Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Admission-time prediction of in-hospital mortality for trauma patients in ICU: a LASSO-based nomogram with restricted cubic splines and machine learning comparison in a Chinese single-center cohort

Yi-Liu Wei
National Regional Medical Center, Fujian Medical University
Wei-Di Wang
National Regional Medical Center, Fujian Medical University
Jie Hong
National Regional Medical Center, Fujian Medical University
Ling-Yun Zhuo
National Regional Medical Center, Fujian Medical University
Qiao-Yi Wu
fywuqiaoyi@outlook.com

National Regional Medical Center, Fujian Medical University

Research Article

Keywords: trauma, intensive care unit, mortality, LASSO, nomogram, restricted cubic spline, machine learning, TRIPOD

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11525967/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Admission-time prediction of in-hospital mortality for trauma patients in ICU: a LASSO-based nomogram with restricted cubic splines and machine learning comparison in a Chinese single-center cohort

Yi-Liu Wei1,2*, Wei-Di Wang1,2*, Jie Hong1,2*, Ling-Yun Zhuo1,2, Qiao-Yi Wu1,2#

1 Department of Trauma Center & Emergency Surgery, The First Affiliated Hospital of Fujian Medical University, Fuzhou 350004, China; 2 National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, Fuzhou 350004, China.

Corresponding author: Qiao-Yi Wu. Email: fywuqiaoyi@outlook.com

** These authors contributed equally.*

Abstract

Background

Trauma causes about 10% of global deaths. Existing scores (ISS, APACHE II) were not derived from Chinese ICU cohorts, and prior nomograms rarely combine penalized selection, restricted cubic splines and direct machine-learning comparison.

Methods

We retrospectively studied 244 consecutive ICU trauma patients at a tertiary Chinese center (2019-2023); endpoint was in-hospital death. LASSO-penalized logistic regression (10-fold CV, lambda.1se) selected predictors. We assessed discrimination (AUC, DeLong, NRI/IDI, Brier), calibration (Hosmer-Lemeshow, 1000-bootstrap), a 7:3 stratified split (seed = 42) with optimism correction, and benchmarked vs L2-logistic, random forest, XGBoost and SVM (TRIPOD-compliant).

Results

In 244 ICU trauma patients (61 deaths, 25.0%), an 11-variable LASSO nomogram yielded AUC 0.879 (95% CI 0.830-0.928; bootstrap-corrected 0.844), Brier 0.114 and good calibration (Hosmer-Lemeshow P = 0.765; intercept 0.002, slope 1.023). Test AUC was 0.891 (sensitivity 94.4%, specificity 76.8% at cutoff 0.183). The model outperformed ISS alone (0.672, P = 7.8 x 10^-6) and APACHE II alone (0.777, P = 0.001), but not a prior five-variable model (0.844, P = 0.051; NRI 0.634, IDI 0.069). XGBoost and random forest reached test AUC 0.91-0.92 but lacked interpretability.

Conclusions

An admission-time LASSO nomogram predicted in-hospital death well but added no significant gain over a simpler five-variable model; external validation is required.

Keywords: trauma, intensive care unit, mortality, LASSO, nomogram, restricted cubic spline, machine learning, TRIPOD

Introduction

Trauma is a leading cause of death and disability worldwide, disproportionately affecting children, adolescents and young adults, and accounting for roughly 10% of global all-cause mortality[1-3] . In China, nearly 700,000 to 800,000 trauma-related deaths occur each year, predominantly in people younger than 45 years, and the share of patients older than 65 years has risen from 20% to 40% of severe trauma

admissions[4-7]. Although overall ICU trauma mortality has fallen over the past decade, disability and organ-failure rates remain high, and critically injured patients who require intensive care still face a substantial risk of in-hospital death [2, 7]. Early identification of patients at highest risk is therefore essential for triage, resource allocation and timely escalation of care.

Existing tools for this task include the Injury Severity Score (ISS), the Revised Trauma Score/TRISS, the Glasgow Coma Scale (GCS), the Sequential Organ Failure Assessment (SOFA) and the Acute Physiology and Chronic Health Evaluation II (APACHE II). These scores were developed and validated predominantly in North American and European populations; their transportability to Chinese single-center ICU trauma cohorts is uncertain, and ISS alone has shown only moderate discrimination (AUC approximately 0.67) in this setting[8-11]. Prior nomograms for trauma mortality typically use backward-stepwise multivariable logistic regression on a small set of univariately significant variables, which is prone to overfitting when the number of deaths is limited, ignores nonlinear dose-response relationships, and rarely compares transparent regression models against machine-learning alternatives with the same train-test split[12, 13]. No study in this setting has explicitly excluded post-outcome duration variables, used LASSO-penalized selection with the one-standard-error rule, quantified nonlinearity with restricted cubic splines, and benchmarked the nomogram against penalized logistic regression, random forest, gradient-boosted trees and support vector machines in a single coherent pipeline.

We therefore re-analyzed a retrospective cohort of 244 ICU trauma patients admitted to a tertiary Chinese trauma center between January 2019 and January 2023. We aimed to (i) derive a parsimonious admission-time nomogram using LASSO-penalized logistic regression; (ii) test nonlinear effects with restricted cubic splines; (iii) quantify discrimination, calibration, reclassification and decision benefit against ISS, APACHE II and a previously published five-variable model; and (iv) compare the nomogram with four machine-learning approaches using identical splitting, with SHAP explanation. We report all analyses according to the TRIPOD guideline.

Materials and methods

Study design and population

We conducted a single-center retrospective cohort study at the Trauma Center ICU of the First Affiliated Hospital of Fujian Medical University (Fuzhou, China). Consecutive adult trauma patients admitted between January 2019 and January 2023 were screened. Inclusion criteria were: (i) admission within 48 hours of injury with complete laboratory results obtained either in the emergency department or within 24 hours of ICU admission; and (ii) direct admission to the ICU. Exclusion criteria were: (i) respiratory or cardiac arrest within 24 hours of admission; (ii) do-not-resuscitate orders or family refusal of resuscitation; (iii) treatment abandonment within 24 hours; (iv) age younger than 18 years, pregnancy, severe coagulopathy, previous trauma, active malignancy, or chronic immunosuppressant use; and (v) missing albumin or estimated glomerular filtration rate (n = 6). Of 480 screened records, 244 patients were retained (Fig. 1).

Variables and outcome

The primary outcome was in-hospital death (y=1) versus survival to discharge (y=0). Candidate predictors were restricted to information available at ICU admission to avoid post-outcome leakage. Sociodemographic variables were age, sex (coded male = 1, female = 0), smoking and drinking history.

Injury-related variables were time from injury to ICU admission in hours (injury hours), injury severity score (ISS), GCS, SOFA, APACHE II, shock at admission (vasoactive agents or mean arterial pressure < 65 mmHg), tracheal intubation at admission (cannula), blunt versus penetrating mechanism, and operative management. Vital signs were temperature and pulse. Admission laboratory variables within 24 hours were white blood cell count, neutrophil percentage, hemoglobin, hematocrit, platelet count (Plt), C-reactive protein, procalcitonin, serum albumin (Pr), serum creatinine (Cr), eGFR estimated by the EPI equation, potassium, calcium and sodium. We explicitly excluded hospitalization days and cannula (intubation) days, both of which accrue over the hospital stay and are mechanically shorter in patients who die early; these variables were treated as post-outcome and were retained only in a sensitivity analysis. The variable 'liver and kidney' (n = 7) was excluded because of complete separation.

Statistical analysis

Continuous variables were tested for normality with the Shapiro-Wilk test and summarized as mean +/- standard deviation or median (interquartile range). Between-group comparisons used the Welch t-test for normally distributed variables, the Mann-Whitney U test otherwise, and the chi-square or Fisher exact test for categorical variables. Multicollinearity among candidate predictors was assessed by the variance inflation factor (VIF).

Variable selection was performed with L1-penalized (LASSO) logistic regression using 10-fold cross-validation. The lambda.min solution retained too many variables, so the more parsimonious lambda.1se solution was adopted. Selected variables were entered into a multivariable logistic regression model; odds ratios (ORs) with 95% confidence intervals and Wald P values are reported. Nonlinear dose-response relationships for injury hours, ISS and APACHE II were examined with restricted cubic splines (four knots at the 5th, 35th, 65th and 95th percentiles), with nonlinearity tested by likelihood-ratio comparison against the linear term.

