

RESEARCH

Open Access



Explainable machine learning model for prediction of 28-day all-cause mortality in immunocompromised patients in the intensive care unit: a retrospective cohort study based on MIMIC-IV database

Zhengqiu Yu^{1,2}, Lexin Fang³ and Yueping Ding^{3*}

Abstract

Objectives This study aimed to develop and validate an explainable machine learning (ML) model to predict 28-day all-cause mortality in immunocompromised patients admitted to the intensive care unit (ICU). Accurate and interpretable mortality prediction is crucial for clinical decision-making and optimal allocation of critical care resources for this vulnerable patient population.

Methods We utilized retrospective clinical data from the MIMIC-IV (version 2.2) database, encompassing ICU admissions at Beth Israel Deaconess Medical Center from 2008 to 2019. Eligible immunocompromised patients, including those with primary immunodeficiencies and chronic acquired conditions, such as hematological malignancies, solid tumors, and organ transplantation, were selected. Data were randomly split into training (80%) and testing (20%) cohorts. Ten ML models (logistic regression, XGBoost, LightGBM, AdaBoost, Random Forest, Gradient Boosting, Gaussian Naive Bayes, Complement Naive Bayes, Multilayer Perceptron, and Support Vector Machine) were developed and evaluated using area under the receiver operating characteristic curve (AUROC), area under the precision–recall curve (AUPRC), sensitivity, specificity, accuracy, and F1 score. Model explainability was achieved through SHapley Additive exPlanations (SHAP), and decision curve analysis (DCA) assessed clinical utility. In addition, Cox proportional hazards regression was conducted to evaluate the impact of predictive factors on time-to-event outcomes.

Results Among the evaluated models, the Support Vector Machine (SVM) demonstrated the highest AUROC of 0.863 (95% CI 0.834–0.890) and a notable AUPRC of 0.678 (95% CI 0.624–0.736). Key predictive factors consistently identified across multiple ML models included 24-h urine output, blood urea nitrogen (BUN) levels, presence of metastatic solid tumors, Charlson Comorbidity Index (CCI), and international normalized ratio (INR). SHAP analyses provided detailed insights into how these features influenced model predictions.

Conclusions The explainable ML models based on various artificial intelligence methods demonstrated promising clinical applicability in predicting 28-day mortality risk among immunocompromised ICU patients. Factors such

*Correspondence:

Yueping Ding

Dingyp0424@zcmu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

as urine output, BUN, metastatic solid tumors, CCI, and INR significantly contributed to prediction outcomes and may serve as important predictors in clinical practice.

Keywords Mortality prediction, Immunocompromised, Intensive care unit, Machine learning, MIMIC-IV, SHAP value

Introduction

The last few decades have witnessed significant advancements in therapeutic interventions for conditions, such as cancer, hematologic malignancies, solid organ transplantation, and autoimmune diseases [1–4]. These advancements have markedly improved the survival rates of patients with these conditions, leading to a growing number of immunocompromised patients in intensive care units (ICUs) [5]. The proportion of critically ill patients with compromised immune systems has risen to approximately one-third of all ICU admissions [6]. These patients face specific challenges and risks, including higher mortality rates and increased complications.

Immunocompromised patients may require ICU admission for various reasons, including severe infections, immune-mediated organ dysfunction, acute gastrointestinal bleeding, acute hypoxemic respiratory failure, and complications related to their primary diseases or treatments [7–13]. Acute respiratory failure (ARF) is a leading cause of ICU admission among immunocompromised patients, often resulting from bacterial (51.2%) and viral infections (25%) [14]. Gastrointestinal bleeding (GIB) is another critical reason for ICU admission, prevalent among those with hematological malignancies and undergoing chemotherapy [7]. Immunocompromised patients' outcomes are influenced not only by their underlying medical conditions but also by factors specific to their ICU stay. Baseline immune dysfunction, the use of broad-spectrum antibiotics, invasive devices, and additional immune-modulating therapies increase their risk of acquiring new infections during their ICU stay [15–18]. These patients are at high risk for secondary complications, such as ventilator-associated pneumonia (VAP) [19, 20], surgical site infections (SSI) [21, 22], and bloodstream infections [23–25]. These complications lead to higher mortality rates, prolonged ICU and hospital stays, and increased healthcare costs. Comprehensive management and effective preventive measures are essential to improve outcomes and reduce mortality in this vulnerable population.

Accurate mortality prediction for immunocompromised ICU patients is crucial for effective patient management and optimal healthcare resource allocation [26]. Traditional predictive models, though beneficial, often fail to account for the intricate

factors affecting outcomes in this high-risk group. Established severity scoring systems such as APACHE, SOFA, or SAPS have demonstrated utility in general ICU populations [27], while specialized models tailored to the unique characteristics of immunocompromised patients may offer additional insights given their distinct clinical presentations and risk profiles. Recent advancements in machine learning (ML) have shown great potential in predicting clinical outcomes in critically ill patients [28]. ML models can integrate a vast array of clinical variables, offering superior predictive accuracy compared to traditional methods.

However, the complexity of these models often results in a “black-box” phenomenon, hindering their acceptance in clinical practice [29]. To address this issue, explainable AI techniques, such as SHapley Additive exPlanations (SHAP) values [30], have emerged to elucidate the contribution of each feature to the model's predictions. The clinical application potential of such explainable prediction models is substantial. Studies have demonstrated that early identification of high-risk immunocompromised patients can positively influence treatment trajectories, with timely interventions potentially improving outcomes in this vulnerable population [31]. In addition, mortality prediction tools have been shown to enhance resource allocation efficiency in ICU settings by helping identify patients who may benefit most from intensive interventions [32], a consideration particularly relevant for immunocompromised patients who often require specialized resources. Beyond direct patient care, prediction models serve as valuable tools for hospital performance evaluation and quality improvement initiatives. Research has shown that risk-adjusted mortality metrics based on accurate prediction models can identify variation in care quality across institutions, leading to targeted improvement efforts and better outcomes [33]. Furthermore, studies examining physician decision-making patterns indicate that objective prognostic information influences treatment choices and management strategies, enabling more personalized and appropriate care for these complex patients [28].

This study aims to develop and validate an explainable ML model to predict 28-day mortality risk in immunocompromised ICU patients. Utilizing data from the MIMIC-IV database, we seek to create a robust tool

that not only predicts outcomes accurately but also offers interpretability through SHAP values. This approach aims to improve clinical decision-making and patient care in critical care settings.

Materials and methods

Study design and data source

The data utilized in this study were derived from the MIMIC-IV (version 2.2) database [34], an extensive repository of clinical information for patients admitted to the ICU at Beth Israel Deaconess Medical Center in the United States between 2008 and 2019. This database encompasses a broad spectrum of data, including demographics, vital signs, laboratory tests, medications, and follow-up information. The MIMIC-IV database is freely accessible to researchers globally upon receiving joint approval from the ethics review boards of MIT and Harvard Medical School. Given the retrospective nature of the study, informed consent was waived. Ethical permission to use the MIMIC-IV database has been secured by the research team (certification no.: 48061114, 38,118,593).

Study population

The study included patients from the MIMIC-IV database who met the criteria for overt immunosuppressive conditions at ICU admissions. We defined the immunosuppressed population to include those with primary immunodeficiencies and chronic acquired immunodeficiencies. Specifically, primary immunodeficiencies encompassed conditions, such as antibody deficiency, cellular deficiency, combined antibody and cellular immune deficiency, phagocytic defects, and complement defects. Chronic acquired immunodeficiencies included hematological malignancies, solid tumors, solid organ transplantation, corticosteroids and other immunosuppressive therapies, hematopoietic stem cell transplantation (HSCT), and HIV [35]. The detailed information will be further elucidated in the supplementary materials. Exclusions were made for patients younger than 18 years and those with ICU stays shorter than 6 h. Only the data from the first ICU treatment during the same hospital admission were included in the analysis.

Feature extraction and data preprocessing

Structured Query Language was used to extract data. We collected variables including demographics, vital signs, laboratory tests, pre-ICU comorbidities, mechanical ventilation, Charlson Comorbidity Index (CCI), Glasgow Coma Scale (GCS) scores, Sequential Organ Failure Assessment (SOFA) scores, and 28-day all-cause mortality. Laboratory tests and vital signs were collected within

the first 24 h after ICU admission. For time-varying measurements during this period, we employed a comprehensive approach to capture different aspects of physiological status. For vital signs (heart rate, respiratory rate, temperature, blood pressure, and oxygen saturation) and laboratory measurements, we calculated three summary statistics: maximum values to identify acute deterioration, minimum values to detect physiological compromise, and mean values to reflect overall trends. Clinical severity scores (GCS and SOFA) were calculated using the worst values of their components within the first 24-h window. For cumulative measurements, such as urine output, we calculated the total sum over the 24-h period. Mechanical ventilation status was defined as any ventilatory support required during the first 24 h of ICU admission. The outcome variable was 28-day all-cause mortality after ICU admission.

Variables with missing values exceeding 20% were excluded from the analysis. For the remaining variables, we applied nearest neighbor imputation algorithms to address any missing values (missing data statistics are provided in Supplementary Table 1). Our feature selection process followed a two-step approach. First, we performed univariate logistic regression analysis for all variables, eliminating features with $P > 0.05$ as they were deemed less likely to be relevant for 28-day mortality prediction. Second, we conducted correlation analyses to reduce multicollinearity: for pairs of variables with correlation coefficients exceeding 0.75 (Spearman correlation for numeric pairs, Cramer's V for categorical pairs, and correlation ratio for mixed pairs), we retained only the variable with the stronger association with mortality (lower p value in univariate logistic regression) and removed the other (correlation patterns are shown in Supplementary Fig. 1). For the final data preprocessing steps, we employed a model-specific approach. For severely skewed numeric features ($|\text{skewness}| > 1.5$), we selectively applied the Yeo–Johnson transformation to improve the performance of linear models, while preserving the original distribution for tree-based models that are naturally less sensitive to feature distributions. All categorical variables were converted to numeric format using one-hot encoding. This tailored preprocessing strategy allowed us to optimize each algorithm's performance while maintaining methodological rigor.

