Received: 24 August 2023 | Revised: 21 April 2024 | Accepted: 30 April 2024

DOI: 10.1002/emp2.13190

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE
Open Access
WILEY

ORIGINAL ARTICLE

Emergency Medical Services

A risk prediction model for efficient intubation in the emergency department: A 4-year single-center retrospective analysis

Hongbo Ding MM[1,2,3] | **Xue Feng PhD**[4] | **Qi Yang MM**[1,2,3] | **Yichang Yang MM**[5] | **Siyi Zhu PhD**[4] | **Xiaozhen Ji MM**[6] | **Yangbo Kang MM**[1,2,3] | **Jiashen Shen MM**[1,2,3] | **Mei Zhao PhD**[7] | **Shanxiang Xu MD**[1,2,3] | **Gangmin Ning PhD**[4] | **Yongan Xu PhD, MD**[1,2,3]

[1]Department of Emergency Medicine, Second Affiliated Hospital & Institute of Emergency Medicine, Zhejiang University School of Medicine, Hangzhou, China

[2]Key Laboratory of the Diagnosis and Treatment of Severe Trauma and Burn of Zhejiang Province, Hangzhou, China

[3]Zhejiang Province Clinical Research Center for Emergency and Critical Care Medicine, Hangzhou, China

[4]Department of Biomedical Engineering, Zhejiang University, Hangzhou, China

[5]Department of Traditional Chinese Medicine, Zhejiang University School of Medicine, Hangzhou, China

[6]Department of Emergency Medicine, Longquan People's Hospital, Longquan, China

[7]Guizhou University of Traditional Chinese Medicine, Guiyang, China

Correspondence
Gangmin Ning, PhD, Department of Biomedical Engineering, Zhejiang University, No. 886, Yuhangtang Road, Xihu District, Hangzhou 310027, China. Email: gmning@zju.edu.cn

Yongan Xu, PhD & MD, Department of Emergency Medicine, Second Affiliated Hospital & Institute of Emergency Medicine, Zhejiang University School of Medicine, No. 88, Jiefang Road, Shangcheng District, Hangzhou 310009, China. Email: xuyongan2000@zju.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: grants 81571916, 81372079, 81201478; Key Research and Development (R&D) Program of Zhejiang Province, Grant/Award Numbers: grants 2022C03G2013326, 2024C3186; Zhejiang Province Public Welfare Foundation Project, Grant/Award Number: grant LGF22H150009; Major Project of National-Zhejiang Provincial

Abstract

Objective: To analyze the risk factors associated with intubated critically ill patients in the emergency department (ED) and develop a prediction model by machine learning algorithms.

Methods: This study was conducted in an academic tertiary hospital in Hangzhou, China. Critically ill patients admitted to the ED were retrospectively analyzed from May 2018 to July 2022. The demographic characteristics, distribution of organ dysfunction, parameters for different organs' examination, and status of mechanical ventilation were recorded. These patients were assigned to the intubation and non-intubation groups according to ventilation support. We used the eXtreme Gradient Boosting (XGBoost) algorithm to develop the prediction model and compared it with other algorithms, such as logistic regression, artificial neural network, and random forest. SHapley Additive exPlanations was used to analyze the risk factors of intubated critically ill patients in the ED.

Results: Of 14,589 critically ill patients, 10,212 comprised the training group and 4377 comprised the test group; 2289 intubated patients were obtained from the electronic

Hongbo Ding, Xue Feng, and Qi Yang contributed equally to this work.

Supervising Editor: Nicholas Johnson, MD

Administration of Traditional Chinese Medicine, Grant/Award Number: grant GZY-ZJ-KJ-24030

medical records. The mean age, mean scores of vital signs, parameters of different organs, and blood oxygen examination results differed significantly between the two groups ($p < 0.05$). The white blood cell count, international normalized ratio, respiratory rate, and pH are the top four risk factors for intubation in critically ill patients. Based on the risk factors in different predictive models, the XGBoost model showed the highest area under the receiver operating characteristic curve (0.84) for predicting ED intubation.

Conclusions: For critically ill patients in the ED, the proposed model can predict potential intubation based on the risk factors in the clinically predictive model.

KEYWORDS
emergency department, machine learning, mechanical ventilation, prediction model, tracheal intubation

1 | INTRODUCTION

1.1 | Background and importance

Invasive mechanical ventilation (IMV) is the most effective therapeutic approach for critically ill patients with acute respiratory distress syndrome or other types of respiratory failure.[1] Mechanical ventilation provides respiratory support and allows adequate time for the recuperation of impaired organs.[2] Annually, more than 240,000 patients require mechanical ventilation in the emergency department (ED), which accounts for 8%–12% of all ventilated patients in the USA.[3,4] Typically, ED serves as the initial therapeutic setting for critically ill patients with acute respiratory failure, severe hemorrhagic shock, septic shock, or multiple organ failure. Therefore, it is imperative to identify the potential need for timely tracheal intubation and mechanical ventilation support in the ED.[5,6]

IMV is indicated for critically ill patients with hypoxemia or hypercapnia, circulatory failure, and unconsciousness requiring airway protection to prevent aspiration.[7] However, stratifying patients for further respiratory support poses a significant challenge. First, some critically ill patients may present with atypical clinical manifestations that could result in delayed intubation.[8] Second, there are differences in the proficiency and implementation of mechanical ventilation among young or inexperienced clinicians.[9] These challenges collectively result in delayed intubation, which can lead to significant risks such as increased in-hospital mortality,[10] prolonged ventilation, and extended hospital stays.[11] Additionally, identifying critically ill patients who are unlikely to derive benefit from non-invasive mechanical ventilation (NMV) is crucial.[12,13] Therefore, developing an intubation prediction tool to guide clinical practice is an urgent requirement.[14]

Electronic medical records offer exponential growth in data, while machine learning (ML) enables the processing of vast amounts of information. By integrating the expertise of professional physicians with advanced algorithms, the accuracy (ACC) of man–machine diagnosis can be enhanced. ML is increasingly utilized in the development of disease prediction models,[15–18] which can offer patients precise, personalized, and timely medical services, thereby improving treatment success rates and mitigating patient suffering and wastage of medical resources.