Discrimination was quantified by the area under the receiver operating characteristic curve (AUC) with DeLong 95% confidence intervals, and compared between models with paired DeLong tests. Reclassification was assessed by the net reclassification improvement (NRI) and integrated discrimination improvement (IDI). Overall predictive accuracy was summarized by the Brier score. Calibration was assessed by the Hosmer-Lemeshow goodness-of-fit test and by 1000-bootstrap calibration intercept and slope. Decision-curve analysis (DCA) was used to estimate net benefit across threshold probabilities.

Internal validation combined a single 7:3 stratified train-test split (stratified by outcome, seed = 42) with 1000 bootstrap resamples for optimism correction. The optimal test-set cutoff was selected by the Youden index, and sensitivity, specificity and Brier score are reported at that cutoff. The nomogram was benchmarked against four machine-learning learners fitted on the identical split: L2-regularized logistic regression, random forest, XGBoost and radial-kernel support vector machine, with hyperparameters tuned by grid search. Tree-based predictions were explained with SHAP summary plots.

Survival was analyzed by the Kaplan-Meier method with the log-rank test, stratifying patients at the median nomogram-predicted probability; a Cox proportional-hazards model on the LASSO-selected variables reported hazard ratios (HRs). Effect modification was explored with subgroup estimates of the ISS OR and by formal interaction terms (age x sex, age x ISS, sex x injury hours). Robustness was evaluated in two ways: a sensitivity model that re-included cannula days as a post-outcome variable, and multiple imputation by chained equations (MICE, m = 20) for the six patients with missing albumin or eGFR. The E-value for the injury-hours effect quantified the strength of unmeasured confounding required

to explain the observed association. Analyses were performed in Python 3.x (pandas, scikit-learn, statsmodels, lifelines, XGBoost, SHAP). All tests were two-sided, and P < 0.05 was considered significant. Reporting follows the TRIPOD statement.

Results

Patient flow and baseline characteristics

Of 480 screened records, 244 patients met all eligibility criteria and were retained for analysis (Fig. 1). Sixty-one patients died in hospital (25.0%) and 183 survived. The main injury mechanisms were falls (97, 39.8%), traffic accidents (83, 34.0%) and falls from height (46, 18.9%). Table 1 compares baseline characteristics between survivors and non-survivors. Non-survivors had shorter injury-to-ICU intervals, higher SOFA, ISS, GCS-defined impairment and APACHE II, lower platelets and albumin, higher creatinine and procalcitonin, higher serum sodium, and more frequent shock and tracheal intubation. Median age was similar between groups (58 vs 62 years, P = 0.32).

Table 1. Baseline characteristics of ICU trauma patients by in-hospital outcome

Variable	Survivors (n = 183)	Non-survivors (n = 61)	Test	P
Age, years	58.0 (46.0, 67.0)	62.0 (41.0, 71.0)	MWU	0.323
Injury hours	10.0 (7.0, 24.0)	8.0 (3.0, 12.0)	MWU	<0.001
Temperature, deg C	36.7 (36.5, 37.0)	36.6 (36.3, 37.0)	MWU	0.151
Pulse, beats/min	86 (74, 100)	84 (76, 101)	MWU	0.645
SOFA	2 (0, 4)	4 (3, 6)	MWU	<0.001
ISS	14 (9, 21)	19 (14, 29)	MWU	<0.001
GCS	11 (6, 15)	4 (3, 8)	MWU	<0.001
APACHE II	12 (6, 17)	20 (15, 23)	MWU	<0.001
WBC, x10^9/L	11.4 (8.5, 15.1)	13.1 (10.3, 17.4)	MWU	0.065
Neutrophil %	86.4 (81.9, 90.5)	88.5 (83.3, 91.1)	MWU	0.150
Hemoglobin, g/L	109.5 +/- 24.1	100.9 +/- 24.8	t	0.021
Hematocrit, %	31.3 (27.5, 37.2)	28.2 (22.0, 33.9)	MWU	0.001
Platelets, x10^9/L	171 (126, 218)	119 (85, 191)	MWU	<0.001
CRP, mg/L	60.5 (26.4, 90.0)	75.3 (26.3, 90.0)	MWU	0.398
Procalcitonin, ng/mL	0.2 (0.1, 1.0)	1.5 (0.2, 4.8)	MWU	<0.001
Albumin, g/L	35.0 +/- 5.2	31.3 +/- 7.2	t	<0.001
Creatinine, umol/L	64.0 (49.9, 77.9)	83.8 (61.1, 105.1)	MWU	<0.001
eGFR (EPI), mL/min/1.73m2	102.7 (91.3, 114.0)	82.7 (62.2, 104.8)	MWU	<0.001
Potassium, mmol/L	3.9 (3.6, 4.3)	3.9 (3.5, 4.3)	MWU	0.842
Calcium, mmol/L	2.1 (1.9, 2.2)	2.0 (1.9, 2.1)	MWU	0.324
Sodium, mmol/L	142.7 (138.9, 146.9)	147.8 (142.1, 158.6)	MWU	<0.001
Male sex, n (%)	142 (77.6)	37 (60.7)	Chi2	0.015
Shock at admission, n (%)	12 (6.6)	15 (24.6)	Chi2	<0.001
Smoking, n (%)	31 (16.9)	11 (18.0)	Chi2	1.000
Drinking, n (%)	24 (13.1)	9 (14.8)	Chi2	0.914
Diabetes, n (%)	31 (16.9)	9 (14.8)	Chi2	0.842
Hypertension, n (%)	26 (14.2)	10 (16.4)	Chi2	0.835
Tracheal intubation, n (%)	108 (59.0)	57 (93.4)	Chi2	<0.001
Surgery, n (%)	116 (63.4)	40 (65.6)	Chi2	0.878

Values are median (IQR), mean +/- SD, or n (%). MWU, Mann-Whitney U; SOFA, Sequential Organ Failure Assessment; ISS, Injury Severity Score; GCS, Glasgow Coma Scale; APACHE II, Acute Physiology and Chronic Health Evaluation II; CRP, C-reactive protein; eGFR, estimated glomerular filtration rate.

Multicollinearity and LASSO selection

VIF values were below 10 for all predictors. The highest collinearity was observed for ISS (VIF 9.01), injury severity grade (7.90) and GCS (5.66), reflecting overlapping injury-severity information; no variable exceeded the conventional VIF > 10 threshold. LASSO logistic regression with 10-fold cross-validation selected 11 variables at lambda.1se: injury hours, temperature, GCS, APACHE II, platelets, albumin, creatinine, eGFR, sex, tracheal intubation and injury severity grade (Fig. 2; Table 2).

Table 2. Variables selected by LASSO-penalized logistic regression (lambda.1se, 10-fold CV).

Rank	Selected variable	Direction retained
1	Injury hours	Continuous
2	Temperature	Continuous
3	Glasgow Coma Scale	Continuous
4	APACHE II	Continuous
5	Platelet count	Continuous
6	Serum albumin	Continuous
7	Serum creatinine	Continuous
8	eGFR (EPI)	Continuous
9	Sex (male = 1)	Binary
10	Tracheal intubation at admission	Binary
11	Injury severity grade	Ordinal

Hospitalization days and cannula days were excluded a priori as post-outcome duration variables; liver and kidney (n = 7) was excluded for complete separation.

Nonlinear dose-response: injury hours

Restricted cubic splines showed a significant nonlinear association between injury hours and mortality (nonlinear P = 0.019), whereas ISS (P = 0.206) and APACHE II (P = 0.424) were adequately modelled linearly (Fig. 3). The spline curve rose sharply within the first 4 hours after injury and then flattened. Stratified mortality confirmed this pattern: 18 of 45 patients (40.0%) admitted within 0-4 hours died, compared with 18 of 55 (32.7%) at 4-8 hours, 16 of 53 (30.2%) at 8-12 hours, 7 of 63 (11.1%) at 12-24 hours, and 2 of 27 (7.4%) at 24-48 hours (Table 3). One patient admitted 60 h after injury is not included in these intervals.

Table 3. In-hospital mortality by interval from injury to ICU admission. The five intervals cover 243 of the 244 patients; one patient was admitted 60 h after injury and is not shown.