Model development and validation

We utilized ten ML models to construct the prediction model, including logistic regression, eXtreme Gradient Boosting (XGBoost), LightGBM, AdaBoost, Random Forest (RF), Gradient Boosting, Gaussian Naive Bayes (GNB), multilayer perceptron neural network (MLP),

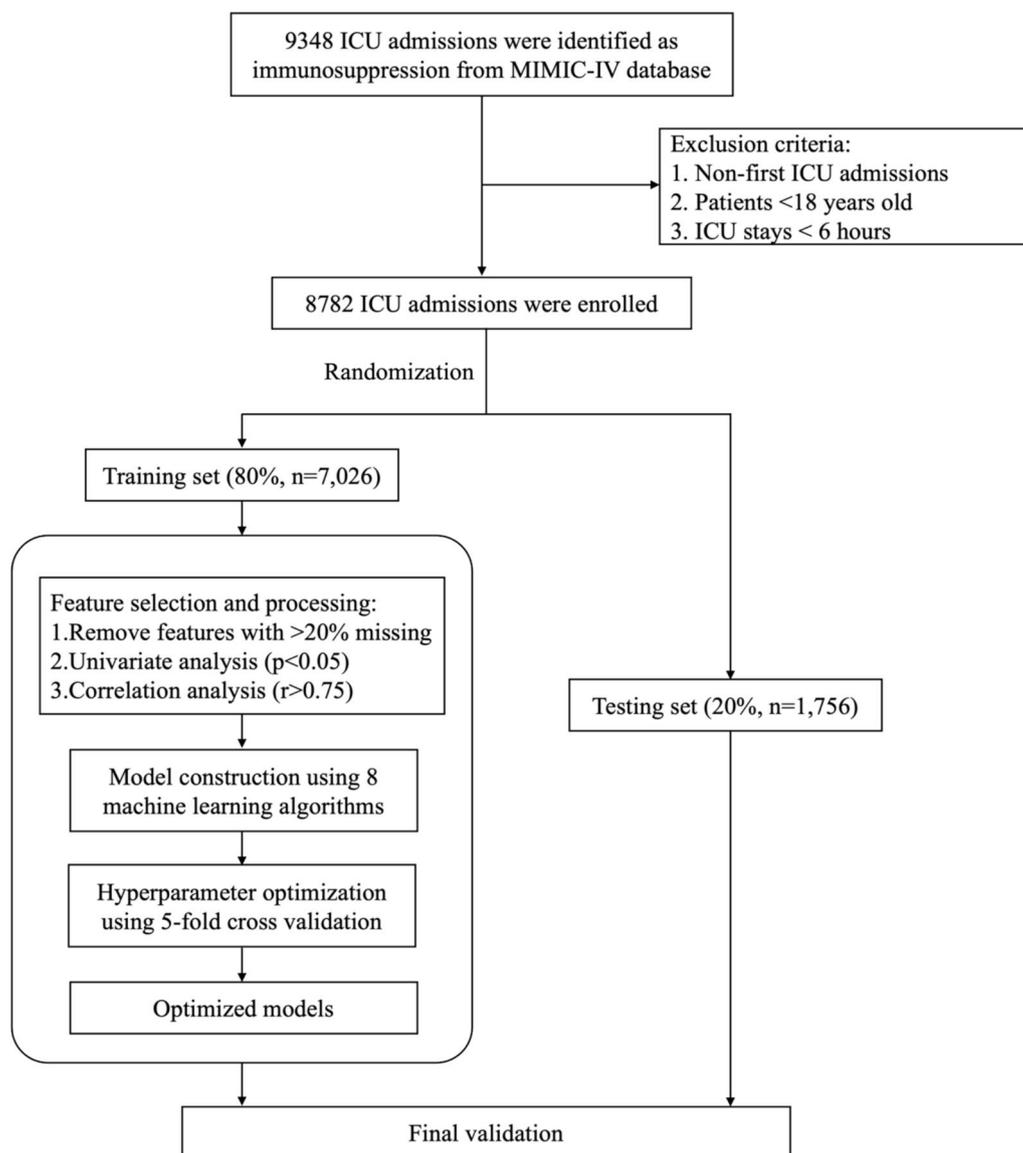


Fig. 1 Flowchart of the study

Complement Naive Bayes (CNB), and support vector machine (SVM). The overall data set was randomly divided into two groups, with 80% in the training cohort and 20% in the testing cohort. To minimize overfitting and identify the optimal hyperparameters, fivefold cross-validation (CV) was performed. Grid search was used to find the optimal hyperparameters for each machine learning model. The models were evaluated using several performance metrics, including the area under the receiver operating characteristic curve (AUROC), area under the precision–recall curve (AUPRC), accuracy, sensitivity (recall), specificity, positive predictive value (PPV), negative predictive value (NPV), and F1 score.

For each model, the optimal probability threshold was determined using the Youden index method, which maximizes the sum of sensitivity and specificity minus one. The 95% confidence intervals for all metrics were derived using bootstrap resampling with 500 iterations. To further interpret the results and understand the contribution of each feature to the predictions, we employed SHapley Additive exPlanations (SHAP) to enhance the transparency and interpretability of the best machine learning model. To assess potential clinical utility, we performed decision curve analysis (DCA) to evaluate net benefits across different threshold probabilities. In addition to the machine learning models, we performed Cox

proportional hazards regression to account for the time-to-event nature of survival data. Both univariable and multivariable Cox regression analyses were conducted, with variables showing significant associations in univariable analysis ($p < 0.05$) included in the multivariable model.

Statistical methods

Continuous variables are presented as median with interquartile ranges (IQR) due to their non-normal distribution, and categorical variables are presented as numbers (percentages). Appropriate statistical tests such as the Mann–Whitney U test, Student's t test, chi-square test, or Fisher's exact test were used to compare baseline characteristic variables. All statistical tests were two-tailed, and a p value less than 0.05 was considered statistically significant. The performance metrics used in this study were calculated as follows:

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Sensitivity (Recall)} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{Positive predictive value (PPV, precision)} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Negative predictive value (NPV)} = \text{TN} / (\text{TN} + \text{FN})$$

$$\text{F1 score} = 2 \times (\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$$

where TP = true positives, TN = true negatives, FP = false positives, and FN = false negatives.

For all model performance metrics (AUC, PRC, accuracy, PPV, sensitivity, specificity, NPV, and F1 score), we estimated 95% confidence intervals using bootstrap resampling with 500 iterations. This involved randomly sampling with replacement from the test data set, calculating each performance metric on these bootstrap samples, and determining the 2.5 th and 97.5 th percentiles of the resulting distribution. Data analysis and model establishment were conducted in Python (version 3.10.0). The complete code used for data preprocessing, model development, and analysis is available at <https://github.com/leanqon/immunocompromised-mortality-ml>.

Results

Baseline characteristics

In this study, we extracted 8,782 eligible patients from MIMIC-IV database, dividing the patients into two groups: a survival group (6,805 cases) and a mortality

group (1,977 cases) according to all-cause 28-day mortality. A flowchart of the study cohort selection process is presented in Fig. 1. The baseline characteristics are presented in Table 1. Among the entire patient population, the median age was 68.00 years (IQR: 58.00, 77.00). The median age was higher in the mortality group (71.00 vs. 67.00 years, $p < 0.001$). Gender distribution was similar between groups, with females comprising 41.8% of the overall population ($p = 0.392$). White patients constituted the majority (72.1%) of the study population, followed by other (11.3%), African–American (9.6%), Asian (4.1%), and Hispanic–American (2.9%).

Vital signs analysis showed that the mortality group had significantly higher heart rate (HR) (mean: 92.89 vs. 84.34 beats per minute, $p < 0.001$), higher respiratory rate (RR) (mean: 20.42 vs. 18.43 breaths per minute, $p < 0.001$), and lower systolic blood pressure (SBP) (mean: 110.05 vs. 116.30 mmHg, $p < 0.001$). Laboratory indicators revealed that the mortality group had higher white blood cell (WBC) counts (mean: 11.48 vs. $10.00 \times 10^9/\text{L}$, $p < 0.001$), lower hemoglobin (mean: 9.52 vs. 10.20 g/dL, $p < 0.001$), and higher blood urea nitrogen (BUN) and creatinine levels (BUN mean: 27.50 vs. 18.50 mg/dL; creatinine mean: 1.10 vs. 0.90 mg/dL, both $p < 0.001$). The mortality group also showed higher international normalized ratio (INR) values (mean: 1.40 vs. 1.23, $p < 0.001$) and significantly lower 24-h urine output (911.00 vs. 1510.00 mL, $p < 0.001$). Comorbidities were more prevalent in the mortality group, including higher incidences of myocardial infarction (14.4% vs. 11.1%, $p < 0.001$), congestive heart failure (24.7% vs. 19.8%, $p < 0.001$), cerebrovascular disease (13.5% vs. 9.5%, $p < 0.001$), chronic pulmonary disease (28.8% vs. 25.1%, $p = 0.001$), and renal disease (22.6% vs. 18.6%, $p < 0.001$). Notably, metastatic solid tumor was significantly more common in the mortality group (52.9% vs. 32.2%, $p < 0.001$). Mechanical ventilation was also more common in the mortality group (29.4% vs. 22.0%, $p < 0.001$).

