

Development and external evaluation of a routine laboratory-based three-ratio score for in-hospital mortality in older critically ill patients with atherosclerotic cardiovascular disease: a MIMIC-IV and eICU study

Shiwen Cui

Yiyang Central Hospital

Xiaoyi Yao

Yiyang Central Hospital

Xiang Zhou

Yiyang Central Hospital

Yi Tao

Yiyang Central Hospital

Liuyin Zhou

Yiyang Central Hospital

Yanqiong Xue

Yiyang Central Hospital

Weiqiong Chen

chenweiqiongxy@qq.com

Yiyang Central Hospital

Research Article

Keywords: atherosclerotic cardiovascular disease, intensive care, blood urea nitrogen, bicarbonate, glucose, three-ratio score, TabM, in-hospital mortality

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10919961/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

Older patients with atherosclerotic cardiovascular disease (ASCVD) often develop concurrent changes in renal handling, acid-base balance, and glucose metabolism after admission to the intensive care unit (ICU). A single laboratory measurement may not capture the combined extent of these early abnormalities.

Methods

We conducted a multidatabase retrospective cohort study using MIMIC-IV 3.1 and eICU 2.0. We included patients aged 65 years or older who had ASCVD, remained in the ICU for at least 24 h, and had the four laboratory measurements required to calculate the three-ratio score (TRS-3). TRS-3 was defined as the equally weighted mean of the blood urea nitrogen-to-creatinine, blood urea nitrogen-to-bicarbonate, and glucose-to-bicarbonate ratios after normalization to clinical reference values. The primary outcome was in-hospital all-cause mortality after the first 24 h of ICU admission. Associations were assessed using multivariable logistic regression, competing risk models, and restricted cubic splines. TabM models underwent patient-level cross-fitting in MIMIC-IV, after which the preprocessing parameters, model parameters, and classification threshold were applied unchanged to eICU.

Results

The MIMIC-IV cohort included 8037 patients, of whom 714 died in hospital. The eICU cohort included 9686 patients, of whom 1002 died in hospital. After full adjustment, the odds ratios (ORs) for in-hospital mortality per 1-standard-deviation increase in TRS-3 were 1.39 (95% CI: 1.28–1.51) in MIMIC-IV and 1.31 (95% CI: 1.22–1.41) in eICU. Compared with the lowest quartile, the highest quartile had ORs of 3.81 (95% CI: 2.80–5.20) and 2.70 (95% CI: 2.10–3.48), respectively. Both cohorts showed continuous nonlinear risk gradients. Adding TRS-3 to the core clinical TabM model increased the area under the receiver operating characteristic curve (AUROC) from 0.788 to 0.829 in MIMIC-IV and from 0.717 to 0.760 in eICU, with paired DeLong $P < 0.001$ in both cohorts. Net reclassification improvement and integrated discrimination improvement were positive in both cohorts.

Conclusions

A higher early TRS-3 was associated with a greater probability of in-hospital death among older critically ill patients with ASCVD and provided incremental discrimination beyond the core clinical TabM model. This fixed score can be calculated from 4 routine laboratory measurements, although its probability calibration and classification threshold require evaluation in prospective cohorts.

Introduction

Atherosclerotic cardiovascular disease (ASCVD) remains a major cause of death and disability worldwide [1]. Older adults with cardiovascular disease frequently have diabetes, heart failure, and other chronic conditions, with substantial variation in physiological reserve and treatment tolerance [2]. Older and multimorbid patients remain underrepresented in clinical research, supporting separate evaluation of routinely monitored data in intensive care populations [3].

After ICU admission, outcomes in older patients with ASCVD reflect the combined effects of the underlying cardiovascular disease, acute circulatory disturbance, and changes in noncardiac organ function. Disease-specific scores may not encompass the full range of risk information in an unselected cardiac critical care population [4]. The blood urea nitrogen-to-creatinine ratio is influenced by renal perfusion, neurohormonal activity, catabolism, and volume status and may complement the renal information provided by creatinine alone [5]. Both blood urea nitrogen and its ratio to creatinine have been associated with mortality in patients with acute ST-segment elevation myocardial infarction [6]. Acid-base disorders are also common during critical illness. Serum bicarbonate reflects metabolic acid-base status, renal handling, fluid therapy, and respiratory compensation and therefore characterizes an important component of the early buffering state [7]. Acid-base status at admission has been associated with in-hospital mortality in cardiac critical care, and lower bicarbonate concentrations are common among high-risk patients [8]. Because a single bicarbonate measurement does not identify the underlying cause, it is best interpreted with other routine laboratory results. Acute

hyperglycemia may reflect both the stress response and altered glucose metabolism. Relative hyperglycemia indices account for antecedent glycemic status, and the stress hyperglycemia ratio has been associated with mortality in patients with cardiovascular disease [9, 10]. Glucose, bicarbonate, and nitrogenous metabolites are all routinely available early after ICU admission, making them suitable components of a fixed, parsimonious score.

Previous MIMIC-IV studies have jointly evaluated the stress hyperglycemia ratio and glycemic variability in critically ill patients with ASCVD [11]. Simplified metabolic indices, including the triglyceride-glucose index, have also been studied in cardiovascular risk assessment, although their interpretation depends on their components, measurement conditions, and target population [12]. A recent study compared the triglyceride-glucose index with body mass index-derived variants in older ICU patients with ASCVD and assessed their incremental discrimination using machine learning [13]. A fixed, parsimonious measure that jointly represents nitrogenous burden, renal handling, bicarbonate, and glucose has not been evaluated in independent multicenter databases. We therefore developed a three-ratio score (TRS-3), examined its association with in-hospital mortality, and compared the discrimination, calibration, and estimated net benefit of TabM models based on TRS-3, core clinical variables, and the 4 original laboratory components.

Methods

Study design and data sources

This multidatabase retrospective cohort study used the Medical Information Mart for Intensive Care IV version 3.1 and the eICU Collaborative Research Database version 2.0. MIMIC-IV contains deidentified hospital and critical care data collected at Beth Israel Deaconess Medical Center from 2008 through 2022 [14]. eICU contains deidentified ICU data from multiple US hospitals collected from 2014 through 2015 [15]. All eligible MIMIC-IV patients formed the MIMIC-IV cohort, which was used for association analyses, model development, and patient-level cross-fitted evaluation. The eICU cohort served as an independent external test cohort for the fixed score and prediction models. The terms internal cross-fitting and external testing were used in accordance with established principles for prediction model validation [16]. The study flow and cohort construction are summarized in the workflow and cohort table (Fig. 1; Table S1).

Study population and cohort construction

The study population comprised patients aged 65 years or older with ASCVD. ASCVD included ischemic heart disease, ischemic stroke, and peripheral artery disease. These conditions were identified in MIMIC-IV using ICD-9 and ICD-10 diagnosis codes. In eICU, the same disease categories were identified from structured diagnoses and standardized diagnostic text in the diagnosis, pastHistory, apacheadmissiondx, and admissionDx fields. A separate principal ASCVD indicator identified admissions in which ASCVD was the principal diagnosis.

Eligibility criteria were age 65 years or older, an ICU stay of at least 24 h, survival through the first 24 h after ICU admission, and at least one positive measurement of blood urea nitrogen, creatinine, bicarbonate, and glucose during the first 24 h. Patients with chronic kidney disease, end-stage renal disease, cirrhosis, or malignancy were excluded. Only the first eligible ICU admission was retained for each patient. The analysis time origin was 24 h after ICU admission, and follow-up for the primary outcome began at that time.

Data extraction and variable harmonization

Demographic variables included age, sex, and race. Comorbidities included hypertension, diabetes mellitus, and heart failure. Clinical data obtained during the first 24 h after ICU admission included cardiogenic shock, mechanical ventilation, vasopressor use, aspirin use, statin use, mean arterial pressure, and heart rate. Vasopressors included norepinephrine, epinephrine, phenylephrine, vasopressin, and dopamine. In MIMIC-IV, vasopressor exposure was identified from actual administration times in inpatientevents, mechanical ventilation from procedureevents, and aspirin and statin exposure from overlap between prescription times and the observation window. Corresponding treatments in eICU were identified from structured records in infusionDrug, apacheApsVar, and medication.

Laboratory variables included blood urea nitrogen, creatinine, bicarbonate, glucose, sodium, anion gap, chloride, white blood cell count, potassium, hemoglobin, and platelet count. MIMIC-IV laboratory results were obtained from labevents and vital signs from chartevents. eICU laboratory results were obtained from lab and vital signs from vitalPeriodic. When mean arterial pressure was unavailable in eICU, meanBP from apacheApsVar was used. Variables were converted to common units, and values outside prespecified physiological ranges were removed. When multiple valid measurements were available within the observation window, their arithmetic mean was calculated. Variable definitions, source fields, units, and time windows are provided in the supplementary material (Table S2).

Derivation of TRS-3

TRS-3 was derived from blood urea nitrogen, creatinine, bicarbonate, and glucose. Its three components were the blood urea nitrogen-to-creatinine ratio, blood urea nitrogen-to-bicarbonate ratio, and glucose-to-bicarbonate ratio. The score structure was fixed before model training and database testing and was calculated as follows:

$$\text{TRS} - 3 = \frac{1}{3} \left[\frac{\frac{\text{BUN}}{\text{Cr}}}{20} + \frac{\frac{\text{BUN}}{\text{HCO}_3^-}}{\frac{20}{24}} + \frac{\frac{\text{Glucose}}{\text{HCO}_3^-}}{\frac{100}{24}} \right]$$

Normalization used clinical reference values of 20 mg/dL for blood urea nitrogen, 1 mg/dL for creatinine, 24 mmol/L for bicarbonate, and 100 mg/dL for glucose. The three components were converted to dimensionless relative values and averaged with equal weights. Neither the reference values nor the component weights were estimated from mortality outcomes. The algebraically simplified equivalent expression was:

$$\text{TRS} - 3 = \frac{\text{BUN}}{60 \times \text{Cr}} + \frac{\text{BUN}}{2.5 \times \text{HCO}_3^-} + \frac{2 \times \text{Glucose}}{25 \times \text{HCO}_3^-}$$

This structure jointly represents nitrogenous burden relative to renal handling, nitrogenous burden relative to acid-base status, and glucose relative to acid-base status and can be calculated directly from 4 routine laboratory tests. Continuous TRS-3 was standardized using the mean and standard deviation of the MIMIC-IV cohort, and effect estimates were reported per 1-standard-deviation increase. For categorical analyses, the 25th, 50th, and 75th percentiles in MIMIC-IV defined quartiles Q1 through Q4; the same cutoffs were applied unchanged to eICU. The blood urea nitrogen-to-creatinine ratio was the prespecified simplified comparator and underwent the same standardization and categorization procedures.