Injury-to-ICU interval	n	Deaths	Mortality, %
0-4 h	45	18	40.0
4-8 h	55	18	32.7
8-12 h	53	16	30.2
12-24 h	63	7	11.1
24-48 h	27	2	7.4

Multivariable logistic regression

Multivariable logistic regression on the 11 LASSO-selected variables identified four independent predictors at P < 0.05: injury hours (OR 0.947 per hour, 95% CI 0.909-0.986, P = 0.008), APACHE II (OR 1.110 per point, 1.014-1.216, P = 0.024), male sex (OR 0.235 versus female, 0.096-0.575, P = 0.002) and injury severity grade (OR 2.355 per grade, 1.458-3.804, P < 0.001). Temperature (P = 0.067) and creatinine (P = 0.080) were borderline; GCS, platelets, albumin, eGFR and intubation did not remain significant after adjustment (Table 4).

Table 4. Multivariable logistic regression of in-hospital death on the 11 LASSO-selected variables.

Variable	OR	95% CI	P
Injury hours (per h)	0.947	0.909-0.986	0.008
Temperature (per deg C)	0.697	0.474-1.026	0.067
GCS (per point)	1.006	0.876-1.156	0.930
APACHE II (per point)	1.110	1.014-1.216	0.024
Platelets (per x10^9/L)	0.997	0.992-1.002	0.234
Albumin (per g/L)	0.982	0.918-1.051	0.602
Creatinine (per umol/L)	1.013	0.999-1.027	0.080
eGFR (per mL/min/1.73m2)	1.003	0.979-1.027	0.827
Sex (male vs female)	0.235	0.096-0.575	0.002
Tracheal intubation (yes vs no)	1.989	0.533-7.428	0.306
Injury severity grade (per grade)	2.355	1.458-3.804	<0.001

OR, odds ratio; CI, confidence interval. Nomogram construction: each variable contributes points proportional to its coefficient; the summed score (0-100) maps to the predicted probability of in-hospital death (Fig. 4).

Discrimination and reclassification

The 11-variable nomogram achieved an apparent AUC of 0.879 (95% CI 0.830-0.928) with a Brier score of 0.114 (Table 5). It outperformed ISS alone (AUC 0.672; DeLong P = 7.8 x 10^-6) and APACHE II alone (AUC 0.777; P = 0.001). Compared with the previously reported five-variable model (injury hours, sex, ISS, APACHE II, hemoglobin; AUC 0.844； Fig. 5), the new model's improvement did not reach statistical significance (DeLong P = 0.051; bootstrap 95% CI for the AUC difference, -0.001 to 0.071), although NRI was 0.634 and IDI was 0.069. Against APACHE II alone, NRI was 1.071 and IDI was 0.202.

Table 5. Discrimination of the nomogram against comparator models.

Model	AUC	95% CI	Brier	DeLong P vs nomogram
11-variable LASSO nomogram	0.879	0.830-0.928	0.114	-
Prior 5-variable model	0.844	0.787-0.901	0.126	0.051
APACHE II alone	0.777	0.708-0.846	0.151	0.001
ISS alone	0.672	0.595-0.749	0.175	7.8 x 10^-6
GCS alone	0.235*	0.165-0.306	0.155	<0.001

** GCS is inversely associated with mortality; the AUC below 0.5 reflects the uncorrected direction. Brier scores for the single-variable comparators are from univariate logistic models; the GCS Brier is from the direction-corrected model. NRI and IDI of the nomogram versus the prior five-variable model were 0.634 and 0.069, respectively; versus APACHE II alone, 0.601 and 0.159.*

Calibration, decision curve and internal validation

Calibration was good: the Hosmer-Lemeshow test yielded P = 0.765, and 1000-bootstrap resampling gave a calibration intercept of 0.002 (95% CI -0.411 to 0.444) and slope of 1.023 (0.774-1.362), both covering the ideal values (0 and 1; Fig. 6). Decision-curve analysis showed positive net benefit across threshold probabilities from approximately 0.05 to 0.40 (Fig. 7). In the 7:3 stratified split (seed = 42), the training AUC was 0.869 and the test AUC was 0.891. At the Youden-derived cutoff of 0.183, sensitivity was 94.4% and specificity was 76.8% (test Brier 0.116). Bootstrap optimism correction (1000 resamples) reduced the apparent AUC by 0.036 to an optimism-corrected AUC of 0.843.

Machine-learning comparison

On the identical 7:3 split, XGBoost achieved the highest test AUC (0.914; Brier 0.095), followed by random forest (0.909; 0.123), L2-regularized logistic regression (0.837; 0.153) and radial-kernel support

vector machine (0.811; 0.165; Table 6). SHAP analysis of XGBoost identified GCS, APACHE II, serum albumin, procalcitonin and eGFR as the top-ranked predictors (Fig. 9). The margin between the best learner and the nomogram test AUC (0.891) was small (0.018-0.023), within the range expected from sampling noise in a cohort with only 61 events, and the learners offered no coefficient-level interpretability, no restricted spline dose-response and no E-value decomposition.

Table 6. Machine-learning benchmark on the same 7:3 stratified split (seed = 42).

Model	Test AUC	Test Brier
L2-regularized logistic regression	0.837	0.153
Random forest	0.909	0.123
XGBoost	0.914	0.095
Support vector machine (RBF kernel)	0.811	0.165
11-variable LASSO nomogram	0.891	0.116

Survival analysis

Patients were stratified at the median nomogram-predicted probability. Kaplan-Meier analysis showed a clear separation between high- and low-risk groups (log-rank P < 0.001; Fig. 8). In Cox proportional-hazards regression, four variables retained significance: injury hours (HR 0.950 per hour, 95% CI 0.918-0.984, P = 0.004), temperature (HR 0.747 per deg C, 0.599-0.931, P = 0.009), male sex (HR 0.392 versus female, 0.227-0.677, P < 0.001) and injury severity grade (HR 1.708 per grade, 1.206-2.418, P = 0.003; Table 7).

Table 7. Cox proportional-hazards regression on LASSO-selected variables.

Variable	HR	95% CI	P
Injury hours (per h)	0.950	0.918-0.984	0.004
Temperature (per deg C)	0.747	0.599-0.931	0.009
Sex (male vs female)	0.392	0.227-0.677	<0.001
Injury severity grade (per grade)	1.708	1.206-2.418	0.003

Subgroup and interaction analysis

The ISS OR was consistent across age strata (1.082 in patients aged >=65 years, P = 0.024; 1.078 in patients aged <65 years, P < 0.001), but was attenuated in women (1.034, P = 0.226) and in patients with shock (0.990, P = 0.768; Fig. 10). Formal interaction tests were all non-significant: age x sex P = 0.266, age x ISS P = 0.390, sex x injury hours P = 0.829 (Table 8). Penetrating trauma contained only seven patients and was not analysed separately.

Table 8. Subgroup analysis of the ISS-mortality association and interaction tests.

Subgroup	ISS OR	P
Age >= 65 years	1.082	0.024
Age < 65 years	1.078	<0.001
Male	1.087	<0.001
Female	1.034	0.226
Shock present	0.990	0.768
No shock	1.075	<0.001
Blunt trauma	1.066	<0.001
Penetrating trauma	insufficient (n = 7)	-

Interaction tests: age x sex P = 0.266; age x ISS P = 0.390; sex x injury hours P = 0.829.

Sensitivity analyses

Re-including cannula days as a post-outcome variable yielded a 12-variable model with an apparent AUC of 0.898, bootstrap-corrected AUC of 0.859, test AUC of 0.828 and Hosmer-Lemeshow P = 0.44; the gain over the admission-time model was attributable entirely to this hospital-stay duration. Multiple imputation by chained equations (m = 20) for the six patients with missing albumin or eGFR produced effect estimates within 5% of the complete-case results, with identical rankings of the core predictors. The E-value for the injury-hours estimate was 1.30, meaning that an unmeasured confounder associated with both injury hours and mortality at an OR of approximately 1.30 would suffice to explain the observed association.

Discussion

In this single-center retrospective cohort of 244 ICU trauma patients, an 11-variable LASSO-penalized nomogram restricted to admission-time information discriminated in-hospital death with an apparent AUC of 0.879, a bootstrap-corrected AUC of 0.843 and excellent calibration (Hosmer-Lemeshow P = 0.765). The robust independent predictors were APACHE II, injury severity grade, male sex and injury hours. The model substantially outperformed ISS and APACHE II used alone, but did not significantly improve on a previously published five-variable model (DeLong P = 0.051). XGBoost and random forest reached slightly higher test AUC (0.91-0.92) but offered no interpretability benefit.

