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.

Early risk stratification for in-hospital mortality after acute diquat poisoning: a retrospective cohort study with external validation

Meiwen Xie
North Sichuan Medical University
Yifan Ye
North Sichuan Medical University
Yuqiang Lin
Twelfth Guangzhou City People's Hospital
Ziyan Wang
Twelfth Guangzhou City People's Hospital
Zhiqian Yang
Twelfth Guangzhou City People's Hospital
Zhi Wang
Twelfth Guangzhou City People's Hospital
Bao Wang
Twelfth Guangzhou City People's Hospital
Yuquan Chen
cyq1142760126.com

Twelfth Guangzhou City People's Hospital

Research Article

Keywords: diquat, acute poisoning, in-hospital mortality, prediction model, external validation, Fine-Gray competing-risks analysis

Posted Date: September 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10838884/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Acute diquat poisoning may deteriorate rapidly, making early risk stratification relevant to monitoring, transfer, and organ-support planning. We aimed to develop and externally validate a parsimonious mortality model based on routinely available admission data.

Methods

This retrospective development cohort included 229 patients admitted within 24 h of diquat ingestion. Missing predictor data were handled using multiple imputation by chained equations. Predictors were selected using least absolute shrinkage and selection operator logistic regression and entered into a multivariable logistic model. Conditional internal validation used 1,000 bootstrap resamples with the final predictor set fixed. The unchanged model was externally validated in an independent public multicenter cohort of 201 patients. Performance was assessed using discrimination, calibration, Brier score, and decision-curve analysis. As a prespecified secondary hospital-course analysis, discharge alive was examined using a Fine-Gray competing-risks model, with in-hospital death as the competing event.

Results

Sixty-three patients (27.5%) died in the development cohort. The final model included HCO_3^-, lactate, estimated ingested volume, and white blood cell count. The optimism-corrected area under the receiver operating characteristic curve was 0.965, with a Brier score of 0.068 and calibration slope of 0.903. In the external cohort, the area under the curve was 0.931 (95% confidence interval 0.896–0.967), the Brier score was 0.105, the observed-to-expected ratio was 1.124, and the calibration slope was 0.639, indicating preserved risk ranking but underestimation and overly extreme predictions. In the secondary competing-risks analysis, greater ingested volume, higher lactate and creatinine, and lower bicarbonate were associated with a lower cumulative incidence of discharge alive.

Conclusion

The four-variable model retained high discrimination in an independent multicenter cohort, but clinically relevant calibration drift limits direct transfer of absolute risk estimates across settings. Local validation and recalibration, followed by prospective impact evaluation, are required before routine implementation.

Background

Acute diquat poisoning can progress within hours to neurological, circulatory, renal, and multiorgan failure [1, 2], and no specific antidote is available [3–6]. Early risk stratification is therefore important for

decisions regarding monitoring intensity, transfer, and organ-support planning [7].

Published mortality models for diquat poisoning vary in prediction timing, outcome definitions, predictor availability, and validation strategies [8–10]. Some rely on specialized measurements, such as plasma diquat concentrations or derived toxicity indices, which may limit their applicability in routine emergency settings [11, 12]. Models based on routinely available admission information therefore remain clinically relevant, particularly when their transportability is evaluated in independent data. External validation should assess both discrimination and calibration, because good risk ranking does not necessarily translate into accurate absolute risk estimates in new populations [13].

We developed a parsimonious model for in-hospital mortality using routinely available clinical and laboratory data obtained early after admission, performed bootstrap internal validation, and subsequently externally validated the model in an independent multicenter cohort. In addition, as a supplement to the primary in-hospital mortality prediction analysis, we conducted a competing-risks analysis, treating in-hospital death as the competing event, to evaluate the associations between early admission factors and discharge alive and thereby further characterize in-hospital clinical outcomes [14]. Reporting followed the TRIPOD statement and, where applicable, the STROBE recommendations [15, 16].

1 Methods

1.1 Study Population

The retrospective development cohort comprised patients diagnosed with acute diquat poisoning at Guangzhou Twelfth People's Hospital between 2018 and 2026. A total of 244 patients were initially screened. Acute diquat poisoning was diagnosed on the basis of a clear history of diquat exposure, clinical manifestations consistent with diquat poisoning, and toxicological supporting evidence available in the medical record, such as qualitative toxicology test results or clinical diagnostic documentation based on toxicological principles.

The inclusion criteria were as follows: (1) admission within 24 hours after ingestion; and (2) a clinical diagnosis of acute diquat poisoning.

Patients were not included if they had co-poisoning with other substances, severe underlying disease such as chronic obstructive pulmonary disease, heart failure, chronic kidney disease, or chronic liver disease, pregnancy or lactation, or an uncertain diagnosis. Patients who met the initial screening criteria but had substantial missing study data were also excluded. After excluding 10 patients with co-poisoning, three who did not present promptly after ingestion, and two with substantial missing data, 229 patients were included in the analysis (Fig. 1).

1.2 Data Collection and Variable Definitions

Demographic characteristics, clinical manifestations, and laboratory measurements were retrospectively extracted from electronic and archived medical records using a prespecified data collection form. Demographic and clinical data included age, sex, estimated ingested volume, length of hospital stay, and early clinical symptoms after admission. Symptoms were coded as binary variables based on medical records and included dizziness or headache, nausea or vomiting, oliguria or anuria, impaired consciousness, and mental or behavioral abnormalities. Laboratory variables were primarily based on the first measurements obtained after admission and included arterial pH, partial pressure of oxygen (PO_2), partial pressure of carbon dioxide (PCO_2), bicarbonate (HCO_3^-), lactate, white blood cell count, alanine aminotransferase (ALT), aspartate aminotransferase (AST), α-amylase, and serum creatinine. Estimated ingested volume was recorded in milliliters and was determined primarily from the history obtained at admission, usually from the patient, family members, or other informants. Because the active ingredient concentration may differ among commercial diquat formulations and precise toxicant concentration or toxicokinetic measurements were not consistently available in clinical practice, this variable was defined as “estimated ingested volume.” It was used to represent the clinically available extent of toxicant exposure rather than an exact measure of the absorbed diquat dose.

1.3 Missing Data and Preprocessing of Candidate Predictors

Predictors with missing values were handled using multiple imputation by chained equations (MICE) [17, 18]. Twenty imputed datasets were generated, with 20 iterations per dataset. Predictive mean matching was used for continuous variables and logistic regression for categorical variables. The outcome was included as an auxiliary predictor when other variables were imputed but was not itself imputed [19]. Parameter estimates and their variances from subsequent regression models were pooled using Rubin’s rules.

Before LASSO analysis, information redundancy and collinearity among candidate variables were rigorously assessed using clinical domain knowledge and statistical criteria. Correlations among laboratory variables were evaluated using Spearman rank correlation coefficients. Within each pathophysiological domain, representative variables were retained after considering the proportion of missing data, measurement stability, clinical interpretability, and availability early after admission. Multicollinearity among the retained candidate predictors was then evaluated using the variance inflation factor (VIF), with VIF < 5 considered acceptable.

1.4 Outcome Definitions and Descriptive Analyses

For the in-hospital mortality prediction model, death during hospitalization was coded as 1 and discharge alive as 0. For the competing-risks analysis of hospital outcomes, admission was the time origin, length of hospital stay was the time scale, discharge alive was the event of interest, and in-hospital death was the competing event. Continuous variables are presented as medians (25th and 75th percentiles), and categorical variables as counts and percentages. The Wilcoxon rank-sum test was used to compare continuous variables between patients who died in hospital and those discharged alive, and

Fisher's exact test was used for categorical variables. Standardized mean differences (SMDs) were also reported. Because most recorded symptoms occurred within 72 hours after admission, they were used only to describe baseline characteristics and were not included as LASSO candidate variables for the mortality prediction model or in the Fine-Gray competing-risks analysis. The prediction model study was reported in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement [15].