Outcome definitions

The primary outcome was in-hospital all-cause mortality after the 24-h time origin. Follow-up began 24 h after ICU admission and ended at in-hospital death or live discharge. In competing risk analyses, in-hospital death was the event of interest and live discharge was the competing event. All-cause mortality at 30 and 90 days after the time origin was also evaluated in MIMIC-IV. Dates of death were derived from in-hospital death times and available out-of-hospital death dates. Death within each time horizon was recorded as an event, and all other patients were censored at 30 or 90 days. Because eICU lacked harmonizable out-of-hospital death dates, only in-hospital mortality was evaluated in that cohort.

Missing data and preprocessing

Complete measurements of the 4 TRS-3 components were required for cohort inclusion. Descriptive analyses used observed values, and missingness was reported for candidate variables. Other continuous covariates in conventional regression models were imputed with the cohort-specific median. Complete-case analyses used the same covariates as the primary fully adjusted model. For prediction modeling, imputation, standardization, and numerical binning were fitted separately within each MIMIC-IV training fold. Parameters estimated in a training fold were applied to its corresponding validation fold and to eICU. eICU data were not used to estimate imputation or standardization parameters, select features, train models, calibrate probabilities, or determine the classification threshold.

Statistical analysis

Continuous variables are presented as medians with interquartile ranges and categorical variables as counts with percentages. Within MIMIC-IV, continuous variables were compared between patients who died in hospital and those discharged alive using the Mann-Whitney U test, and categorical variables were compared using the chi-square or Fisher exact test. Standardized mean differences were used to describe differences in variable distributions between the two cohorts.

The association between TRS-3 and in-hospital mortality was evaluated using generalized linear models with a binomial distribution, and results are reported as odds ratios (ORs) with 95% confidence intervals (CIs). Model 1 was unadjusted. Model 2 adjusted for age and sex. Model 3 additionally adjusted for hypertension, diabetes mellitus, heart failure, cardiogenic shock, principal ASCVD diagnosis, vasopressor use, mechanical ventilation, mean arterial pressure, heart rate, aspirin use, statin use, sodium, anion gap, chloride, white blood cell count, potassium, hemoglobin, and platelet count. Covariates were selected a priori according to their clinical relationships (Fig S1). TRS-3 entered the models either per 1-standard-deviation increase based on MIMIC-IV or as quartiles, with a test for trend across quartiles. Models in eICU used hospital-clustered robust standard errors.

Restricted cubic splines were incorporated into Model 3. The displayed range was fixed at the 1st through 99th percentiles of TRS-3 in MIMIC-IV, with knots at the 1st, 27.5th, 50th, 72.5th, and 99th percentiles and the median as the reference. The same boundaries, knots, and reference value were used in eICU. Overall association and nonlinearity were tested by likelihood ratio tests of nested models. The 95% CIs for the eICU curve were derived from a hospital-clustered robust covariance matrix.

Prespecified subgroups were defined by age, sex, hypertension, diabetes mellitus, heart failure, mechanical ventilation, vasopressor use, and aspirin use. The corresponding stratifying variable was removed from each subgroup model. Interaction was tested by adding a product term between TRS-3 and the subgroup variable. Within each cohort, interaction P values were adjusted using the Benjamini-Hochberg procedure. Effects were not estimated for strata with fewer than 60 patients or 20 events.

Mortality at 30 and 90 days was analyzed using Cox proportional hazards models, with hazard ratios (HRs) and 95% CIs reported. The proportional hazards assumption was assessed using Schoenfeld residuals. The cumulative incidence of in-hospital death was estimated in a competing risk framework, and differences across TRS-3 quartiles were assessed using Gray's test. All tests were two-sided, with $P < 0.05$ denoting statistical significance.

TabM model development and evaluation

The prediction outcome was in-hospital mortality after the 24-h time origin. The 4 prespecified input sets were TRS-3 alone, core clinical variables, core clinical variables plus TRS-3, and core clinical variables plus the 4 original laboratory components. Core clinical variables comprised age, sex, hypertension, diabetes mellitus, heart failure, cardiogenic shock, principal ASCVD diagnosis, vasopressor use, mechanical ventilation, mean arterial pressure, heart rate, aspirin use, and statin use. For an expanded clinical comparator, L1-penalized logistic regression with 10-fold cross-validation was used to select shared cross-database covariates from Model 3, with the regularization parameter chosen by the 1-standard-error rule.

All prediction models used TabM. The MIMIC-IV cohort underwent stratified, patient-level 3-fold cross-fitting. Within each training fold, median imputation, standardization, numerical binning, model fitting, and early stopping were performed independently, and out-of-fold predictions were generated only for the corresponding validation fold. Each of the 3 fold-specific models generated probabilities for all eICU patients, and the mean of these probabilities formed the output of the frozen ensemble. TabM used PiecewiseLinearEmbeddings version B with 48 numerical bins and 16-dimensional embeddings. The model contained 32 parallel submodels, 2 residual blocks, 512 hidden units, and a dropout rate of 0.1. AdamW optimization used a learning rate of 0.002 and weight decay of 0.0003. Training continued for up to 140 epochs, with early stopping based on validation-fold area under the precision-recall curve (AUPRC). MIMIC-IV out-of-fold probabilities were calibrated by 5-fold cross-fitted logistic calibration so that no patient's outcome contributed to the calibration parameters applied to that patient. A final calibrator was then fitted using all MIMIC-IV out-of-fold probabilities and applied unchanged to eICU. The classification threshold was selected from calibrated MIMIC-IV out-of-fold probabilities using the Youden index and then applied directly to eICU. Performance measures included AUROC, AUPRC, Brier score, calibration intercept, calibration slope, accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and F1 score. Confidence intervals were estimated using 1000 bootstrap resamples; eICU estimates used hospitals as the resampling unit in a cluster bootstrap. Model AUROCs were compared using paired DeLong tests, and the eICU AUC difference and 95% CI were additionally estimated using a hospital-clustered bootstrap. Calibration curves were constructed by deciles of predicted probability.

Decision curves compared each model with treat-all and treat-none strategies over threshold probabilities from 0.01 to 0.30. Incremental information was evaluated using the change in AUC, continuous net reclassification improvement, and integrated discrimination improvement. Model development and reporting followed TRIPOD + AI guidance [17].

Model interpretation, software, and ethics

SHapley Additive exPlanations (SHAP) were calculated for the model combining core clinical variables with TRS-3. In each cohort, 480 patients were sampled with stratification by outcome. For MIMIC-IV patients, GradientExplainer values were calculated using the model from the validation fold to which each patient belonged. For eICU patients, SHAP values were calculated using each of the 3 frozen fold-specific models and then averaged. Global explanations included mean absolute SHAP values and the direction of feature values. Partial dependence and SHAP locally weighted scatterplot smoothing (LOWESS) curves were also generated for TRS-3 and age. Partial dependence analyses used a stratified sample of 320 patients from each cohort, displayed the 2nd through 98th percentiles, and obtained 95% CIs from 500 patient-level bootstrap resamples. SHAP values were interpreted as contributions to model output and were not used to infer causal effects or define treatment thresholds.

Statistical and machine-learning analyses were conducted in Python and R. MIMIC-IV and eICU are deidentified databases accessed through PhysioNet after credentialing. This study used only existing database records and involved no recontact with patients or modification of clinical care.

Results

Study population and baseline characteristics

In MIMIC-IV, 11395 patients met the age, ASCVD, ICU length-of-stay, and 24-h time-origin criteria. After patients with chronic kidney disease, end-stage renal disease, cirrhosis, or malignancy were excluded, 8136 met the disease-related criteria, of whom 8037 had complete positive measurements for all TRS-3 components and entered the MIMIC-IV cohort. In-hospital death occurred in 714 patients (8.9%). The eICU cohort included 9686 patients from 190 hospitals, selected from 139367 first ICU admissions; 1002 patients died in hospital (10.3%). The two cohorts comprised 17723 patients and 1716 primary outcome events (Fig. 1; Table S1). The median age in MIMIC-IV was 76.00 years, and 59.6% of patients were men. Median age was 80.00 years among patients who died in hospital and 75.00 years among those discharged alive. Corresponding median values in these two groups were 25.33 and 17.00 mg/dL for blood urea nitrogen, 1.08 and 0.90 mg/dL for creatinine, 22.00 and 23.00 mmol/L for bicarbonate, and 145.88 and 134.00 mg/dL for glucose. Median TRS-3 was 1.45 among patients who died in hospital and 1.11 among those discharged alive ($P < 0.001$). Heart failure, cardiogenic shock, and mechanical ventilation were more frequent among patients who died, whereas aspirin and statin use were less frequent (Table 1). The 4 TRS-3 components had no missing values by design. In eICU, the missingness rate was 22.83% for anion gap and no greater than 5.56% for all other candidate variables (Table S3). Median TRS-3 was 1.12 in MIMIC-IV and 1.15 in eICU, with a standardized mean difference of 0.089. Several treatment variables and laboratory measures differed in distribution between the cohorts (Table S4).