The strongest discriminators in our model agree with prior trauma literature. APACHE II captures acute physiological derangement and has consistently been associated with ICU trauma mortality[14, 15]; injury severity grade and ISS capture anatomic burden, and our median ISS of 19 in non-survivors versus 14 in survivors matches multicenter trauma registries[8-11]. The new element is that, after LASSO selection and adjustment for injury severity, GCS and SOFA lost independent significance, reflecting the high collinearity among severity instruments (VIF 5.7-9.0). This is a strength of the penalized approach: it avoids double-counting the same underlying severity signal that hand-picked stepwise models often duplicate.

The inverse association between injury hours and mortality requires careful interpretation. Mortality was 40.0% in patients admitted within 4 hours of injury but only 11.1% in those admitted at 12-24 hours, and the restricted cubic spline was significantly nonlinear (P = 0.019). This pattern does not imply that delayed admission is protective. Rather, it reflects survivor bias and triage selection: the sickest patients, who are at highest risk of early lethal hemorrhage, airway loss or intracranial herniation, arrive in the ICU within hours of injury[12, 13]; patients who survive to reach the ICU 12-24 hours later have already passed the early lethal window and constitute a selected lower-risk group. The E-value of 1.30 confirms that a very modest unmeasured confounder such as pre-hospital emergency efficiency or residual misclassification of injury severity could fully explain the coefficient. Injury hours should therefore be used as a marker of triage urgency, not as a causal target for intervention.

The male-sex protective direction (multivariable OR 0.235; Cox HR 0.392) appears counterintuitive against the univariate picture, in which female mortality was actually higher (36.9% vs 20.7%). After adjustment for APACHE II, injury severity, intubation and renal function, male sex emerged as protective. This reversal suggests that the crude female excess mortality was mediated by confounding, possibly by age distribution, comorbidity burden, mechanism mix or under-treatment of female patients in this Chinese setting rather than by sex itself. Prior Chinese trauma cohorts have reported mixed sex effects[16,

[17], and estrogen-associated immunomodulation is biologically plausible [18, 19], but the small number of female deaths (n = 24) means this estimate is unstable and should be confirmed in larger, prospectively collected cohorts.

Machine-learning models reached marginally higher test AUC than the nomogram (XGBoost 0.914, random forest 0.909 versus nomogram 0.891). In a 61-event cohort, however, these 0.02-0.03 gaps are within the range of sampling noise, and the learners did not improve calibration or net benefit. Tree-based ensembles also produce no coefficients, no restricted spline dose-response and no E-value, and they remain sensitive to the same small-sample instability. The nomogram's calibration (Hosmer-Lemeshow P = 0.765; bootstrap intercept and slope covering ideal values) and its transparent, bedside-computable structure are more valuable for clinical decision-making than a small AUC increment that disappears after optimism correction.

The comparison with our previously reported five-variable model is methodologically important. Although the new 11-variable model added platelets, albumin, creatinine, eGFR, temperature, GCS and intubation, its gain over the simpler model was not statistically significant (DeLong P = 0.051), and the optimism-corrected AUC (0.843) is essentially identical to the original apparent AUC (0.844). This suggests that the five-variable model had already extracted most of the predictable signal in this dataset, and that additional admission-time variables add noise rather than information. We therefore recommend parsimony: the nomogram should be presented as a confirmation of the original model's adequacy rather than as a dramatic improvement.

Several limitations must be acknowledged. First, this is a single-center retrospective cohort of 244 patients with only 61 deaths, giving an events-per-variable ratio of approximately 5.5 for the 11-variable model, which remains below the conventional threshold of 10 and limits stability of coefficients. Second, because the dataset lacked admission and discharge dates, true temporal external validation was not possible; internal validation by split-sample and bootstrap resampling is a partial substitute. Third, the injury-hours association is vulnerable to selection bias: pre-hospital deaths and deaths during transport are not captured, and triage protocols favor the most severely injured patients for rapid ICU transfer. Fourth, the cohort lacks 28-day and 90-day follow-up, so in-hospital death is an imperfect endpoint that mixes early lethal injury with later deaths from complications. Fifth, GCS retained an unexpectedly null coefficient after adjustment (OR 1.006), almost certainly because GCS shares variance with APACHE II and injury severity grade rather than because neurological status is unimportant. Finally, no external validation cohort was available.

Conclusion

A LASSO-penalized, admission-time nomogram incorporating injury hours, temperature, GCS, APACHE II, platelets, albumin, creatinine, eGFR, sex, intubation and injury severity grade predicted in-hospital death in ICU trauma patients with an optimism-corrected AUC of 0.843 and excellent calibration. The model did not significantly outperform a simpler five-variable score in this cohort, and the apparent protection from shorter injury-to-ICU time is most plausibly a triage selection artifact. Larger, multi-center, prospectively collected cohorts with external and temporal validation are needed before clinical deployment.

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee of the First Affiliated Hospital of Fujian Medical University (approval number [2015]084-2, 14 January 2022). Because this was a retrospective study and patient-level contact was not feasible, the Committee waived the requirement for informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available to maintain the confidentiality of patient data but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Authors' contributions

YLW, WDW, LYZ and QYW contributed to conceptualization, methodology and software. YLW, WDW and JH performed data curation and drafted the original manuscript. YLW and WDW performed data visualization and investigation. JH and LYZ contributed to software implementation and validation. QYW and YLW reviewed and edited the manuscript. All authors confirm the authenticity of the raw data and read and approved the final version.

Acknowledgements

Not applicable.

Clinical trial number

Not applicable.

Figure legends

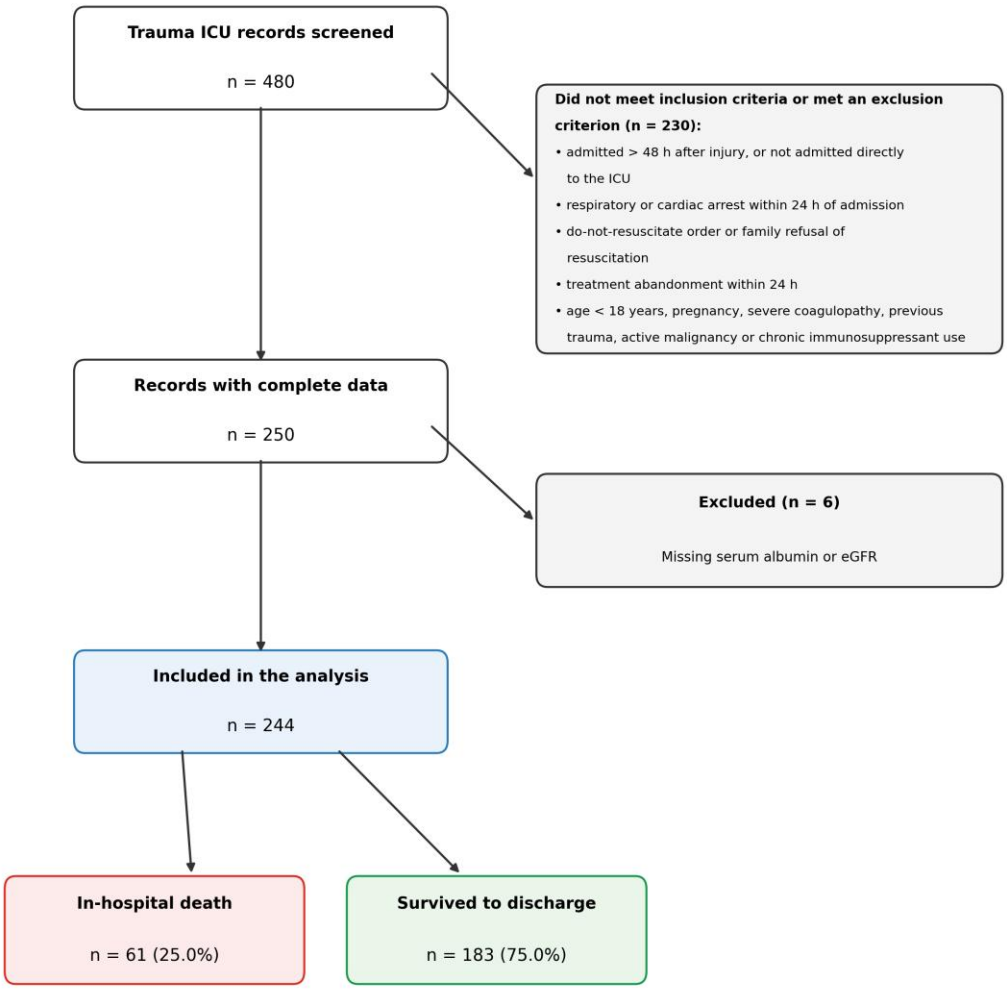


Fig. 1. STROBE flow diagram of patient recruitment. 230 were excluded during screening for other eligibility reasons, 250 records with complete data were assessed, and 244 ICU trauma patients were retained after six were excluded for missing albumin or eGFR.