Selected predictive features

Our systematic feature selection process yielded a final set of 44 features that demonstrated significant association with 28-day mortality. These included demographic characteristics (age, weight, ethnicity), clinical severity and comorbidity indices (Charlson Comorbidity Index [CCI], GCS, SOFA score), vital sign measurements (HR mean, RR minimum and mean, temperature minimum and mean, SBP minimum, mean arterial pressure [MAP] minimum and maximum, SpO2 minimum, maximum and mean), laboratory values (BUN minimum, chloride minimum, calcium minimum, sodium

Table 1 Baseline characteristics of the patients

Variable	Overall (n = 8782)	Survival group (n = 6805)	Mortality group (n = 1977)	p value
Age (years)	68.00 (58.00, 77.00)	67.00 (58.00, 76.00)	71.00 (61.00, 81.00)	< 0.001 ^c
Gender	3669 (41.8%)	2826 (41.5%)	843 (42.6%)	0.392 ^a
White	6333 (72.1%)	4978 (73.2%)	1355 (68.5%)	< 0.001 ^a
Asian	362 (4.1%)	274 (4.0%)	88 (4.5%)	0.440 ^a
African–American	840 (9.6%)	645 (9.5%)	195 (9.9%)	0.639 ^a
Hispanic–American	254 (2.9%)	204 (3.0%)	50 (2.5%)	0.308 ^a
Other	993 (11.3%)	704 (10.3%)	289 (14.6%)	< 0.001 ^a
Weight (kg)	75.80 (64.00, 89.43)	76.80 (65.00, 90.20)	72.00 (61.10, 85.50)	< 0.001 ^c
Glasgow Coma Scale	15.00 (15.00, 15.00)	15.00 (15.00, 15.00)	15.00 (14.00, 15.00)	< 0.001 ^c
Vital Signs				
HR min (bpm)	72.00 (62.00, 84.00)	71.00 (61.00, 82.00)	77.50 (65.00, 90.00)	< 0.001 ^c
HR max (bpm)	103.00 (90.00, 118.00)	101.00 (88.00, 115.00)	111.00 (96.00, 125.88)	< 0.001 ^c
HR mean (bpm)	85.99 (74.98, 98.06)	84.34 (73.69, 95.62)	92.89 (80.60, 105.25)	< 0.001 ^c
RR min (bpm)	13.00 (11.00, 15.00)	13.00 (10.00, 15.00)	14.00 (11.00, 17.00)	< 0.001 ^c
RR max (bpm)	26.50 (23.00, 31.00)	26.00 (23.00, 30.00)	28.00 (24.00, 33.00)	< 0.001 ^c
RR mean (bpm)	18.82 (16.48, 21.83)	18.43 (16.25, 21.24)	20.42 (17.46, 23.89)	< 0.001 ^c
Temperature min (°C)	36.44 (36.11, 36.67)	36.44 (36.17, 36.67)	36.39 (36.00, 36.61)	< 0.001 ^c
Temperature max (°C)	37.17 (36.89, 37.67)	37.17 (36.94, 37.61)	37.11 (36.83, 37.67)	< 0.001 ^c
Temperature mean (°C)	36.80 (36.58, 37.06)	36.81 (36.60, 37.06)	36.73 (36.49, 37.04)	< 0.001 ^c
SBP min (mmHg)	93.00 (84.00, 103.00)	94.00 (86.00, 104.00)	89.00 (80.00, 99.00)	< 0.001 ^c
SBP max (mmHg)	141.00 (127.00, 156.00)	142.00 (128.00, 157.00)	137.00 (122.50, 152.00)	< 0.001 ^c
SBP mean (mmHg)	115.04 (105.07, 126.73)	116.30 (106.21, 127.94)	110.05 (101.16, 122.20)	< 0.001 ^c
DBP min (mmHg)	48.00 (41.00, 55.00)	48.00 (42.00, 55.00)	46.00 (39.00, 54.00)	< 0.001 ^c
DBP max (mmHg)	82.00 (72.00, 94.00)	82.00 (72.00, 94.00)	81.50 (71.00, 94.00)	0.036 ^c
DBP mean (mmHg)	62.11 (55.58, 69.68)	62.50 (56.02, 69.96)	60.69 (53.92, 68.54)	< 0.001 ^c
MAP min (mmHg)	61.00 (54.00, 68.00)	61.00 (55.00, 69.00)	58.00 (51.00, 65.00)	< 0.001 ^c
MAP max (mmHg)	97.00 (87.00, 108.00)	97.00 (88.00, 109.00)	95.00 (84.50, 106.00)	< 0.001 ^c
MAP mean (mmHg)	76.25 (69.75, 84.35)	77.11 (70.52, 84.98)	73.69 (67.34, 81.51)	< 0.001 ^c
SpO2 min (%)	90.50 (89.00, 92.00)	90.50 (89.00, 92.00)	90.00 (87.00, 91.50)	< 0.001 ^c
SpO2 max (%)	100.00 (99.00, 100.00)	100.00 (99.00, 100.00)	100.00 (99.00, 100.00)	0.025 ^c
SpO2 mean (%)	96.20 (94.83, 97.52)	96.30 (95.00, 97.57)	95.83 (94.25, 97.33)	< 0.001 ^c
Laboratory values				
WBC min ($\times 10^9/L$)	9.40 (6.20, 13.70)	9.20 (6.20, 13.10)	10.50 (6.38, 16.40)	< 0.001 ^c
WBC max ($\times 10^9/L$)	11.00 (7.30, 16.10)	10.70 (7.20, 15.40)	12.50 (7.80, 19.10)	< 0.001 ^c
WBC mean ($\times 10^9/L$)	10.25 (6.90, 14.80)	10.00 (6.80, 14.11)	11.48 (7.19, 17.96)	< 0.001 ^c
Hemoglobin min (g/dL)	9.70 (8.34, 11.30)	9.90 (8.50, 11.50)	9.20 (8.00, 10.70)	< 0.001 ^c
Hemoglobin max (g/dL)	10.40 (9.10, 11.90)	10.60 (9.20, 12.00)	9.90 (8.70, 11.30)	< 0.001 ^c
Hemoglobin mean (g/dL)	10.05 (8.75, 11.55)	10.20 (8.90, 11.70)	9.52 (8.40, 10.95)	< 0.001 ^c
Platelet count min ($\times 10^9/L$)	181.00 (116.00, 259.00)	185.00 (124.00, 257.00)	163.00 (86.00, 267.00)	< 0.001 ^c
Platelet count max ($\times 10^9/L$)	198.00 (134.00, 276.00)	201.00 (141.00, 273.00)	186.00 (104.00, 289.00)	< 0.001 ^c
Platelet count mean ($\times 10^9/L$)	189.67 (125.33, 266.73)	193.00 (133.50, 264.50)	174.00 (95.38, 277.00)	< 0.001 ^c
MCH min (pg)	30.00 (28.40, 31.50)	30.00 (28.40, 31.40)	29.80 (28.10, 31.60)	0.295 ^c
MCH max (pg)	30.30 (28.70, 31.80)	30.30 (28.70, 31.80)	30.30 (28.50, 32.00)	0.891 ^c
MCH mean (pg)	30.15 (28.52, 31.65)	30.20 (28.60, 31.60)	30.00 (28.33, 31.76)	0.526 ^c
MCHC min (g/dL)	32.80 (31.70, 33.90)	32.90 (31.90, 34.00)	32.30 (31.20, 33.40)	< 0.001 ^c
MCHC max (g/dL)	33.30 (32.20, 34.40)	33.40 (32.30, 34.50)	32.90 (31.70, 34.00)	< 0.001 ^c
MCHC mean (g/dL)	33.07 (31.95, 34.10)	33.20 (32.10, 34.20)	32.67 (31.50, 33.70)	< 0.001 ^c
MCV min (fL)	90.00 (86.00, 95.00)	90.00 (86.00, 94.00)	91.00 (86.00, 96.00)	< 0.001 ^c
MCV max (fL)	91.00 (87.00, 96.00)	91.00 (87.00, 95.00)	92.00 (88.00, 98.00)	< 0.001 ^c
MCV mean (fL)	91.00 (86.50, 95.00)	90.33 (86.00, 94.75)	91.50 (87.00, 97.00)	< 0.001 ^c
Bicarbonate min (mEq/L)	22.00 (19.00, 25.00)	23.00 (20.00, 25.00)	21.00 (17.00, 24.00)	< 0.001 ^c

Table 1 (continued)