Table 1
Baseline characteristics of the MIMIC-IV cohort according to in-hospital survival

Variable	Level	Overall	Survivor	Non-survivors	P value
Age, years		76.00 [70.00, 82.00]	75.00 [70.00, 82.00]	80.00 [73.00, 87.00]	< 0.001
Heart rate, beats/min		79.75 [72.07, 88.41]	79.48 [71.92, 87.74]	84.66 [73.28, 97.17]	< 0.001
Mean arterial pressure, mmHg		75.44 [70.77, 82.16]	75.37 [70.84, 82.00]	76.27 [69.63, 83.46]	0.614
Blood urea nitrogen, mg/dL		17.50 [13.50, 24.00]	17.00 [13.33, 23.00]	25.33 [17.00, 38.31]	< 0.001
Creatinine, mg/dL		0.90 [0.75, 1.10]	0.90 [0.75, 1.10]	1.08 [0.80, 1.53]	< 0.001
Bicarbonate, mmol/L		23.00 [21.00, 25.00]	23.00 [21.50, 25.00]	22.00 [19.00, 25.00]	< 0.001
Glucose, mg/dL		134.62 [116.33, 156.62]	134.00 [116.00, 154.20]	145.88 [119.69, 186.49]	< 0.001
TRS-3		1.12 [0.97, 1.36]	1.11 [0.96, 1.32]	1.45 [1.16, 1.91]	< 0.001
BUN/creatinine ratio		19.55 [15.93, 24.38]	19.38 [15.79, 24.00]	22.64 [17.62, 29.34]	< 0.001
Anion gap, mmol/L		12.50 [10.50, 15.00]	12.50 [10.50, 14.50]	14.67 [12.62, 17.00]	< 0.001
Sodium, mmol/L		137.80 [135.75, 140.00]	137.67 [135.75, 140.00]	139.00 [136.00, 141.95]	< 0.001
Potassium, mmol/L		4.20 [3.90, 4.48]	4.20 [3.90, 4.48]	4.13 [3.80, 4.50]	0.090
Chloride, mmol/L		105.50 [102.50, 108.00]	105.50 [102.60, 108.00]	104.50 [101.00, 108.00]	< 0.001
Hemoglobin, g/dL		10.50 [9.25, 11.90]	10.50 [9.25, 11.87]	10.70 [9.27, 12.00]	0.406
Red blood cell count, 10 ⁶ /uL		3.51 [3.13, 3.97]	3.50 [3.13, 3.96]	3.56 [3.11, 4.07]	0.235
White blood cell count, 10 ³ /uL		11.40 [8.67, 14.80]	11.27 [8.60, 14.60]	12.75 [9.54, 17.00]	< 0.001
Platelet count, 10 ³ /uL		173.67 [136.00, 225.50]	172.00 [135.10, 223.00]	196.17 [146.92, 258.08]	< 0.001
Calcium, mg/dL		8.35 [8.00, 8.80]	8.38 [8.00, 8.80]	8.30 [7.90, 8.75]	0.009
Male sex	Yes	4792 (59.6%)	4450 (60.8%)	342 (47.9%)	< 0.001
Hypertension	Yes	6487 (80.7%)	5939 (81.1%)	548 (76.8%)	0.005
Diabetes mellitus	Yes	2722 (33.9%)	2480 (33.9%)	242 (33.9%)	0.988
Heart failure	Yes	2730 (34.0%)	2382 (32.5%)	348 (48.7%)	< 0.001
Cardiogenic shock	Yes	461 (5.7%)	341 (4.7%)	120 (16.8%)	< 0.001
Ischemic heart disease	Yes	6595 (82.1%)	6066 (82.8%)	529 (74.1%)	< 0.001
Ischemic stroke	Yes	1381 (17.2%)	1182 (16.1%)	199 (27.9%)	< 0.001
Peripheral artery disease	Yes	1107 (13.8%)	1000 (13.7%)	107 (15.0%)	0.325
Data are median [interquartile range] for continuous variables and n (%) for categorical variables. P values were obtained using the Wilcoxon rank-sum test for continuous variables and the chi-square or Fisher exact test for categorical variables, as appropriate. TRS-3, three-ratio score; ASCVD, atherosclerotic cardiovascular disease; BUN, blood urea nitrogen.					

Variable	Level	Overall	Survivor	Non-survivors	P value
ASCVD as admission diagnosis	Yes	2997 (37.3%)	2789 (38.1%)	208 (29.1%)	< 0.001
Mechanical ventilation	Yes	3709 (46.1%)	3347 (45.7%)	362 (50.7%)	0.011
Vasopressor use	Yes	3580 (44.5%)	3252 (44.4%)	328 (45.9%)	0.432
Aspirin use	Yes	5659 (70.4%)	5303 (72.4%)	356 (49.9%)	< 0.001
Statin use	Yes	5642 (70.2%)	5296 (72.3%)	346 (48.5%)	< 0.001
Race/ethnicity	White	5684 (70.7%)	5254 (71.7%)	430 (60.2%)	< 0.001
	Black	420 (5.2%)	375 (5.1%)	45 (6.3%)	
	Hispanic	144 (1.8%)	135 (1.8%)	9 (1.3%)	
	Asian	173 (2.2%)	155 (2.1%)	18 (2.5%)	
	Other/unknown	1616 (20.1%)	1404 (19.2%)	212 (29.7%)	
Data are median [interquartile range] for continuous variables and n (%) for categorical variables. P values were obtained using the Wilcoxon rank-sum test for continuous variables and the chi-square or Fisher exact test for categorical variables, as appropriate. TRS-3, three-ratio score; ASCVD, atherosclerotic cardiovascular disease; BUN, blood urea nitrogen.					

Association between TRS-3 and in-hospital mortality

With live discharge treated as a competing event, cumulative incidence curves for in-hospital death separated progressively across TRS-3 quartiles in both cohorts. The 30-day cumulative incidences in quartiles Q1 through Q4 were approximately 3.3%, 4.3%, 7.3%, and 19.5% in MIMIC-IV and 3.7%, 6.5%, 8.2%, and 19.1% in eICU. Gray's test yielded $P < 0.001$ in both cohorts (Fig. 2A, B; Table S5). After full adjustment, each 1-standard-deviation increase in TRS-3 was associated with an OR for in-hospital mortality of 1.39 (95% CI: 1.28–1.51; $P < 0.001$) in MIMIC-IV and 1.31 (95% CI: 1.22–1.41; $P < 0.001$) in eICU. Compared with Q1, Q4 had an OR of 3.81 (95% CI: 2.80–5.20) in MIMIC-IV and 2.70 (95% CI: 2.10–3.48) in eICU. Tests for trend across quartiles yielded $P < 0.001$ in both cohorts (Table 2). For each 1-standard-deviation increase in the blood urea nitrogen-to-creatinine ratio, the fully adjusted ORs were 1.17 (95% CI: 1.09–1.26) and 1.10 (95% CI: 1.03–1.18), respectively. Associations for the blood urea nitrogen-to-bicarbonate and glucose-to-bicarbonate ratios were also positive in both cohorts (Table S6).

Table 2
Associations of TRS-3 and the BUN-to-creatinine ratio with in-hospital mortality

Cohort	Index and scale	Model 1	Model 2	Model 3
MIMIC-IV cohort	TRS-3, Per 1 MIMIC-IV SD	1.77 (1.66–1.88), P < 0.001	1.71 (1.61–1.82), P < 0.001	1.39 (1.28–1.51), P < 0.001
	TRS-3, Q1	1.00 (reference)	1.00 (reference)	1.00 (reference)
	TRS-3, Q2	1.46 (1.05–2.02), P = 0.023	1.44 (1.04–2.00), P = 0.027	1.49 (1.07–2.09), P = 0.020
	TRS-3, Q3	2.47 (1.83–3.33), P < 0.001	2.32 (1.72–3.13), P < 0.001	1.91 (1.39–2.61), P < 0.001
	TRS-3, Q4	7.72 (5.89–10.14), P < 0.001	6.68 (5.08–8.79), P < 0.001	3.81 (2.80–5.20), P < 0.001
	TRS-3, P for trend	< 0.001	< 0.001	< 0.001
MIMIC-IV cohort	BUN/creatinine ratio, Per 1 MIMIC-IV SD	1.40 (1.32–1.49), P < 0.001	1.31 (1.23–1.39), P < 0.001	1.17 (1.09–1.26), P < 0.001
	BUN/creatinine ratio, Q1	1.00 (reference)	1.00 (reference)	1.00 (reference)
	BUN/creatinine ratio, Q2	0.96 (0.74–1.24), P = 0.755	0.91 (0.70–1.18), P = 0.467	0.98 (0.75–1.29), P = 0.900
	BUN/creatinine ratio, Q3	1.43 (1.13–1.81), P = 0.003	1.21 (0.95–1.54), P = 0.119	1.18 (0.91–1.53), P = 0.207
	BUN/creatinine ratio, Q4	2.55 (2.05–3.17), P < 0.001	1.94 (1.54–2.43), P < 0.001	1.50 (1.17–1.92), P = 0.001
	BUN/creatinine ratio, P for trend	< 0.001	< 0.001	< 0.001
eICU cohort	TRS-3, Per 1 MIMIC-IV SD	1.60 (1.51–1.69), P < 0.001	1.59 (1.50–1.68), P < 0.001	1.31 (1.22–1.41), P < 0.001
	TRS-3, Q1	1.00 (reference)	1.00 (reference)	1.00 (reference)
	TRS-3, Q2	1.81 (1.43–2.30), P < 0.001	1.77 (1.39–2.24), P < 0.001	1.41 (1.10–1.81), P = 0.007
	TRS-3, Q3	2.28 (1.81–2.88), P < 0.001	2.20 (1.74–2.78), P < 0.001	1.49 (1.15–1.92), P = 0.002
	TRS-3, Q4	6.08 (4.95–7.48), P < 0.001	5.81 (4.71–7.16), P < 0.001	2.70 (2.10–3.48), P < 0.001
	TRS-3, P for trend	< 0.001	< 0.001	< 0.001
eICU cohort	BUN/creatinine ratio, Per 1 MIMIC-IV SD	1.21 (1.15–1.28), P < 0.001	1.19 (1.12–1.26), P < 0.001	1.10 (1.03–1.18), P = 0.004
	BUN/creatinine ratio, Q1	1.00 (reference)	1.00 (reference)	1.00 (reference)
	BUN/creatinine ratio, Q2	0.97 (0.79–1.20), P = 0.783	0.94 (0.76–1.16), P = 0.571	0.88 (0.69–1.13), P = 0.318
	BUN/creatinine ratio, Q3	1.43 (1.20–1.71), P < 0.001	1.35 (1.14–1.61), P < 0.001	1.29 (1.06–1.56), P = 0.010
	BUN/creatinine ratio, Q4	1.94 (1.64–2.29), P < 0.001	1.80 (1.51–2.13), P < 0.001	1.44 (1.18–1.74), P < 0.001

Values are odds ratios (95% confidence intervals) unless otherwise indicated. Continuous estimates are reported per 1 standard deviation of the MIMIC-IV cohort. Quartile boundaries were derived in the MIMIC-IV cohort and applied unchanged to the eICU cohort. Model 1 was unadjusted. Model 2 included age and sex. Model 3 additionally included hypertension, diabetes, heart failure, cardiogenic shock, principal ASCVD admission, vasopressor use, mechanical ventilation, mean arterial pressure, heart rate, aspirin use, statin use, sodium, anion gap, chloride, white blood cell count, potassium, haemoglobin, and platelet count. eICU estimates used hospital-cluster robust variance.