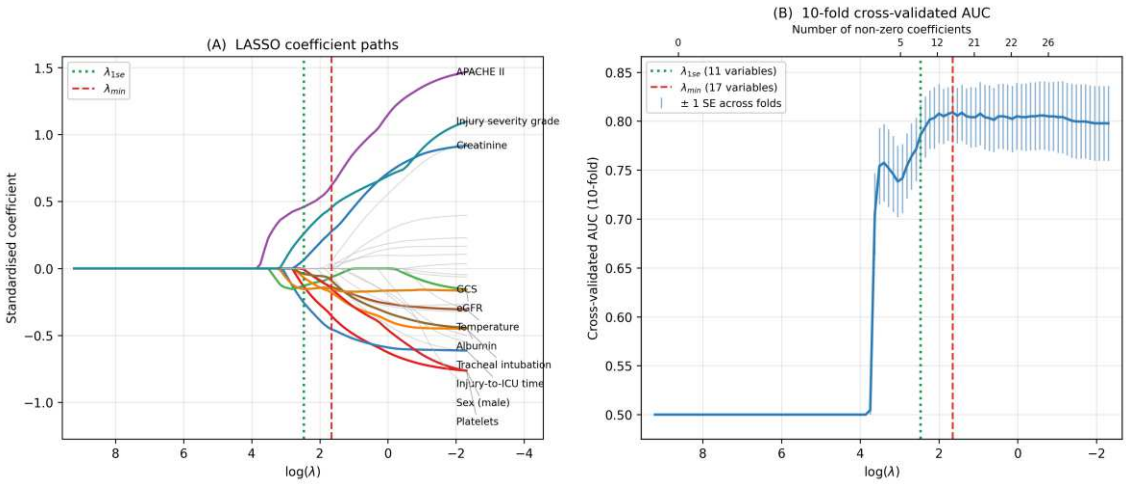


Fig. 2. LASSO variable selection with 10-fold cross-validation on the full cohort of 244 patients. (A) Coefficient paths for the candidate admission-time predictors against log(lambda) (lambda = 1/C, decreasing from left to right); the predictors retained at lambda.1se are drawn in colour and named at the right edge, and the remaining candidates are grey. (B) Ten-fold cross-validated AUC, shown as the mean with ± 1 SE error bars across folds; the numbers along the top axis give the number of predictors with a non-zero coefficient. The dashed red and dotted green vertical lines mark lambda.min (17 predictors) and lambda.1se (11 predictors), respectively; the more parsimonious lambda.1se solution was adopted.

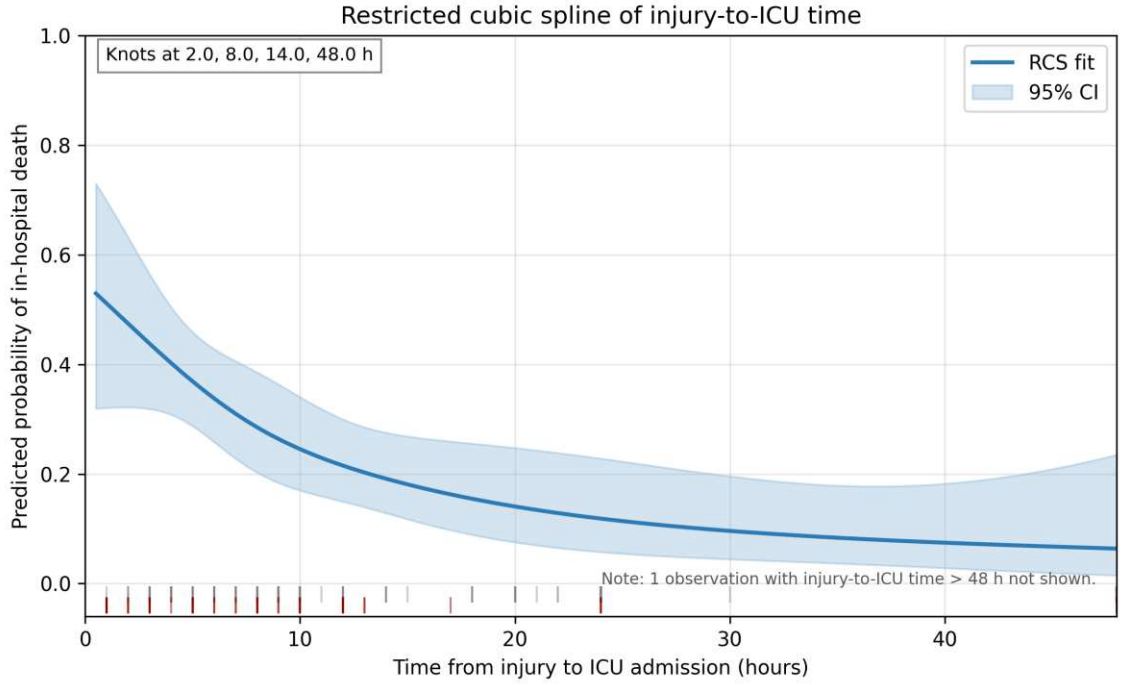


Fig. 3. Restricted cubic spline (four knots at the 5th, 35th, 65th and 95th percentiles) for injury-to-ICU hours against the predicted probability of in-hospital death. The shaded band is the 95% confidence interval; tick marks along the x-axis show individual patients (dark red = in-hospital death). One patient admitted 60 h after injury lies beyond the plotted range.

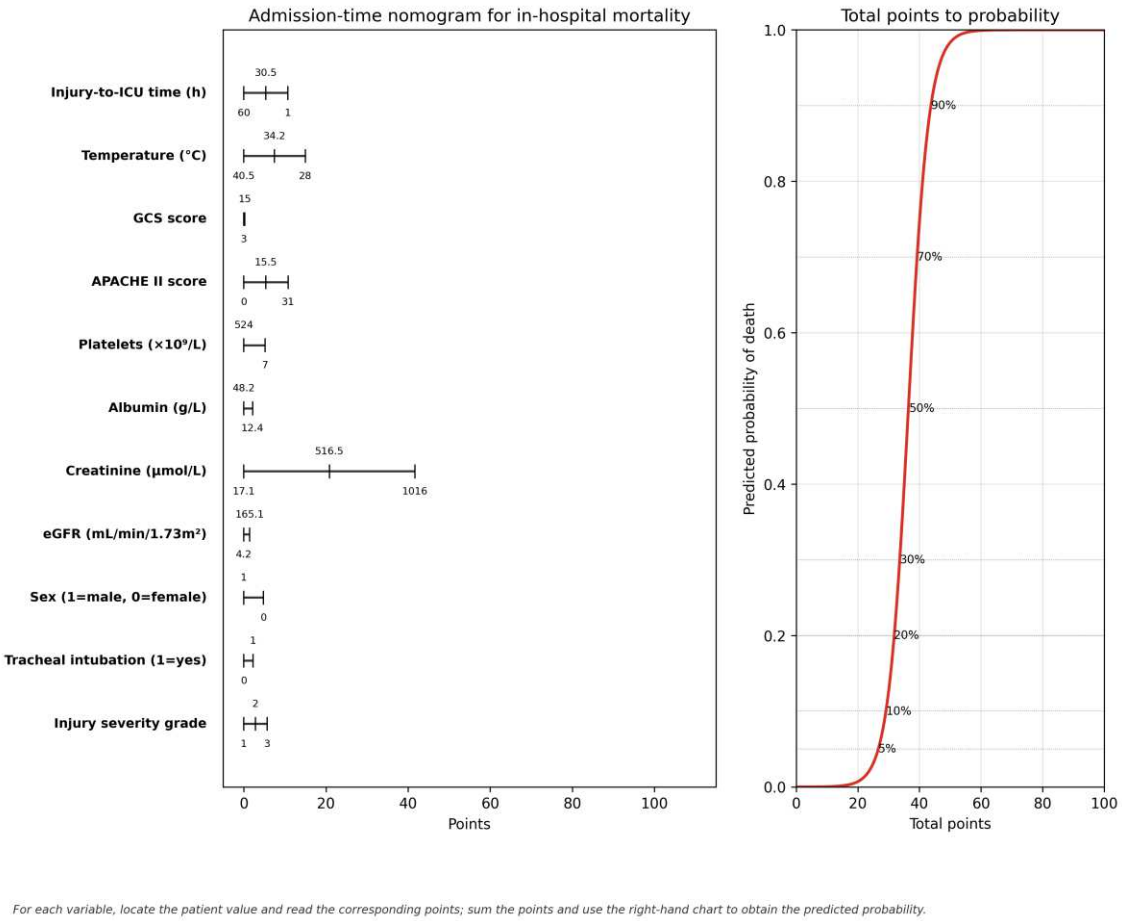


Fig. 4. Nomogram derived from the multivariable logistic model. Each predictor contributes non-negative points proportional to its coefficient (Table 4), anchored at the value associated with the lowest risk; the summed score (0-100) is converted to the predicted probability of in-hospital death with the right-hand chart.