Variable	Overall (n = 8782)	Survival group (n = 6805)	Mortality group (n = 1977)	p value
Bicarbonate max (mEq/L)	24.00 (21.00, 26.00)	24.00 (22.00, 26.00)	23.00 (19.00, 26.00)	< 0.001 ^c
Bicarbonate mean (mEq/L)	23.00 (20.33, 25.50)	23.00 (21.00, 26.00)	21.67 (18.00, 25.00)	< 0.001 ^c
BUN min (mg/dL)	19.00 (13.00, 30.00)	18.00 (12.00, 26.00)	26.00 (17.00, 43.00)	< 0.001 ^c
BUN max (mg/dL)	21.00 (14.00, 33.00)	19.00 (13.00, 29.00)	29.00 (18.00, 47.00)	< 0.001 ^c
BUN mean (mg/dL)	20.00 (13.50, 31.00)	18.50 (13.00, 27.71)	27.50 (17.00, 45.50)	< 0.001 ^c
Creatinine min (mg/dL)	0.90 (0.70, 1.30)	0.90 (0.70, 1.20)	1.10 (0.70, 1.80)	< 0.001 ^c
Creatinine max (mg/dL)	1.00 (0.70, 1.50)	0.90 (0.70, 1.40)	1.20 (0.80, 2.00)	< 0.001 ^c
Creatinine mean (mg/dL)	0.95 (0.70, 1.40)	0.90 (0.70, 1.30)	1.10 (0.75, 1.90)	< 0.001 ^c
INR min	1.20 (1.10, 1.50)	1.20 (1.10, 1.40)	1.40 (1.20, 1.70)	< 0.001 ^c
INR max	1.30 (1.10, 1.60)	1.30 (1.10, 1.50)	1.50 (1.20, 1.90)	< 0.001 ^c
INR mean	1.30 (1.10, 1.50)	1.23 (1.10, 1.40)	1.40 (1.20, 1.80)	< 0.001 ^c
Glu min (mg/dL)	115.00 (96.00, 140.00)	115.00 (98.00, 140.00)	111.00 (92.00, 139.00)	< 0.001 ^c
Glu max (mg/dL)	139.00 (113.00, 178.00)	139.00 (114.00, 176.00)	138.00 (110.00, 185.00)	0.825 ^c
Glu mean (mg/dL)	128.00 (107.00, 158.27)	128.67 (108.00, 157.50)	125.00 (103.00, 162.12)	0.016 ^c
Sodium min (mEq/L)	137.00 (134.00, 140.00)	137.00 (135.00, 140.00)	136.00 (133.00, 140.00)	< 0.001 ^c
Sodium max (mEq/L)	139.00 (136.00, 141.00)	139.00 (136.00, 141.00)	138.00 (135.00, 142.00)	< 0.001 ^c
Sodium mean (mEq/L)	138.00 (135.33, 140.50)	138.00 (136.00, 140.50)	137.00 (133.67, 141.00)	< 0.001 ^c
Potassium min (mEq/L)	4.00 (3.60, 4.40)	4.00 (3.60, 4.30)	4.00 (3.60, 4.50)	< 0.001 ^c
Potassium max (mEq/L)	4.30 (3.90, 4.70)	4.30 (3.90, 4.70)	4.40 (4.00, 5.00)	< 0.001 ^c
Potassium mean (mEq/L)	4.13 (3.80, 4.50)	4.10 (3.80, 4.50)	4.20 (3.80, 4.70)	< 0.001 ^c
Calcium min (mg/dL)	8.20 (7.60, 8.70)	8.20 (7.70, 8.70)	8.00 (7.40, 8.60)	< 0.001 ^c
Calcium max (mg/dL)	8.40 (8.00, 8.90)	8.40 (8.00, 8.90)	8.30 (7.80, 8.90)	< 0.001 ^c
Calcium mean (mg/dL)	8.30 (7.80, 8.75)	8.30 (7.85, 8.80)	8.15 (7.63, 8.70)	< 0.001 ^c
Magnesium min (mEq/L)	1.90 (1.70, 2.10)	1.90 (1.70, 2.10)	1.90 (1.70, 2.10)	< 0.001 ^c
Magnesium max (mEq/L)	2.10 (1.90, 2.30)	2.10 (1.90, 2.30)	2.10 (1.90, 2.30)	< 0.001 ^c
Magnesium mean (mEq/L)	2.00 (1.80, 2.15)	1.98 (1.80, 2.13)	2.00 (1.80, 2.20)	< 0.001 ^c
Chloride min (mEq/L)	103.00 (99.00, 106.00)	103.00 (100.00, 106.00)	102.00 (97.00, 106.00)	< 0.001 ^c
Chloride max (mEq/L)	105.00 (101.00, 108.00)	105.00 (102.00, 108.00)	104.00 (99.00, 108.00)	< 0.001 ^c
Chloride mean (mEq/L)	104.00 (100.40, 107.00)	104.00 (101.00, 107.00)	103.00 (98.00, 107.00)	< 0.001 ^c
Urine output (mL/24 h)	1377.00 (805.00, 2170.00)	1510.00 (953.00, 2300.00)	911.00 (400.00, 1565.00)	< 0.001 ^c
Comorbidities				
Myocardial infarction	1041 (11.9%)	757 (11.1%)	284 (14.4%)	< 0.001 ^a
Congestive heart failure	1837 (20.9%)	1349 (19.8%)	488 (24.7%)	< 0.001 ^a
Peripheral vascular disease	668 (7.6%)	505 (7.4%)	163 (8.2%)	0.243 ^a
Cerebrovascular disease	910 (10.4%)	644 (9.5%)	266 (13.5%)	< 0.001 ^a
Dementia	184 (2.1%)	129 (1.9%)	55 (2.8%)	0.020 ^a
Chronic pulmonary disease	2280 (26.0%)	1710 (25.1%)	570 (28.8%)	0.001 ^a
Rheumatic disease	256 (2.9%)	194 (2.9%)	62 (3.1%)	0.557 ^a
Mild liver disease	1168 (13.3%)	818 (12.0%)	350 (17.7%)	< 0.001 ^a
Paraplegia	363 (4.1%)	248 (3.6%)	115 (5.8%)	< 0.001 ^a
Renal disease	1715 (19.5%)	1269 (18.6%)	446 (22.6%)	< 0.001 ^a
Malignant cancer	6827 (77.7%)	5250 (77.1%)	1577 (79.8%)	0.015 ^a
Severe liver disease	523 (6.0%)	325 (4.8%)	198 (10.0%)	< 0.001 ^a
Metastatic solid tumor	3234 (36.8%)	2189 (32.2%)	1045 (52.9%)	< 0.001 ^a
AIDS	301 (3.4%)	266 (3.9%)	35 (1.8%)	< 0.001 ^a
Mechanical ventilation	2078 (23.7%)	1496 (22.0%)	582 (29.4%)	< 0.001 ^a

Categorical data were presented as frequency (percentage), continuous data were presented as median (interquartile ranges). None of the continuous variables followed normal distribution; therefore, no variables were presented as mean ± standard deviation. Vital signs were recorded as minimum, maximum, and mean values during the first 24 h after ICU admission. Laboratory tests were recorded as minimum, maximum, and mean values during the first 24 h after ICU admission. Statistical tests: ^a chi-square test; ^b Fisher's exact test; ^c Mann-Whitney U test

SBP Systolic Blood Pressure, DBP Diastolic Blood Pressure, MAP Mean Arterial Pressure, SpO2 Peripheral Oxygen Saturation, HR Heart Rate, RR Respiratory Rate, WBC White Blood Cell Count, MCH Mean Corpuscular Hemoglobin, MCHC Mean Corpuscular Hemoglobin Concentration, MCV Mean Corpuscular Volume, BUN Blood Urea Nitrogen, INR International Normalized Ratio, GCS Glasgow Coma Scale, SOFA Sequential Organ Failure Assessment

minimum, potassium maximum, bicarbonate minimum, hemoglobin minimum, platelet count [PLT] minimum, WBC count minimum, MCV maximum, mean corpuscular hemoglobin concentration [MCHC] minimum, INR mean, magnesium mean, phosphate maximum, glucose [GLU] maximum), urine output, and comorbidities (myocardial infarction, congestive heart failure, cerebrovascular disease, chronic pulmonary disease, mild and severe liver disease, paraplegia, renal disease, metastatic solid tumor, acquired immune deficiency syndrome [AIDS]), and mechanical ventilation status. The

correlation structure among these features is visualized in Supplementary Fig. 1, demonstrating the effective reduction of multicollinearity while preserving clinically informative variables. Detailed information about these final features is provided in the supplementary materials.

Model performance

The performance of all models on the testing data set is summarized in Table 2 and Fig. 2. The Logistic Regression model achieved an AUC of 0.857 (95% CI 0.826–0.887) and PRC of 0.662 (95% CI 0.597–0.727), while

Table 2 Model performance metrics of ML models in validation data set (95% CI)

Models	AUC	PRC	Accuracy	PPV	Sensitivity	Specificity	NPV	F1 Score
LR	0.857 (0.826–0.887)	0.662 (0.597–0.727)	0.817 (0.791–0.842)	0.578 (0.512–0.643)	0.675 (0.603–0.737)	0.858 (0.828–0.883)	0.901 (0.883–0.916)	0.623 (0.561–0.672)
GNB	0.772 (0.734–0.808)	0.480 (0.415–0.552)	0.695 (0.665–0.722)	0.403 (0.353–0.452)	0.751 (0.695–0.806)	0.679 (0.641–0.714)	0.904 (0.884–0.919)	0.525 (0.476–0.574)
CNB	0.834 (0.801–0.867)	0.621 (0.554–0.691)	0.820 (0.797–0.846)	0.602 (0.531–0.666)	0.584 (0.519–0.652)	0.889 (0.865–0.909)	0.881 (0.862–0.897)	0.593 (0.533–0.654)
SVM	0.863 (0.834–0.890)	0.678 (0.624–0.736)	0.800 (0.772–0.826)	0.536 (0.485–0.592)	0.787 (0.729–0.839)	0.804 (0.774–0.831)	0.929 (0.912–0.943)	0.638 (0.589–0.684)
MLP	0.859 (0.826–0.887)	0.687 (0.622–0.744)	0.823 (0.800–0.848)	0.590 (0.519–0.651)	0.680 (0.612–0.745)	0.864 (0.841–0.886)	0.903 (0.885–0.917)	0.632 (0.584–0.686)
AdaBoost	0.851 (0.823–0.880)	0.670 (0.605–0.730)	0.810 (0.785–0.837)	0.563 (0.509–0.620)	0.680 (0.612–0.740)	0.848 (0.822–0.872)	0.902 (0.882–0.917)	0.616 (0.561–0.671)
RF	0.779 (0.745–0.817)	0.547 (0.481–0.622)	0.743 (0.718–0.772)	0.448 (0.392–0.507)	0.635 (0.572–0.701)	0.774 (0.742–0.802)	0.880 (0.858–0.898)	0.525 (0.471–0.573)
Gradient Boosting	0.831 (0.798–0.862)	0.626 (0.564–0.692)	0.802 (0.774–0.829)	0.552 (0.482–0.618)	0.619 (0.555–0.691)	0.855 (0.829–0.878)	0.886 (0.866–0.902)	0.584 (0.529–0.640)
LightGBM	0.815 (0.787–0.846)	0.590 (0.516–0.668)	0.769 (0.741–0.801)	0.489 (0.434–0.542)	0.685 (0.617–0.744)	0.793 (0.763–0.821)	0.897 (0.879–0.913)	0.571 (0.522–0.618)
XGBoost	0.815 (0.779–0.849)	0.601 (0.533–0.668)	0.779 (0.753–0.804)	0.506 (0.445–0.573)	0.650 (0.585–0.724)	0.817 (0.789–0.843)	0.890 (0.871–0.906)	0.569 (0.510–0.620)

