

RESEARCH

Open Access



An artificial intelligence application to predict prolonged dependence on mechanical ventilation among patients with critical orthopaedic trauma: an establishment and validation study

Weigang Jiang¹, Tao Liu², Baisheng Sun^{3,4}, Lixia Zhong⁵, Zhencan Han⁶, Minhua Lu^{1*} and Mingxing Lei^{7,8*}

Abstract

Background Prolonged dependence on mechanical ventilation is a common occurrence in clinical ICU patients and presents significant challenges for patient care and resource allocation. Predicting prolonged dependence on mechanical ventilation is crucial for improving patient outcomes, preventing ventilator-associated complications, and guiding targeted clinical interventions. However, specific tools for predicting prolonged mechanical ventilation among ICU patients, particularly those with critical orthopaedic trauma, are currently lacking. The purpose of the study was to establish and validate an artificial intelligence (AI) platform to assess the prolonged dependence on mechanical ventilation among patients with critical orthopaedic trauma.

Methods This study analyzed 1400 patients with critical orthopaedic trauma who received mechanical ventilation, and the prolonged dependence on mechanical ventilation was defined as not weaning from mechanical ventilation for ≥ 7 days. Patients were randomly classified into a training cohort and a validation cohort based on the ratio of 8:2. Patients in the training cohort were used to establish models using machine learning techniques, including logistic regression (LR), extreme gradient boosting machine (eXGBM), decision tree (DT), random forest (RF), support vector machine (SVM), and light gradient boosting machine (LightGBM), whereas patients in the validation cohort were used to validate these models. The prediction performance of these models was evaluated using discrimination and calibration. A scoring system was used to comprehensively assess and compare the prediction performance of the models, based on ten evaluation metrics. External validation of the model was performed in 122 patients with critical orthopaedic trauma from a university teaching hospital. Furthermore, the optimal model was deployed as an AI calculator, which was accessible online, to assess the risk of prolonged dependence on mechanical ventilation.

*Correspondence:

Minhua Lu

lmh910307@126.com

Mingxing Lei

leimingxing2@sina.com

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



© The Author(s) 2024. **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/>.

Results Among the developed models, the eXGBM model had the highest score of 50, followed by the LightGBM model (48) and the RF model (37). In detail, the eXGBM model outperformed other models in terms of recall (0.892), Brier score (0.088), log loss (0.291), and calibration slope (0.999), and the model was the second best in terms of area under the curve value (0.949, 95%: 0.933–0.961), accuracy (0.871), F1 score (0.873), and discrimination slope (0.647). The SHAP revealed that the most important five features were respiratory rate, lower limb fracture, glucose, PaO₂, and PaCO₂. External validation of the eXGBM model also demonstrated favorable prediction performance, with an AUC value of 0.893 (95%CI: 0.819–0.967). The eXGBM model was successfully deployed as an AI platform, which was at <https://prolongedmechanicalventilation-lqsfm6ecky6dpd4ybkvohu.streamlit.app/>. By simply clicking the link and inputting features, users were able to obtain the risk of experiencing prolonged dependence on mechanical ventilation for individuals. Based on the risk of prolonged dependence on mechanical ventilation, patients were stratified into the high-risk or the low-risk groups, and corresponding therapeutic interventions were recommended, accordingly.

Conclusions The AI model shows potential as a valuable tool for stratifying patients with a high risk of prolonged dependence on mechanical ventilation. The AI model may offer a promising approach for optimizing patient care and resource allocation in critical care settings.

Clinical trial number Not applicable.

Keywords Artificial intelligence, Mechanical ventilation, Critical orthopaedic trauma, Machine learning, A comprehensive evaluation system

Introduction

Prolonged dependence on mechanical ventilation is a significant concern in intensive care units (ICUs) and poses various challenges for patient care [1]. The epidemiological impact of prolonged mechanical ventilation in ICU patients includes increased susceptibility to ventilator-associated complications, such as ventilator-associated pneumonia [2], lung and diaphragm injury [3], inspiratory muscle weakness [4], and longer ICU and hospital stays [1, 5], leading to a substantial healthcare burden. Patients requiring prolonged mechanical ventilation may also experience poor survival outcomes [6, 7]. Identifying factors contributing to prolonged mechanical ventilation is crucial for improving patient outcomes and optimizing resource allocation in critical care settings.

Several factors have been identified as influencing the duration of mechanical ventilation among ICU patients [1, 8, 9]. These factors include pre-existing comorbidities such as chronic lung diseases, hematologic malignancies, and previous endotracheal intubation, cause of the respiratory failure [10], as well as critical illness-related factors like poor oxygenation, high peak inspiratory pressures, and the need for renal replacement therapy [1, 8, 9]. Understanding these factors is essential for risk stratification and the development of targeted interventions to prevent and manage prolonged mechanical ventilation. In recent years, the application of artificial intelligence (AI) in critical care medicine has shown promise in predicting and managing critical outcomes [11–13]. Machine learning techniques have been utilized to develop predictive models for various clinical outcomes, including ventilator-associated complications [14], extubation readiness [15], and mechanical ventilator dependence [16]. These

models have the potential to improve risk stratification and guide individualized patient care in critical care settings. However, prediction models for prolonged dependence on mechanical ventilation particularly among patients with critical orthopaedic trauma are scarce. Patients suffering from severe orthopedic trauma often exhibit distinct clinical features compared to other critically ill populations. These individuals frequently endure multiple injuries, substantial blood loss [17], and the development of complications such as compartment syndrome, which often necessitate complex and prolonged surgical interventions [18]. The cumulative effect of systemic stress, compounded by extended immobilization, significantly increases the risk of complications like infections, deep vein thrombosis, and respiratory failure. Consequently, mechanical ventilation becomes a crucial aspect of their critical care management. Given these challenges, early recognition of patients at risk for prolonged mechanical ventilation dependency is essential to improving outcomes and reducing morbidity [19]. However, there is a lack of specific tools in clinical practice for predicting prolonged dependence on mechanical ventilation in ICU patients. This may be due to the absence of large-scale, trauma-specific datasets, as most research has focused on broader critical care populations, such as those with sepsis or acute respiratory distress syndrome. This absence of predictive tools makes it challenging for healthcare professionals to accurately assess the risk of prolonged mechanical ventilation for patients in the ICU, hindering the timely implementation of targeted clinical interventions.

Therefore, the objective of this study is to establish and validate an AI platform to predict prolonged dependence

on mechanical ventilation among patients with critical orthopaedic trauma. Building on the identified risk factors and leveraging machine learning techniques, our hypothesis is that the developed AI model will accurately stratify patients based on their risk of prolonged mechanical ventilation, providing a valuable tool for optimizing patient care and resource allocation in critical care settings. Additionally, the study aims to further deploy the optimal model as an accessible and user-friendly AI application to support clinical decision-making and improve patient outcomes in the critical care setting.

Methods

Database and study design

This study analyzed 1400 patients with critical orthopaedic trauma who received mechanical ventilation from the Medical Information Mart for Intensive Care III (MIMIC-III) database, which is a freely accessible critical care database developed by the Massachusetts Institute of Technology. The MIMIC III database is widely used for clinical research and has been utilized to develop artificial intelligence models for various medical applications [20–24], and it is an invaluable resource for researchers to study critical care medicine, develop predictive models, and conduct observational studies. Access to the MIMIC III database is granted after completing the National Institutes of Health's web-based training course on Protecting Human Research Participants and the Good Clinical Practice course. Once approved, researchers can access and download the de-identified data for their research studies. The use of the MIMIC-III database was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center (32128436) [25]. Since the data in the MIMIC-III database has been de-identified, patient consent was not required.

In the context of this study, the MIMIC III database was used to extract data on 1400 patients with critical orthopaedic trauma who required mechanical ventilation. This rich and extensive dataset provided a large observational cohort for the study, allowing for the development and validation of a predictive model for prolonged dependence on mechanical ventilation among these patients. In the previous study, we extracted 2,662 critically ill patients with orthopedic trauma to develop a model for predicting 30-day mortality [26]. In this study, we further excluded 26 individuals under the age of 18 and 1,236 individuals who did not undergo mechanical ventilation. Hence, in this study, we obtained 1,400 patients for analysis.