Cohort	Index and scale	Model 1	Model 2	Model 3
	BUN/creatinine ratio, P for trend	< 0.001	< 0.001	< 0.001

Values are odds ratios (95% confidence intervals) unless otherwise indicated. Continuous estimates are reported per 1 standard deviation of the MIMIC-IV cohort. Quartile boundaries were derived in the MIMIC-IV cohort and applied unchanged to the eICU cohort. Model 1 was unadjusted. Model 2 included age and sex. Model 3 additionally included hypertension, diabetes, heart failure, cardiogenic shock, principal ASCVD admission, vasopressor use, mechanical ventilation, mean arterial pressure, heart rate, aspirin use, statin use, sodium, anion gap, chloride, white blood cell count, potassium, haemoglobin, and platelet count. eICU estimates used hospital-cluster robust variance.

Secondary outcomes and complete-case analyses

In MIMIC-IV, each 1-standard-deviation increase in TRS-3 was associated with fully adjusted HRs of 1.25 (95% CI: 1.19–1.30) for 30-day all-cause mortality and 1.24 (95% CI: 1.19–1.29) for 90-day all-cause mortality; both P values were < 0.001 (Table S7). Complete-case analyses included 7974 MIMIC-IV patients and 6935 eICU patients, with fully adjusted ORs of 1.38 (95% CI: 1.28–1.50) and 1.29 (95% CI: 1.19–1.41), respectively, consistent in direction with the primary analyses (Table S8). Schoenfeld residual tests found no evidence that the TRS-3 term violated the proportional hazards assumption in the 30-day or 90-day models; P = 0.957 in the primary 90-day model (Table S9).

Continuous associations and subgroup analyses

Restricted cubic spline analyses showed an overall association between TRS-3 and in-hospital mortality in both MIMIC-IV and eICU (P < 0.001 for each). Evidence of nonlinearity was detected in both cohorts, with P < 0.001 and P = 0.026, respectively. The MIMIC-IV curve rose with increasing TRS-3 across the main data range, and its 95% CI widened in the high-value tail as the number of observations decreased. The eICU curve increased throughout the displayed range (Fig. 3A, B; Table S10). All prespecified subgroup estimates exceeded 1. In MIMIC-IV, the ORs were 1.23 (95% CI: 1.10–1.37) in women and 1.62 (95% CI: 1.44–1.83) in men, with an FDR-adjusted interaction P = 0.018. The ORs were 1.25 (95% CI: 1.11–1.40) in patients who did not receive vasopressors and 1.46 (95% CI: 1.30–1.64) in those who did, with an FDR-adjusted interaction P = 0.011. The corresponding ORs were 1.14 (95% CI: 1.03–1.27) and 1.60 (95% CI: 1.42–1.80) for patients without and with aspirin exposure, with an FDR-adjusted interaction P < 0.001. All stratum-specific ORs in eICU exceeded 1, and all FDR-adjusted interaction P values were > 0.05 (Fig. 4A, B; Table S11).

Discrimination of the TabM models

L1-penalized logistic regression retained 18 of 20 shared candidate variables; coefficients for diabetes mellitus and vasopressor use were shrunk to 0. These selected variables formed the expanded clinical comparator (Fig. 5A, B; Table S12). In MIMIC-IV, AUROCs were 0.711 (95% CI: 0.689–0.730) for TRS-3 alone, 0.788 (95% CI: 0.771–0.804) for the core clinical model, 0.829 (95% CI: 0.813–0.842) for the core clinical model plus TRS-3, and 0.833 (95% CI: 0.817–0.845) for the core clinical model plus the 4 original components. Corresponding AUPRCs were 0.209, 0.271, 0.312, and 0.327. In eICU, AUROCs for the 4 models were 0.707 (95% CI: 0.687–0.726), 0.717 (95% CI: 0.701–0.734), 0.760 (95% CI: 0.744–0.775), and 0.769 (95% CI: 0.753–0.786), with AUPRCs of 0.220, 0.236, 0.303, and 0.301, respectively (Fig. 5C, D; Table 3; Fig S2A, B; Table S13).

Table 3
Performance of prespecified TabM input sets in the MIMIC-IV and eICU cohorts

Cohort	Model	N/events	AUROC (95% CI)	AUPRC (95% CI)	Brier score (95% CI)	Calibration intercept	Calibration slope	Sensitivity	Specificity
MIMIC-IV cohort	TRS-3-only TabM	8037/714	0.711 (0.689–0.730)	0.209 (0.190–0.231)	0.076 (0.075–0.077)	-0.005	0.997	0.562	0.789
MIMIC-IV cohort	Core clinical TabM	8037/714	0.788 (0.771–0.804)	0.271 (0.246–0.297)	0.072 (0.071–0.074)	-0.005	0.997	0.780	0.675
MIMIC-IV cohort	Core clinical + TRS-3 TabM	8037/714	0.829 (0.813–0.842)	0.312 (0.285–0.343)	0.069 (0.068–0.071)	-0.004	0.998	0.779	0.749
MIMIC-IV cohort	Core clinical + raw components TabM	8037/714	0.833 (0.817–0.845)	0.327 (0.297–0.358)	0.069 (0.067–0.071)	-0.001	0.999	0.787	0.736
eICU cohort	TRS-3-only TabM	9686/1002	0.707 (0.687–0.726)	0.220 (0.196–0.246)	0.087 (0.080–0.093)	-0.015	0.956	0.562	0.743
eICU cohort	Core clinical TabM	9686/1002	0.717 (0.701–0.734)	0.236 (0.211–0.261)	0.088 (0.082–0.093)	-0.276	1.011	0.822	0.435
eICU cohort	Core clinical + TRS-3 TabM	9686/1002	0.760 (0.744–0.775)	0.303 (0.268–0.337)	0.084 (0.078–0.089)	-0.130	1.101	0.825	0.538
eICU cohort	Core clinical + raw components TabM	9686/1002	0.769 (0.753–0.786)	0.301 (0.266–0.333)	0.085 (0.079–0.090)	-0.204	1.129	0.865	0.493

The MIMIC-IV estimates use patient-level out-of-sample probabilities generated by prespecified cross-fitting. The eICU estimates use predictions from the frozen MIMIC-IV ensemble without external retuning. Threshold-dependent measures used thresholds selected within the MIMIC-IV fitting procedure and then applied unchanged. AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve.

Calibration, net benefit, and incremental information

For the core clinical model plus TRS-3, Brier scores were 0.069 in MIMIC-IV and 0.084 in eICU. Calibration intercepts were -0.004 and -0.130 , and calibration slopes were 0.998 and 1.101, respectively. The MIMIC-IV calibration curve closely followed the ideal line, whereas deviation was observed at higher predicted probabilities in eICU (Fig. 6A, C; Table 3). Decision curve analysis showed that the core clinical model plus TRS-3 had greater net benefit than treat-none and treat-all strategies over the prespecified threshold range and greater net benefit than the core clinical model over most of that range (Fig. 6B, D). Adding TRS-3 to the core clinical model increased AUROC from 0.788 to 0.829 in MIMIC-IV, a difference of 0.041 (paired DeLong $P < 0.001$). Continuous net reclassification improvement was 0.531 (95% CI: 0.456–0.605), and integrated discrimination improvement was 0.038 (95% CI: 0.031–0.045). In eICU, AUROC increased from 0.717 to 0.760, a difference of 0.042 (paired DeLong $P < 0.001$). Continuous net reclassification improvement was 0.492 (95% CI: 0.418–0.565), and integrated discrimination improvement was 0.030 (95% CI: 0.025–0.036) (Fig. 6E, F; Table 4). Compared with the core clinical model plus TRS-3, the model containing the 4 original components had an AUROC difference of 0.004 in MIMIC-IV (paired DeLong $P = 0.160$) and 0.010 in eICU (paired DeLong $P = 0.003$), with a hospital-clustered bootstrap 95% CI of 0.004–0.015 in eICU (Fig. 6E, F; Table S14). At the MIMIC-IV-derived threshold of 0.100, the core clinical model plus TRS-3 had a sensitivity of 77.9%, specificity of 74.9%, positive predictive value of 23.3%, and negative predictive value of 97.2% in MIMIC-IV. In eICU, sensitivity was 82.5%, specificity 53.8%, positive predictive value 17.1%, and negative predictive value 96.4% (Fig. 6G, H; Fig S3; Table 3).

Table 4
Incremental discrimination and reclassification relative to the core clinical TabM model

Cohort	Model	AUROC (95% CI)	Delta AUROC	DeLong P	Continuous NRI (95% CI)	NRI P	IDI (95% CI)	IDI P
MIMIC-IV cohort	Core clinical + TRS-3 TabM	0.829 (0.813–0.842)	0.041	< 0.001	0.531 (0.456–0.605)	0.002	0.038 (0.031–0.045)	0.002
MIMIC-IV cohort	Core clinical + raw components TabM	0.833 (0.817–0.845)	0.044	< 0.001	0.549 (0.474–0.622)	0.002	0.046 (0.038–0.054)	0.002
eICU cohort	Core clinical + TRS-3 TabM	0.760 (0.744–0.775)	0.042	< 0.001	0.492 (0.418–0.565)	0.002	0.030 (0.025–0.036)	0.002
eICU cohort	Core clinical + raw components TabM	0.769 (0.753–0.786)	0.052	< 0.001	0.506 (0.442–0.571)	0.002	0.043 (0.037–0.049)	0.002
All models were compared in the same patients within each cohort. Paired AUROC differences were tested using DeLong tests. Continuous net reclassification improvement and integrated discrimination improvement were estimated by patient-level bootstrap in MIMIC-IV and hospital-cluster bootstrap in eICU. NRI, net reclassification improvement; IDI, integrated discrimination improvement.								

Interpretation of the TabM model

In the core clinical model plus TRS-3, TRS-3 had the highest mean absolute SHAP value in both MIMIC-IV (0.0431) and eICU (0.0433). Mechanical ventilation ranked second, with mean absolute SHAP values of 0.0250 and 0.0387, respectively. Higher TRS-3 values were concentrated in the positive SHAP region, whereas lower TRS-3 values were concentrated in the negative SHAP region (Fig. 7A, B). Mean absolute SHAP values across features were strongly correlated between the two cohorts (Spearman rho = 0.89; P < 0.001; Fig S4). SHAP LOWESS and partial dependence curves both showed increasing predicted probabilities of in-hospital death with higher TRS-3 across the main data range. SHAP LOWESS and partial dependence curves for age also increased in both cohorts (Fig. 7C-F).