AUC, area under curve; PRC, precision–recall curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR, logistic regression; GNB, Gaussian Naive Bayes; CNB, Complement Naive Bayes; SVM, support vector machine; MLP, multilayer perceptron; RF, Random Forest; XGBoost, extreme gradient boosting

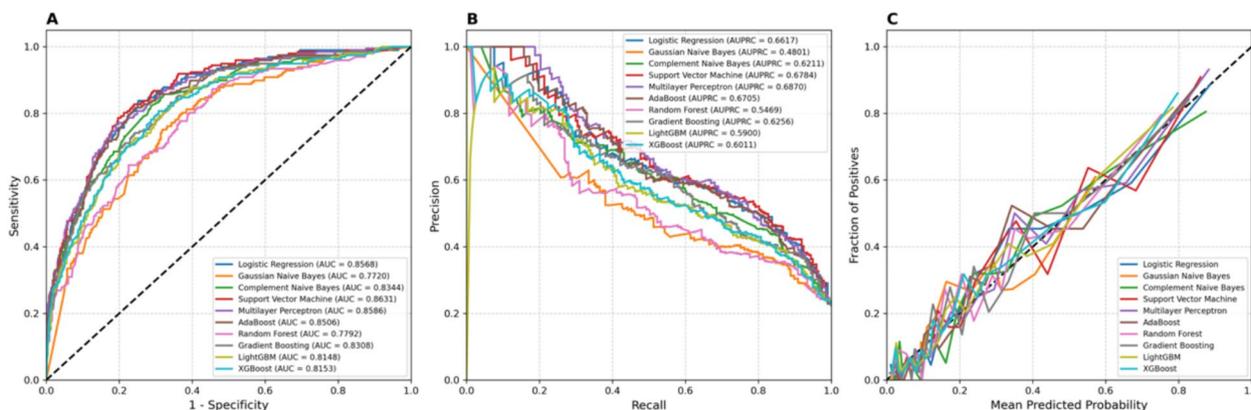


Fig. 2 Area under the receiver operating characteristic curve of models, precision–recall curve, and calibration plot in testing set. **A** ROC curves of models. **B** PRC curves of models. **C** Calibration curves of models

the support vector machine (SVM) model showed an AUC of 0.863 (95% CI 0.834–0.890) and PRC of 0.678 (95% CI 0.624–0.736). The Multilayer Perceptron (MLP) model performed similar to an AUC of 0.859 (95% CI 0.826–0.887) and the highest PRC of 0.687 (95% CI 0.622–0.744).

All models showed varying strengths across different metrics. The SVM model demonstrated the highest sensitivity (0.787, 95% CI 0.729–0.839) but moderate precision. The Complement Naive Bayes (CNB) model achieved the highest specificity (0.889, 95% CI 0.865–0.909) but lower sensitivity. The Gaussian Naive Bayes (GNB) model, while having a lower overall AUC (0.772, 95% CI 0.734–0.808), showed high sensitivity (0.751, 95% CI 0.695–0.806) that could be valuable in scenarios where identifying all potential high-risk patients is prioritized. Calibration plots of all models in different data sets are shown in Fig. 2.

To better understand the importance of features in prediction, we analyzed three of our tree-based models (Gradient Boosting, LightGBM, and XGBoost) using SHAP values, as shown in Fig. 3. Across all three models, similar key predictors emerged. The 24-h urine output consistently appeared as the most important predictor, followed by blood urea nitrogen minimum value (bun_min) and presence of metastatic solid tumors. Additional important predictors included the Charlson Comorbidity Index (CCI), international normalized ratio mean value (inr_pt_mean), and heart rate mean (hr_mean). Figure 3 not only ranks these features by importance but also illustrates how each feature's values impact predictions. For example, lower urine output (blue values in the SHAP summary plots) is associated with higher predicted mortality risk, while the presence of metastatic solid tumors (red values) similarly increases predicted risk.

Cox regression analysis further supported our findings from machine learning models. The multivariable Cox analysis revealed that metastatic solid tumor (HR 1.146, 95% CI 1.097–1.198, $p < 0.001$), severe liver disease (HR 1.127, 95% CI 1.029–1.234, $p = 0.01$), and paraplegia (HR 1.105, 95% CI 0.993–1.230, $p = 0.068$) were associated

with increased mortality risk. Several physiological parameters, including heart rate (HR 1.074, 95% CI 1.051–1.097, $p < 0.001$) and respiratory rate (HR 1.066, 95% CI 1.043–1.089, $p < 0.001$), also showed significant associations with mortality. Detailed results of the Cox regression analysis are presented in Supplementary Table 2.

Decision curve analysis

To evaluate the clinical utility of our prediction models, we performed decision curve analysis (Fig. 4). As shown, all models demonstrated positive net benefits across a wide range of threshold probabilities (10–90%) compared to the default strategies of treating all patients or treating none. The greatest incremental benefit over the reference strategies was observed in the lower threshold probability range (10–40%), where the vertical separation between model curves and reference lines was maximal. All models performed similar to one another, with comparable net benefit curves throughout most threshold ranges, suggesting equivalent clinical utility for mortality risk prediction in immunocompromised ICU patients. These findings indicate that implementing any of these models could support clinical decision-making by helping identify high-risk patients who might benefit from more intensive monitoring or interventions.

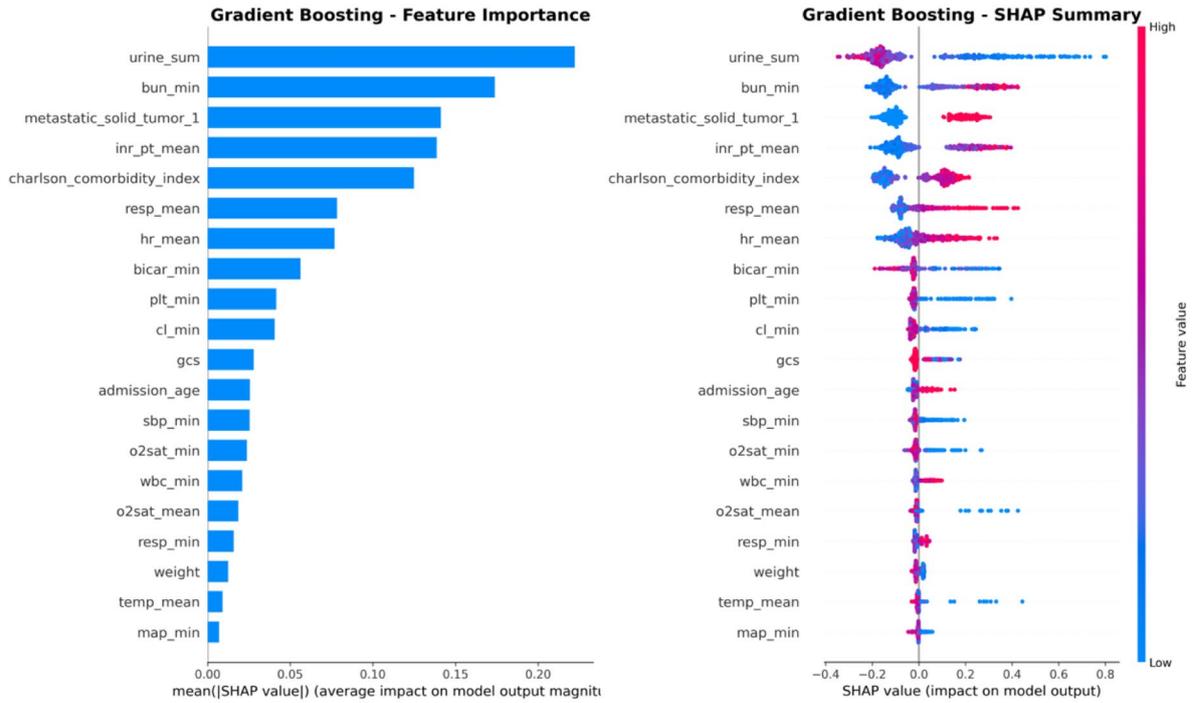
Discussion

Compared with other ICU patients, immunocompromised patients face unique challenges that significantly impact their prognosis. This study leverages a large data set and machine learning algorithms with the primary objective of developing and validating an explainable ML model to predict 28-day all-cause mortality risk in immunocompromised patients admitted to the ICU. Based on large amounts of data and ML algorithms, this study had the following new findings: (i) several factors were consistently associated with the mortality of immunocompromised ICU patients across multiple models, including 24-h urine output, BUN levels, INR levels, the presence of metastatic solid tumors, and CCI and (ii) multiple ML

(See figure on next page.)

Fig. 3 SHAP analysis of key predictive models. Feature importance and SHAP summary plots for (A) gradient boosting, (B) LightGBM, and (C) XGBoost models. Left panels show feature importance ranked by mean absolute SHAP value. Right panels illustrate the impact of feature values on model output, with red indicating higher feature values and blue indicating lower values. The horizontal position shows whether the effect of that value is associated with higher or lower prediction of mortality risk urine_sum, total 24-h urine output; bun_min, minimum blood urea nitrogen (BUN) level; metastatic_solid_tumor, presence of metastatic solid tumor (binary indicator); inr_pt_mean, mean international normalized ratio (INR); charlson_comorbidity_index, Charlson Comorbidity Index; resp_mean, mean respiratory rate; hr_mean, mean heart rate; bicar_min, minimum bicarbonate level; plt_min, minimum platelet count; cl_min, minimum chloride level; o2_sat_min, minimum oxygen saturation (SpO₂); resp_min, minimum respiratory rate; sbp_min, minimum systolic blood pressure; wbc_min, minimum white blood cell (WBC) count; o2_sat_mean, mean oxygen saturation (SpO₂); gcs, Glasgow Coma Scale; map_min, minimum mean arterial pressure (MAP); na_min, minimum sodium level; sofa_score, SOFA score (sequential organ failure assessment); temp_mean, mean body temperature

A.