The patients were randomly classified into a training cohort and a validation cohort based on the ratio of 8:2. Patients in the training cohort were used to establish models using machine learning techniques, whereas patients in the validation cohort were used to validate

these models (Fig. 1). External validation of the model was performed in 122 patients with critical orthopaedic trauma from a university teaching hospital, and the Ethics Committee of our hospital approved the study. We commit to strictly adhere to ethical guidelines and legal regulations at all stages of the research and ensure the appropriate use and protection of patient privacy.

Primary outcome

The primary outcome of the study was to assess the prolonged dependence on mechanical ventilation. In this study, prolonged dependence on mechanical ventilation was defined as the inability to wean from mechanical ventilation for a period of at least 7 days [8]. This outcome was chosen as it is a critical factor in the recovery and prognosis of patients requiring mechanical ventilation.

Clinical characteristics

We analyzed various clinical characteristics from the patients, including demographics such as age and gender, addiction history including alcohol abuse and tobacco use disorder, laboratory findings such as hemoglobin, glucose, hematocrit, and platelet count, information on injuries like spine, pelvic, low limb, and upper limb fractures, vital signs including PaO₂, PaCO₂, pH, heart rate, and respiratory rate, details on vasopressor use, severity of disease scores (Oxford Acute Severity of Illness Score [OASIS], Simplified Acute Physiology Score II [SAPSII], Sequential Organ Failure Assessment [SOFA]), and comorbidities like hypertension, diabetes, and obesity. These characteristics are crucial for understanding the patients' health status and analyzing their condition. All features were extracted at the first instance after the patient was admitted to the ICU. In addition, the selection of these parameters was guided by previous research, expert recommendations, and the availability of variables. Additionally, the proportion of missing data was a critical factor in our decision-making process.

Data preparation

In the data preparation phase for machine learning, we utilized the SMOTETomek resampling technique to balance the data. This method combines the SMOTE (Synthetic Minority Over-sampling Technique) and Tomek links to address class imbalance. Additionally, a stratified strategy was applied to ensure that the two outcome variables were consistent and adequately represented in the training and testing sets. This approach could help maintain the proportion of different classes in both the training and testing datasets. In detail, after applying SMOTETomek resampling, we obtained a total of 2,782 samples, with the incidence of prolonged dependence on mechanical ventilation at 50%. This resulted in a completely balanced outcome following the SMOTETomek

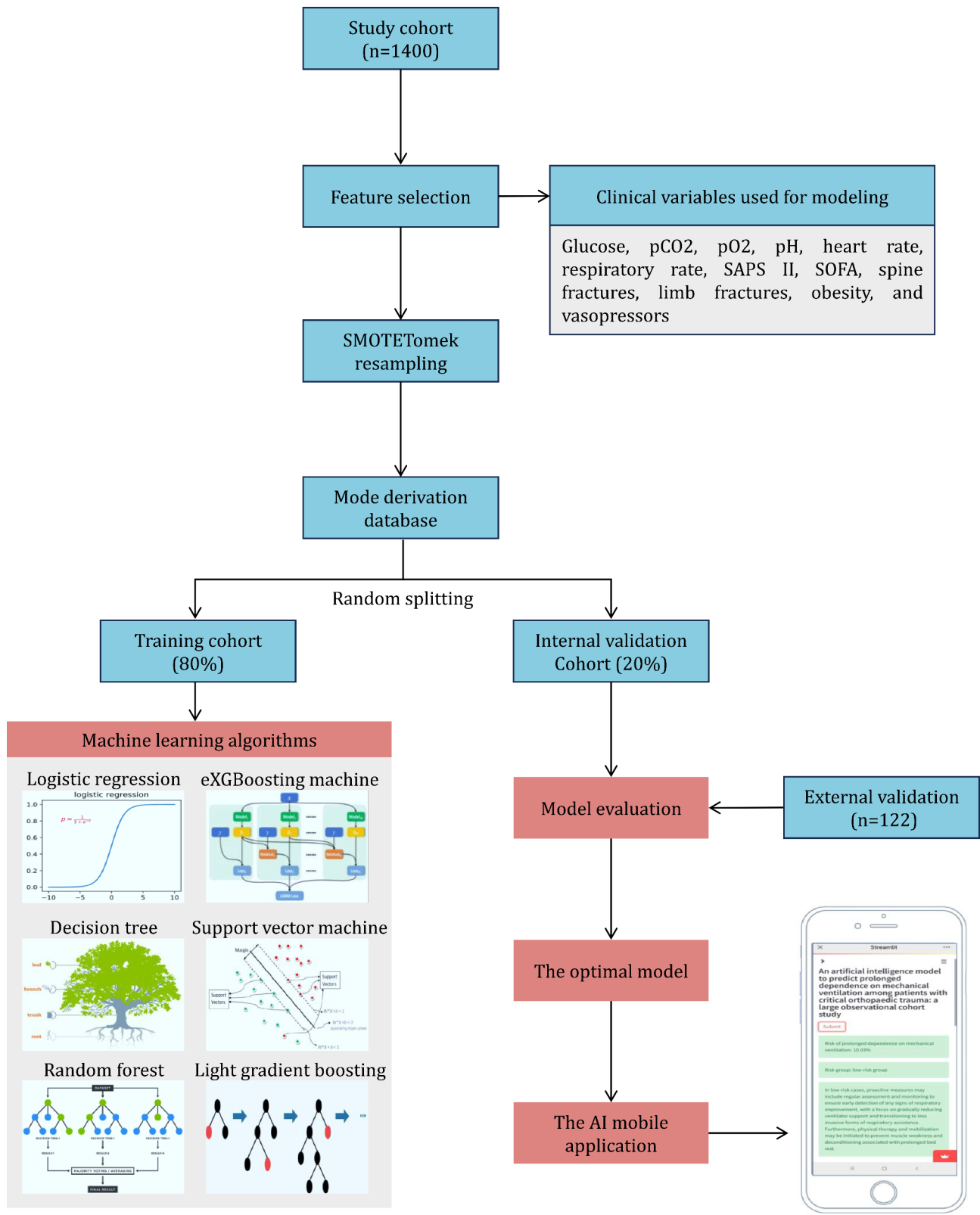


Fig. 1 Study design and machine learning techniques used in the study

resampling. Furthermore, in the model development cohort, the incidence of prolonged dependence on mechanical ventilation was 50%, and the model internal validation cohort showed an incidence of 50% for prolonged dependence on mechanical ventilation as well. Furthermore, data preprocessing pipelines were utilized to prepare the data in a format that could be fitted into scikit-learn algorithms. These pipelines included various steps such as scaling, encoding categorical variables, and data standardization to ensure the data is in a suitable format for training machine learning models.

Modeling and validation

In the study, we employed various machine learning models, including logistic regression (LR), extreme gradient boosting machine (eXGBM), decision tree (DT), random forest (RF), support vector machine (SVM), and light gradient boosting machine (LightGBM). The introduction to the above machine learning algorithms is summarized in Supplementary Table 1. The hyperparameters for these models were optimized using grid search or random grid search techniques, involving an exhaustive search over a specified parameter grid to identify the best hyperparameters for the models. This parameter tuning process was essential for enhancing the models' predictive performance. To evaluate the prediction performance of these models, we employed methods to assess discrimination and calibration. Discrimination and calibration are important metrics for evaluating the models' ability to distinguish between different classes and the agreement between predicted probabilities and actual outcomes. Furthermore, a scoring system [27, 28] was utilized to comprehensively assess and compare the prediction performance of the models. This scoring system was based on ten evaluation metrics, including area under the curve (AUC), accuracy, precision, recall, F1 score, Brier score, log loss, discrimination slope, calibration slope, and intercept-in-large value. These metrics provided a comprehensive evaluation of the models' performance, taking into account various aspects of prediction accuracy, calibration, and model fit.

Model explanation of feature importance

Model explanation using SHAP assessment of variable importance was conducted to understand the impact of different features on the predictions made by the machine learning models. The SHAP (SHapley Additive exPlanations) values provide insights into the contribution of each feature in influencing the model's output, thus offering a comprehensive understanding of variable importance [29].