1.5 Development and Internal Validation of the In-Hospital Mortality Prediction Model

LASSO logistic regression was performed separately in each imputed dataset for variable selection [20]. Ten-fold cross-validation with binomial deviance was used, and the penalty parameter was selected by the one-standard-error rule (λ1SE) [21]. Selected variables were entered into the final multivariable logistic regression model without further significance-based removal. Model estimates were pooled using Rubin's rules, and pooled coefficients were used to derive the prediction equation and clinical score.

Model performance was assessed by AUC with 95% CI [22], calibration plot, calibration intercept and slope [23], Brier score [24], and decision curve analysis [25]. The Hosmer-Lemeshow test was used only as a supplementary goodness-of-fit measure. Internal validation used 1,000 bootstrap resamples with the final predictor set fixed to estimate optimism-corrected performance; out-of-bag (OOB) estimates were also reported [26].

1.6 External Validation

External validation used the independent publicly available multicenter cohort reported by Li et al., comprising 201 patients with the same outcome of in-hospital death. The original intercept and all four predictor coefficients were kept fixed, with no variable reselection, coefficient re-estimation, or threshold optimization. Because bicarbonate was not directly reported externally, HCO_3^- was calculated from admission pH and PCO_2 using the Henderson-Hasselbalch equation: HCO_3^- (mmol/L) = 0.03 ×PCO_2(mmHg) × 10^(pH − 6.1).

External performance was assessed using AUC with DeLong 95% CI [22], Brier and scaled Brier scores, O:E ratio, CITL, calibration slope, ICI [27], calibration plots, and decision curve analysis. Confidence intervals for the Brier score and calibration measures were obtained from 2,000 nonparametric bootstrap resamples. Sensitivity analysis recalculated HCO_3^- using alternative blood-gas constants. Intercept-only and intercept-plus-slope updates were performed solely as exploratory recalibration analyses [28] and did not replace evaluation of the fixed original model.

1.7 Fine-Gray Competing-Risks Analysis of Hospital Outcomes

Cumulative incidence functions (CIFs) were used to describe discharge alive and in-hospital death over time. A prespecified multivariable Fine-Gray model included age, estimated ingested volume, HCO_3^-, lactate, white blood cell count, creatinine, and alanine aminotransferase [29], with effects reported as subdistribution hazard ratios (sHRs) and 95% confidence intervals. Proportional subdistribution hazards were evaluated using covariate-by-log-time interactions with Holm correction. Full univariable estimates and diagnostic results are provided in Supplementary Tables S1-S2 and Supplementary Figure S1.

Analyses were performed in R version 4.6.1. Tests were two-sided, and $P < 0.05$ was considered statistically significant unless multiplicity-adjusted. The competing-risks analysis was prespecified as secondary and associational; sHRs were not interpreted as causal effects or direct changes in length of stay.

1.8 Ethics Review

This retrospective study was approved by the Ethics Committee of Guangzhou Twelfth People's Hospital (approval no. 2024039). The study was conducted in strict accordance with institutional requirements and the Declaration of Helsinki. All medical records were reviewed in de-identified form, and study data were accessible only to the investigators. The requirement for written informed consent was waived by the ethics committee because the study used anonymized retrospective clinical data.

2 Results

2.1 Study Population and Baseline Characteristics

The development cohort included 229 patients, of whom 63 (27.5%) died in hospital and 166 (72.5%) were discharged alive. The median age was 24.0 years (19.0, 30.0), and the median estimated ingested volume was 50.0 mL (20.0, 100.0). Compared with patients discharged alive, those who died had greater estimated ingested volumes and higher lactate, white blood cell count, ALT, AST, α-amylase, and creatinine levels, but lower pH, PCO_2, and HCO_3^- levels. Impaired consciousness, mental or behavioral abnormalities, and oliguria or anuria were also more frequent among patients who died (Table 1). These between-group comparisons were used only to describe cohort characteristics and did not inform variable screening before LASSO.

Table 1
Baseline characteristics and early post-admission symptoms in the development cohort

Variable	Overall (n = 229)	Discharged alive (n = 166)	Died in hospital (n = 63)	P value	SMD
Age (years)	24.00 [19.00, 30.00]	24.00 [19.00, 30.00]	24.00 [19.00, 30.50]	0.800	-0.042
Estimated ingested volume (mL)	50.00 [20.00, 100.00]	40.00 [15.00, 87.50]	150.00 [100.00, 200.00]	< 0.001	1.341
pH	7.40 [7.36, 7.43]	7.40 [7.38, 7.43]	7.35 [7.25, 7.41]	< 0.001	-0.529
PO_2 (kPa)	13.20 [11.70, 14.80]	13.15 [11.72, 14.53]	13.20 [11.46, 16.65]	0.366	0.340
PCO_2 (kPa)	4.50 [3.70, 5.13]	4.70 [4.20, 5.28]	3.51 [2.95, 4.45]	< 0.001	-0.930
HCO_3^- (mmol/L)	21.25 [17.38, 24.50]	22.55 [19.45, 25.22]	14.65 [11.80, 19.38]	< 0.001	-1.258
Lactate (mmol/L)	1.70 [1.20, 2.82]	1.50 [1.10, 2.10]	4.25 [2.73, 7.88]	< 0.001	1.226
White blood cell count ($\times 10^9$/L)	14.49 [10.25, 22.05]	12.07 [9.39, 17.43]	24.19 [19.45, 29.65]	< 0.001	1.487
ALT (U/L)	32.70 [13.45, 126.50]	20.00 [11.50, 60.40]	225.25 [66.98, 531.60]	< 0.001	0.818
AST (U/L)	37.15 [19.07, 112.43]	27.35 [17.38, 47.42]	282.00 [75.70, 499.75]	< 0.001	0.761
α-Amylase (U/L)	88.60 [57.60, 156.60]	74.80 [54.05, 115.60]	160.20 [79.97, 264.65]	< 0.001	0.565
Creatinine (μmol/L)	165.50 [80.00, 313.75]	130.00 [69.00, 304.50]	228.00 [152.50, 336.00]	0.002	0.226
Male	111/229 (48.5%)	79/166 (47.6%)	32/63 (50.8%)	0.767	0.064
Dizziness or headache	51/229 (22.3%)	39/166 (23.5%)	12/63 (19.0%)	0.594	-0.109

Note: Continuous variables are presented as median (25th percentile, 75th percentile), and categorical variables as n/N (%). P values are provided only for descriptive between-group comparisons and were not used for variable screening before LASSO. SMD, standardized mean difference. PO_2 and PCO_2 are both reported in kPa.

Variable	Overall (n = 229)	Discharged alive (n = 166)	Died in hospital (n = 63)	P value	SMD
Nausea or vomiting	189/229 (82.5%)	132/166 (79.5%)	57/63 (90.5%)	0.054	0.311
Oliguria or anuria	41/229 (17.9%)	21/166 (12.7%)	20/63 (31.7%)	0.002	0.472
Impaired consciousness	58/229 (25.3%)	14/166 (8.4%)	44/63 (69.8%)	< 0.001	1.619
Mental or behavioral abnormalities	31/229 (13.5%)	7/166 (4.2%)	24/63 (38.1%)	< 0.001	0.912

Note: Continuous variables are presented as median (25th percentile, 75th percentile), and categorical variables as n/N (%). P values are provided only for descriptive between-group comparisons and were not used for variable screening before LASSO. SMD, standardized mean difference. PO_2 and PCO_2 are both reported in kPa.

2.2 Development of the Mortality Prediction Model

2.2.1 Preprocessing of Candidate Variables

Thirteen variables were initially considered. After redundancy assessment, PCO_2 and AST were excluded, leaving 11 candidate predictors were included: sex, age, estimated ingested volume, pH, PO_2, HCO_3^-, lactate, white blood cell count, ALT, α-amylase, and creatinine. Following the information redundancy assessment, PCO_2 and AST were excluded, and HCO_3^- and ALT were retained as representative variables for their respective clinical domains. All 11 preprocessed candidate predictors had VIF values < 2.0. The maximum VIF was 1.949 for HCO_3^-, indicating no substantial multicollinearity (Fig. 2).

