which led to simple and effective tools still lacking in clinical work. The APACHE II score involved multiple variables and complex computation. In addition, the generalizability of APACHE II score across diseases was poor because it was not designed specifically for AP. The BISAP score involved a subjective assessment of mental status.[11] The SOFA score consisted of five variables scored within 24 hours of admission, while it was simplistic to use.[25]
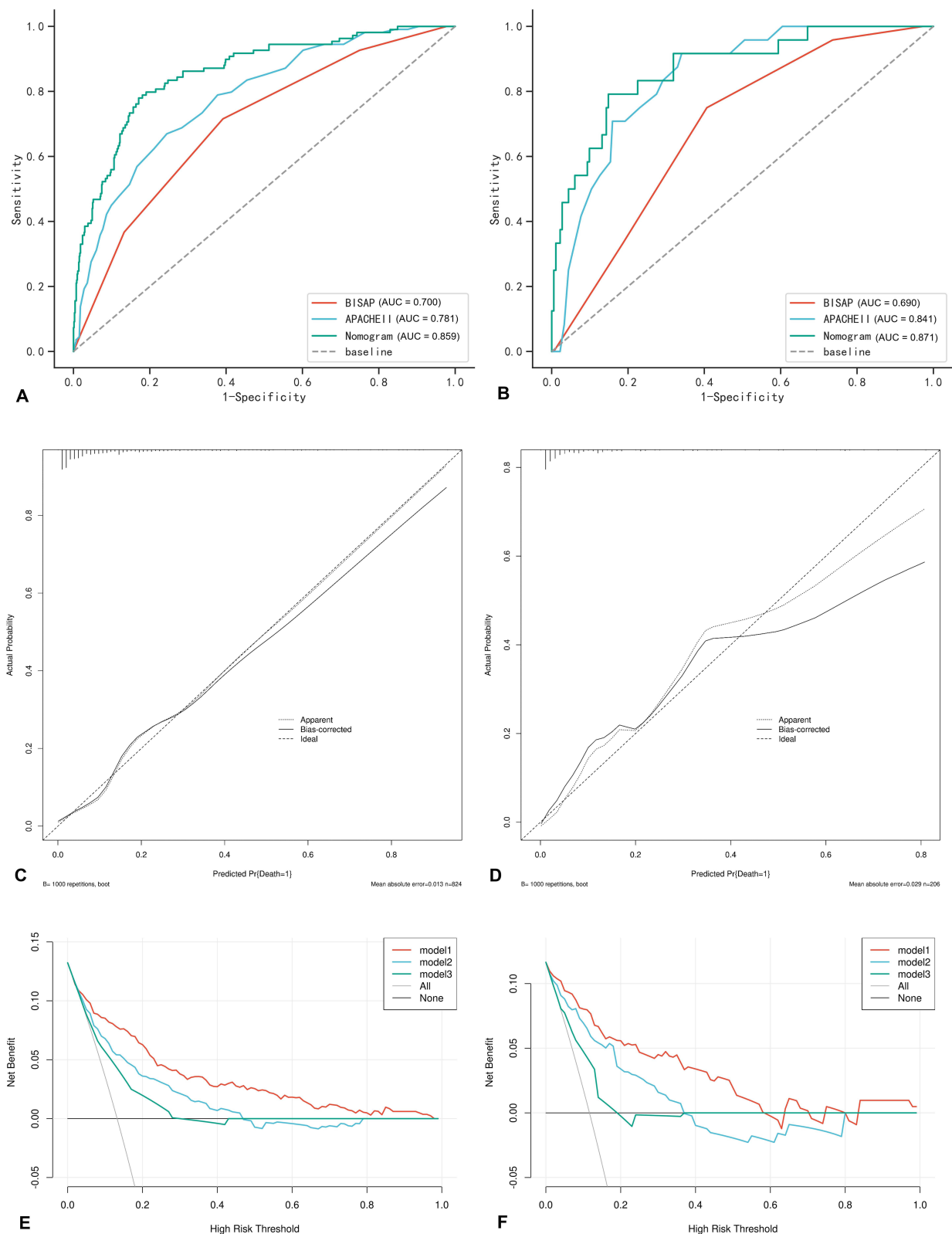
Figure 4 The discrimination, calibration, and clinical value of the nomogram. (**A** and **B**) Receiver operating characteristic curves in the development and validation group by Nomogram and Score system. **A**, Development group. **B**, Validation group. (**C**) Calibration curves of development group. (**D**) Calibration curves of validation group (Briser Score = 0.08). (**E** and **F**) Decision curve analysis of Nomogram and Score system. **E**, Development group. **F**, Validation group. Model 1, nomogram; model 2, APACHE II score; model 3, BISAP score.