Design of the application

The optimal model was deployed as an AI application, accessible online, to assess the risk of prolonged dependence on mechanical ventilation. This AI application aimed to provide a user-friendly interface for healthcare professionals to input patient data, obtain risk predictions, and receive personalized treatment recommendations, thereby supporting clinical decision-making for patients requiring mechanical ventilation. The AI application was constructed in the following four main parts based on previous studies [30]. (1) Data input: Users can input patient data, including demographics, clinical characteristics, and laboratory findings. (2) Risk prediction of prolonged dependence on mechanical ventilation and patient's stratification: The application utilizes the deployed model to predict the risk of prolonged dependence on mechanical ventilation and stratify patients based on their risk levels. (3) Variable importance ranking based on SHAP: The application provides a ranking of risk factors based on their importance in influencing the predicted risk of prolonged dependence on mechanical ventilation, leveraging SHAP values to offer transparency and interpretability. (4) Personalized treatment recommendations: Based on the risk assessment and variable importance ranking, the application generates personalized treatment recommendations tailored to individual patient profiles, aiding clinicians in making informed decisions about patient care.

Statistical analysis

The statistical analysis section of the study involved the presentation of quantitative data in the form of median [IQR] and qualitative data in percentage format. Group comparisons for quantitative data were conducted using the Wilcoxon rank-sum test, while the Chi-square test was employed for qualitative data. Quantitative data were imputed using multiple imputation techniques, while qualitative data were filled in using the mode. Variables with more than 20% missing data were not included for analysis in this study. The machine learning modeling and validation were performed using the Python software, version 3.9.7, which included the implementation of various machine learning algorithms for predictive modeling and the evaluation of model performance. In addition, standard statistical analyses were carried out using the R programming language, version 4.1.2, for tasks such as descriptive statistics and traditional inferential statistical tests. A significance level of $p < 0.05$ was used to determine statistical significance.

Results

Patient clinical characteristics

A total of 1400 patients were included in the study, with a median age of 55 years. The majority of patients,

constituting 65.0% of the total, were male. The median values for glucose, hematocrit, and platelet count were observed to be 137.70 mg/dL, 36.25%, and 237.50 K/uL, respectively. The spine was identified as the most common site of fracture, accounting for 50.4% of all patients. Analysis of vital signs revealed median values of 177.50 mmHg for PaO₂, 42.30 mmHg for PaCO₂, 89.00 BPM for heart rate, and 17.00 BPM for respiratory rate. Hypertension emerged as the most prevalent comorbidity among the patients, with a prevalence of 31.9%, followed by diabetes at 17.3%. A more comprehensive overview of disease severity can be found in Table 1. The incidence rate of prolonged dependence on mechanical ventilation was 15.64%.

Subgroup analysis of patients stratified by prolonged dependence on mechanical ventilation

Subgroup analysis revealed significant differences between patients with prolonged dependence on mechanical ventilation and those without in multiple variables. Patients with prolonged dependence on mechanical ventilation exhibited elevated glucose levels ($P < 0.001$), increased PaCO₂ levels ($P < 0.001$), lower PaO₂ levels ($P = 0.006$), decreased pH levels ($P < 0.001$), elevated heart rate ($P = 0.006$) and respiratory rate ($P < 0.001$), higher SAPS II ($P < 0.001$) and SOFA ($P < 0.001$) scores, a higher incidence of spine fractures ($P < 0.001$), a lower incidence of lower limb fractures ($P = 0.025$), a higher prevalence of obesity ($P = 0.023$), and a greater usage of vasopressors ($P < 0.001$). Thus, the

Table 1 Patient’s clinical characteristics and a comparison according to the presence of prolonged dependence on mechanical ventilation among critically ill patients with orthopaedic trauma

Clinical characteristics	Overall	Prolonged dependence on mechanical ventilation		P
		No	Yes	
<i>n</i>	1400	1181	219	
Demographics				
Age (years, median [IQR])	55.00 [36.00, 73.00]	55.00 [35.00, 73.20]	56.00 [40.00, 72.00]	0.832
Gender (male/female, %)	910/490 (65.0/35.0)	759/422 (64.3/35.7)	151/68 (68.9/31.1)	0.209
Addiction				
Alcohol abuse (no/yes, %)	1246/154 (89.0/11.0)	1053/128 (89.2/10.8)	193/26 (88.1/11.9)	0.740
Tobacco use disorder (no/yes, %)	1268/132 (90.6/9.4)	1070/111 (90.6/9.4)	198/21 (90.4/9.6)	1.000
Laboratory examination				
Hemoglobin (g/dL, median [IQR])	12.30 [10.80, 13.83]	12.40 [10.80, 13.80]	12.30 [10.80, 13.90]	0.765
Glucose (mg/dL, median [IQR])	137.70 [115.00, 171.00]	134.00 [114.00, 165.00]	150.00 [128.00, 188.00]	< 0.001
Hematocrit (% , median [IQR])	36.25 [31.90, 40.12]	36.30 [31.90, 40.10]	36.10 [31.95, 40.70]	0.723
Platelet count (K/uL, median [IQR])	237.50 [187.00, 296.00]	238.00 [186.00, 296.00]	232.00 [190.50, 287.00]	0.581
Injury information				
Spine fracture (no/yes, %)	694/706 (49.6/50.4)	612/569 (51.8/48.2)	82/137 (37.4/62.6)	< 0.001
Pelvic fracture (no/yes, %)	1125/275 (80.4/19.6)	957/224 (81.0/19.0)	168/51 (76.7/23.3)	0.166
Low limb fracture (no/yes, %)	881/519 (62.9/37.1)	728/453 (61.6/38.4)	153/66 (69.9/30.1)	0.025
Upper limb fracture (no/yes, %)	906/494 (64.7/35.3)	775/406 (65.6/34.4)	131/88 (59.8/40.2)	0.115
Vital signs				
pO ₂ (mmHg, median [IQR])	177.50 [96.00, 300.25]	183.00 [99.00, 306.00]	149.00 [81.50, 267.00]	0.006
pCO ₂ (mmHg, median [IQR])	42.30 [38.00, 49.00]	42.00 [37.00, 49.00]	44.80 [39.00, 53.00]	< 0.001
pH (unit, median [IQR])	7.34 [7.27, 7.40]	7.34 [7.28, 7.40]	7.31 [7.23, 7.38]	< 0.001
Heart rate (BPM, median [IQR])	89.00 [76.00, 104.00]	88.00 [75.00, 103.00]	94.00 [77.00, 110.00]	0.006
Respiratory rate (BPM, median [IQR])	17.00 [14.00, 21.00]	17.00 [14.00, 20.00]	19.00 [15.00, 24.00]	< 0.001
Vasopressors (no/yes, %)	847/553 (60.5/39.5)	763/418 (64.6/35.4)	84/135 (38.4/61.6)	< 0.001
Severity of disease				
OASIS (median [IQR])	35.00 [31.00, 41.00]	35.00 [30.00, 41.00]	36.00 [31.00, 41.00]	0.150
SAPSII (median [IQR])	33.00 [24.00, 43.00]	32.00 [23.00, 43.00]	36.00 [27.00, 45.00]	< 0.001
SOFA (median [IQR])	4.00 [2.00, 6.00]	3.00 [2.00, 6.00]	5.00 [2.00, 7.00]	< 0.001
Comorbidities				
Hypertension (no/yes, %)	953/447 (68.1/31.9)	799/382 (67.7/32.3)	154/65 (70.3/29.7)	0.485
Diabetes (no/yes, %)	1158/242 (82.7/17.3)	987/194 (83.6/16.4)	171/48 (78.1/21.9)	0.061
Obesity (no/yes, %)	1331/69 (95.1/4.9)	1130/51 (95.7/4.3)	201/18 (91.8/8.2)	0.023

IQR, Interquartile range; BPM, Beats per minute; OASIS, Oxford acute severity of illness score; SAPS II, Simplified acute physiology score II; SOFA, Sequential organ failure assessment score