2.2.2 LASSO Selection and Model Development

Using λ1SE, LASSO selected HCO_3^-, lactate, estimated ingested volume, and white blood cell count in all 20 imputed datasets. The final multivariable logistic regression model for the primary analysis therefore included these four stable predictors. The linear predictor was: LP = − 2.146207 − 0.206240 × HCO_3^- + 0.220543 × lactate + 0.031436 × estimated ingested volume + 0.082689 × white blood cell count. The predicted probability of in-hospital death was P = 1/[1 + exp(− LP)], and a clinical scoring system was also developed. In the multivariable model, lower HCO_3^-, greater estimated ingested volume, and higher white blood cell count were statistically associated with a higher risk of in-hospital death. The regression coefficient for lactate was positive, but the 95% CI for its OR included 1 (P = 0.122). Because the objective was prediction rather than causal inference, lactate was retained in the final model in accordance with the prespecified LASSO stability criterion (Table 2 and Fig. 3).

Table 2
Multivariable logistic regression results for the in-hospital mortality prediction model

Variable	β	OR	95% CI	P value
Intercept	-2.146	—	—	—
HCO_3^- (per 1 mmol/L)	-0.206	0.814	0.722–0.917	< 0.001
Lactate (per 1 mmol/L)	0.221	1.247	0.942–1.650	0.122
Estimated ingested volume (per 1 mL)	0.031	1.032	1.019–1.045	< 0.001
White blood cell count (per 1 × 10^9/L)	0.083	1.086	1.011–1.167	0.025

Note: Regression coefficients and their variances were pooled across 20 imputed datasets using Rubin's rules. Final predictors were determined using the prespecified LASSO stability criterion and were not subjected to secondary selection based on multivariable regression P values.

Note

A, ten-fold cross-validation; B, standardized coefficient paths; C–D, four-variable consensus scoring system and the relationship between total score and predicted risk of in-hospital death.

Figure 3. LASSO selection and scoring system for the in-hospital mortality prediction model

2.2.3 Apparent Performance and Bootstrap Internal Validation of the Mortality Prediction Model

The apparent AUC was 0.971 (95% CI, 0.953–0.989), and the Brier score was 0.060. The apparent calibration intercept was close to 0, and the calibration slope was 1.054. The Hosmer-Lemeshow test statistic was 6.689 (df = 8, P = 0.571), providing no evidence of substantial lack of fit. DCA showed that the model provided greater net benefit than the treat-all and treat-none strategies across the analyzed threshold range of 0.01–0.80 (Fig. 4A–C). After 1,000 bootstrap resamples, the optimism-corrected AUC was 0.965, the Brier score was 0.068, the calibration intercept was 0.011, and the calibration slope was 0.903. The corresponding OOB estimates were 0.962, 0.070, 0.004, and 0.941, respectively (Table 3 and Fig. 4D). These findings indicate that the model retained high discrimination after internal validation and had limited overall optimism. However, the calibration slope below 1 suggested mild overfitting.

Table 3
Bootstrap internal validation of the in-hospital mortality prediction model

Measure	Apparent value	Mean optimism	Corrected value	OOB estimate
AUC	0.971	0.006	0.965	0.962
Brier score	0.060	-0.007	0.068	0.070
Calibration intercept	0.000	-0.011	0.011	0.004
Calibration slope	1.054	0.151	0.903	0.941

Note: All 1,000 bootstrap resamples were completed successfully. OOB, out-of-bag estimate. Optimism was defined as the difference between apparent and bootstrap test performance. Because lower Brier scores indicate better performance, mean optimism in the Brier score may be negative.

A, receiver operating characteristic curve; B, calibration plot; C, decision curve analysis; D, bootstrap out-of-bag calibration.

2.2.4 External Validation of the Mortality Prediction Model

The external cohort included 201 patients, of whom 100 (49.8%) died in hospital and 101 survived to discharge. No data were missing for pH, PCO_2, lactate, estimated ingested volume, white blood cell count, or outcome, all of which were required by the original model. PCO_2 ranged from 12.00 to 44.19 mmHg. The median HCO_3^- concentration calculated from pH and PCO_2 was 18.31 mmol/L, with a range of 4.75–32.40 mmol/L. With the original intercept and all regression coefficients kept fully fixed, external validation yielded an AUC of 0.931 (95% CI, 0.896–0.967), a Brier score of 0.105 (bootstrap 95% CI, 0.074–0.138), and a scaled Brier score of 0.580. The observed mortality rate was 49.8%, whereas the mean predicted risk of death was 44.3%. The O:E ratio was 1.124 (95% CI, 1.022–1.239), and CITL was 0.806 (95% CI, 0.155–1.424), indicating overall underestimation of mortality risk in the external cohort. The calibration slope was 0.639 (95% CI, 0.460–0.956), and the ICI was 0.068. A calibration slope below 1 indicated overly extreme predicted probabilities and calibration drift in the external population. DCA showed that the fixed original model provided greater net benefit than the treat-all and treat-none strategies at threshold probabilities of 0.10–0.80. At lower thresholds of 0.02–0.09, its net benefit was slightly lower than that of the treat-all strategy. Recalculation of HCO_3^- using alternative blood gas constants yielded an AUC of 0.932, a Brier score of 0.106, and a calibration slope of 0.638. These results were consistent with the primary analysis and indicated that the external validation findings were robust to the choice of blood gas conversion constants within a reasonable range (Table 4 and Fig. 5).

Table 4
External validation of the fixed original in-hospital mortality prediction model

Measure	Estimate	95% CI
Sample size/deaths	201/100	—
AUC	0.931	0.896–0.967
Brier score	0.105	0.074–0.138
Scaled Brier score	0.580	—
O:E	1.124	1.022–1.239
CITL	0.806	0.155–1.424
Calibration slope	0.639	0.460–0.956
ICI	0.068	—

Note: The 95% CI for the AUC was calculated using the DeLong method. The 95% CIs for the Brier score, O:E ratio, CITL, and calibration slope were obtained from 2,000 nonparametric bootstrap resamples.

A, receiver operating characteristic curve; B, calibration plot; C, decision curve analysis; D, distribution of predicted risk by outcome.

2.2.5 Exploratory Recalibration of the Mortality Prediction Model

Exploratory recalibration showed that updating the intercept alone changed it from − 2.146 to − 1.340 and increased the mean predicted risk of death from 44.3% to 49.8%, matching the observed mortality rate in the external cohort. The Brier score decreased from 0.105 to 0.098. Jointly updating the intercept and calibration slope yielded a Brier score of 0.099. Because these updates used outcome information from the external cohort, they were performed only to examine potential recalibration strategies and did not replace the external validation results for the fixed original model. The updated model still requires validation in a new independent cohort.

2.3 Cumulative Incidence of Hospital Outcomes and Fine-Gray Analysis

2.3.1 Cumulative Incidence of Hospital Outcomes

Among 229 patients, 166 were discharged alive and 63 died. The cumulative incidences of discharge alive at 7, 14, and 28 days were 10.9%, 38.0%, and 62.4%; corresponding mortality incidences were 25.3%, 26.2%, and 27.5% (Fig. 6).

Discharge alive was the event of interest, and in-hospital death was the competing event.