Discussion

This study included 17723 older ICU patients with ASCVD. Higher TRS-3 was associated with a greater probability of in-hospital death in both MIMIC-IV and eICU, and the quartile and spline analyses showed a continuous risk gradient. Adding TRS-3 to the core clinical TabM model increased AUROC by 0.041 in MIMIC-IV and 0.042 in eICU, with concordant net reclassification improvement and integrated discrimination improvement. Models containing the 4 original components had AUROCs that were 0.004 and 0.010 higher than those containing TRS-3, indicating that the fixed score retained most, but not all, of the discriminative information in its components.

Previous studies of ICU patients with coronary artery disease have associated both the stress hyperglycemia ratio and glycemic variability with mortality [18]. Cohorts of critically ill patients with acute myocardial infarction in the United States and China have reported similar associations for the stress hyperglycemia ratio [19]. The glucose-to-bicarbonate component of TRS-3 jointly represents glucose burden and buffering status at the time of measurement, which differs from relative hyperglycemia measures that require glycated hemoglobin. The present findings do not define a target for glucose-lowering treatment. A meta-analysis in heart failure associated a higher blood urea nitrogen-to-creatinine ratio with all-cause mortality [20]. In clinical trial data spanning the spectrum of ejection fraction, this ratio retained prognostic information after adjustment for estimated glomerular filtration rate and N-terminal pro-B-type natriuretic peptide [21]. In the present study, the standardized effect estimate for the blood urea nitrogen-to-creatinine ratio was smaller than that for TRS-3, and the blood urea nitrogen-to-bicarbonate and glucose-to-bicarbonate components were positively associated with mortality in both cohorts. TRS-3 therefore captured information beyond the blood urea nitrogen-to-creatinine ratio alone.

Large ICU cohort studies have associated early metabolic acidosis with in-hospital mortality [22]. Among ICU patients with heart failure, bicarbonate also modifies the relationship between serum chloride and mortality [23]. By placing bicarbonate in the denominator of two components, TRS-3 expresses nitrogenous burden and glucose relative to buffering status. Lactate

accumulation, renal acidosis, chloride load, and respiratory compensation may all affect the score. TRS-3 is therefore suited to integrated risk description but does not replace evaluation of the specific cause of an acid-base disturbance.

Studies of the stress hyperglycemia ratio in critically ill patients have reported both linear and nonlinear risk relationships [24]. A large emergency care cohort also associated abnormal admission glucose with mortality and cardiovascular outcomes [25]. TRS-3 uses prespecified reference values and equal component weights rather than weights estimated from mortality outcomes. The nonlinear curves in both cohorts indicate that a single cutoff cannot adequately summarize the score-risk relationship, and they do not establish biological synergy among the three components. Existing cardiac critical care scores use blood urea nitrogen, anion gap, and clinical status to estimate in-hospital mortality [26]. Another parsimonious score combining blood urea nitrogen, oxygen saturation, blood pressure, mechanical ventilation, age, and anion gap has undergone external evaluation [27]. TRS-3 instead provides a fixed formula that condenses 4 routine laboratory results. The modestly higher discrimination of the model containing all 4 original components indicates that a full clinical model retains more patient-level information than the summary score.

Electronic health record models often show changes in discrimination and calibration across data sources [28]. Independent hospital data can test a frozen model under differences in case mix, documentation, and clinical practice [29]. The eICU cohort included 190 hospitals, and the core clinical model plus TRS-3 retained an AUROC increment, with a calibration slope of 1.101. At the threshold fixed in MIMIC-IV, sensitivity was 82.5% in eICU, whereas specificity declined to 53.8%, indicating that probability calibration and classification thresholds require reassessment across databases. Discrimination and calibration describe distinct properties of a prediction model and should both be reported during external evaluation [30]. The present decision curves quantify model-derived net benefit over prespecified thresholds but do not show that the model changes treatment decisions or patient outcomes. SHAP describes the computational contribution of variables to predictions from electronic health record models [31]. TRS-3 ranked first by mean absolute SHAP value in both cohorts, and feature-importance rankings were correlated across databases.

The analysis used a fixed TRS-3 formula, patient-level cross-fitting in MIMIC-IV, independent multicenter evaluation in eICU, and a common 24-h time origin for association and prediction analyses. Several limitations arise from the retrospective electronic health record design. ASCVD ascertainment differed between databases, cohort inclusion required complete measurements of the four score components, treatment timing could not be fully harmonized, and residual confounding remained possible. Several covariates had substantial missingness in eICU, and the fixed threshold was less specific than in MIMIC-IV. Interactions with sex, vasopressor use, and aspirin use in MIMIC-IV were not reproduced in eICU and should be regarded as exploratory. Prospective multicenter studies should retain the fixed score formula while evaluating probability calibration, classification thresholds, and effects on clinical workflow.

Conclusions

A higher TRS-3 at 24 h after ICU admission was associated with a greater probability of in-hospital death among older critically ill patients with ASCVD, with concordant associations in MIMIC-IV and eICU. TRS-3 condenses information on nitrogenous burden, renal handling, bicarbonate, and glucose from 4 routine laboratory tests and adds discriminative information to a core clinical TabM model. The model containing the original components showed slightly greater discrimination, and specificity at the fixed threshold declined in eICU; prospective calibration and clinical impact assessment are therefore required before TRS-3 is used as a simplified risk score.

Abbreviations

ASCVD
atherosclerotic cardiovascular disease
AUPRC
area under the precision-recall curve
AUROC
area under the receiver operating characteristic curve
BUN
blood urea nitrogen

CI
confidence interval
Cr
creatinine
DCA
decision curve analysis
HCO₃⁻
bicarbonate
IDI
integrated discrimination improvement
ICU
intensive care unit
NRI
net reclassification improvement
OR
odds ratio
RCS
restricted cubic spline
SHAP
SHapley Additive exPlanations
TRS-3
three-ratio score.

Declarations

Ethics approval and consent to participate

The original establishment of the MIMIC-IV and eICU databases was approved by the Institutional Review Boards of Beth Israel Deaconess Medical Center and the Massachusetts Institute of Technology, with the requirement for informed consent waived. This study used publicly available deidentified data and involved no recontact with participants. The Medical Ethics Committee of Yiyang Central Hospital reviewed the study and determined that it met the criteria for exemption from ethical review (Ethics Opinion No. 2026053); the exemption opinion letter is provided in the supplementary material. This study was reported in accordance with the STROBE statement.

Consent for publication

Not applicable. This manuscript reports aggregate analyses of de-identified data and contains no individual participant details.

Availability of data and materials

MIMIC-IV version 3.1 and the eICU Collaborative Research Database version 2.0 are available through PhysioNet after completion of credentialing and a data use agreement. Patient-level source data cannot be redistributed under the database use agreements. Variable definitions, model configurations, and analysis scripts used to generate the results are available from the corresponding author subject to the applicable database use agreements.

Competing interests

The authors declare that they have no competing interests.

Funding

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

Authors' contributions

Shiwen Cui drafted the manuscript and performed the data analysis. Xiaoyi Yao contributed to data curation, variable harmonization, literature review, and preparation of tables and figures. Xiang Zhou, Yi Tao and Liuyin Zhou assisted with data checking, visualization, and interpretation of the results. Yanqiong Xue and Weiqiong Chen contributed to study design, supervised the study, and critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the teams responsible for maintaining the MIMIC-IV and eICU databases and providing access to deidentified critical care data.

Authors' information

Not applicable.

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, Barengo NC, Beaton AZ, Benjamin EJ, Benziger CP, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990–2019: Update From the GBD 2019 Study. *J Am Coll Cardiol*. 2020;76(25):2982–3021.
2. Forman DE, Maurer MS, Boyd C, Brindis R, Salive ME, Horne FM, Bell SP, Fulmer T, Reuben DB, Zieman S, et al. Multimorbidity in Older Adults With Cardiovascular Disease. *J Am Coll Cardiol*. 2018;71(19):2149–61.
3. Rich MW, Chyun DA, Skolnick AH, Alexander KP, Forman DE, Kitzman DW, Maurer MS, McClurken JB, Resnick BM, Shen WK, et al. Knowledge Gaps in Cardiovascular Care of the Older Adult Population: A Scientific Statement From the American Heart Association, American College of Cardiology, and American Geriatrics Society. *Circulation*. 2016;133(21):2103–22.
4. Jentzer JC, Rossello X. Past, present, and future of mortality risk scores in the contemporary cardiac intensive care unit. *Eur Heart J Acute Cardiovasc Care*. 2021;10(8):940–6.
5. Paulus MC, Melchers M, van Es A, Kouw IWK, van Zanten ARH. The urea-to-creatinine ratio as an emerging biomarker in critical care: a scoping review and meta-analysis. *Crit Care*. 2025;29(1):175.
6. Aronson D, Hammerman H, Beyar R, Yalonetsky S, Kapeliovich M, Markiewicz W, Goldberg A. Serum blood urea nitrogen and long-term mortality in acute ST-elevation myocardial infarction. *Int J Cardiol*. 2008;127(3):380–5.
7. Achanti A, Szerlip HM. Acid-Base Disorders in the Critically Ill Patient. *Clin J Am Soc Nephrol*. 2023;18(1):102–12.
8. Canova TJ, Lipps K, Dahiya G, Hillerson DB, Kashani KB, Jentzer JC. Admission Acid-Base Status and Mortality in Cardiac Intensive Care Unit Patients. *J Intensive Care Med*. 2026;41(7):670–82.
9. Roberts GW, Quinn SJ, Valentine N, Alhawassi T, O'Dea H, Stranks SN, Burt MG, Doogue MP. Relative Hyperglycemia, a Marker of Critical Illness: Introducing the Stress Hyperglycemia Ratio. *J Clin Endocrinol Metab*. 2015;100(12):4490–7.
10. Esdaile H, Khan S, Mayet J, Oliver N, Reddy M, Shah ASV. The association between the stress hyperglycaemia ratio and mortality in cardiovascular disease: a meta-analysis and systematic review. *Cardiovasc Diabetol*. 2024;23(1):412.
11. Wang F, Guo Y, Tang Y, Zhao S, Xuan K, Mao Z, Lu R, Hou R, Zhu X. Combined assessment of stress hyperglycemia ratio and glycemic variability to predict all-cause mortality in critically ill patients with atherosclerotic cardiovascular diseases across different glucose metabolic states: an observational cohort study with machine learning. *Cardiovasc Diabetol*. 2025;24(1):199.
12. Tao LC, Xu JN, Wang TT, Hua F, Li JJ. Triglyceride-glucose index as a marker in cardiovascular diseases: landscape and limitations. *Cardiovasc Diabetol*. 2022;21(1):68.
13. Abdu FA, Mohammed AQ, Zhang W, Alifu J, Liu L, Yin G, Liao Y, Zhu G, Che W. Head-to-head comparison of fasting-insulin-based versus non-insulin-based insulin resistance surrogates for predicting all-cause and cardiovascular mortality in hypertensive adults across the glycemic continuum: a prospective cohort study. *Cardiovasc Diabetol* 2026, 25(1).
14. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, Pollard TJ, Hao S, Moody B, Gow B, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10(1):1.
15. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci Data*. 2018;5:180178.