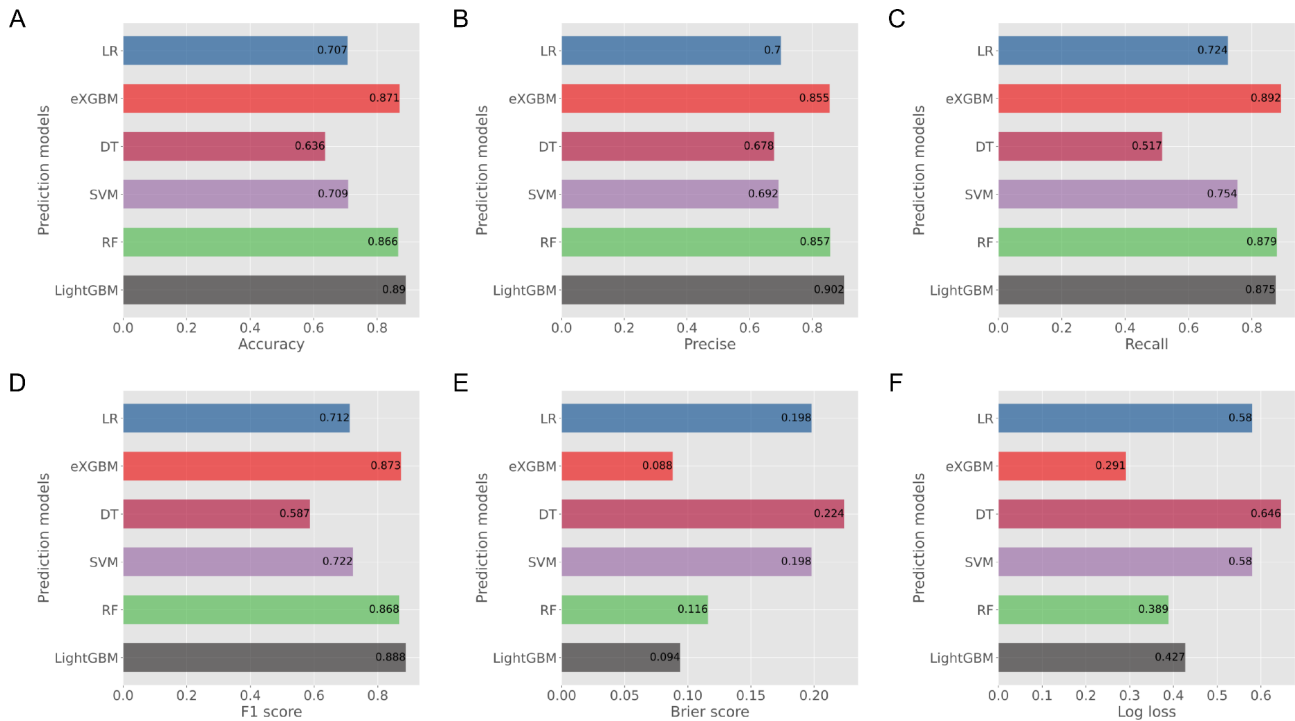


Fig. 2 Prediction performance of machine learning models developed in the study. (A) Accuracy; (B) Precise; (C) Recall; (D) F1 score; (E) Brier score; (F) Log loss

Table 2 Prediction performance for each model in the validation cohort

Metrics	LR	eXGBM	DT	SVM	RF	LightGBM
AUC (95%CI)	0.769 (0.727–0.804)	0.949 (0.933–0.961)	0.710, (0.670–0.756)	0.771 (0.726–0.813)	0.936 (0.912–0.956)	0.953 (0.938–0.965)
Accuracy	0.707	0.871	0.636	0.709	0.866	0.890
Precise	0.700	0.855	0.678	0.692	0.857	0.902
Recall	0.724	0.892	0.517	0.754	0.879	0.875
F1 score	0.712	0.873	0.587	0.722	0.868	0.888
Brier score	0.198	0.088	0.224	0.198	0.116	0.094
Log loss	0.580	0.291	0.646	0.580	0.389	0.427
Discrimination slope	0.188	0.647	0.171	0.180	0.422	0.749
Calibration slope	1.169	0.999	0.625	1.275	2.074	0.375
Intercept-in-large value	0.049	-0.108	0.047	0.037	-0.112	0.100
Total score	28	50	17	32	37	48

AUC, area under the curve; CI, confident interval

above variables were used as model features for further analysis.

Prediction performance

The eXGBM model outperformed other models in terms of recall (0.892), Brier score (0.088), and log loss (0.291) (Fig. 2; Table 2), and the model was the second best in terms of AUC value (0.949, 95%: 0.933–0.961) (Fig. 3), accuracy (0.871), F1 score (0.873). Calibration curve showed that the majority of the models had favorable calibration ability (Supplementary Fig. 1), with the eXGBM model had the best calibration slope (0.999), which was very close to 1 (Supplementary Fig. 2). Probability density curve analysis depicted that the models, particularly the

eXGBM model, the RF model, and the LightGBM model, had very favorable discrimination ability, as the peaks of the density curves were well separated between patients with and without prolonged dependence on mechanical ventilation (Fig. 4). Violin plots also demonstrated that the LightGBM model had the highest discrimination slope (0.749), followed by the eXGBM model (0.647) and the RF model (0.422) (Supplementary Fig. 3). Based on the evaluation scoring system, the eXGBM model had the highest score of 50, closely followed by the LightGBM model (48) and the RF model (37) (Fig. 5). Regarding the external validation cohort, the incidence rate of prolonged dependence on mechanical ventilation was 14.75%. In addition,

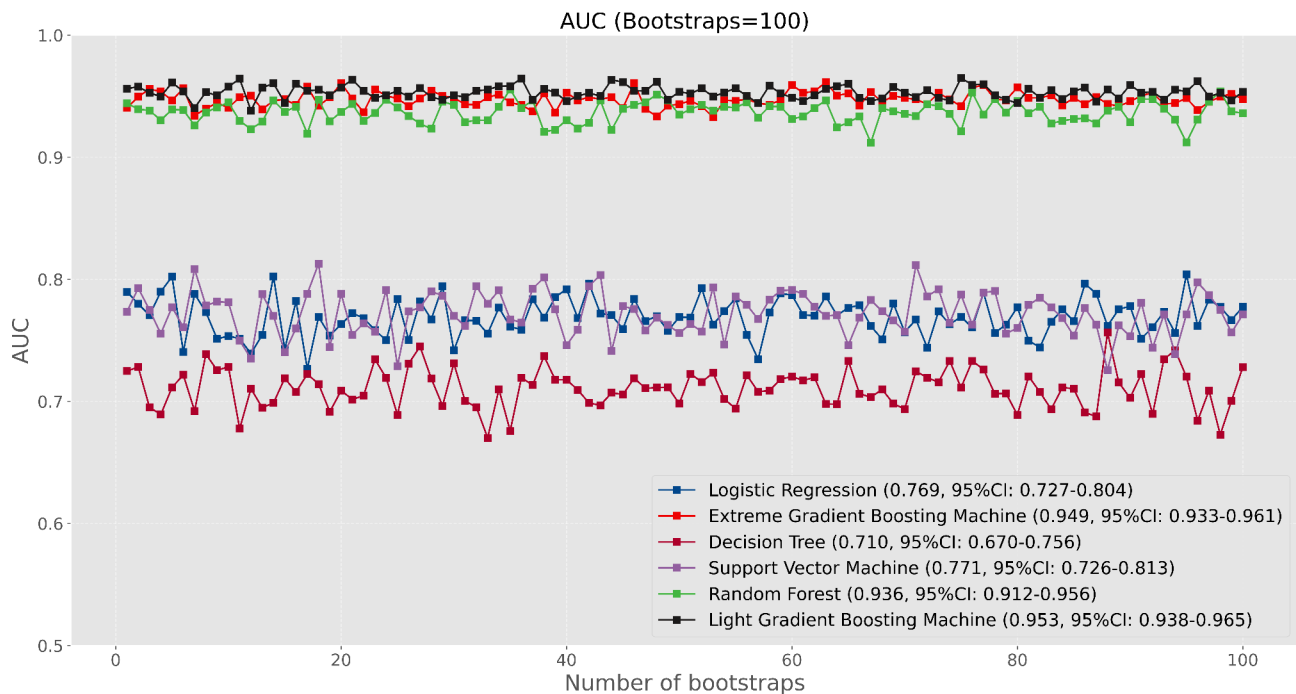


Fig. 3 The area under the curve of machine learning models

external validation of the eXGBM model also demonstrated favorable prediction performance, with an AUC value of 0.893 (95%CI: 0.819–0.967) (Supplementary Fig. 4), and it had the highest AUC value in comparison to other models. Therefore, the eXGBM model was considered optimal due to its superior performance across both internal and external validation. The feature importance analysis and the development of the AI application were both carried out using the eXGBM model.