Abbreviations: AUC, the area under the receiver operating characteristic curve; APACHE II, Acute Physiology and Chronic Health Evaluation II; BISAP, Bedside Index of Severity in Acute Pancreatitis.

Table 3 Comparison of the Nomogram Model and Score System for Predicting the in-Hospital Mortality of Acute Pancreatitis in the Development, Internal Validation and External Validation Set

	AUC	Sensitivity	Specificity	Yoden Index
Development group				
Nomogram	0.859	0.780	0.829	0.609
APACHE II	0.781[a]	0.670	0.755	0.425
BISAP	0.700[b]	0.716	0.608	0.324
Validation group				
Nomogram	0.871	0.792	0.852	0.643
APACHE II	0.841[c]	0.917	0.659	0.576
BISAP	0.690[d]	0.750	0.593	0.343
External validation	0.847	0.872	0.684	0.556

Notes: [a]$p = 0.001$, the comparison between the nomogram and APACHE II in the development group. [b]$p < 0.001$, the comparison between the nomogram and BISAP in the development group. [c]$p = 0.537$, the comparison between the nomogram and APACHE II in the validation group. [d]$p = 0.001$, the comparison between the nomogram and BISAP in the validation group.
Abbreviations: AUC, the area under the receiver operating characteristic curve; APACHE II, Acute Physiology and Chronic Health Evaluation II; BISAP, Bedside Index of Severity in Acute Pancreatitis.

Both MIMIC and eICU-CRD are publicly accessible and well-established research databases, analysis can be accessed without restriction once a data use agreement is accepted. It is obvious that these medical critical care databases were widely used in areas of clinical research and education all over the world.[26] In the past several years, a large number of researchers put efforts to probe prognostic analysis,[27,28] length of stay,[29] prediction of severity,[17] and treatment strategy[30,31] of AP based on the two databases. Currently, several studies have developed nomograms for the prediction of in-hospital mortality in AP based on critical care database. Liu[15] and Li[16] constructed nomograms to predict in-hospital mortality of AP based on Medical Information Mart for IntensiveCare-III (MIMIV-III) database and eICU-CRD, respectively. However, both two traditional nomograms lack external validation. A study by Xu[17] also developed an
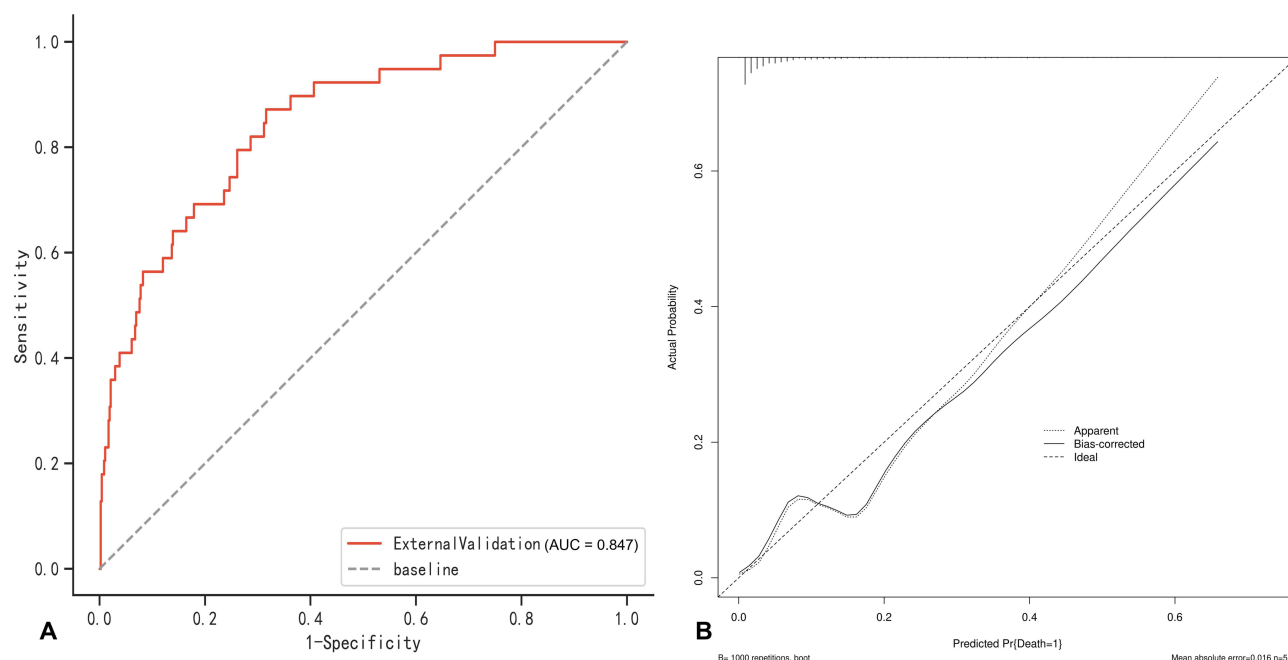