B.

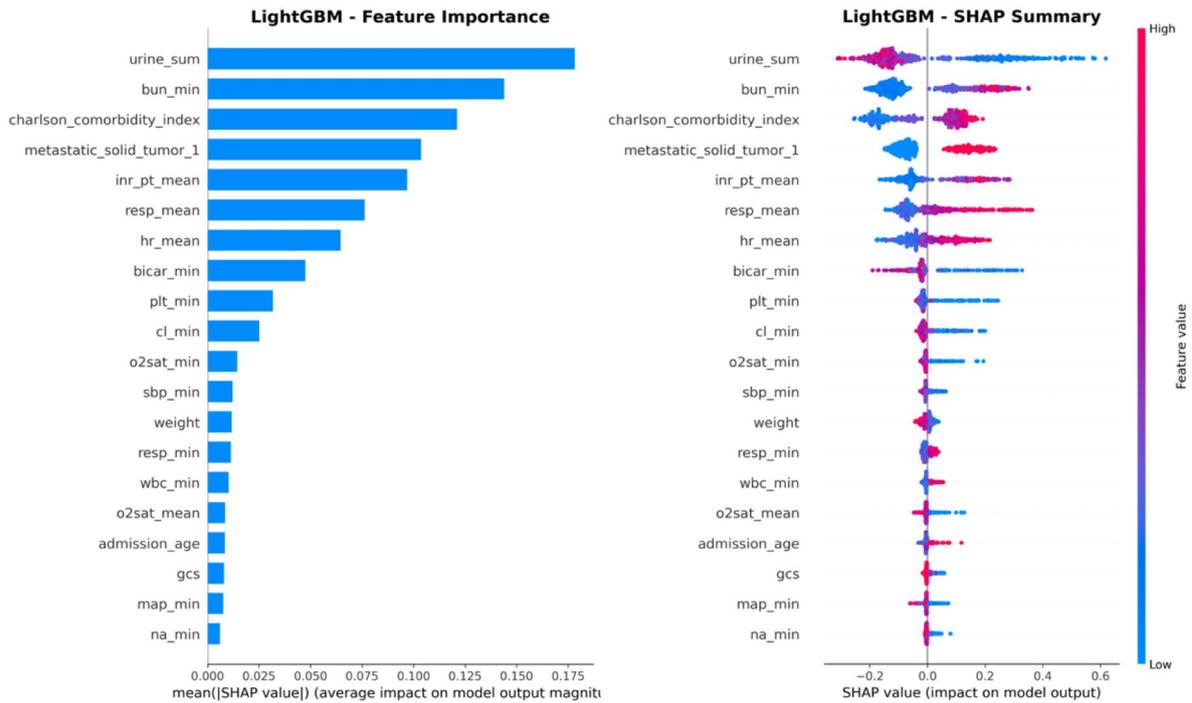


Fig. 3 (See legend on previous page.)

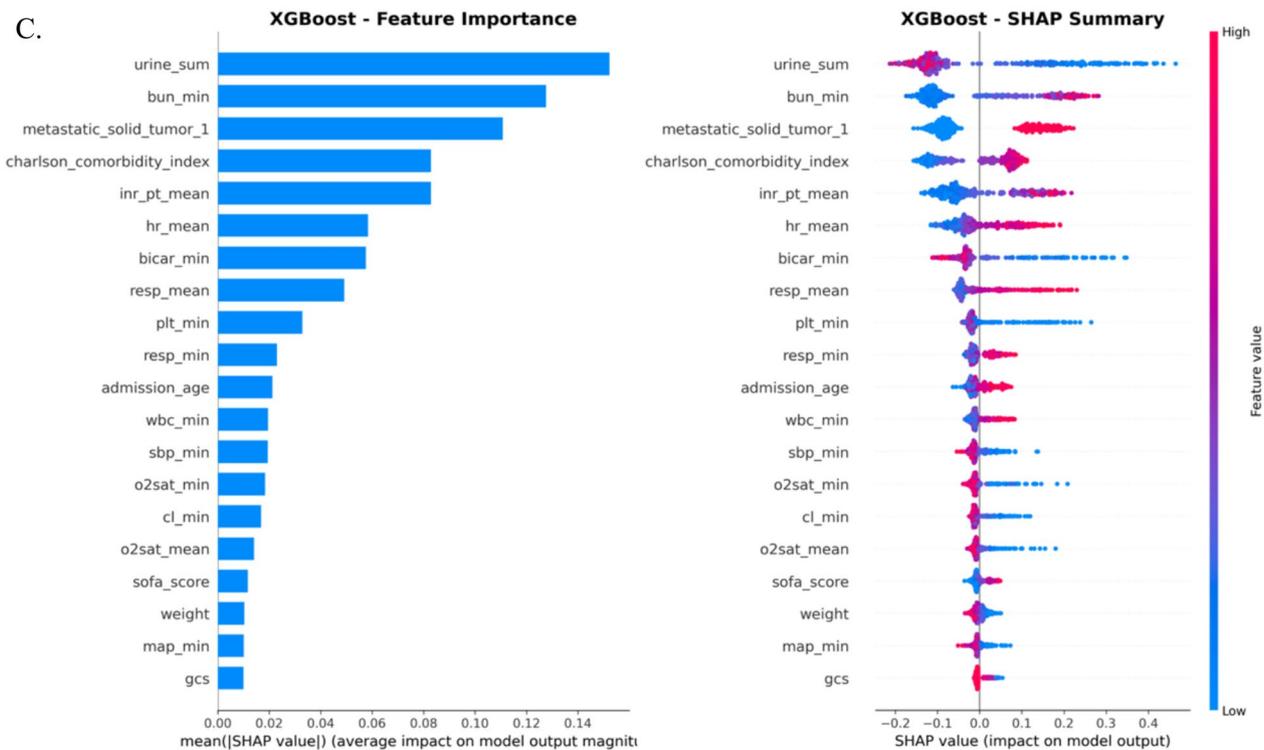


Fig. 3 continued

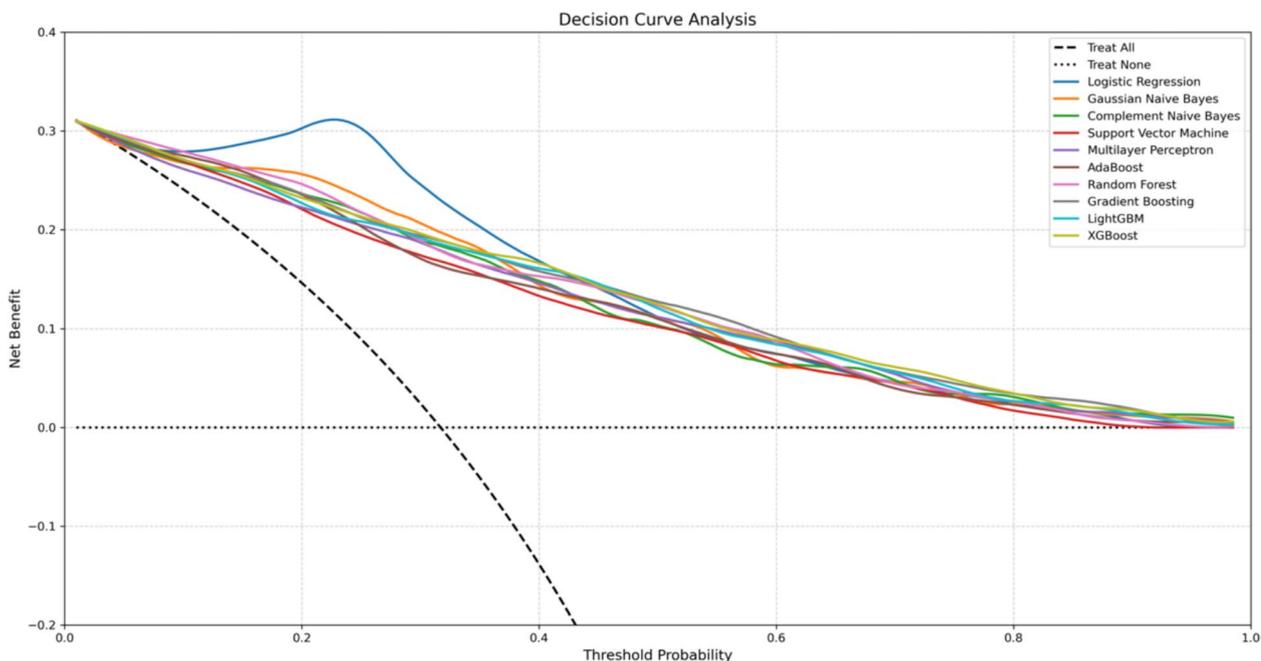


Fig. 4 Decision curve analysis comparing net benefits of different models across threshold probabilities. The horizontal dotted line represents treating no patients, while the sloped dashed line represents treating all patients. All models showed positive net benefits compared to default strategies. XGBoost, extreme gradient boosting

models demonstrated promising performance in predicting mortality, with several models achieving comparable results.