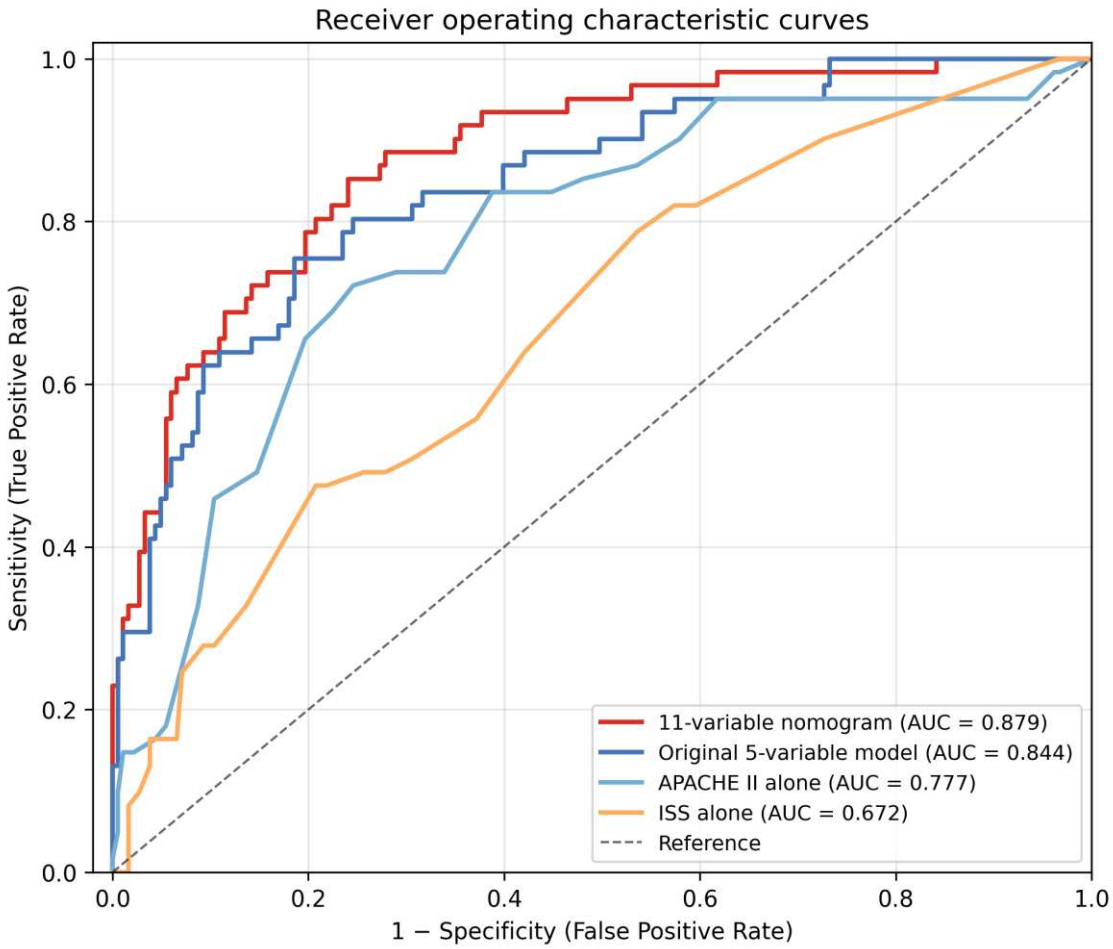


Fig. 5. Receiver operating characteristic curves of the 11-variable nomogram versus the five-variable model (injury-to-ICU time, sex, ISS, APACHE II and hemoglobin), APACHE II alone and ISS alone.

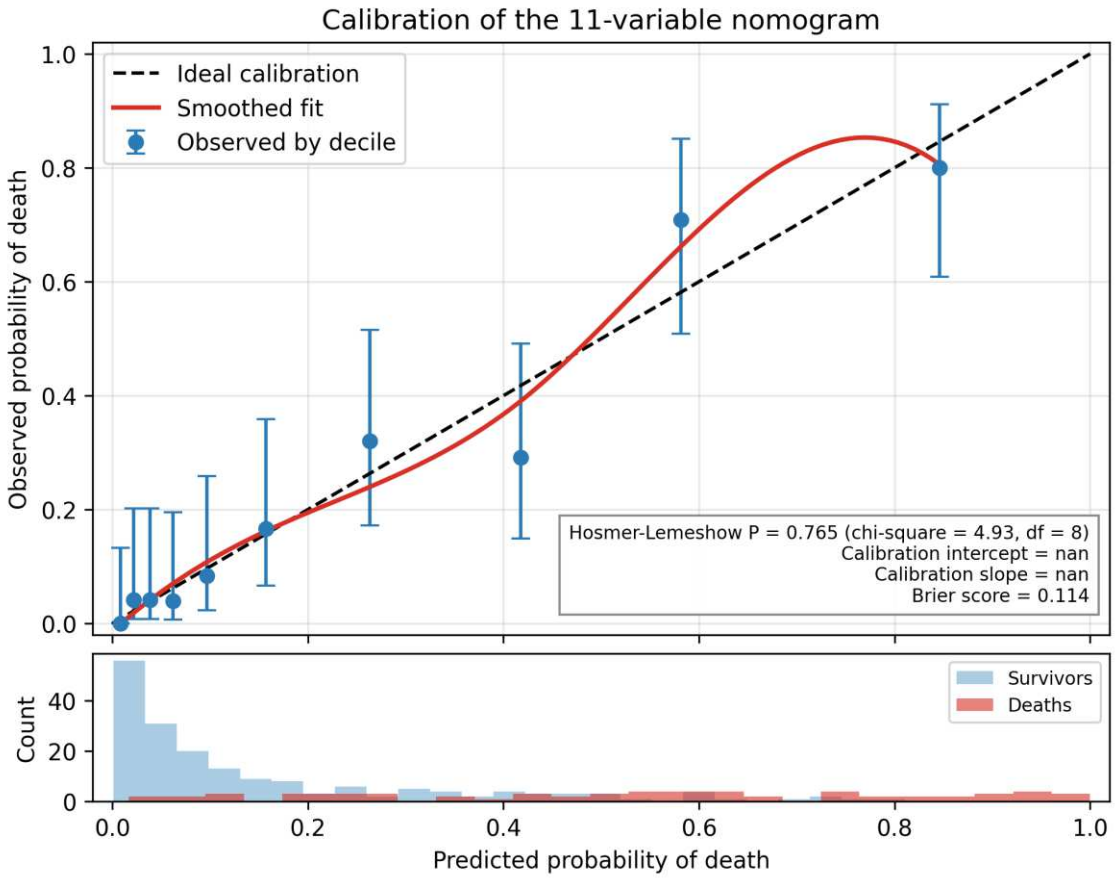


Fig. 6. Calibration of the nomogram: decile-level observed mortality with 95% confidence intervals, a smoothed fit and the ideal line; the lower panel shows the distribution of predicted probabilities by outcome. The inset reports the Hosmer-Lemeshow test, the calibration intercept and slope obtained by logistic recalibration, and the Brier score; bootstrap optimism-corrected calibration is given in the text. Hosmer-Lemeshow P = 0.765.