Feature importance

The SHAP revealed that the most important five features were respiratory rate, lower limb fracture, glucose, PaO₂, and PaCO₂ in the both training (Supplementary Fig. 5A) and validation (Supplementary Fig. 5B) cohorts. The respiratory rate and SOFA score both exhibited linear relationships with their SHAP values (Supplementary Fig. 6), indicating larger value represented greater contribution to the outcome. Interestingly, the SAPS II score and its corresponding SHAP values exhibited a curve that initially rises and then declines. To elaborate, the relationship between the SAPS II score and its SHAP values was characterized by an “n” shape curve. Specifically, as the SAPS II score increased, the negative SHAP values gradually decreased and approached zero, followed by an increase in negative SHAP values as the SAPS II score continued to rise. This curve suggested that both low and high SAPS II scores could significantly impact the outcome variable.

AI platform and prediction

Based on the optimal model, we further deployed it online as an AI platform, and the code is available at <https://github.com/Starxueshu/ventilationfor7days>. In this platform (Fig. 6), the risk of prolonged dependence on mechanical ventilation for individuals can be obtained after inputting features and submitting the data. In addition, risk stratification was conducted based on the threshold of the model, and stratified interventions were recommended based on the risk stratification. For example, a case report was summarized in the following. The patient, who had a spine fracture and low limb fracture, presented with a glucose level of 139 mg/dL, PaO₂ of 149 mmHg, PaCO₂ of 36 mmHg, pH of 7.49, heart rate of 93 beats per minute, respiratory rate of 8 breaths per minute, SAPS-II of 36, and SOFA of 6, and was classified into the low-risk group based on the risk of prolonged dependence on mechanical ventilation (10.03%).

Discussion

Main findings

The main finding of the study was the successful development and validation of an AI model for predicting prolonged dependence on mechanical ventilation among patients with critical orthopaedic trauma. The study analyzed data from 1400 patients and used machine learning techniques to establish and validate predictive models. The eXGBM model demonstrated the highest performance scores, outperforming other models in various metrics such as recall, Brier score, log loss, and

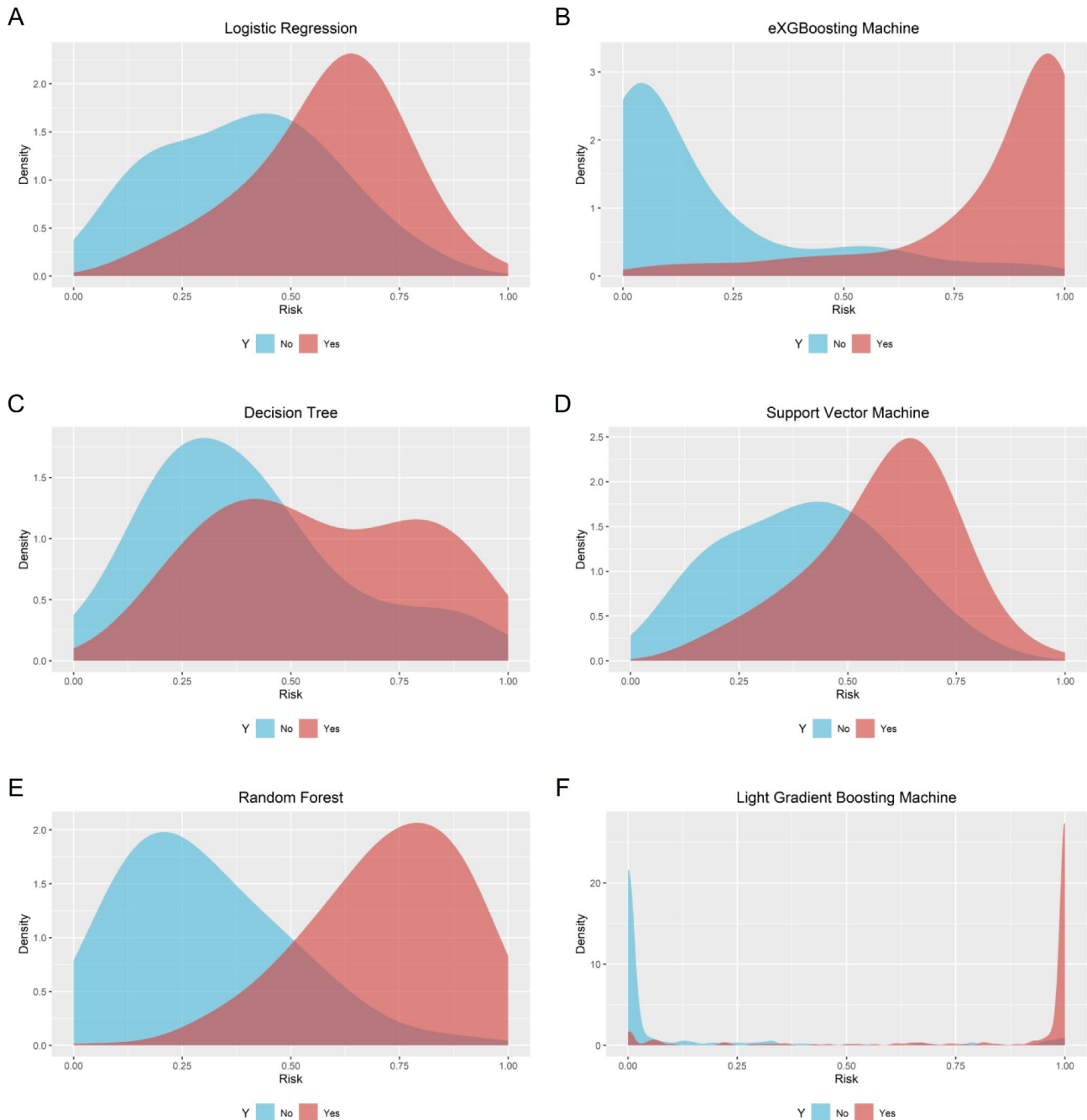


Fig. 4 Density curve of machine learning models. (A) logistic regression; (B) eXGBoosting machine; (C) decision tree; (D) support vector machine; (E) random forest; (F) light gradient boosting machine

calibration slope. In addition, external validation of the model also showed favorable prediction performance. In addition, the AI tool was designed with user-friendly features to facilitate clinical use. Healthcare professionals could input patient data and receive risk predictions, which stratified patients into high- or low-risk categories. The platform also provided personalized treatment recommendations based on individual risk profiles, helping optimize care strategies. This seamless integration

of machine learning into clinical decision-making represents a promising step toward improving patient outcomes and resource allocation in ICU settings.