Our findings align with existing research emphasizing the importance of comorbidities and laboratory values in predicting mortality among critically ill patients. In immunosuppressed patients, the presence of metastatic solid tumors significantly worsens prognosis. Vigneron et al. [36] demonstrated in their multivariate analysis that metastatic disease is an independent factor associated with increased ICU mortality, with a cause-specific hazard (CSH) ratio of 1.78 (95% CI 1.38–2.30, $p < 0.001$), and patients with cancer in progression have a similarly increased risk, with a CSH ratio of 1.62 (95% CI 1.28–2.05, $p < 0.001$). Among the 1279 patients with complete follow-up, the 1-year survival rate was 33.2%, with lung and gastrointestinal cancers being the most common tumor sites. Patients requiring intensive care interventions, such as mechanical ventilation or vasopressors, have higher in-hospital mortality rates due to severe complications of metastatic cancer [37]. Immunosuppression complicates management further by impairing the body's ability to combat infections and recover from critical illnesses [38]. In addition, previous studies have shown a consistent association between elevated BUN levels and poor prognosis in critically ill patients [39–41]. In immunosuppressed patients, higher BUN levels, indicative of renal dysfunction, are associated with poorer outcomes. Xia et al. demonstrated that higher BUN/ALB ratio is positively related to 30-day mortality in pneumonia patients receiving glucocorticoids [42]. INR levels, reflecting coagulation abnormalities, also correlate with higher mortality rates among critically ill and immunosuppressed patients [43]. INR is an important predictor for the formation of microthrombi at an advanced stage of septic shock [44, 45]. A retrospective observational study showed that a higher INR was associated with a higher risk of mortality after ICU admission in patients with rheumatoid arthritis [46]. The Charlson Comorbidity Index [47], an assessment tool designed specifically to predict long-term mortality, is a well-established and validated tool to evaluate comorbidities, enabling the early identification of a constellation of symptoms and syndromes in individual patients, and improving prognostic estimations of health risks [48]. CCI is strongly associated with both mortality and length of stay and is used as a prognostic marker in the ICU [49, 50]. In immunosuppressed patients, CCI also shows good performance, such as in cases of kidney transplant [51], inflammatory bowel diseases [52], systemic lupus erythematosus [53], colorectal cancer [54]. Furthermore, urine output within the first 24 h is a vital indicator of renal function and

fluid status, with decreased urine output associated with higher mortality in this patient population [55].

Machine learning has increasingly been integrated into ICU settings to enhance predictive modeling. Numerous studies have shown that ML models can effectively predict critical outcomes, such as sepsis, mortality, and hemodynamic deterioration [56]. ML models significantly outperform traditional methods by leveraging complex interactions between various clinical variables to provide early warnings and improve patient management [57, 58]. The ability of ML models to handle high-dimensional and nonlinear relationships among clinical features makes them particularly suited for the ICU environment, where patient data are vast and complex.

In this study, we observed that several ML models, including Logistic Regression, SVM, and MLP, demonstrated comparable performance in mortality prediction with overlapping confidence intervals in their AUC and PRC values. The use of SHAP values allowed us to decompose the predictions and understand the influence of individual features, thereby addressing the “black-box” issue commonly associated with ML models. The SHAP analysis across multiple models consistently revealed that lower 24-h urine output, elevated BUN levels, presence of metastatic solid tumors, higher CCI, and elevated INR values were associated with increased mortality risk. This consistency across different modeling approaches strengthens our confidence in these findings and can guide clinicians in prioritizing patients who may benefit from more intensive monitoring and early interventions. The decision curve analysis further confirmed the clinical utility of our models, demonstrating positive net benefits across a wide range of threshold probabilities. This analysis complements traditional performance metrics by directly quantifying the clinical value of implementing these models in decision-making processes. The comparable performance of different models in terms of net benefit aligns with their similar discriminative ability shown by conventional metrics, reinforcing the robustness of our findings.

Despite the promising results, several limitations need to be acknowledged. First, the retrospective nature of the study and reliance on the MIMIC-IV database may limit the generalizability of the findings. The patient population in this database may not fully represent all immunocompromised patients in diverse clinical settings. Second, while SHAP values enhance the interpretability of the ML models, they are still subject to the limitations of the underlying model's assumptions and the quality of the input data. Another

limitation is the potential for selection bias due to the exclusion of certain patients based on missing data or other criteria. This could result in a data set that is not fully representative of the broader patient population. In addition, some important clinical variables with a high proportion of missing values were not included in the analysis, which might affect the model's predictive performance. We also acknowledge that while we implemented fivefold cross-validation and rigorous evaluation methods, there remains a degree of performance difference between training and test sets, particularly in tree-based models, highlighting the challenge of developing models that generalize perfectly. Finally, external validation of the models was not performed, which is crucial for assessing their applicability in different clinical environments. Future studies should include external validation cohorts to enhance the robustness and generalizability of the findings.

Conclusions

The development of explainable machine learning models for predicting 28-day all-cause mortality in immunocompromised ICU patients represents a significant advancement in the field of critical care. Our comprehensive evaluation of multiple modeling approaches revealed that several models demonstrate promising and comparable performance in mortality prediction. The interpretability of these models through SHAP values consistently identified key predictors of mortality risk, with 24-h urine output, BUN levels, and metastatic solid tumors emerging as particularly important factors. These insights can help guide clinical decision-making and resource allocation in the management of this vulnerable patient population. Continued research, including external validation studies across diverse clinical settings, is essential to further refine these models and fully realize their potential benefits in improving patient outcomes.

Competing interests

The authors declare no competing interests.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40001-025-02622-3>.

Additional file 1

Author contributions

Z.Y. and L.F. wrote the main manuscript text and prepared all figures. Y.D. reviewed the manuscript. All authors reviewed and approved the final version of the manuscript.

Funding

Not Applicable.

Data availability

The data that support the findings of this study are derived from the MIMIC-IV database (<https://physionet.org/content/mimic-iv>), a publicly available critical care dataset maintained by the Massachusetts Institute of Technology (MIT). Access to MIMIC-IV requires completion of the Collaborative Institutional Training Initiative (CITI) program and adherence to the data use agreement (DUA).

Code availability

The code and analysis used in this study are available at <https://github.com/leanqon/immunocompromised-mortality-ml>.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Institutional Review Board of MIT and Harvard Medical School. Given the retrospective nature of the study using the MIMIC-IV database, the requirement for informed consent was waived. The researchers obtained necessary permissions to access the MIMIC-IV database (certification numbers: 48061114, 38118593).

Author details

¹School of Medicine, Xiamen University, 422 South Siming Road, Xiamen 361005, Fujian, China. ²National Institute for Data Science in Health and Medicine, Xiamen University, 422 South Siming Road, Xiamen 361005, Fujian, China. ³Department of Critical Care Medicine, The Second Affiliated Hospital of Zhejiang Chinese Medical University, 318 Chaowang Road, Hangzhou 31000, Zhejiang, China.

Received: 17 February 2025 Accepted: 21 April 2025

Published online: 03 May 2025

References

1. Mereiter S, Balmaña M, Campos D, Gomes J, Reis CA. Glycosylation in the era of cancer-targeted therapy: where are we heading? *Cancer Cell*. 2019;36(1):6–16. <https://doi.org/10.1016/j.ccell.2019.06.006>.
2. Tang L, Huang Z, Mei H, Hu Y. Immunotherapy in hematologic malignancies: achievements, challenges and future prospects. *Signal Transduct Target Ther*. 2023;8(1):306. <https://doi.org/10.1038/s41392-023-01521-5>.
3. Limaye AP, Babu TM, Boeckh M. Progress and Challenges in the Prevention, Diagnosis, and Management of Cytomegalovirus Infection in Transplantation. *Clin Microbiol Rev*. 2020. <https://doi.org/10.1128/cmr.00043-19>.
4. Schett G, Mackensen A, Mougiakakos D. CAR-T-cell therapy in autoimmune diseases. *Lancet*. 2023;402(10416):2034–44. [https://doi.org/10.1016/s0140-6736\(23\)01126-1](https://doi.org/10.1016/s0140-6736(23)01126-1).
5. Azoulay E, Schellongowski P, Darmon M, et al. The intensive care medicine research agenda on critically ill oncology and hematology patients. *Intensive Care Med*. 2017;43(9):1366–82. <https://doi.org/10.1007/s00134-017-4884-z>.
6. Azoulay E, Russell L, Van de Louw A, et al. Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med*. 2020;46(2):298–314. <https://doi.org/10.1007/s00134-019-05906-5>.
7. Catano J, Sacloux SC, Gornet JM, et al. Gastrointestinal bleeding in critically ill immunocompromised patients. *Ann Intensive Care*. 2021;11(1):130. <https://doi.org/10.1186/s13613-021-00913-6>.
8. Baucher L, Lemiale V, Joseph A, et al. Severe infections requiring intensive care unit admission in patients receiving ibrutinib for hematological malignancies: a groupe de recherche respiratoire en réanimation onco-hématologique (GRRR-OH) study. *Ann Intensive Care*. 2023;13(1):123. <https://doi.org/10.1186/s13613-023-01219-5>.