16. Steyerberg EW, Harrell FE Jr.. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol.* 2016;69:245–7.
17. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, et al. TRIPOD + AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ.* 2024;385:e078378.
18. He HM, Zheng SW, Xie YY, Wang Z, Jiao SQ, Yang FR, Li XX, Li J, Sun YH. Simultaneous assessment of stress hyperglycemia ratio and glycemic variability to predict mortality in patients with coronary artery disease: a retrospective cohort study from the MIMIC-IV database. *Cardiovasc Diabetol.* 2024;23(1):61.
19. Liu J, Zhou Y, Huang H, Liu R, Kang Y, Zhu T, Wu J, Gao Y, Li Y, Wang C, et al. Impact of stress hyperglycemia ratio on mortality in patients with critical acute myocardial infarction: insight from american MIMIC-IV and the chinese CIN-II study. *Cardiovasc Diabetol.* 2023;22(1):281.
20. Zhou Y, Zhao Q, Liu Z, Gao W. Blood urea nitrogen/creatinine ratio in heart failure: Systematic review and meta-analysis. *PLoS ONE.* 2024;19(5):e0303870.
21. Tolomeo P, Butt JH, Kondo T, Campo G, Desai AS, Jhund PS, Kober L, Lefkowitz MP, Rouleau JL, Solomon SD, et al. Independent prognostic importance of blood urea nitrogen to creatinine ratio in heart failure. *Eur J Heart Fail.* 2024;26(2):245–56.
22. Mochizuki K, Fujii T, Paul E, Anstey M, Pilcher DV, Bellomo R. Early metabolic acidosis in critically ill patients: a binational multicentre study. *Crit Care Resusc.* 2021;23(1):67–75.
23. Tan Z, Liu Y, Hong K. The association between serum chloride and mortality in ICU patients with heart failure: The impact of bicarbonate. *Int J Cardiol.* 2024;399:131672.
24. Song G, Liu X, Lu Z, Guan J, Chen X, Li Y, Liu G, Wang G, Ma F. Relationship between stress hyperglycaemic ratio (SHR) and critical illness: a systematic review. *Cardiovasc Diabetol.* 2025;24(1):188.
25. Djupsjo C, Kuhl J, Andersson T, Lundback M, Holzmann MJ, Nystrom T. Admission glucose as a prognostic marker for all-cause mortality and cardiovascular disease. *Cardiovasc Diabetol.* 2022;21(1):258.
26. Jentzer JC, Anavekar NS, Bennett C, Murphree DH, Keegan MT, Wiley B, Morrow DA, Murphy JG, Bell MR, Barsness GW. Derivation and Validation of a Novel Cardiac Intensive Care Unit Admission Risk Score for Mortality. *J Am Heart Assoc.* 2019;8(17):e013675.
27. Yamga E, Mantena S, Rosen D, Bucholz EM, Yeh RW, Celi LA, Ustun B, Butala NM. Optimized Risk Score to Predict Mortality in Patients With Cardiogenic Shock in the Cardiac Intensive Care Unit. *J Am Heart Assoc.* 2023;12(13):e029232.
28. Guan C, Gong A, Zhao Y, Yin C, Geng L, Liu L, Yang X, Lu J, Xiao B. Interpretable machine learning model for new-onset atrial fibrillation prediction in critically ill patients: a multi-center study. *Crit Care.* 2024;28(1):349.
29. Luo X, Li B, Zhu R, Tai Y, Wang Z, He Q, Zhao Y, Bi X, Wu C. Development and validation of an interpretable machine learning model for predicting in-hospital mortality for ischemic stroke patients in ICU. *Int J Med Inf.* 2025;198:105874.
30. Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW. Topic Group 'Evaluating diagnostic t, prediction models' of the Si: Calibration: the Achilles heel of predictive analytics. *BMC Med.* 2019;17(1):230.
31. Payrovnaziri SN, Chen Z, Rengifo-Moreno P, Miller T, Bian J, Chen JH, Liu X, He Z. Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. *J Am Med Inf Assoc.* 2020;27(7):1173–85.

Figures

Cohort construction and analysis framework

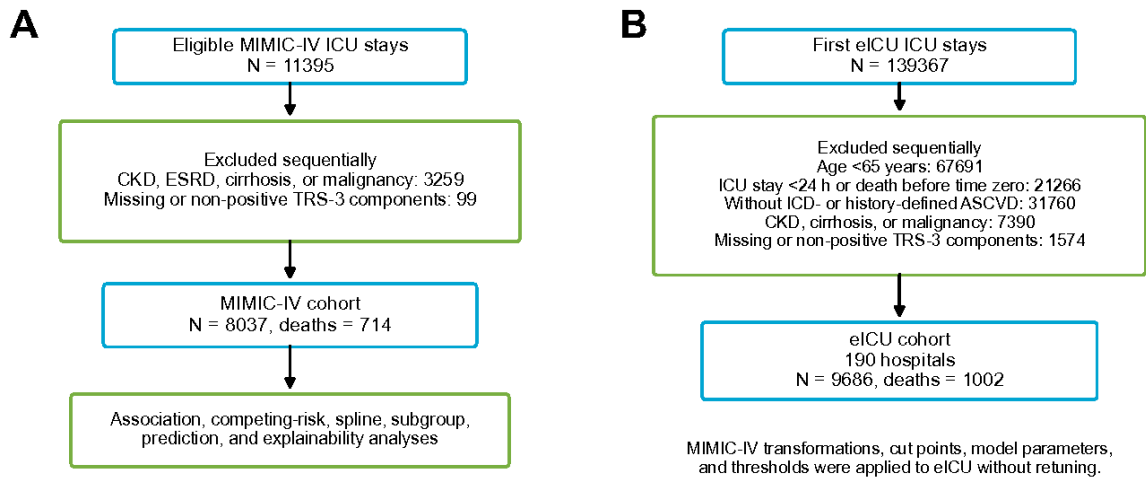


Figure 1

Cohort selection and analysis workflow

A, selection of the MIMIC-IV cohort. B, selection of the eICU cohort and cross-database application of frozen parameters. The TRS-3 components and percentiles, preprocessing parameters, model parameters, probability calibrator, and classification threshold were all determined in MIMIC-IV and kept unchanged in eICU. CKD, chronic kidney disease; ESRD, end-stage renal disease.

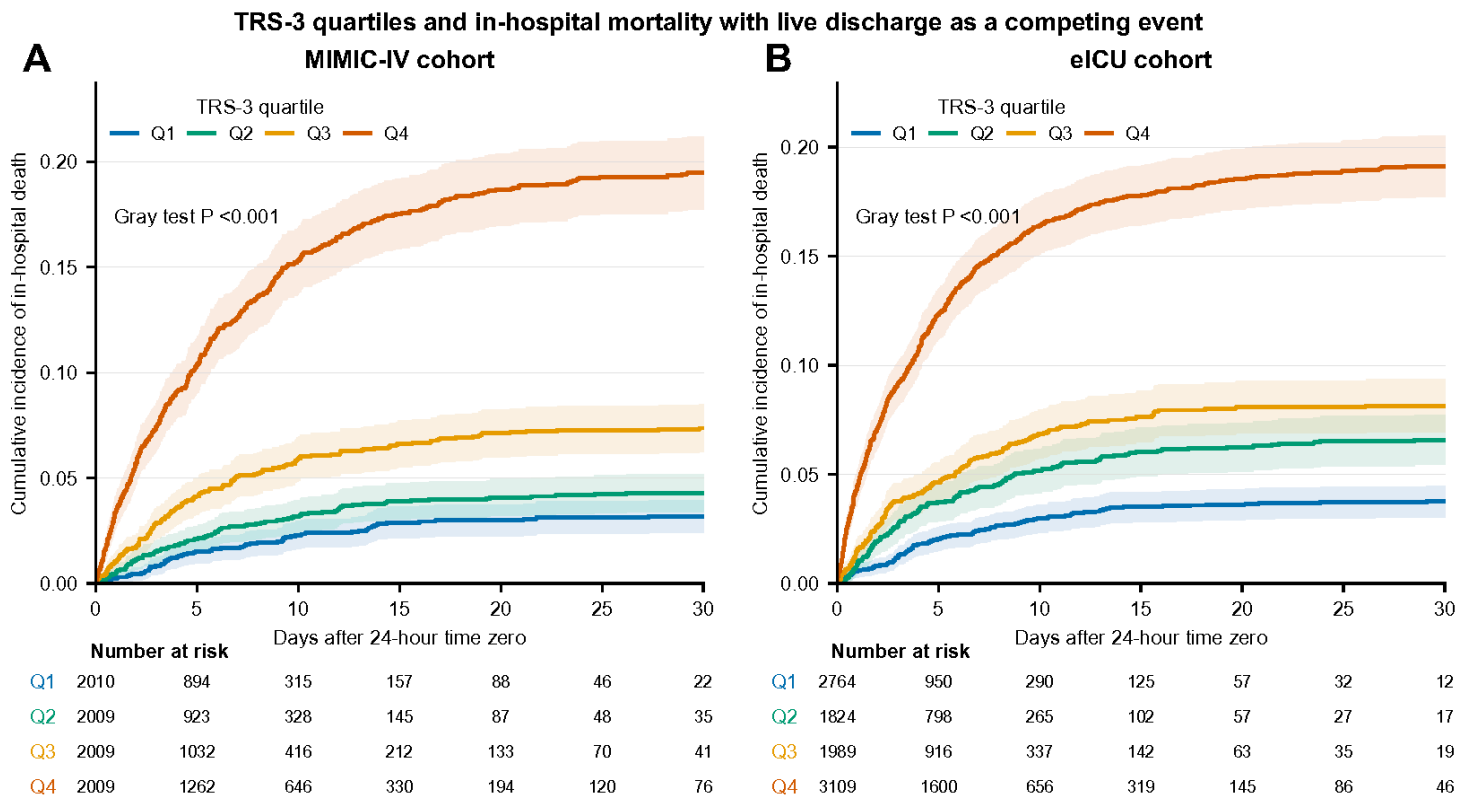


Figure 2

TRS-3 quartiles and the cumulative incidence of in-hospital death

A, MIMIC-IV cohort. B, eICU cohort. In-hospital death was the event of interest, and live discharge was the competing event. Curves were grouped using TRS-3 quartile cutoffs derived from MIMIC-IV. P values were obtained using Gray's test. Numbers at risk are shown below the plots.