Risk factors for prolonged dependence on mechanical ventilation

In the study, we found that the incidence of prolonged dependence on mechanical ventilation was 15.6%, and it was consistent with other studies. For instance, a

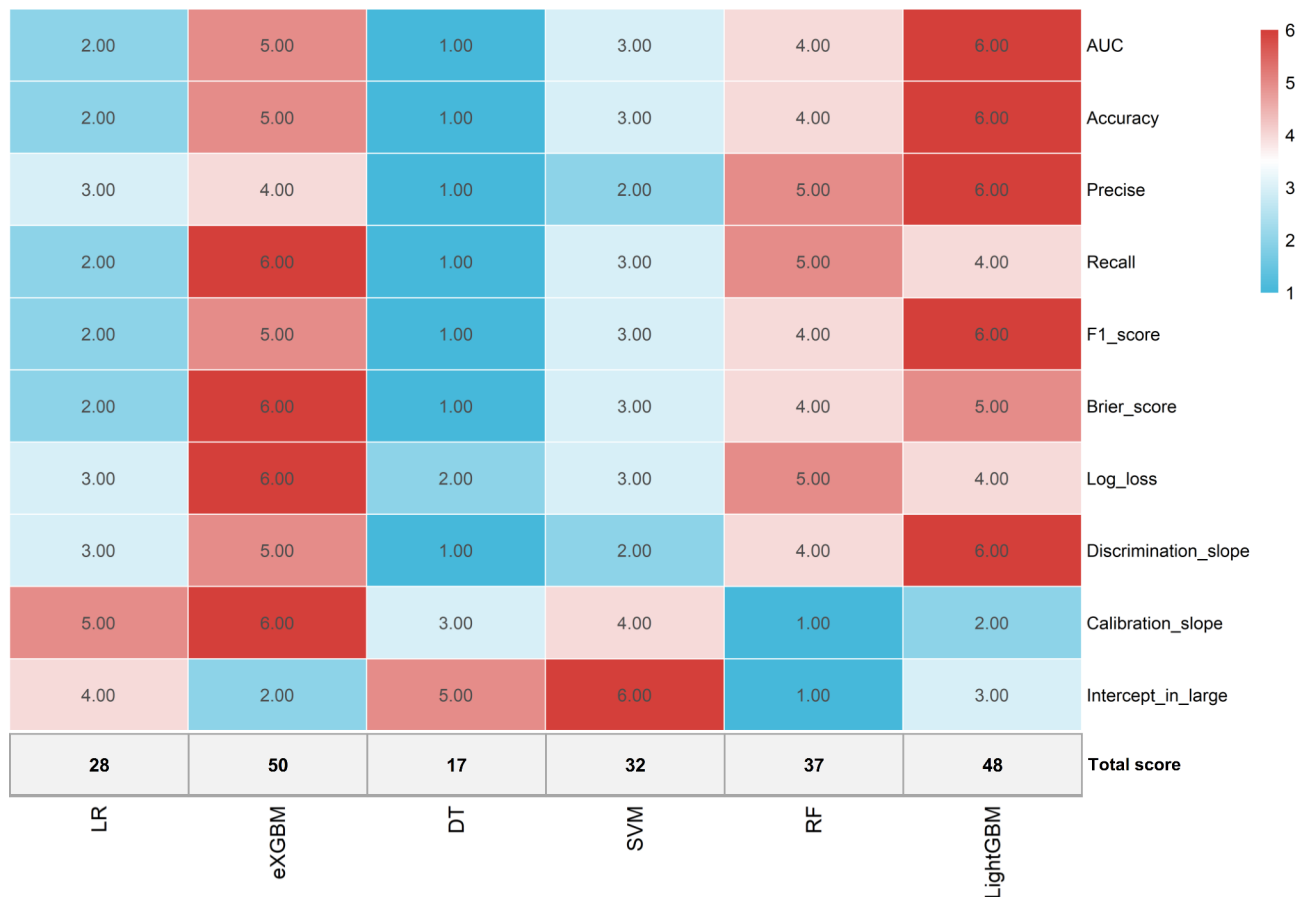


Fig. 5 Heatmap of comprehensively evaluating the prediction performance of machine learning models based on the scoring system

prospective study enrolled a total of 915 patients receiving mechanical ventilation in medical intensive care units, with 18.8% classified as experiencing prolonged weaning [8]. In addition, several risk factors have been found to be associated with prolonged dependence on mechanical ventilation. In the study conducted by Na et al. [8], the prolonged weaning group had higher proportions of patients with previous endotracheal intubation, chronic lung disease, and hematologic malignancies. Additionally, they had a longer median duration of mechanical ventilation before the first spontaneous breathing trial, a higher rate of tracheostomy, higher peak inspiratory pressures, and lower PaO₂/FiO₂ ratios compared to the non-prolonged weaning group [8]. Multivariate analyses revealed that the duration of mechanical ventilation before the first spontaneous breathing trial, tracheostomy status, poor oxygenation, and the need for renal replacement therapy were independently associated with prolonged weaning [8]. Notably, a study pointed out that there were no gender differences in weaning status and ventilator dependence among patients requiring prolonged mechanical ventilation [9]. More recently, a review summarized that systemic comorbidities, including chronic respiratory diseases (COPD, bronchiectasis,

and pulmonary fibrosis), heart failure, cerebrovascular diseases, neuromuscular diseases, end-stage renal disease, Liver cirrhosis, malignancy Infection (sepsis and multi-drug resistant infection), malnutrition, ventilator-induced diaphragm dysfunction, critical illness neuromyopathy, and critical illness encephalopathy were risk factors for prolonged dependence on mechanical ventilation [1]. In comparison to previous studies, our study found similar risk factors for prolonged dependence on mechanical ventilation, such as poor oxygenation. However, our study also identified additional risk factors such as elevated glucose levels, higher SAPS II and SOFA scores, and a higher incidence of spine fractures. These findings suggest that a combination of physiological and clinical variables may contribute to the prolonged dependence on mechanical ventilation, and further research is needed to explore the interplay between these factors and their impact on patient outcomes. Additionally, our study did not find gender differences in weaning status and ventilator dependence, which is consistent with previous research [9]. Overall, the identification of these risk factors can aid in the development of targeted interventions to improve outcomes for patients with prolonged dependence on mechanical ventilation. One of the innovations

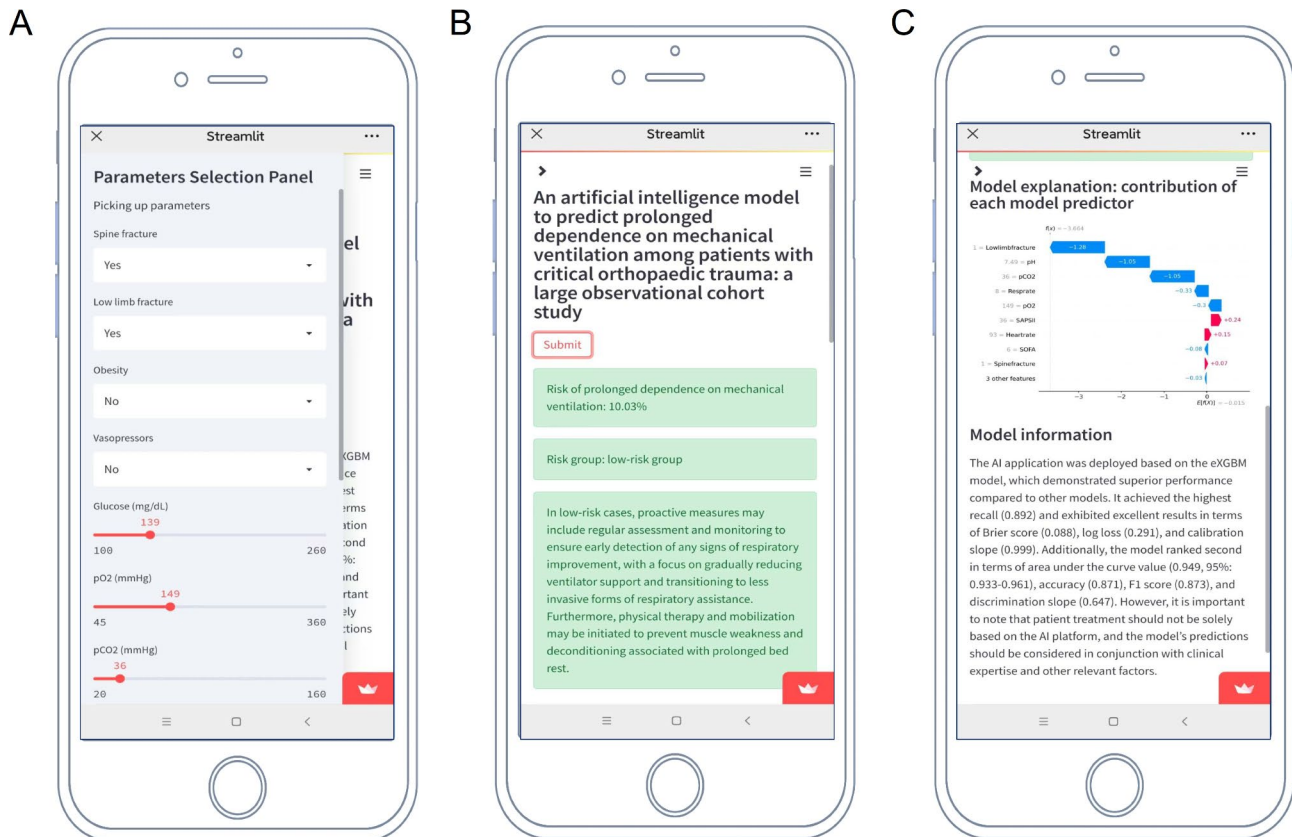


Fig. 6 AI application. (A) Feature input; (B) Risk calculation; (C) Model information

of this study was the integration of SHAP values to explain the model's predictions. This feature enhanced the interpretability of the model, allowing clinicians to understand which variables (such as respiratory rate, glucose levels, and fracture types) had the most significant impact on prolonged mechanical ventilation.