2.3.2 Fine-Gray Analysis

In the prespecified multivariable model, greater estimated ingested volume (per 10 mL: sHR 0.895, 95% CI 0.863–0.930), higher lactate (per 1 mmol/L: sHR 0.752, 95% CI 0.578–0.979), and higher creatinine (per 100 μmol/L: sHR 0.902, 95% CI 0.832–0.979) were associated with a lower subdistribution hazard of discharge alive. Higher bicarbonate was associated with a higher subdistribution hazard of discharge alive (per 5 mmol/L: sHR 1.352, 95% CI 1.118–1.635). Age, white blood cell count, and alanine aminotransferase were not independently associated. No clear violation of the proportional subdistribution hazards assumption was detected after Holm correction (smallest adjusted P = 0.058). Full estimates and diagnostics are provided in Supplementary Tables S1-S2 and Supplementary Figure S1.

3 Discussion

3.1 Principal Findings

We developed an early four-variable mortality model using routinely available admission data and evaluated the unchanged model in an independent multicenter cohort. Discrimination remained high after conditional bootstrap validation and external validation. External calibration, however, showed underestimation of mortality and a calibration slope of 0.639, indicating that absolute risk estimates were not directly transportable across settings without local evaluation and recalibration. The principal contribution of the model is therefore its parsimonious structure and preserved risk discrimination rather than immediate use as a treatment decision rule. As a prespecified supplementary analysis, the competing-risks analysis identified admission factors associated with discharge alive while accounting for in-hospital death as a competing event, thereby providing complementary information on in-hospital prognosis rather than constituting a second prediction model.

3.2 Clinical Interpretation of the Predictors

Estimated ingested volume reflects the early toxicant burden and is an important clinical indicator of the severity of acute diquat poisoning. As the ingested volume increases, redox cycling and reactive oxygen species production intensify, which may further promote tissue injury and multiple organ dysfunction. Previous studies have also evaluated ingested quantity as a prognostic marker in diquat poisoning [7, 8, 30]. Although this estimate is susceptible to patient recall, vomiting, and differences in formulation concentration, it can be obtained rapidly without specialized testing and therefore remains clinically useful.

HCO_3^- and lactate primarily reflect acid-base and metabolic status early in poisoning. Severe diquat poisoning can cause inadequate circulatory perfusion, tissue hypoxia, and metabolic disturbances, leading to lactate accumulation and metabolic acidosis. Reduced HCO_3^- may reflect both an increased acid load and impaired compensatory capacity. Elevated lactate more directly reflects tissue

hypoperfusion, impaired oxygen utilization, and systemic metabolic stress. Although the two measures are correlated to some extent, they capture different pathophysiological information. Their combined use may therefore provide a more comprehensive characterization of early disease severity [6, 31].

White blood cell count primarily reflects acute stress and inflammation. Oxidative stress induced by diquat can activate inflammatory responses [32], accompanied by neutrophil mobilization and an increased peripheral white blood cell count. Patients with severe poisoning also frequently have tissue injury, a risk of infection, and multiple organ dysfunction. White blood cell count is therefore best regarded as a nonspecific marker of systemic stress. It is inexpensive and widely available, and several previous prognostic studies of diquat poisoning have also found that it provides relatively stable prognostic information [33].

The four variables in the model represent distinct pathophysiological dimensions: toxicant burden, acid-base disturbance, metabolic stress, and systemic inflammation. Unlike models that rely on plasma diquat concentrations or complex organ function scores, all four measures can be obtained routinely early after admission. The model may therefore offer bedside utility while remaining parsimonious.

3.3 Comparison with Previous Models and Implications of External Validation

Previous prognostic and mortality prediction studies of diquat poisoning have differed substantially in study design, variable selection, and validation methods [8–10, 34, 35]. Some studies relied mainly on internal validation or random splitting of a single data source. Compared with previous studies, the present study placed greater emphasis on parsimony and independent validation. The final model included only four routinely available measures obtained early after admission and did not require plasma diquat concentrations or complex scoring systems. It may therefore be suitable for emergency settings with limited testing capacity. During external validation, the original intercept and regression coefficients were held fixed, avoiding performance overestimation caused by variable reselection or parameter adjustment in the validation data. In addition to discrimination, we evaluated overall prediction error, calibration, and decision curves, providing a more comprehensive assessment of performance in a new cohort.

Discrimination was well preserved in the external validation, whereas calibration deteriorated. Thus, the model continued to distinguish patients at relatively high and low risk, but the predicted probabilities did not fully agree with the observed risks in the external population. This discrepancy may reflect between-center differences in disease severity, referral patterns, time to presentation, treatment strategies, and laboratory conditions [36]. It may also indicate heterogeneity in predictor effects across populations.

Exploratory recalibration suggested that correcting the baseline event risk in the external population could improve probability estimates to some extent. However, because these updates depended on outcome information from the external cohort, they cannot replace true external validation. Before the model is implemented at a particular hospital or in a specific region, the fixed model should first be

evaluated in the target population. The need for model updating can then be determined from its calibration [37]. In addition, HCO_3^- in the external cohort was derived from pH and PCO_2. Although the sensitivity analysis produced stable results, the difference in measurement methods may still have affected calibration and should be standardized in future prospective validation.

3.4 Complementary Value of the Fine-Gray Competing-Risks Analysis

The in-hospital mortality prediction model estimates the eventual probability of death but does not capture the timing of hospital outcomes. In patients with acute diquat poisoning, discharge alive and in-hospital death are competing outcomes because death precludes subsequent discharge alive. A competing-risks approach can therefore describe the cumulative incidence of discharge alive more appropriately while accounting for in-hospital death. The analysis further suggested that toxicant burden, acid-base and metabolic disturbances were associated not only with mortality risk but also with progression toward the favorable outcome of discharge alive. Notably, the role of renal function differed between the mortality prediction model and the competing-risks analysis. Some clinical factors may therefore have different prognostic implications for whether a patient dies and whether the patient recovers sufficiently for discharge. This distinction broadens the focus from the terminal outcome of death to the overall course of recovery during hospitalization.

The Fine-Gray analysis served as a supplementary prognostic analysis of discharge alive while accounting for competing in-hospital death. The sHR reflects the cumulative incidence of discharge alive rather than hospital stay itself [38]. Because discharge is also influenced by institutional practice and other non-biological factors, these associations should be interpreted cautiously.

3.5 Strengths and Limitations

This study had several strengths. Predictors were largely restricted to routinely available information obtained early after admission. Symptoms that could have been recorded up to 72 hours after admission were not included as candidate predictors, reducing the introduction of information collected after the intended prediction time point. Missing data were addressed using 20 multiple imputations, and LASSO selection was performed in each imputed dataset. Final predictors were determined using a prespecified stability criterion. The final model was internally validated using 1,000 bootstrap resamples and externally validated with all parameters fixed in an independent, publicly available multicenter cohort. Its discrimination, calibration, overall error, and potential clinical utility were comprehensively assessed using the AUC, Brier score, calibration measures, and DCA. The Fine-Gray analysis also provided a complementary competing-risks assessment of discharge alive during hospitalization.

This study also had limitations. The development cohort was derived retrospectively from a single center, and the sample size and number of events limited further exploration of complex nonlinear relationships and interactions [39, 40]. The long enrollment period may also have encompassed changes in clinical practice. Although external validation used an independent multicenter database, the two

cohorts differed in case mix, baseline mortality risk, sampling, and management. In addition, HCO_3^- was derived from pH and PCO_2 in the external cohort, which may have affected calibration. Estimated ingested volume was based on the clinical history and was subject to measurement error. Despite high external discrimination, calibration drift indicates that absolute risk predictions require further validation and, where necessary, local recalibration in the target population. We did not perform a prospective clinical impact assessment [41], and the Fine-Gray results should be interpreted as prognostic associations rather than causal effects. Prospective multicenter external validation with adequate sample size [42–44] and subsequent clinical impact evaluation are therefore required before routine implementation.

4 Conclusions

This early four-variable model retained high discrimination in an independent multicenter cohort but showed clinically relevant calibration drift. It may serve as an early risk stratification tool to help identify patients at higher risk of in-hospital mortality after acute diquat poisoning. However, absolute risk estimates should not be directly used for clinical decision-making across centers without local validation and recalibration. Prospective validation and clinical impact evaluation are required before routine implementation.