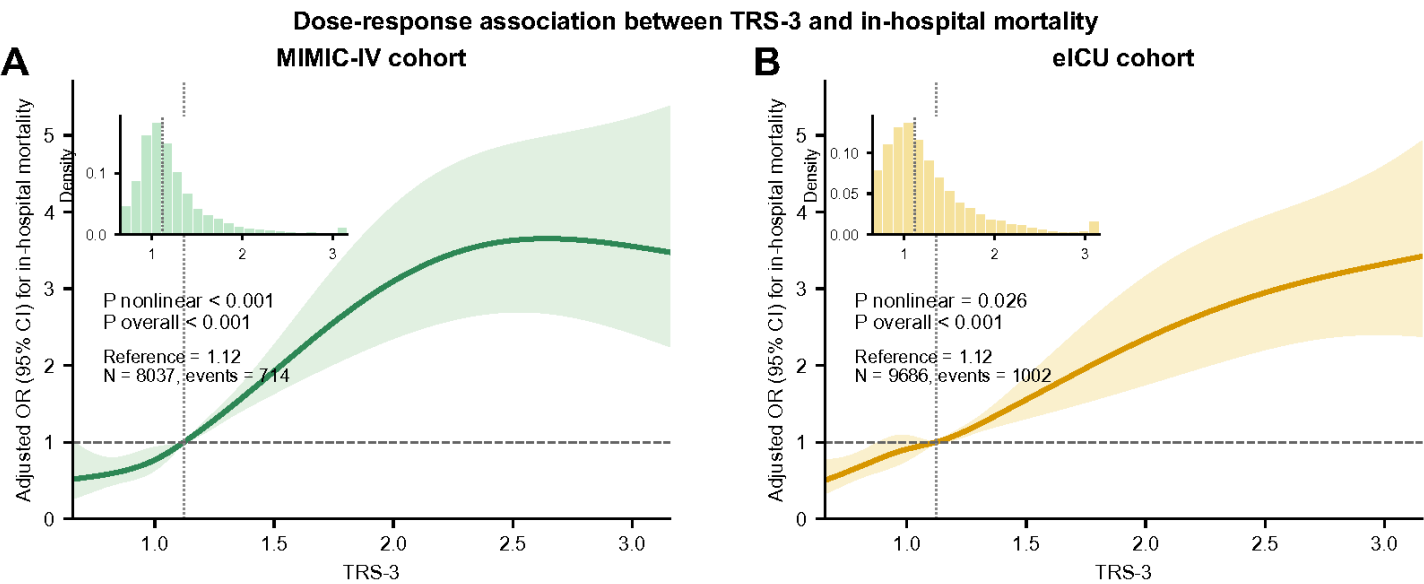


Figure 3

Dose-response association between TRS-3 and in-hospital mortality

A, MIMIC-IV cohort. B, eICU cohort. Curves were estimated using fully adjusted restricted cubic spline models. Solid lines indicate adjusted ORs, shaded areas indicate 95% CIs, and horizontal dashed lines indicate OR=1. The reference value, boundaries, and knots were determined in MIMIC-IV and applied directly to eICU. Insets show the distribution of TRS-3. Overall and nonlinear P values are displayed in each panel.

Association between TRS-3 and in-hospital mortality across prespecified subgroups

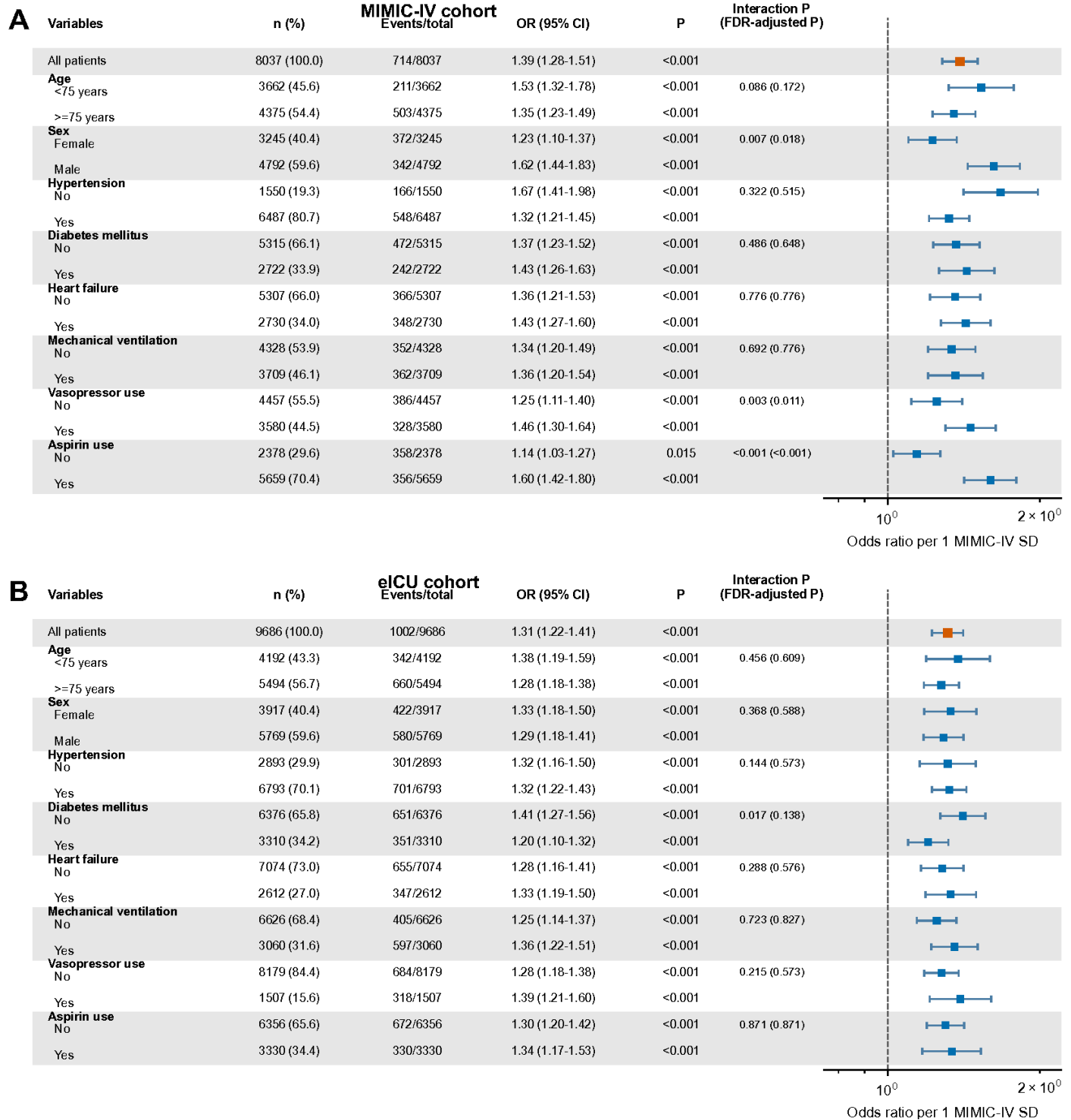


Figure 4

Prespecified subgroup analyses of the association between TRS-3 and in-hospital mortality

A, MIMIC-IV cohort. B, eICU cohort. ORs are reported per 1-standard-deviation increase in TRS-3 based on MIMIC-IV. Subgroup models included covariates from the primary fully adjusted model, excluding the corresponding stratifying variable. Interaction P values were obtained from product terms, with Benjamini-Hochberg-adjusted P values shown in parentheses.