Prediction of prolonged dependence on mechanical ventilation

Several studies have used machine learning to develop models for early detecting extubation readiness [15], mechanical ventilation-associated severe acute kidney injury [14], and mechanical ventilator dependence among patients who survived sepsis/septic shock with respiratory failure [16]. To elaborate, a study developed a decision support model using convolutional neural networks to predict extubation readiness, and by analyzing historical ICU data extracted from MIMIC-III, the developed model achieved an accuracy of 86% and AUC value of 0.94 [15]. The model incorporated 25 features, among which inspired O₂ fraction, ventilator model, and peak inspiratory pressure emerged as the most significant positive factors. Conversely, the Richmond-RAS scale, plateau pressure, and spontaneous breathing traits were identified as the most important negative factors

influencing the outcome. Yan's model differed from the model in this study. Our study specifically focused on orthopedic trauma patients, while Yan's model targeted all ICU patients. Additionally, this study applied various machine learning algorithms to identify the best-performing model, whereas Yan's model only used Convolutional Neural Networks. Furthermore, the predicted outcomes differed: our study aimed to predict prolonged dependence on mechanical ventilation, while Yan's model predicted successful weaning without considering time. Therefore, the model in this study was fundamentally different from Yan's model, and thus the CNN model was not tested in this study.

In addition, a study developed and validated a clinical prediction model for the early detection of mechanical ventilation-associated severe acute kidney injury in the ICU settings [14]. The study found that the random forest algorithm outperformed the logistic regression algorithm in model development. The developed models were internally and externally validated, demonstrating good predictive performance, with the AUC value being around 0.80 in the both internal and external validation cohorts. Lastly, a study aimed to develop a scoring system to predict mechanical ventilator dependence among patients who survived sepsis/septic shock with

respiratory failure. After evaluating 251 adult patients in medical ICUs over a period of two years, the research identified risk factors for ventilator dependence and constructed a ventilator dependence risk score based on four variables. The ventilator dependence risk score demonstrated good predictive performance, with the area under the curve being 0.725 in the derivation group and 0.658 in the validation group. This suggests that the ventilator dependence risk score could be effectively applied to predict prolonged mechanical ventilation in patients who survive sepsis/septic shock [16]. To our best knowledge, in comparison to previous studies, our study was the first to develop machine learning models to predict prolonged dependence on mechanical ventilation specifically among patients with critical orthopaedic trauma. Our study found that the eXGBM model outperformed other models, demonstrating high accuracy, recall, and area under the curve value. Additionally, the model was deployed as an AI calculator, providing a user-friendly interface for healthcare professionals to input patient data, receive risk predictions, and access personalized treatment recommendations tailored to individual patient profiles. Furthermore, the AI application was deployed online, allowing clinicians to assess the risk of prolonged dependence on mechanical ventilation, thereby supporting clinical decision-making for patients requiring mechanical ventilation. The AI application allows for the classification of patients into high-risk or low-risk categories, enabling tailored interventions to be implemented. For low-risk patients, proactive measures could involve regular assessment and monitoring to facilitate early identification of respiratory improvement, with a focus on gradually reducing ventilator support and transitioning to less invasive forms of respiratory assistance. Additionally, initiating physical therapy and mobilization may help prevent muscle weakness and deconditioning associated with prolonged bed rest. Conversely, high-risk patients may necessitate more aggressive interventions, including close monitoring and frequent assessments to promptly identify and address potential complications. Advanced ventilator management strategies, such as lung-protective ventilation techniques, prone positioning, and consideration of extracorporeal membrane oxygenation in refractory cases, may be warranted. Furthermore, close collaboration between multidisciplinary teams, including critical care specialists, respiratory therapists, and physical therapists, is essential to tailor individualized management plans and provide comprehensive support for high-risk patients. This approach facilitates the optimization of patient care and resource allocation in critical care settings. However, it is important to note that the AI model developed in this study should not replace clinical judgment. Rather, it should serve as a supplementary tool to assist healthcare providers in risk stratification

and treatment planning. Clinical expertise and patient-specific factors should still be considered when making treatment decisions for individual patients.

Limitations

Limitations of this study may include the following aspects. Firstly, the cohort used for model development and validation was specific to patients with critical orthopaedic trauma. The generalizability of the AI model to patients with different types of injuries or medical conditions may require further investigation. Secondly, the study relied on retrospective data from a single observational cohort, which may have inherent limitations such as missing data, potential biases, or unmeasured confounders. Prospective data collection and verification in a multi-center setting could enhance the reliability of the model. Thirdly, while the AI model demonstrated promising prediction performance, the actual clinical impact and utility of the model in decision-making processes, patient outcomes, and resource allocation were not directly assessed. Clinical impact assessment requires extensive validation through clinical trials or real-world studies to ensure that the model's predictions lead to improved outcomes. Thus, further research on the practical implementation and impact of the AI platform in clinical settings is extremely warranted. Addressing these limitations in future research could further enhance the applicability and reliability of the AI model in predicting prolonged dependence on mechanical ventilation.

Conclusions

Our study successfully establishes and validates an AI platform for assessing the risk of prolonged dependence on mechanical ventilation among patients with critical orthopaedic trauma. The eXGBM model demonstrates superior predictive performance, outperforming other models in terms of various evaluation metrics. Notably, the model's deployment as an online AI calculator provides a user-friendly tool for individual risk assessments, enabling the stratification of patients into high-risk or low-risk groups. This approach has the potential to optimize patient care and resource allocation in critical care settings, offering a promising avenue for tailored therapeutic interventions based on individual risk assessments. The AI model represents a valuable tool for identifying patients at high risk of prolonged mechanical ventilation and may contribute to improving patient outcomes in critical care settings.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12891-024-08245-9>.

Supplementary Figure 1: Calibration curve and histogram of machine learning models. **Supplementary Figure 2:** Calibration curve of machine

learning models. (A) logistic regression; (B) eXGBoosting machine; (C) decision tree; (D) support vector machine; (E) random forest; (F) light gradient boosting machine. **Supplementary Figure 3:** Volin plot of showing discrimination slope of machine learning models. (A) logistic regression; (B) eXGBoosting machine; (C) decision tree; (D) support vector machine; (E) random forest; (F) light gradient boosting machine. **Supplementary Figure 4:** Area under the curve analysis in the external validation cohort. **Supplementary Figure 5:** Feature importance analysis based on the SHAP analysis. (A) Training cohort; (B) Validation cohort. **Supplementary Figure 6:** Association analysis between variables and SHAP values. (A) Respiratory rate in the training cohort; (B) SOFA in the training cohort; (C) SAPS II in the training cohort; (D) Respiratory rate in the validation cohort; (E) SOFA in the validation cohort; (F) SAPS II in the validation cohort

Supplementary Table 1: A brief introduction to the machine learning models in the study

Supplementary Table 2: The baseline characteristics of patients in the external validation cohort

Author contributions

Weigang Jiang, Tao Liu, Baisheng Sun, and Mingxing Lei participated in designing and writing the manuscript. Zhenkan Han, Baisheng Sun, Mingxing Lei, and Lixia Zhong were involved in data extraction, collection, and analysis. Minhua Lu and Mingxing Lei supervised the study. All authors read and approved the final manuscript. All authors took part in designing and writing the manuscript.