Figure 5 The discrimination and calibration in external validation set based on the data of Electronic Intensive Care Unit Collaborative Research Database. (**A**) Receiver operating characteristic curve. (**B**) Calibration curve (Briser Score = 0.077).

ordinary nomogram combining serum TBIL and albumin with other significant features based on MIMIC-III and eICU-CRD, which also show favorable discrimination. But the predictive variables selected by Xu X mainly focused on endogenous antioxidant indexes and the SOFA score included in the nomogram needs further calculation. In addition, MIMIC-III includes data more than 50,000 patients from 2001 to 2012 while over 70,000 ICU admissions in MIMIV-IV from 2008 to 2019.[18,26] Compared with MIMIC-III, MIMIC-IV has a great boost in the number of cases and the data of MIMIC-IV are closer to the medical level of the last decade. Therefore, we used MIMIC-IV database to develop a dynamic nomogram and exerted internal validation. The predictors in our nomogram only involved vital signs and common laboratory indicators which were easily accessed after admission. It can guide physicians to exert risk stratification and make clinical decisions for AP patients.

Older age was regarded as a well-established marker of poor prognosis[32,33] in AP and was included in the APACHE II score, Ranson score, the BISAP score, and the Japanese severity score as a predictor. The likely explanation was that there were higher odds of comorbid diseases with increasing age.[34] Age was also demonstrated as an independent risk factor in the present study. RDW is a laboratory parameter widely used to evaluate the extent of erythrocyte anisocytosis. In addition to being a traditional marker of anemia,[35] studies suggested that it can be used as an indicator of inflammation and was related to multiple links of inflammatory response in recent years.[36–38] Therefore, there was a significant association between RDW and the mortality of patients with AP.[39,40] With a cutoff value of 14.8 for RDW, approximately 77% of the deaths could be correctly predicted.[41] Similarly, RDW was also correlated with the in-hospital mortality of AP and weighed heavily in our nomogram. BUN, albumin, and TBIL were significant predictors for mortality of AP, which were already confirmed by numerous studies.[11,17,42–44] Current nomograms which were established for mortality of AP by other researchers usually included one or more of them as predictors,[15,16,45] which were consistent with our nomogram. AS for SBP, respiratory rate, and oxygen saturation as a reflection of patients' vital signs, it can surely be used as a predictor of mortality.

This retrospective analysis of a large sample from multiple centers explored risk factors of mortality in AP and developed a nomogram that can predict in-hospital mortality in patients with AP precisely. The predictive variables in the nomogram were easily accessible within 24 hours after ICU admission. In addition, we improved the limitations of lack of external validation and small sample size commonly found in previous studies. Finally, on the basis of the traditional ordinary nomogram, we further constructed a dynamic nomogram, which can facilitate rapid and accurate risk stratification and clinical decision-making for clinicians. Therefore, our nomogram model should be reliable and user-friendly. However, there were also several limitations in the present study. First, selection bias was inevitable in this retrospective study. Prospective studies are needed to improve the level of evidence. Second, we excluded the variables with more than 30% of observations missing to ensure the accuracy of our study. It led to some important laboratory index such as C-reactive protein, procalcitonin being excluded due to significant missing data. Moreover, in the acute setting, albumin may not be a reliable indicator and its levels can be influenced by nutritional or hepatic dysfunction issues, which may not necessarily reflect the severity of AP. Third, the mortality between the two databases was different. There were only about 40 deaths in the eICU-CRD. However, we included 8 variables in nomogram, which meant the principle of 10-fold Events per variable was not followed in external validation set. Finally, data was extracted from the medical records of patients and the performance of the nomogram was closely related to the accuracy of the recording.