Declarations

Data Availability Statement

The datasets analyzed during the current study are not publicly available because they contain clinical information that may permit re-identification, but de-identified data may be available from the corresponding author on reasonable request and subject to institutional approval.

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of Guangzhou Twelfth People's Hospital (approval No. 2024039).

Author Contributions

YQC conceived and supervised the study and was responsible for project administration. MX, YY, and YQC contributed to the study methodology. MX, YY, YQL, ZiW, ZQY, ZhW, and BW conducted the investigation. MX, YY, YQL, and ZhW curated the data. YY and MX performed the formal analyses, and YY prepared the visualizations. MX and YY drafted the original manuscript. YQC, MX, YY, ZiW, ZQY, ZhW, and BW critically reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Guangzhou Key Medical Discipline (2025–2027), the Second Batch of Guangzhou Major Difficult and Rare Disease Diagnosis and Treatment Centers (Grant No. 2026ZLZX-17), and the Guangzhou City-University-Institute-Enterprise Joint Funding Program (Guangzhou Key Laboratory of Occupational Environment and Health, Grant No. 2023A03J0502).

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to publish.

Acknowledgments

We thank our colleagues in the Department of Occupational Diseases and Poisoning for their support.

References

1. Jones GM, Vale JA. Mechanisms of toxicity, clinical features, and management of diquat poisoning: a review. J Toxicol Clin Toxicol. 2000;38:123–8. 10.1081/CLT-100100926.
2. Zhang Y, Chen X, Du H, Zhao M, Jiang X. Association between initial diquat plasma concentration, severity index and in-hospital mortality in patients with acute diquat poisoning: a retrospective cohort study. Clin Toxicol (Phila). 2024;62:557–63. 10.1080/15563650.2024.2387649.
3. Zhou JN, Lu YQ. Lethal diquat poisoning manifests as acute central nervous system injury and circulatory failure: a retrospective cohort study of 50 cases. EClinicalMedicine. 2022;52:101609. 10.1016/j.eclinm.2022.101609.
4. Yu G, Wang J, Jian T, Shi L, Zhao L, Li Y, et al. Case series: diquat poisoning with acute kidney failure, myocardial damage, and rhabdomyolysis. Front Public Health. 2022;10:991587. 10.3389/fpubh.2022.991587.
5. Zeng DH, Chen XH, Li Y, Zhu XY, Wang BL, Ma XG, et al. Clinical and pathological characteristics of acute kidney injury caused by diquat poisoning. Clin Toxicol (Phila). 2023;61:705–8. 10.1080/15563650.2023.2262113.
6. Magalhães N, Carvalho F, Dinis-Oliveira RJ. Human and experimental toxicology of diquat poisoning: toxicokinetics, mechanisms of toxicity, clinical features, and treatment. Hum Exp Toxicol. 2018;37:1131–60. 10.1177/0960327118765330.
7. Ling Y, Mao Z, Liu W, Zhou C, Li J, Jiang L, et al. Severity indices of diquat poisoning for triage and prognosis in acute diquat poisoning: a multicenter prospective cohort study. Ann Emerg Med. 2025;85:512–20. 10.1016/j.annemergmed.2025.02.022.
8. Li H, Liu Z, Sun W, Li T, Dong X. Interpretable machine learning for the prediction of death risk in patients with acute diquat poisoning. Sci Rep. 2024;14:16101. 10.1038/s41598-024-67257-6.
9. Lv M, Du Y. Construction of a mortality risk prediction model for patients with acute diquat poisoning based on clinically accessible data. J Occup Med Toxicol. 2024;19:20. 10.1186/s12995-024-00416-

7.

10. Zhang Y, Chen X, Zhao M, Du H, Jiang X, Ma Y. Development and validation of a prediction model for in-hospital mortality among patients with acute diquat poisoning: a retrospective cohort study. Med (Baltim). 2025;104:e45613. 10.1097/MD.0000000000045613.
11. Meng N, Sun Y, Dong Y, Lv B, Yao D, Gao H, et al. Prognostic value of plasma diquat concentration in patients with acute oral diquat poisoning: a retrospective study. Front Public Health. 2024;12:1333450. 10.3389/fpubh.2024.1333450.
12. Zhang Y, Chen X, Du H, Zhao M, Jiang X, Ma Y. Association between toxicity-index of diquat and in-hospital mortality in patients with acute diquat poisoning: a retrospective cohort study. Biomarkers. 2024;29:427–33. 10.1080/1354750X.2024.2410238.
13. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: validating a prognostic model. BMJ. 2009;338:b605. 10.1136/bmj.b605.
14. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation. 2016;133:601–9. 10.1161/CIRCULATIONAHA.115.017719.
15. Collins GS, Reitsma JB, Altman DG, Moons KGM, TRIPOD Group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162:55–63. 10.7326/M14-0697.
16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet. 2007;370:1453–7. 10.1016/S0140-6736(07)61602-X.
17. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30:377–99. 10.1002/sim.4067.
18. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45:1–67. 10.18637/jss.v045.i03.
19. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393. 10.1136/bmj.b2393.
20. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Stat Methodol. 1996;58:267–88. 10.1111/j.2517-6161.1996.tb02080.x.
21. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33:1–22. 10.18637/jss.v033.i01.
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44:837–45. 10.2307/2531595.
23. Van Calster B, Nieboer D, Vergouwe Y, De Cock B, Pencina MJ, Steyerberg EW. A calibration hierarchy for risk models was defined: from utopia to empirical data. J Clin Epidemiol. 2016;74:167–76. 10.1016/j.jclinepi.2015.12.005.

24. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models. Epidemiology. 2010;21:128–38. 10.1097/EDE.0b013e3181c30fb2.
25. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak. 2006;26:565–74. 10.1177/0272989X06295361.
26. Efron B, Tibshirani R. Improvements on cross-validation: the.632 + bootstrap method. J Am Stat Assoc. 1997;92:548–60. 10.1080/01621459.1997.10474007.
27. Austin PC, Steyerberg EW. The Integrated Calibration Index (ICI) and related metrics for quantifying the calibration of logistic regression models. Stat Med. 2019;38:4051–65. 10.1002/sim.8281.
28. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Calibration: the Achilles heel of predictive analytics. BMC Med. 2019;17:230. 10.1186/s12916-019-1466-7.
29. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc. 1999;94:496–509. 10.1080/01621459.1999.10474144.
30. Tang Q, Wang H, Wang H, Xu J, Luo X, Hua S, et al. Risk factors for death in patients with acute diquat poisoning. World J Emerg Med. 2025;16:225–30. 10.5847/wjem.j.1920-8642.2025.040.
31. Yan M, Liu H, Yang Y, Cheng X, Sun W, Ma T, et al. Clinical characteristics of survivors versus non-survivors after acute diquat poisoning: a comparative study. Intern Emerg Med. 2024;19:307–12. 10.1007/s11739-023-03460-z.
32. Lai K, Wang J, Lin S, Chen Z, Lin G, Ye K, et al. Sensing of mitochondrial DNA by ZBP1 promotes RIPK3-mediated necroptosis and ferroptosis in response to diquat poisoning. Cell Death Differ. 2024;31:635–50. 10.1038/s41418-024-01279-5.
33. Wang P, Lin LY, Song CY, Fang JJ, Zhang BW, Li WS, et al. Early immune-inflammatory profiling and prognostic biomarkers in patients with acute diquat poisoning: a multicenter study with exploratory bioinformatics analysis. J Inflamm Res. 2025;18:14245–60. 10.2147/JIR.S547369.
34. He C, Liang L, Zhang Y, Wang T, Wang R. Prognosis prediction of procalcitonin within 24 h for acute diquat poisoning. BMC Emerg Med. 2024;24:61. 10.1186/s12873-024-00975-2.
35. Jiang L, Li M, Fang Y, Yang X. Construction and validation of a nomogram model to predict mortality risk in patients with acute diquat poisoning. Am J Transl Res. 2025;17:1901–9. 10.62347/XXIL7895.
36. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130:515–24. 10.7326/0003-4819-130-6-199903160-00016.
37. Janssen KJM, Moons KGM, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. J Clin Epidemiol. 2008;61:76–86. 10.1016/j.jclinepi.2007.04.018.
38. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med. 2017;36:4391–400. 10.1002/sim.7501.
39. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE Jr, Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: part II - binary and time-to-event outcomes. Stat Med.