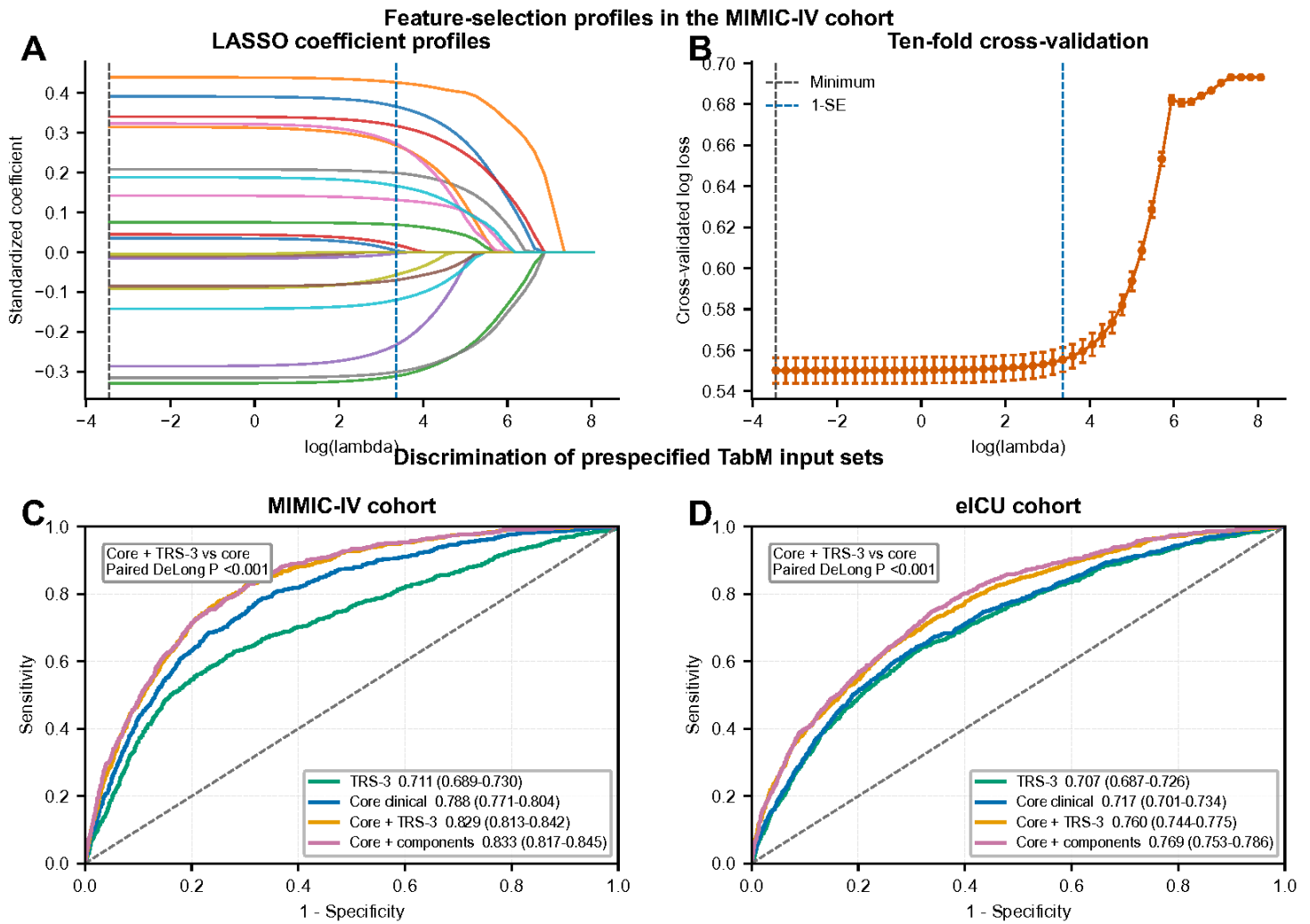


Figure 5

Feature selection and discrimination of the prespecified TabM input sets

A, coefficient paths from L1-penalized logistic regression. B, 10-fold cross-validation error with regularization parameters corresponding to the minimum error and the 1-standard-error rule. C, receiver operating characteristic curves in MIMIC-IV. D, receiver operating characteristic curves in eICU. AUROCs with 95% CIs are listed in the legends. DeLong P values compare the core clinical model with the core clinical model plus TRS-3.

Calibration, clinical utility, discrimination, and locked-threshold performance

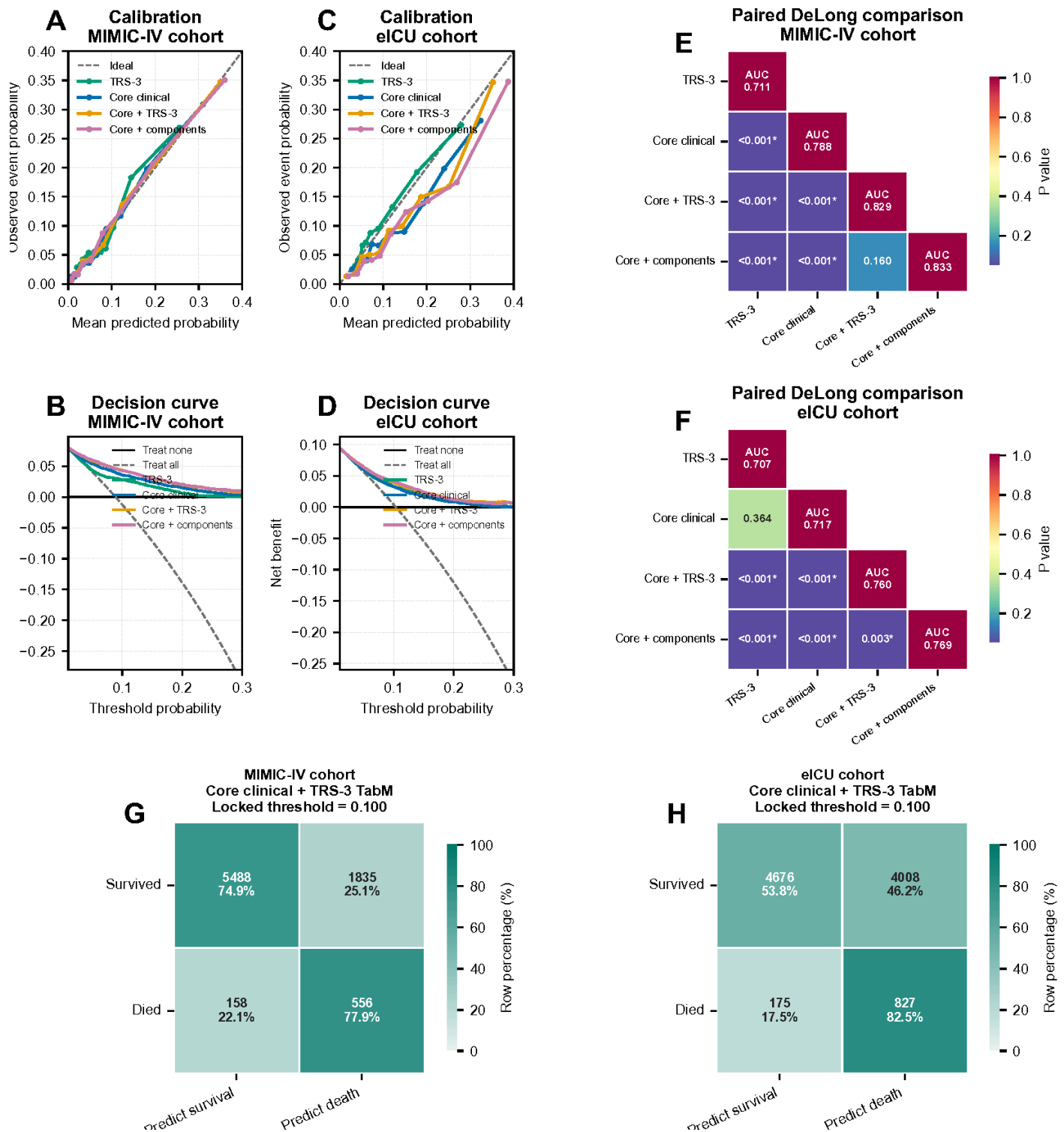


Figure 6

Calibration, decision curves, paired DeLong comparisons, and performance at the fixed threshold for the TabM models

A and C, calibration curves in MIMIC-IV and eICU, respectively. B and D, decision curves. E and F, matrices of paired DeLong tests between model AUROCs. G and H, confusion matrices for the core clinical model plus TRS-3 at the fixed threshold of 0.100. eICU probabilities were generated by the frozen MIMIC-IV models and calibrator, without reselection of the threshold in eICU.

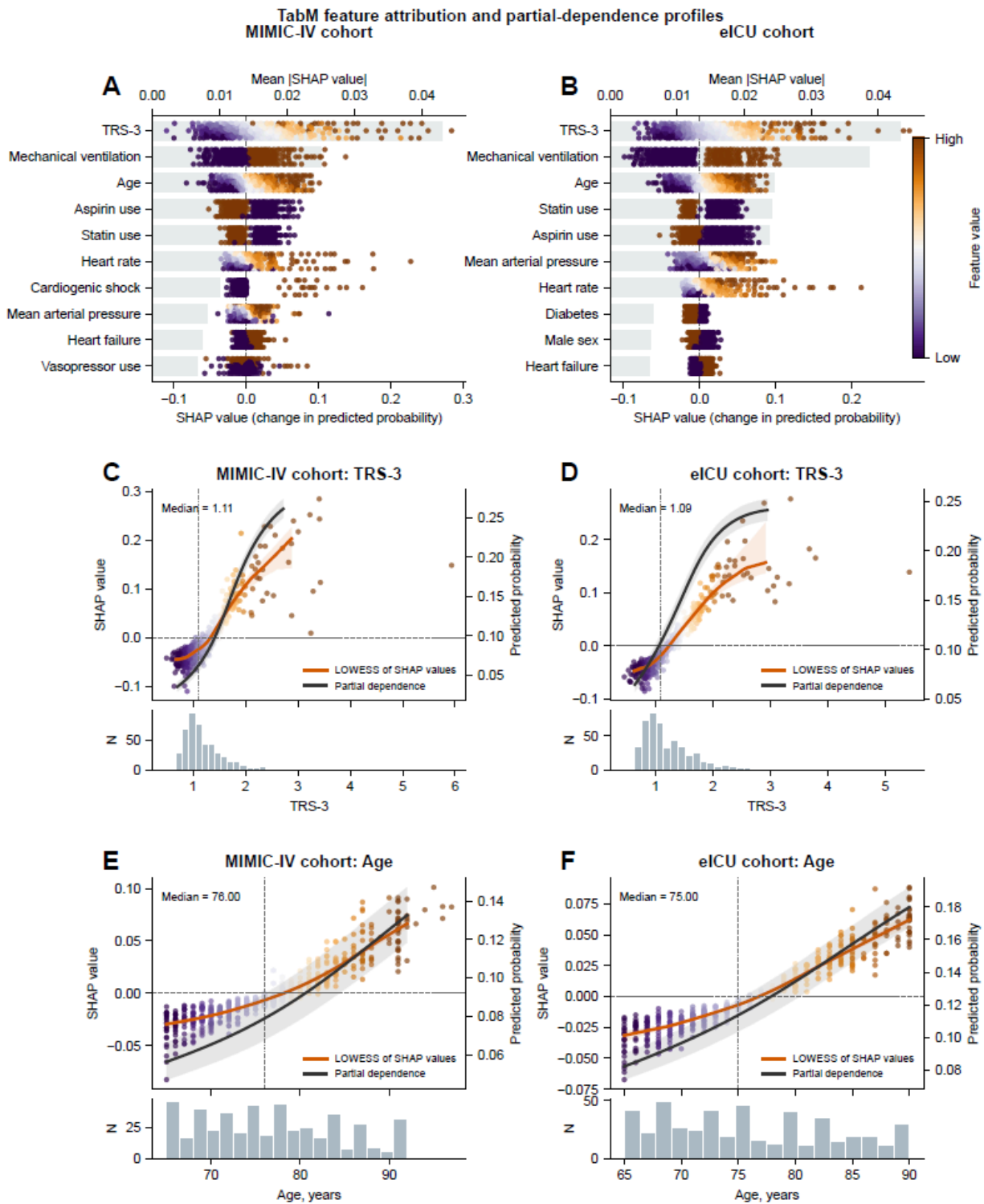


Figure 7

TabM feature attribution and partial dependence analyses

A and B, mean absolute SHAP value bar plots with superimposed beeswarm plots in MIMIC-IV and eICU, respectively. Each point represents 1 sampled patient, and color indicates the standardized feature value. C and D, SHAP LOWESS curves, partial dependence curves, and data distributions for TRS-3. E and F, corresponding analyses for age. SHAP values describe the computational contribution of each feature to model output.

Supplementary Files

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

- [SupplementaryFigure1.pdf](#)
- [SupplementaryFigure4.pdf](#)
- [SupplementaryTables.docx](#)
- [SupplementaryFigure3.pdf](#)
- [SupplementaryFigure2.pdf](#)