Conclusion

We constructed a dynamic nomogram and make internal and external validation based on MIMIC-IV and eICU-CRD. The link of the dynamic nomogram is https://model.sci-inn.com/KangZou/. The nomogram showed good discrimination, calibration, and clinical practicability in predicting the in-hospital mortality of AP. However, prospective studies are still necessary in order to bring such predictive models into clinical application earlier.

Data Sharing Statement

The data are available on the website at https://physionet.org/.

Ethics Approval and Informed Consent

The requirement for individual patient consent and an ethical approval statement was exempted in the present study as this retrospective study was based on public databases and all patient privacy information provided in the databases was de-identified.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

The authors report no conflicts of interest in this work.

References

1. Xiao AY, Tan ML, Wu LM, et al. Global incidence and mortality of pancreatic diseases: a systematic review, meta-analysis, and meta-regression of population-based cohort studies. *Lancet Gastroenterol Hepatol*. 2016;1(1):45–55. doi:10.1016/S2468-1253(16)30004-8
2. Boxhoorn L, Voermans RP, Bouwense SA, et al. Acute pancreatitis. *Lancet*. 2020;396(10252):726–734. doi:10.1016/S0140-6736(20)31310-6
3. Warndorf MG, Kurtzman JT, Bartel MJ, et al. Early fluid resuscitation reduces morbidity among patients with acute pancreatitis. *Clin Gastroenterol Hepatol*. 2011;9(8):705–709. doi:10.1016/j.cgh.2011.03.032
4. Jabaudon M, Belhadj-Tahar N, Rimmelé T, et al. Thoracic epidural analgesia and mortality in acute pancreatitis: a multicenter propensity analysis. *Crit Care Med*. 2018;46(3):e198–e205. doi:10.1097/CCM.0000000000002874
5. Bakker OJ, van Brunschot S, van Santvoort HC, et al. Early versus on-demand nasoenteric tube feeding in acute pancreatitis. *N Engl J Med*. 2014;371(21):1983–1993. doi:10.1056/NEJMoa1404393
6. Gliem N, Ammer-Herrmenau C, Ellenrieder V, Neesse A. Management of severe acute pancreatitis: an update. *Digestion*. 2021;102(4):503–507. doi:10.1159/000506830
7. Kiat TTJ, Gunasekaran SK, Junnarkar SP, Low JK, Woon W, Shelat VG. Are traditional scoring systems for severity stratification of acute pancreatitis sufficient? *Ann Hepatobiliary Pancreat Surg*. 2018;22(2):105–115. doi:10.14701/ahbps.2018.22.2.105
8. Ranson JHC, Pasternack BS. Statistical methods for quantifying the severity of clinical acute pancreatitis. *J Surg Res*. 1977;22(2):79–91. doi:10.1016/0022-4804(77)90045-2
9. Blamey SL, Imrie CW, O'Neill J, Gilmour WH, Carter DC. Prognostic factors in acute pancreatitis. *Gut*. 1984;25(12):1340–1346. doi:10.1136/gut.25.12.1340
10. Hagjer S, Kumar N. Evaluation of the BISAP scoring system in prognostication of acute pancreatitis - A prospective observational study. *Int J Surg*. 2018;54(Pt A):76–81. doi:10.1016/j.ijsu.2018.04.026
11. Wu BU, Johannes RS, Sun X, Tabak Y, Conwell DL, Banks PA. The early prediction of mortality in acute pancreatitis: a large population-based study. *Gut*. 2008;57(12):1698–1703. doi:10.1136/gut.2008.152702
12. Halonen KI, Pettilä V, Leppäniemi AK, Kemppainen EA, Puolakkainen PA, Haapiainen RK. Multiple organ dysfunction associated with severe acute pancreatitis. *Crit Care Med*. 2002;30(6):1274–1279. doi:10.1097/00003246-200206000-00019
13. Larvin M, McMahon MJ. APACHE-II score for assessment and monitoring of acute pancreatitis. *Lancet*. 1989;2(8656):201–205. doi:10.1016/S0140-6736(89)90381-4
14. Wilson C, Heath DI, Imrie CW. Prediction of outcome in acute pancreatitis: a comparative study of APACHE II, clinical assessment and multiple factor scoring systems. *Br J Surg*. 1990;77(11):1260–1264. doi:10.1002/bjs.1800771120
15. Liu Z, Yang Y, Song H, Luo J. A prediction model with measured sentiment scores for the risk of in-hospital mortality in acute pancreatitis: a retrospective cohort study. *Ann Transl Med*. 2022;10(12):676. doi:10.21037/atm-22-1613
16. Li C, Ren Q, Wang Z, Wang G. Early prediction of in-hospital mortality in acute pancreatitis: a retrospective observational cohort study based on a large multicentre critical care database. *BMJ Open*. 2020;10(12):e041893. doi:10.1136/bmjopen-2020-041893
17. Xu X, Ai F, Huang M. Deceased serum bilirubin and albumin levels in the assessment of severity and mortality in patients with acute pancreatitis. *Int J Med Sci*. 2020;17(17):2685–2695. doi:10.7150/ijms.49606
18. Johnson A, Bulgarelli L, Pollard T, Horng S, Celi LA, Mark R. MIMIC-IV (version 2.0). PhysioNet; 2022.
19. Peng S, Huang J, Liu X, et al. Interpretable machine learning for 28-day all-cause in-hospital mortality prediction in critically ill patients with heart failure combined with hypertension: a retrospective cohort study based on medical information mart for intensive care database-IV and eICU databases. *Front Cardiovasc Med*. 2022;9:994359. doi:10.3389/fcvm.2022.994359
20. Pollard T, Johnson A, Raffa J, Celi LA, Badawi O, Mark R. eICU collaborative research database (version 2.0). PhysioNet; 2019.
21. Cosgriff CV, Celi LA, Ko S, et al. Developing well-calibrated illness severity scores for decision support in the critically ill. *NPJ Digit Med*. 2019;2:76. doi:10.1038/s41746-019-0153-6