9. Dumas G, Bertrand M, Lemiale V, et al. Prognosis of critically ill immunocompromised patients with virus-detected acute respiratory failure. *Ann Intensive Care*. 2023;13(1):101.
10. Teh BW, Mikulska M, Averbuch D, et al. Consensus position statement on advancing the standardised reporting of infection events in immunocompromised patients. *Lancet Infect Dis*. 2024;24(1):e59–68.
11. Fizza Haider S, Sloss R, Jhanji S, Nicholson E, Creagh-Brown B. Management of adult patients with haematological malignancies in critical care. *Anaesthesia*. 2023;78(7):874–83.
12. Mallick S, Anila K, Sivaprasadan S, Sudhindran S. Immunosuppression in liver transplant recipients in the setting of sepsis. *J Clin Exp Hepatol*. 2023;13(4):682–90.
13. Giacobbe DR, Dettori S, Di Pilato V, et al. Pneumocystis jirovecii pneumonia in intensive care units: a multicenter study by ESGCIP and EFISG. *Crit Care*. 2023;27(1):323. <https://doi.org/10.1186/s13054-023-04608-1>.
14. Secreto C, Chean D, Van de Louw A, et al. Characteristics and outcomes of patients with acute myeloid leukemia admitted to intensive care unit with acute respiratory failure: a post-hoc analysis of a prospective multicenter study. *Ann Intensive Care*. 2023;13(1):79.
15. Feys S, Lagrou K, Lauwers HM, et al. High burden of COVID-19-associated pulmonary aspergillosis in severely immunocompromised patients requiring mechanical ventilation. *Clin Infect Dis*. 2024;78(2):361–70.
16. Dumas G, Arabi YM, Bartz R, et al. Diagnosis and management of autoimmune diseases in the ICU. *Intensive Care Med*. 2024;50(1):17–35.
17. Mokrani D, Chommeloux J, Pineton de Chambrun M, Hékimian G, Luyt C-E. Antibiotic stewardship in the ICU: time to shift into overdrive. *Ann Intensive Care*. 2023;13(1):39.
18. Dumas G, Lemiale V, Rathi N, et al. Survival in immunocompromised patients ultimately requiring invasive mechanical ventilation: a pooled individual patient data analysis. *Am J Respir Crit Care Med*. 2021;204(2):187–96. <https://doi.org/10.1164/rccm.202009-3575OC>.
19. Luo W, Xing R, Wang C. The effect of ventilator-associated pneumonia on the prognosis of intensive care unit patients within 90 days and 180 days. *BMC Infect Dis*. 2021;21(1):684. <https://doi.org/10.1186/s12879-021-06383-2>.
20. Kreitmann L, Gaudet A, Nseir S. Ventilator-associated pneumonia in immunosuppressed patients. *Antibiotics*. 2023. <https://doi.org/10.3390/antibiotics12020413>.
21. Seidelman JL, Mantyh CR, Anderson DJ. Surgical site infection prevention: a review. *JAMA*. 2023;329(3):244–52.
22. Alverdy JC, Hyman N, Gilbert J. Re-examining causes of surgical site infections following elective surgery in the era of asepsis. *Lancet Infect Dis*. 2020;20(3):e38–43.
23. Timsit J-F, Ruppé E, Barbier F, Tabah A, Bassetti M. Bloodstream infections in critically ill patients: an expert statement. *Intensive Care Med*. 2020;46(2):266–84.
24. Li Z, Zhuang H, Wang G, Wang H, Dong Y. Prevalence, predictors, and mortality of bloodstream infections due to methicillin-resistant *Staphylococcus aureus* in patients with malignancy: systemic review and meta-analysis. *BMC Infect Dis*. 2021;21:1–10.
25. Zhang L, Zhen S, Shen Y, et al. Bloodstream infections due to Carbapenem-resistant enterobacteriaceae in hematological patients: assessment of risk factors for mortality and treatment options. *Ann Clin Microbiol Antimicrob*. 2023;22(1):41.
26. Lindell RB, Nishisaki A, Weiss SL, Traynor DM, Fitzgerald JC. Risk of mortality in immunocompromised children with severe sepsis and septic shock. *Crit Care Med*. 2020;48(7):1026–33. <https://doi.org/10.1097/ccm.0000000000004329>.
27. Reddy V, Reddy H, Gemnani R, Kumar S, Acharya S, Reddy V. Navigating the complexity of scoring systems in sepsis management: a comprehensive review. *Cureus*. 2024. <https://doi.org/10.7759/cureus.54030>.
28. Hong N, Liu C, Gao J, et al. State of the art of machine learning-enabled clinical decision support in intensive care units: literature review. *JMIR Med Inform*. 2022;10(3): e28781.
29. Quinn TP, Jacobs S, Senadeera M, Le V, Coghlan S. The three ghosts of medical AI: can the black-box present deliver? *Artif Intell Med*. 2022;124: 102158.
30. Nohara Y, Matsumoto K, Soejima H, Nakashima N. Explanation of machine learning models using improved shapley additive explanation. 2019:546-546.
31. Azoulay E, Russell L, Van de Louw A, et al. Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med*. 2020;46:298–314.
32. Johnson AE, Ghassemi MM, Nemati S, Niehaus KE, Clifton DA, Clifford GD. Machine learning and decision support in critical care. *Proc IEEE*. 2016;104(2):444–66.
33. Endo H, Uchino S, Hashimoto S, et al. Development and validation of the predictive risk of death model for adult patients admitted to intensive care units in Japan: an approach to improve the accuracy of healthcare quality measures. *J Intensive Care*. 2021;9:1–11.
34. Johnson AE, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10(1):1.
35. Kreitmann L, Helms J, Martin-Loeches I, et al. ICU-acquired infections in immunocompromised patients. *Intensive Care Med*. 2024;50:1–18.
36. Vigneron C, Charpentier J, Valade S, et al. Patterns of ICU admissions and outcomes in patients with solid malignancies over the revolution of cancer treatment. *Ann Intensive Care*. 2021;11:1–10.
37. Geijteman ECT, Kuip EJM, Oskam J, Lees D, Bruera E. Illness trajectories of incurable solid cancers. *Bmj*. 2024;384: e076625. <https://doi.org/10.1136/bmj-2023-076625>.
38. Gonzalez F, Starka R, Ducros L, et al. Critically ill metastatic cancer patients returning home after unplanned ICU stay: an observational, multicentre retrospective study. *Ann Intensive Care*. 2023;13(1):73. <https://doi.org/10.1186/s13613-023-01170-5>.
39. Weng J, Hou R, Zhou X, et al. Development and validation of a score to predict mortality in ICU patients with sepsis: a multicenter retrospective study. *J Transl Med*. 2021;19:1–12.
40. Huang D, Yang H, Yu H, et al. Blood urea nitrogen to serum albumin ratio (BAR) predicts critical illness in patients with coronavirus disease 2019 (COVID-19). *Int J General Med*. 2021: 4711-4721.
41. Yue S, Li S, Huang X, et al. Machine learning for the prediction of acute kidney injury in patients with sepsis. *J Transl Med*. 2022;20(1):215.
42. Xia B, Song B, Zhang J, Zhu T, Hu H. Prognostic value of blood urea nitrogen-to-serum albumin ratio for mortality of pneumonia in patients receiving glucocorticoids: secondary analysis based on a retrospective cohort study. *J Infect Chemother*. 2022;28(6):767–73.
43. Fei A, Lin Q, Liu J, Wang F, Wang H, Pan S. The relationship between coagulation abnormality and mortality in ICU patients: a prospective, observational study. *Sci Rep*. 2015;5:9391. <https://doi.org/10.1038/srep09391>.
44. Peltan ID, Vande Vusse LK, Maier RV, Watkins TR. An international normalized ratio-based definition of acute traumatic coagulopathy is associated with mortality, venous thromboembolism, and multiple organ failure after injury. *Crit Care Med*. 2015;43(7):1429–38. <https://doi.org/10.1097/ccm.0000000000000981>.
45. Kanda N, Ohbe H, Nakamura K. Effects of antithrombin on persistent inflammation, immunosuppression, and catabolism syndrome among patients with sepsis-induced disseminated intravascular coagulation. *J Clin Med*. 2023;12(11):3822.
46. Fujiwara T, Tokuda K, Momii K, et al. Prognostic factors for the short-term mortality of patients with rheumatoid arthritis admitted to intensive care units. *BMC Rheumatol*. 2020;4(1):64. <https://doi.org/10.1186/s41927-020-00164-1>.
47. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373–83.
48. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: a critical review of clinimetric properties. *Psychother Psychosom*. 2022;91(1):8–35.
49. Yıldız A, Yiğit A, Benli AR. The prognostic role of Charlson comorbidity index for critically ill elderly patients. *Eur Res J*. 2020. <https://doi.org/10.18621/eurj.451391>.
50. Aronsson Dannewitz A, Svennblad B, Michaëlsson K, Lipcsey M, Gedeberg R. Optimized diagnosis-based comorbidity measures for all-cause mortality prediction in a national population-based ICU population. *Crit Care*. 2022;26(1):306.
51. Levine MA, Schuler T, Gourishankar S. Complications in the 90-day postoperative period following kidney transplant and the relationship of the Charlson comorbidity index. *Can Urol Assoc J*. 2017;11(12):388–93. <https://doi.org/10.5489/cuaj.4378>.

52. Kochar B, Cai W, Cagan A, Ananthakrishnan AN. Pretreatment frailty is independently associated with increased risk of infections after immunosuppression in patients with inflammatory bowel diseases. *Gastroenterology*. 2020;158(8): e2.
53. Kim S-K, Choe J-Y, Lee S-S. Charlson comorbidity index is related to organ damage in systemic lupus erythematosus: data from KORean lupus Network (KORNET) registry. *J Rheumatol*. 2017;44(4):452–8.
54. Marventano S, Grosso G, Mistretta A, et al. Evaluation of four comorbidity indices and Charlson comorbidity index adjustment for colorectal cancer patients. *Int J Colorectal Dis*. 2014;29:1159–69.
55. Vincent J-L, Ferguson A, Pickkers P, et al. The clinical relevance of oliguria in the critically ill patient: analysis of a large observational database. *Crit Care*. 2020;24:1–14.
56. Bomrah S, Uddin M, Upadhyay U, et al. A scoping review of machine learning for sepsis prediction- feature engineering strategies and model performance: a step towards explainability. *Crit Care*. 2024;28(1):180. <https://doi.org/10.1186/s13054-024-04948-6>.
57. Lim L, Gim U, Cho K, Yoo D, Ryu HG, Lee H-C. Real-time machine learning model to predict short-term mortality in critically ill patients: development and international validation. *Crit Care*. 2024;28(1):76.
58. Deng H-F, Sun M-W, Wang Y, et al. Evaluating machine learning models for sepsis prediction: a systematic review of methodologies. *Iscience*. 2022;25(1):103651.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.