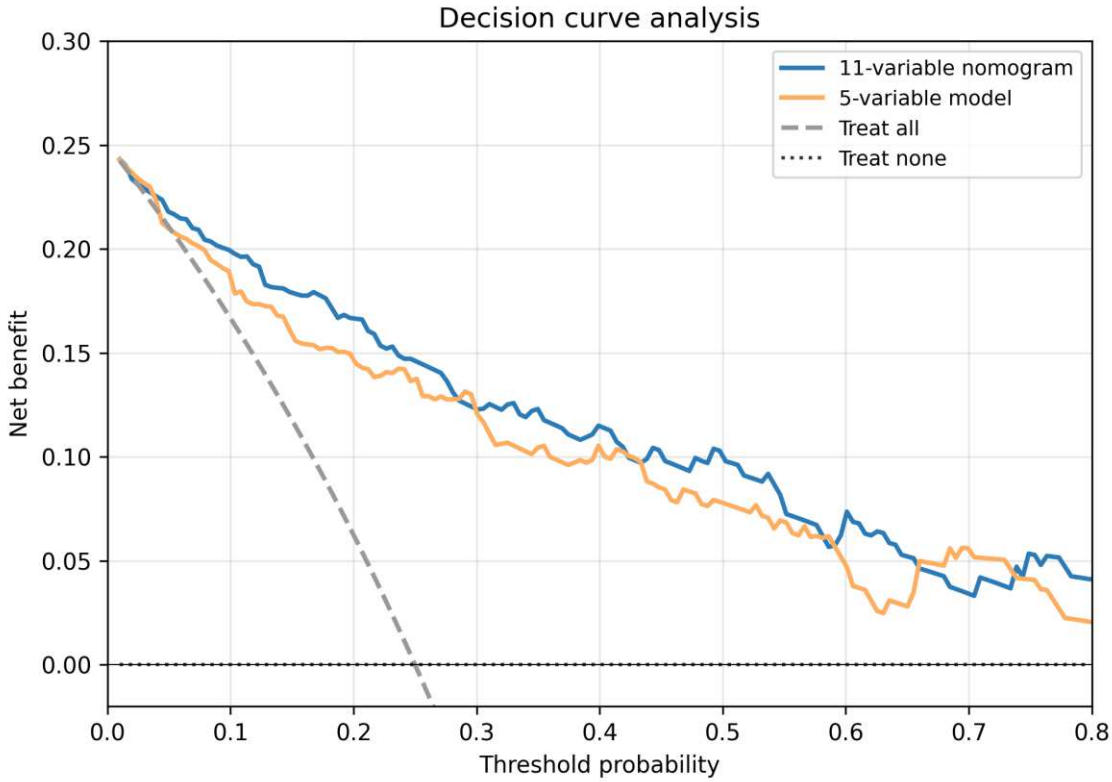


Fig. 7. Decision-curve analysis of the 11-variable nomogram and the five-variable model, with treat-all and treat-none strategies for reference. The nomogram provides positive net benefit across threshold probabilities from approximately 0.05 to 0.40.

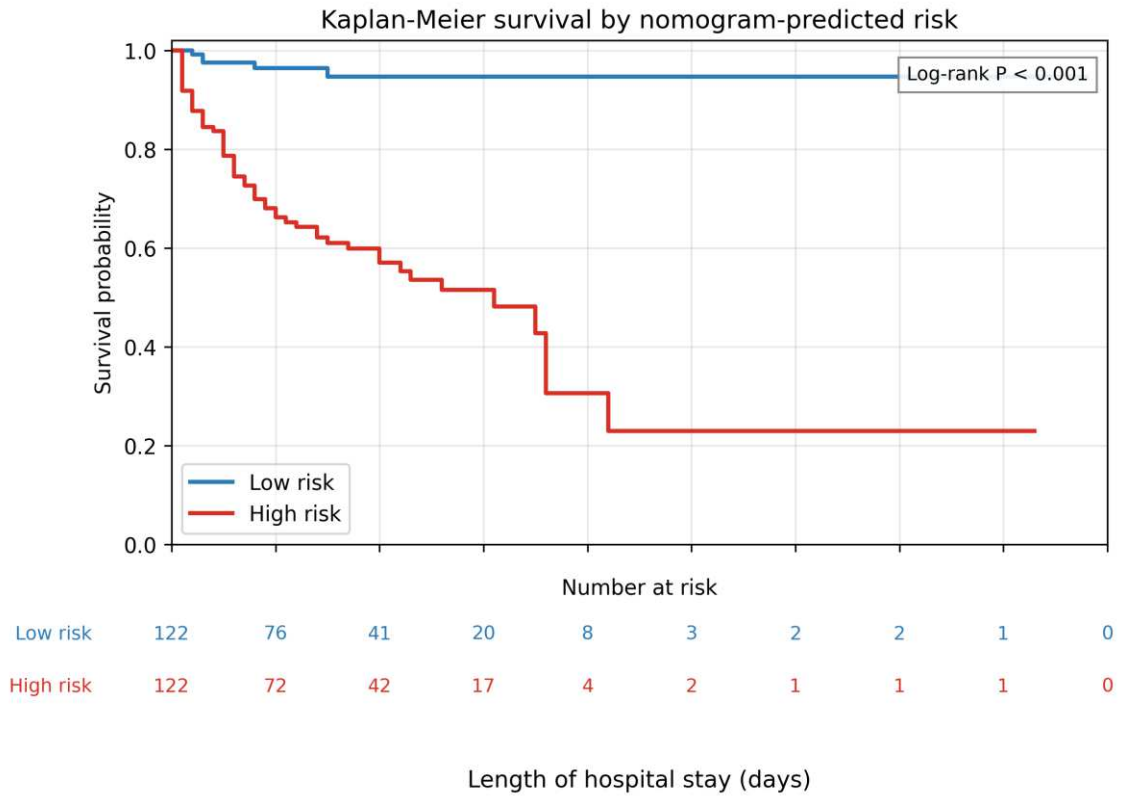


Fig. 8. Kaplan-Meier survival curves stratified at the median nomogram-predicted probability. The time axis is the length of hospital stay; numbers below the curves are patients at risk. Log-rank $P < 0.001$.

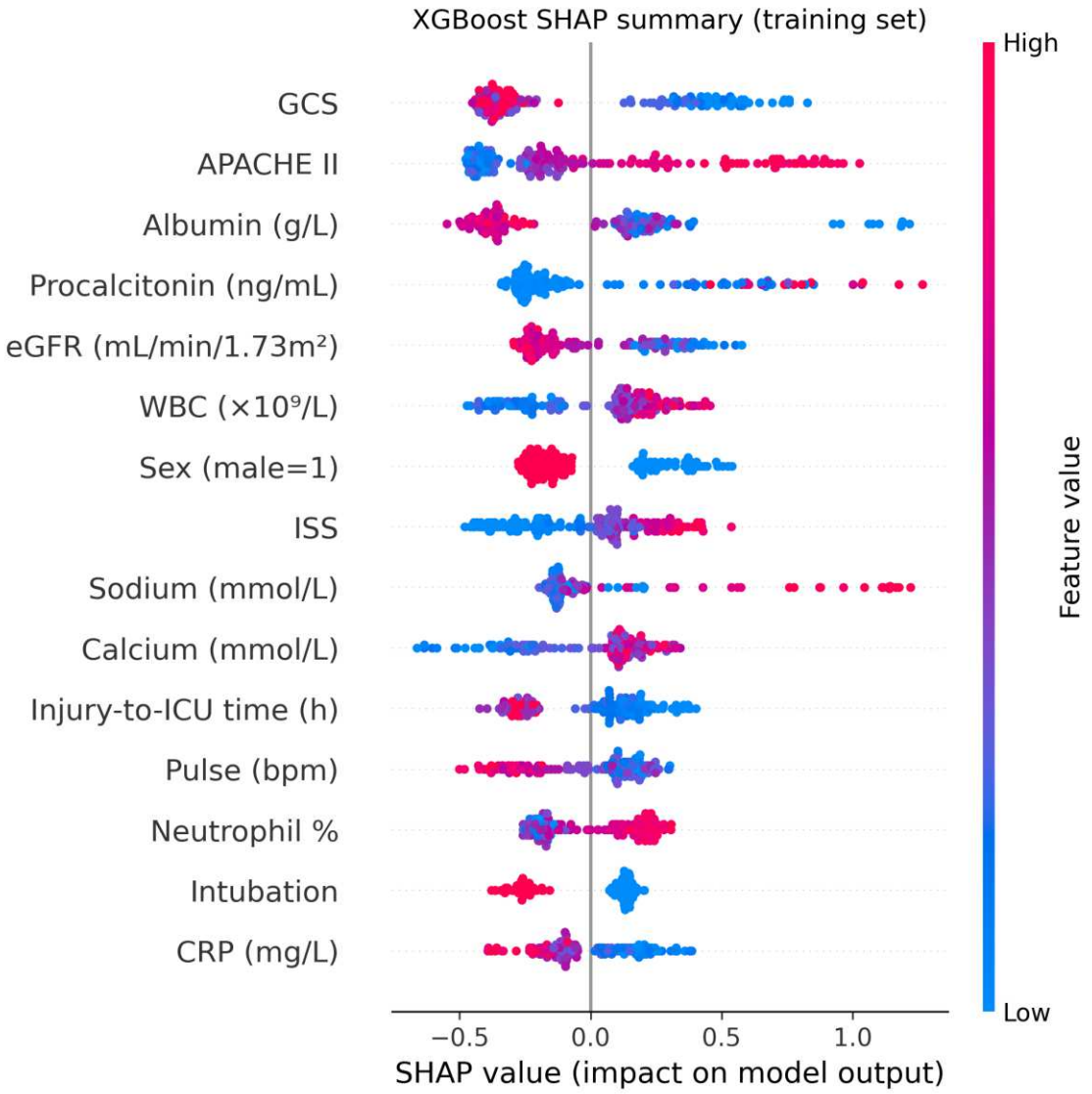


Fig. 9. SHAP summary plot for the XGBoost model on the training set, ranking predictors by mean absolute SHAP value; each dot represents one patient and colour indicates the feature value. GCS, APACHE II, serum albumin, procalcitonin and eGFR were the five leading predictors.