Funding

This study was supported by Hainan Province Clinical Medical Center and National Clinical Research Center for Orthopedics, Sports Medicine & Rehabilitation.

Data availability

The datasets of the current study are available under reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate

The use of the MIMIC-III database was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center (32128436). Since the data in the MIMIC-III database has been de-identified, patient consent was not required. External validation of the model was performed in 122 patients with critical orthopaedic trauma from a university teaching hospital, and the Ethics Committee of our hospital approved the study. We commit to strictly adhere to ethical guidelines and legal regulations at all stages of the research and ensure the appropriate use and protection of patient privacy.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Orthopedics, The Fourth Affiliated Hospital of Soochow University (Suzhou Dushu Lake Hospital), Suzhou City 215000, Jiang Su Province, People's Republic of China

²Department of Orthopedics, The 9th Medical Centre of Chinese PLA General Hospital, Beijing, China

³Department of Critical Care Medicine, The First Medical Centre of Chinese PLA General Hospital, Beijing, China

⁴Chinese PLA Medical School, Beijing, China

⁵Department of Intensive Care Medicine, Beijing Friendship Hospital, Capital Medical University, Beijing, China

⁶Department of Orthopedics, Peking University Third Hospital, Beijing, China

⁷Department of Orthopedics, Hainan Hospital of PLA General Hospital, Hainan, China

⁸Department of Orthopedics, National Clinical Research Center for Orthopedics, Sports Medicine & Rehabilitation, PLA General Hospital, 51 Fucheng Road, Haidian District, Beijing 100142, People's Republic of China

Received: 15 February 2024 / Accepted: 23 December 2024

Published online: 30 December 2024

References

- Huang HY, Huang CY, Li LF. Prolonged mechanical ventilation: outcomes and management. *J Clin Med* 2022, 11(9).
- Papazian L, Klompas M, Luyt CE. Ventilator-associated pneumonia in adults: a narrative review. *Intensive Care Med*. 2020;46(5):888–906.
- Goligher EC, Dres M, Patel BK, Sahetya SK, Beitler JR, Telias I, Yoshida T, Vaporidi K, Grieco DL, Schepens T, et al. Lung- and diaphragm-protective ventilation. *Am J Respir Crit Care Med*. 2020;202(7):950–61.
- Ratti LSR, Tonella RM, Figueiredo LCd, Saad IAB, Falcão ALE. Oliveira PPMd: Inspiratory Muscle Training Strategies in Tracheostomized critically ill individuals. *Respir Care*. 2022;67(8):939–48.
- Loss SH, de Oliveira RP, Maccari JG, Savi A, Boniatti MM, Hetzel MP, Dallegre DM, Balzano Pde C, Oliveira ES, Höher JA, et al. The reality of patients requiring prolonged mechanical ventilation: a multicenter study. *Rev Bras Ter Intensiva*. 2015;27(1):26–35.
- Huang C. The survival outcomes of patients requiring prolonged mechanical ventilation. *Med (Kaunas Lithuania)* 2023, 59(3).
- Damuth E, Mitchell J, Bartock J, Roberts B, Trzeciak S. Long-term survival of critically ill patients treated with prolonged mechanical ventilation: a systematic review and meta-analysis. *Lancet Respir Med*. 2015;3(7):544–53.
- Na SJ, Ko RE, Nam J, Ko MG, Jeon K. Factors associated with prolonged weaning from mechanical ventilation in medical patients. *Ther Adv Respir Dis*. 2022;16:17534666221117005.
- Huang C. Gender differences in prolonged mechanical Ventilation patients - A Retrospective Observational Study. *Int J Gen Med*. 2022;15:5615–26.
- Huang C. How prolonged mechanical ventilation is a neglected disease in chest medicine: a study of prolonged mechanical ventilation based on 6 years of experience in Taiwan. *Ther Adv Respir Dis*. 2019;13:1753466619878552.
- Yoon JH, Pinsky MR, Clermont G. Artificial Intelligence in critical Care Medicine. *Crit Care*. 2022;26(1):75.
- Gutierrez G. Artificial Intelligence in the Intensive Care Unit. *Crit Care*. 2020;24(1):101.
- Tang R, Zhang S, Ding C, Zhu M, Gao Y. Artificial Intelligence in Intensive Care Medicine: bibliometric analysis. *J Med Internet Res*. 2022;24(11):e42185.
- Huang S, Teng Y, Du J, Zhou X, Duan F, Feng C. Internal and external validation of machine learning-assisted prediction models for mechanical ventilation-associated severe acute kidney injury. *Aust Crit Care*. 2023;36(4):604–12.
- Jia Y, Kaul C, Lawton T, Murray-Smith R, Habli I. Prediction of weaning from mechanical ventilation using Convolutional neural networks. *Artif Intell Med*. 2021;117:102087.
- Chang YC, Huang KT, Chen YM, Wang CC, Wang YH, Tseng CC, Lin MC, Fang WF. Ventilator dependence risk score for the prediction of prolonged mechanical ventilation in patients who survive Sepsis/Septic shock with respiratory failure. *Sci Rep*. 2018;8(1):5650.
- Dilday J, Lewis MR. Transfusion management in the trauma patient. *Curr Opin Crit Care*. 2022;28(6):725–31.
- Tisherman SA, Stein DM. ICU management of Trauma patients. *Crit Care Med*. 2018;46(12):1991–7.
- Kung SC, Lin WT, Tsai TC, Lin MH, Chang CH, Lai CC, Chao CM. Epidemiologic characteristics and outcomes of major trauma patients requiring prolonged mechanical ventilation. *Med (Baltim)*. 2017;96(52):e9487.
- Zhang Z, Ho KM, Hong Y. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care. *Crit Care*. 2019;23(1):112.
- Yue S, Li S, Huang X, Liu J, Hou X, Zhao Y, Niu D, Wang Y, Tan W, Wu J. Machine learning for the prediction of acute kidney injury in patients with sepsis. *J Transl Med*. 2022;20(1):215.
- Li F, Xin H, Zhang J, Fu M, Zhou J, Lian Z. Prediction model of in-hospital mortality in intensive care unit patients with heart failure: machine learning-based, retrospective analysis of the MIMIC-III database. *BMJ Open*. 2021;11(7):e044779.

23. Lei M, Han Z, Wang S, Han T, Fang S, Lin F, Huang T. A machine learning-based prediction model for in-hospital mortality among critically ill patients with hip fracture: an internal and external validated study. *Injury*. 2023;54(2):636–44.
24. Lei M, Han Z, Wang S, Guo C, Zhang X, Song Y, Lin F, Huang T. Biological signatures and prediction of an immunosuppressive status-persistent critical illness-among orthopedic trauma patients using machine learning techniques. *Front Immunol*. 2022;13:979877.
25. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3:160035.
26. Han T, Xiong F, Sun B, Zhong L, Han Z, Lei M. Development and validation of an artificial intelligence mobile application for predicting 30-day mortality in critically ill patients with orthopaedic trauma. *Int J Med Inf*. 2024;184:105383.
27. Cui Y, Shi X, Wang S, Qin Y, Wang B, Che X, Lei M. Machine learning approaches for prediction of early death among lung cancer patients with bone metastases using routine clinical characteristics: an analysis of 19,887 patients. *Front Public Health*. 2022;10:1019168.
28. Shi X, Cui Y, Wang S, Pan Y, Wang B, Lei M. Development and validation of a web-based AI prediction model to assess intraoperative massive blood loss among metastatic spinal disease using machine learning techniques. *The Spine J* 2023;In Press.
29. Xiong F, Cao X, Shi X, Long Z, Liu Y, Lei M. A machine learning-based model to predict early death among bone metastatic breast cancer patients: a large cohort of 16,189 patients. *Front Cell Dev Biol*. 2022;10:1059597.
30. Shi X, Cui Y, Wang S, Pan Y, Wang B, Lei M. Development and validation of a web-based artificial intelligence prediction model to assess massive intraoperative blood loss for metastatic spinal disease using machine learning techniques. *Spine J* 2023.

Publisher's note

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