2019;38:1276–96. 10.1002/sim.7992.
40. van Smeden M, Moons KGM, de Groot JAH, Collins GS, Altman DG, Eijkemans MJC, et al. Sample size for binary logistic prediction models: beyond events per variable criteria. Stat Methods Med Res. 2019;28:2455–74. 10.1177/0962280218784726.
41. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ. 2009;338:b606. 10.1136/bmj.b606.
42. Riley RD, Archer L, Snell KIE, Ensor J, Dhiman P, Martin GP, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. BMJ. 2024;384:e074820. 10.1136/bmj-2023-074820.
43. Riley RD, Snell KIE, Archer L, Ensor J, Debray TPA, van Calster B, et al. Evaluation of clinical prediction models (part 3): calculating the sample size required for an external validation study. BMJ. 2024;384:e074821. 10.1136/bmj-2023-074821.
44. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med. 2021;40:4230–51. 10.1002/sim.9025.

Figures

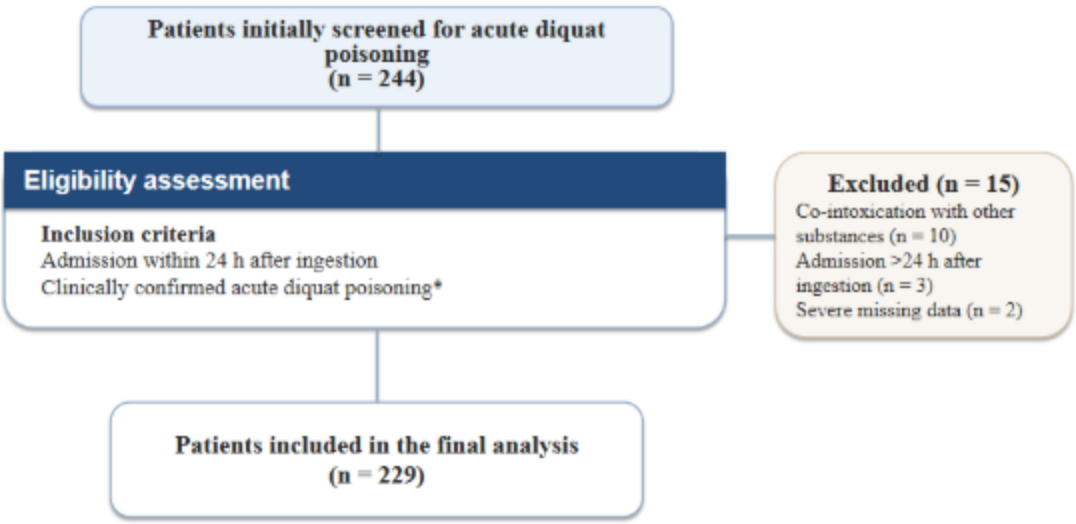


Figure 1

Patient selection flowchart

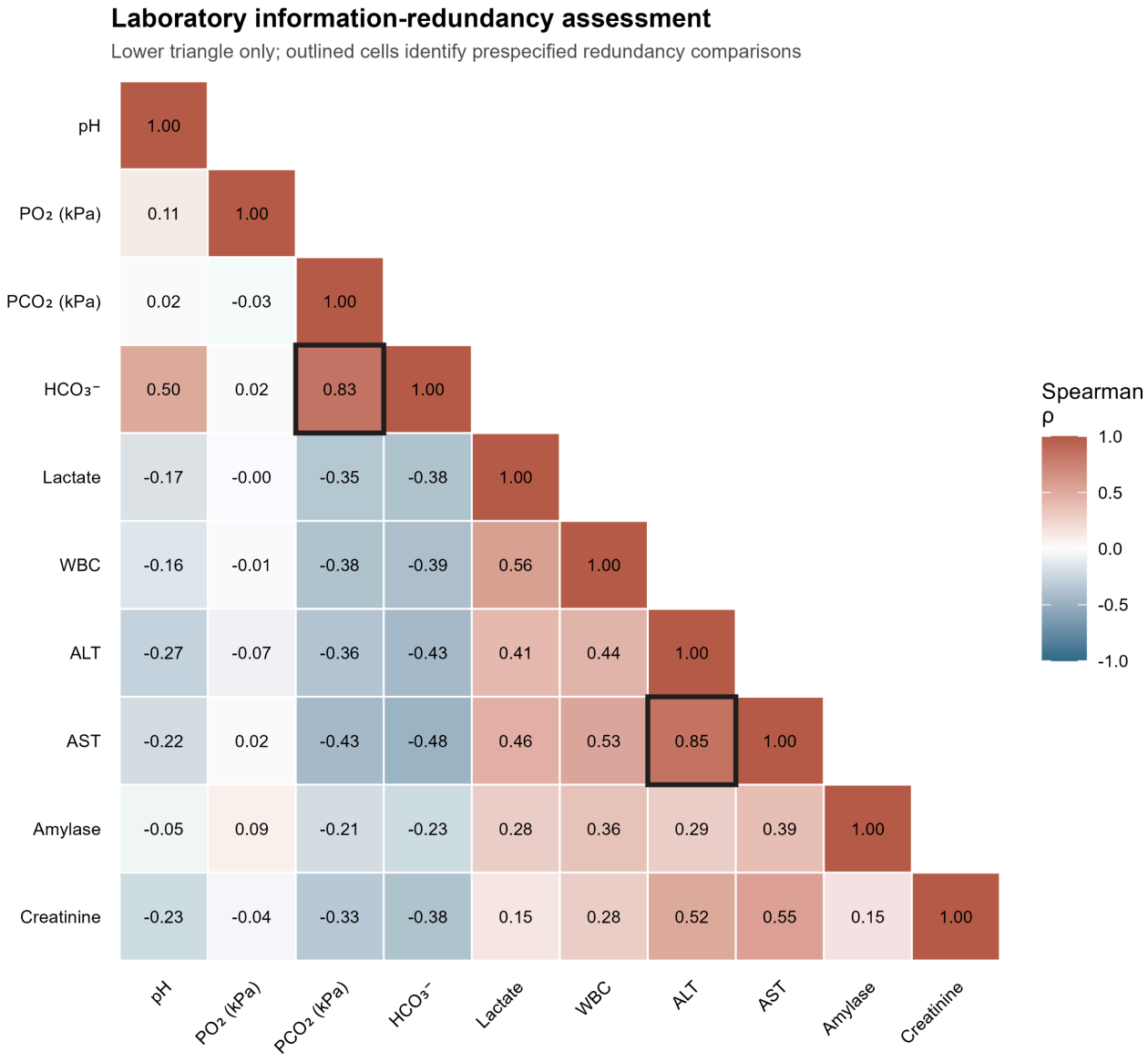


Figure 2

Spearman correlations and information redundancy assessment among candidate laboratory variables