22. Papachristou GI, Muddana V, Yadav D, et al. Comparison of BISAP, Ranson's, APACHE-II, and CTSI scores in predicting organ failure, complications, and mortality in acute pancreatitis. *Am J Gastroenterol*. 2010;105(2):435–442. doi:10.1038/ajg.2009.622

23. Kumar A, Singh Griwan M. A comparison of APACHE II, BISAP, Ran- son's score and modified CTSI in predicting the severity of acute pancreati- tis based on the 2012 revised Atlanta Classification. *Gastroenterol Rep*. 2018;6:127–131. doi:10.1093/gastro/gox029

24. Zhang J, Shahbaz M, Fang R, et al. Comparison of the BISAP scores for predicting the severity of acute pancreatitis in Chinese patients accord- ing to the latest Atlanta classification. *J Hepatobiliary Pancreat Sci*. 2014;21:689–694. doi:10.1002/jhbp.118

25. Teng TZJ, Tan JKT, Baey S, et al. Sequential organ failure assessment score is superior to other prognostic indices in acute pancreatitis. *World J Crit Care Med*. 2021;10(6):355–368. doi:10.5492/wjccm.v10.i6.355

26. Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035. doi:10.1038/sdata.2016.35

27. Liu Q, Zheng HL, Wu MM, et al. Association between lactate-to-albumin ratio and 28-days all-cause mortality in patients with acute pancreatitis: a retrospective analysis of the MIMIC-IV database. *Front Immunol*. 2022;13:1076121. doi:10.3389/fimmu.2022.1076121

28. Hameed MAB, Alamgir Z. Improving mortality prediction in Acute Pancreatitis by machine learning and data augmentation. *Comput Biol Med*. 2022;150:106077. doi:10.1016/j.compbiomed.2022.106077