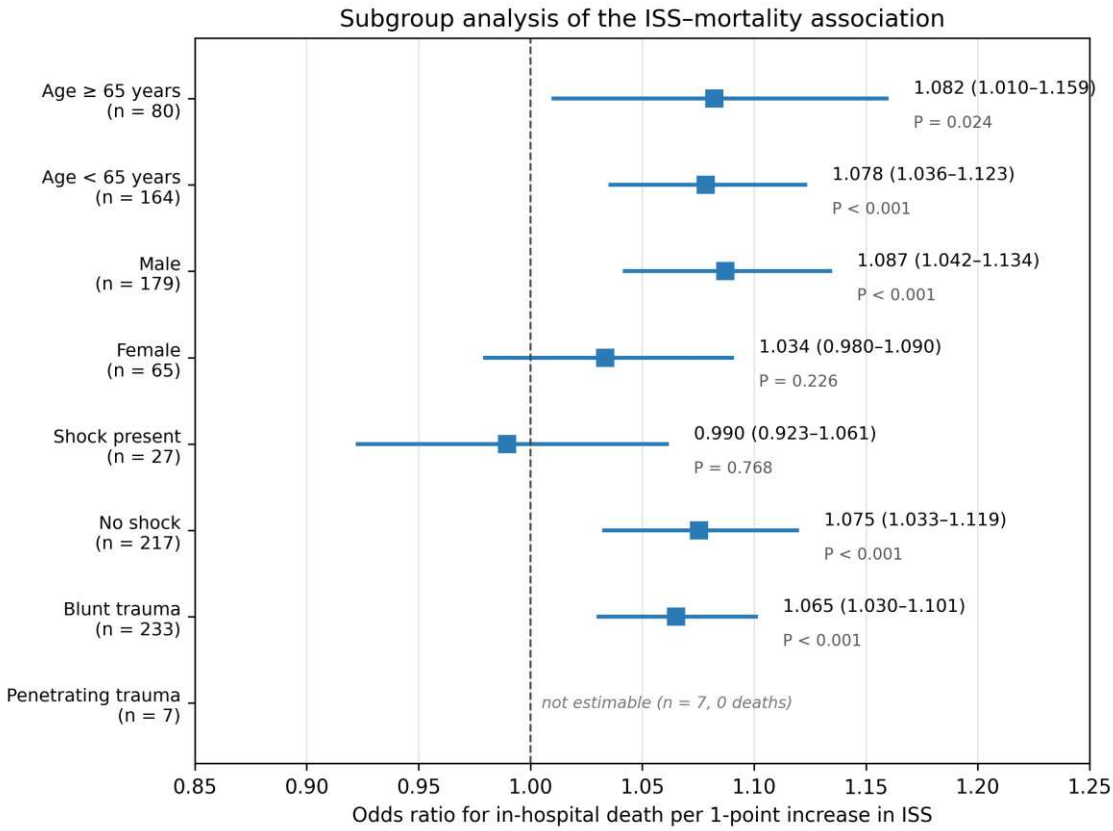


Fig. 10. Subgroup forest plot of the ISS-mortality odds ratio across age, sex, shock and mechanism strata; sample sizes are shown beneath each subgroup label; the penetrating-trauma stratum (n = 7, no deaths) was too small for estimation and is listed as not estimable.

References

[1] NAGHAVI M, ABAJOBIR A A, ABBAFATI C, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016 [J]. The Lancet, 2017, 390(10100): 1151-210.

[2] DE MUNTER L, POLINDER S, LANSINK K W W, et al. Mortality prediction models in the general trauma population: A systematic review [J]. Injury, 2017, 48(2): 221-9.

[3] VOS T, LIM S S, ABBAFATI C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019 [J]. The Lancet, 2020, 396(10258): 1204-22.

[4] ZHOU XIN X Y, SUN YUANSONG, ET AL. Early diagnosis and prognostic evaluation value of HMGB1, suPAR, WBC and PCT in traumatic sepsis [J]. Chinese Journal of Emergency Medicine, 2021, 30（8）：(1015-1018).

[5] JIANG L, ZHENG Z, ZHANG M. The incidence of geriatric trauma is increasing and comparison of different scoring tools for the prediction of in-hospital mortality in geriatric trauma patients [J]. World Journal of Emergency Surgery, 2020, 15(1).

[6] HAKKOYMAZ H, GEDIK M S, KILCI A İ, et al. Predictors of mortality in patients with geriatric trauma in the emergency service [J]. Revista da Associação Médica Brasileira, 2023, 69(7).

[7] MORAN C G, LECKY F, BOUAMRA O, et al. Changing the System - Major Trauma Patients and Their Outcomes in the NHS (England) 2008–17 [J]. EClinicalMedicine, 2018, 2-3: 13-21.

[8] HOLEŠ D J R, BURŠA F. Comparison of Trauma Triage System in Prehospital Care Based on the Clinical Parameters with the ISS Score in Severely Injured Patients in the Czech Republic [J]. Acta Chir Orthop Traumatol Cech, 2020, 87(2): 120-6.

[9] DE VRIES R, REININGA I H F, DE GRAAF M W, et al. Older polytrauma: Mortality and complications [J]. Injury, 2019, 50(8): 1440-7.
[10] COLNARIC J M, EL SIBAI R H, BACHIR R H, et al. Injury severity score as a predictor of mortality in adult trauma patients by injury mechanism types in the United States: A retrospective observational study [J]. Medicine, 2022, 101(28).
[11] TAN T-W, TOIDA C, MUGURUMA T, et al. Ten-year in-hospital mortality trends among Japanese injured patients by age, injury severity, injury mechanism, and injury region: A nationwide observational study [J]. Plos One, 2022, 17(8).
[12] YANG L B X W L. Construction and study of nomograph model for prognosis of multiple trauma patients [J]. Chin J Emerg Med, 2023, 32（4）: 540-5.
[13] MIAO ZJ Z D, LIANG YP, ET AL. Analysis of independent risk factors and establishment and validation of a prediction model for in-hospital mortality of multiple trauma patients [J]. Chin J Trauma, 2023, 39（7）: 643-51.
[14] ZHU GL L Q. Analysis of prognostic factors related to severe trauma patients in the intensive care unit [J]. Chin J Emerg Med, 2018, 27（2）: 194-9.
[15] YU Q J L, GAO JB. Meta-analysis of APACHE II score and TRISS score in predicting prognosis of ICU trauma patients [J]. Chin Crit Care Med, 2022, 34(1): 59-63.
[16] ZHU Z, SHANG X, QI P, et al. Sex-based differences in outcomes after severe injury: an analysis of blunt trauma patients in China [J]. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 2017, 25(1).
[17] NAGATA I, ABE T, UCHIDA M, et al. Ten-year inhospital mortality trends for patients with trauma in Japan: a multicentre observational study [J]. BMJ Open, 2018, 8(2).
[18] WENIGER M, ANGELE M K, CHAUDRY I H. The Role and Use of Estrogens Following Trauma [J]. Shock, 2016, 46(3S): 4-11.
[19] KAWASAKI T, CHAUDRY I H. The effects of estrogen on various organs: therapeutic approach for sepsis, trauma, and reperfusion injury. Part 1: central nervous system, lung, and heart [J]. Journal of Anesthesia, 2012, 26(6): 883-91.