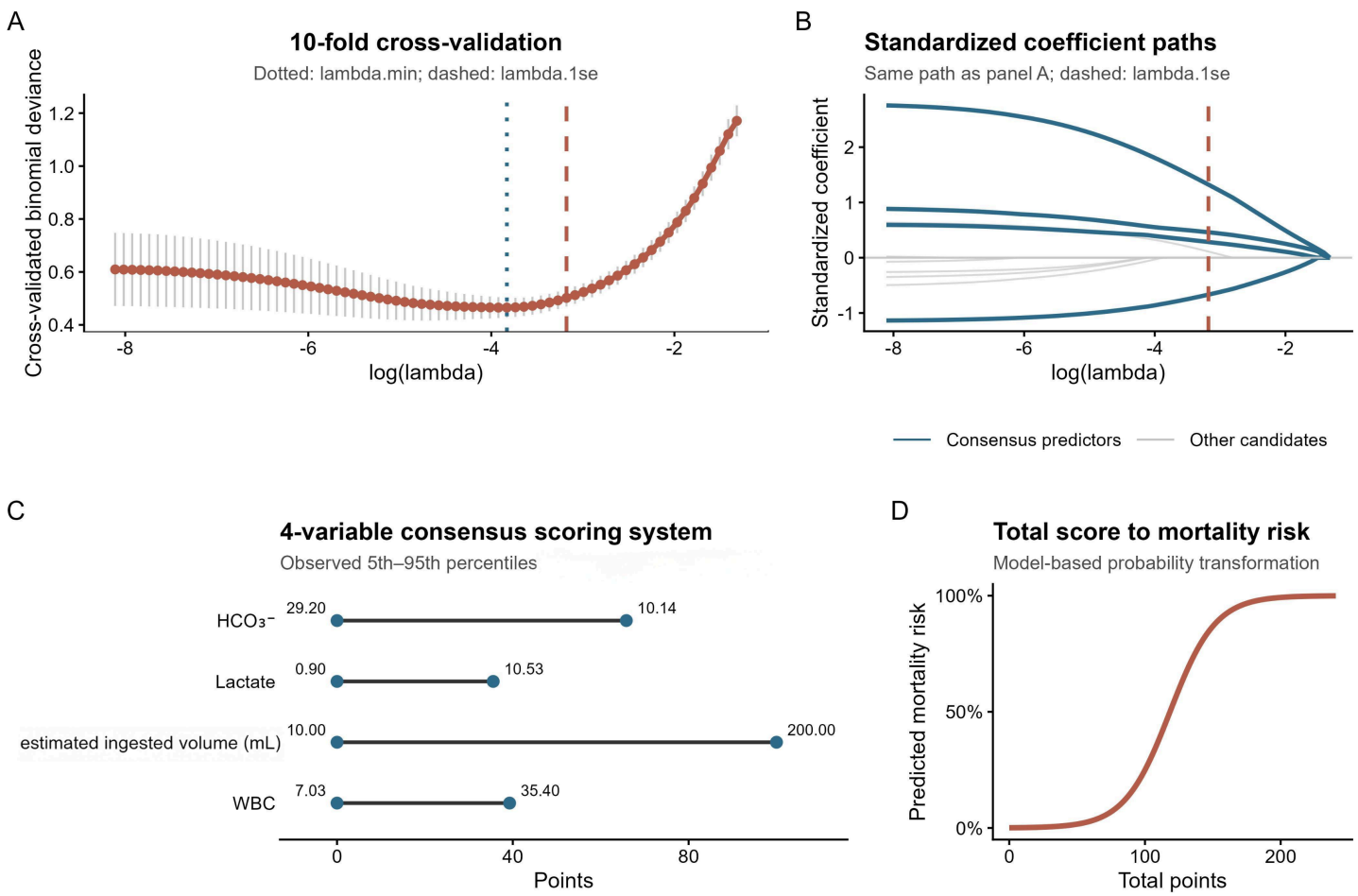


Figure 3

LASSO selection and scoring system for the in-hospital mortality prediction model

Note: A, ten-fold cross-validation; B, standardized coefficient paths; C–D, four-variable consensus scoring system and the relationship between total score and predicted risk of in-hospital death.

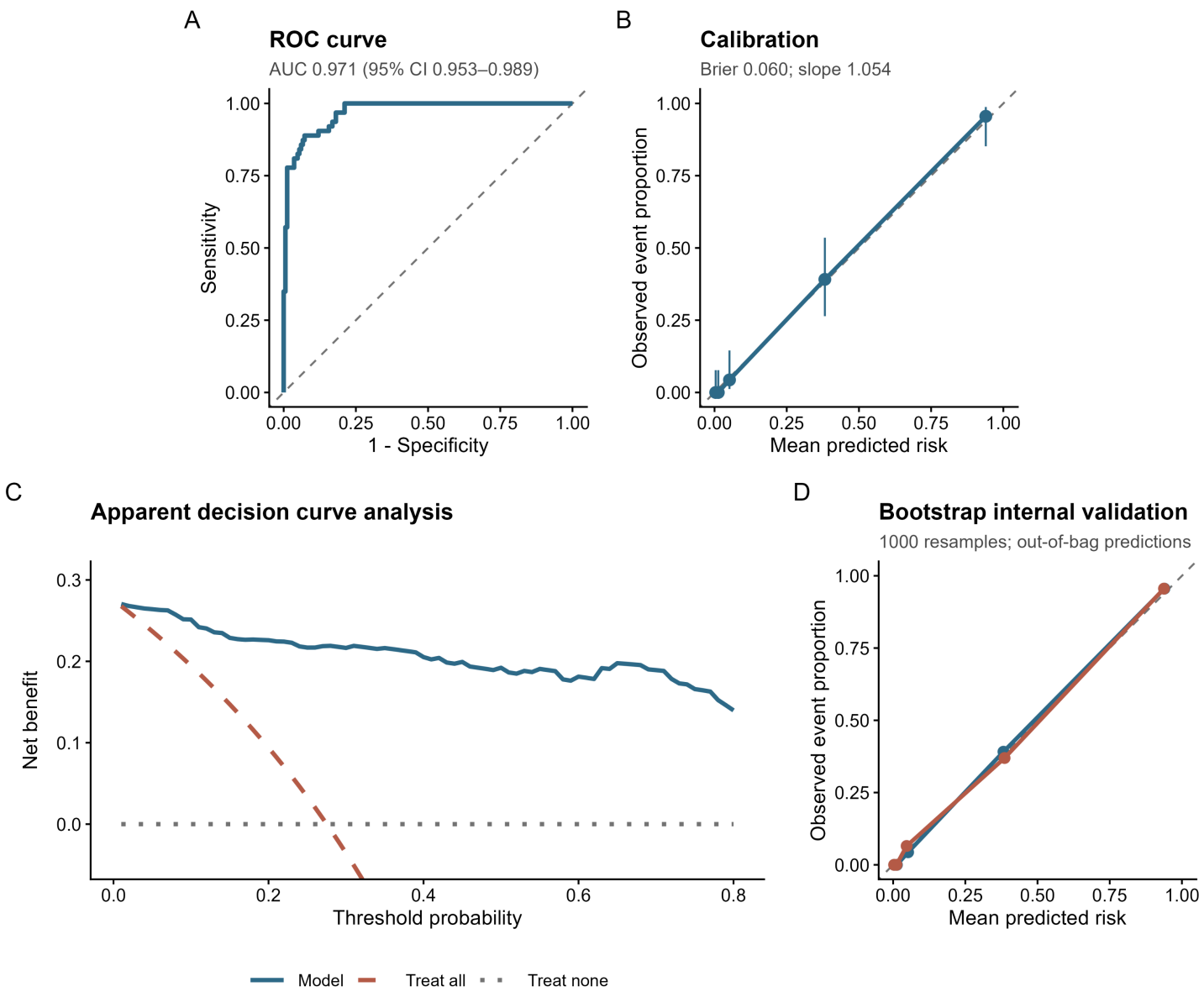


Figure 4

Performance and internal validation of the in-hospital mortality prediction model

A, receiver operating characteristic curve; B, calibration plot; C, decision curve analysis; D, bootstrap out-of-bag calibration.