29. Wang D, Lu J, Zhang P, Hu Z, Shi Y. Relationship between blood glucose levels and length of hospital stay in patients with acute pancreatitis: an analysis of MIMIC-III database. *Clin Transl Sci*. 2023;16(2):246–257. doi:10.1111/cts.13445

30. Ma Y, Yan T, Xu F, et al. Infusion of human albumin on acute pancreatitis therapy: new tricks for old dog? *Front Pharmacol*. 2022;13:842108. doi:10.3389/fphar.2022.842108

31. Gaitanidis A, Breen K, Mendoza A, et al. Enteral nutrition is associated with high rates of pneumonia in intensive care unit (ICU) patients with acute pancreatitis. *J Crit Care*. 2022;69:154012. doi:10.1016/j.jcrc.2022.154012

32. Yasuda H, Horibe M, Sanui M, et al. Etiology and mortality in severe acute pancreatitis: a multicenter study in Japan. *Pancreatology*. 2020;20 (3):307–317. doi:10.1016/j.pan.2020.03.001

33. Di MY, Liu H, Yang ZY, Bonis PA, Tang JL, Lau J. Prediction models of mortality in acute pancreatitis in adults: a systematic review. *Ann Intern Med*. 2016;165(7):482–490. doi:10.7326/M16-0650

34. Moran RA, García-Rayado G, de la Iglesia-García D, et al. Influence of age, body mass index and comorbidity on major outcomes in acute pancreatitis, a prospective nation-wide multicentre study. *United European Gastroenterol J*. 2018;6(10):1508–1518. doi:10.1177/2050640618798155

35. Miyamoto K, Inai K, Takeuchi D, Shinohara T, Nakanishi T. Relationships among red cell distribution width, anemia, and interleukin-6 in adult congenital heart disease. *Circ J*. 2015;79(5):1100–1106. doi:10.1253/circj.CJ-14-1296

36. Pierce CN, Larson DF. Inflammatory cytokine inhibition of erythropoiesis in patients implanted with a mechanical circulatory assist device. *Perfusion*. 2005;20(2):83–90. doi:10.1191/0267659105pf793oa

37. Lorente L, Martín MM, Abreu-González P, et al. Red blood cell distribution width during the first week is associated with severity and mortality in septic patients. *PLoS One*. 2014;9(8):e105436. doi:10.1371/journal.pone.0105436

38. Ghaffari S. Oxidative stress in the regulation of normal and neoplastic hematopoiesis. *Antioxid Redox Signal*. 2008;10(11):1923–1940. doi:10.1089/ars.2008.2142

39. Zhang FX, Li ZL, Zhang ZD, Ma XC. Prognostic value of red blood cell distribution width for severe acute pancreatitis. *World J Gastroenterol*. 2019;25(32):4739–4748. doi:10.3748/wjg.v25.i32.4739

40. Yao J, Lv G. Association between red cell distribution width and acute pancreatitis: a cross-sectional study. *BMJ Open*. 2014;4(8):e004721. doi:10.1136/bmjopen-2013-004721

41. Şenol K, Saylam B, Kocaay F, Tez M. Red cell distribution width as a predictor of mortality in acute pancreatitis. *Am J Emerg Med*. 2013;31 (4):687–689. doi:10.1016/j.ajem.2012.12.015

42. Wu BU, Johannes RS, Sun X, Conwell DL, Banks PA. Early changes in blood urea nitrogen predict mortality in acute pancreatitis. *Gastroenterology*. 2009;137(1):129–135. doi:10.1053/j.gastro.2009.03.056

43. Ocskay K, Vinkó Z, Németh D, et al. Hypoalbuminemia affects one third of acute pancreatitis patients and is independently associated with severity and mortality. *Sci Rep*. 2021;11(1):24158. doi:10.1038/s41598-021-03449-8

44. Ni T, Wen Y, Wang Y, et al. Association between albumin or prealbumin levels at different stages and prognosis in severe acute pancreatitis: a 5-year retrospective study. *Sci Rep*. 2022;12(1):16792. doi:10.1038/s41598-022-21278-1

45. Jiang X, Su Z, Wang Y, et al. Prognostic nomogram for acute pancreatitis patients: an analysis of publicly electronic healthcare records in intensive care unit. *J Crit Care*. 2019;50:213–220. doi:10.1016/j.jcrc.2018.10.030