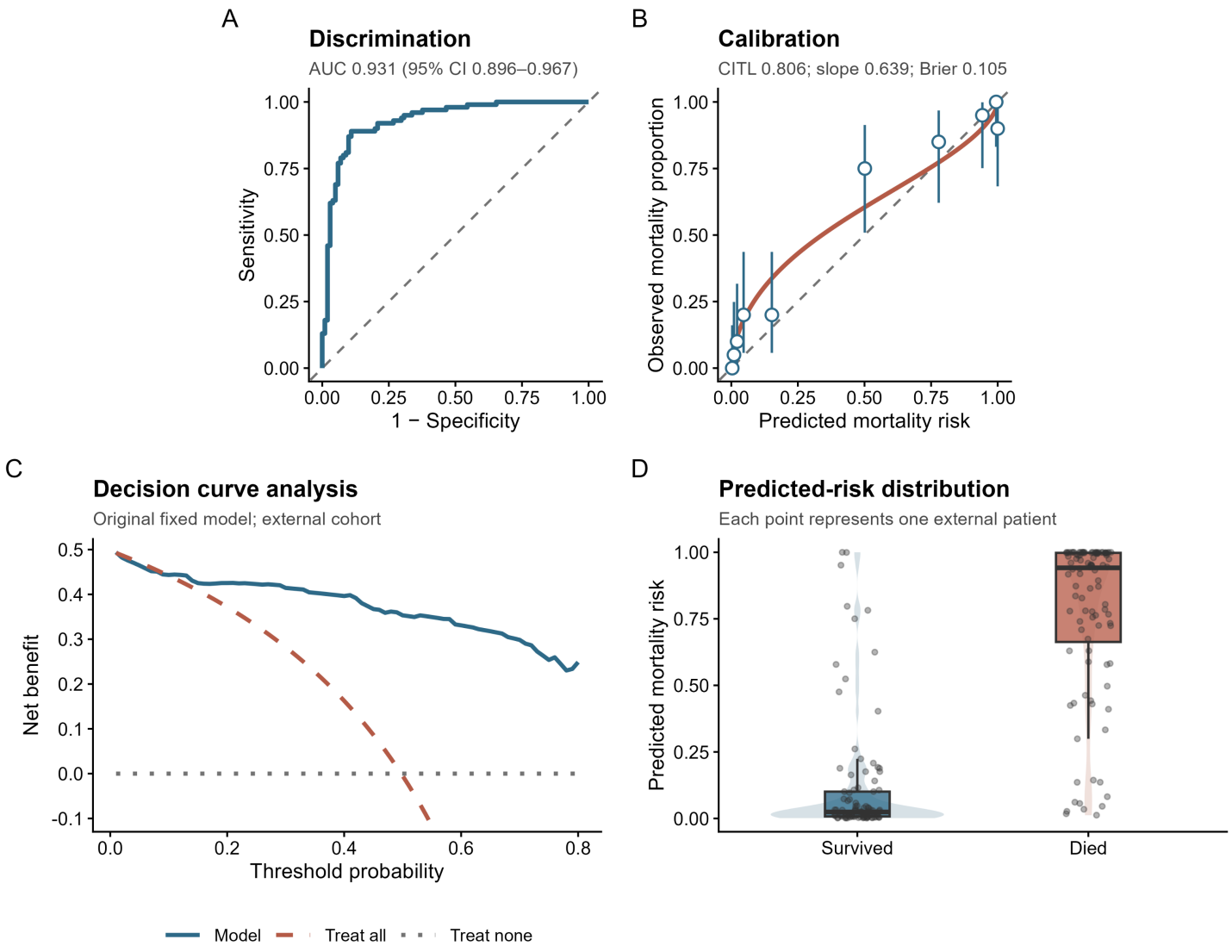


Figure 5

External validation of the fixed original in-hospital mortality prediction model
A, receiver operating characteristic curve; B, calibration plot; C, decision curve analysis; D, distribution of predicted risk by outcome.

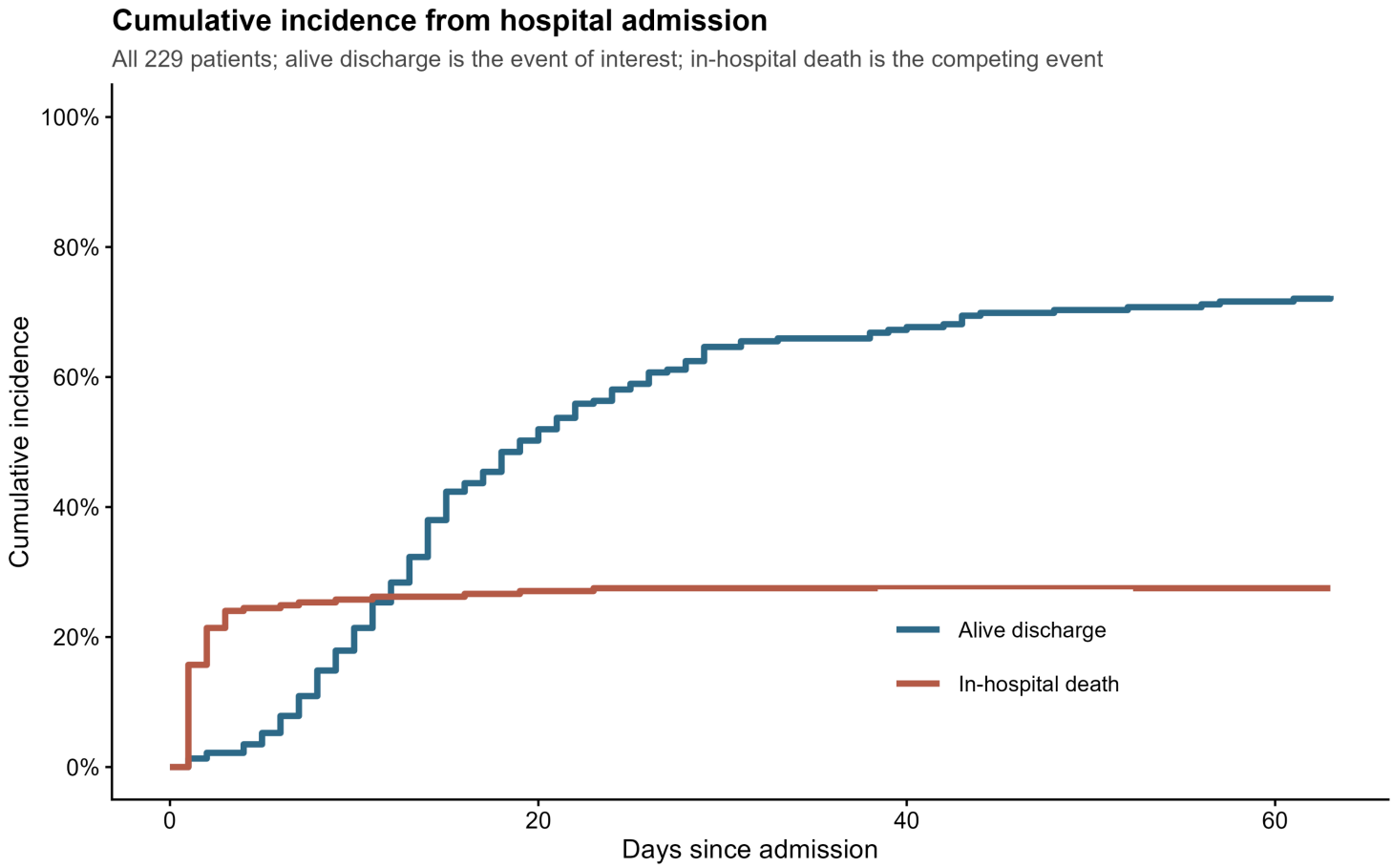


Figure 6

Cumulative incidence functions for discharge alive and in-hospital death
Discharge alive was the event of interest, and in-hospital death was the competing event.

Supplementary Files

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

- FineGraycompetingrisksanalysis.docx