1.2 | Goals of this investigation

The application of ML in the field of emergency medicine is not well-elucidated, especially for ventilation prediction models. The present study applied eXtreme Gradient Boosting (XGBoost) and developed an intubation risk prediction model for critically ill patients in ED based on the available indicators.[19,20] This model classifies critically ill patients into different risk levels while reducing the negative impact of the subjective cognition of clinicians on clinical decision making regarding timely tracheal intubation. Therefore, this retrospective study analyzed the clinical data from critically ill patients in the ED of a tertiary teaching hospital and developed a multifactorial prediction model to guide the decision making.

2 | METHODS

2.1 | Data source and extraction

The study was conducted in the Second Affiliated Hospital Zhejiang University School of Medicine (SAHZU), Hangzhou, China. The Institutional Review Board approved the present study and waived the requirement of consent for patient data due to the retrospective nature of the study. Endotracheal intubation was defined as the endotracheal tube inserted into the trachea through the glottis.[5] The intubated patients were supported by a mechanical ventilator until they were admitted into the intensive care unit (ICU) for further

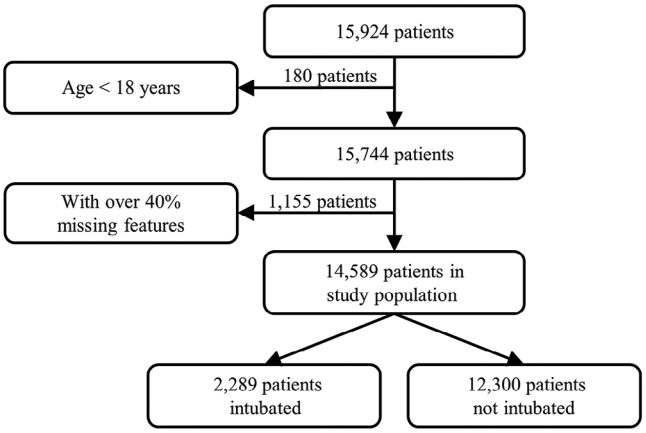

FIGURE 1 Flowchart of the patient screening process.

therapy. Patients' information was extracted from the electronic health records of SAHZU from May 2018 to July 2022.

The patient screening process is illustrated in Figure 1. Patients who did not meet the following screening criteria were excluded: those with no clear intubation status, younger than 18 years to eliminate age interference, and patients with a missing data rate of >40% to ensure the stability of the model.

The model framework is illustrated in Figure 2. The data source, patient cohort screening, and feature selection process are introduced, followed by the data preprocessing process and statistical analysis. Then, the structure and implementation of the model are described. Finally, the model evaluation metrics are stated.

Subsequently, the mean value of all other patients was used to fill the gaps in order to reduce the sample loss caused by missing data. The normalization process is to standardize all data into values between 0 and 1 to avoid the impact of inconsistent data dimensionality and value range.

2.2 | Model development

The model consists of three parts: an unsupervised cluster model, a group of supervised base models, and attention layers. First, the unsupervised model was used to classify patients into different subgroups. Then, different base models were used to predict intubation risk in different subgroups of patients. The specific construction process of the model is described in detail in the Supporting Information S1.

2.3 | Evaluation

2.3.1 | Control models setting

In order to compare the performance of the models, several commonly used ML models were chosen as control models, such as logistic regression, artificial neural network, and random forest algorithm. All the control models were hyperparametric tuned.

The Bottom Line

This retrospective analysis of 15 years of adult out of hospital cardiac arrest (OHCA) data from Salt Lake City found that when compared to women, men have a higher incidence of OHCA, higher rates of characteristics associated with improved survival, and higher unadjusted survival. However, adjusted analysis showed no difference in survival between men and women.

In addition, data without clustering partitioning were fed directly into XGBoost to verify the effect of model structure on performance.

2.3.2 | Model evaluation metrics

The performance of the established models was evaluated by area under the receiver operating characteristic curve (AUROC), area under the precision recall curve, ACC, sensitivity (SEN), and specificity (SPE). The greater the AUROC value, the better the classification performance.

2.3.3 | Statistical analysis of data

The data were analyzed using statistical software (SPSS, version 17.0). The differences between the intubation and non-intubation groups were analyzed by Student's *t*-test or Mann–Whitney *U*-test, as shown in Table 1. $p < 0.05$ indicates a significant difference. Analysis of variance test or Kruskal–Wallis *H*-test was used to test the differences among the three subgroups. Median and quartile differences described the distribution of continuous variables.

3 | RESULTS

3.1 | General information

A significant difference analysis of all data was conducted between the intubation and non-intubation groups, as shown in Table 1. The results showed that most of the features had significant differences.

According to literature and physician recommendations,[21–23] the characteristics, including demographics, vital signs, and laboratory tests used to construct the model, are summarized in Table 1. Considering the impact of data missing on model performance, characteristics with a missing rate >40% were excluded. Finally, a study cohort of 14,589 patients with 33 features was determined, with a intubation:non-intubation ratio of 2289:12,300. The 33 features are summarized in the Supplementary Material S2. The intubation decisions were made by experienced senior clinicians in SAHZU and practiced strictly according to the intubation indication. The data were

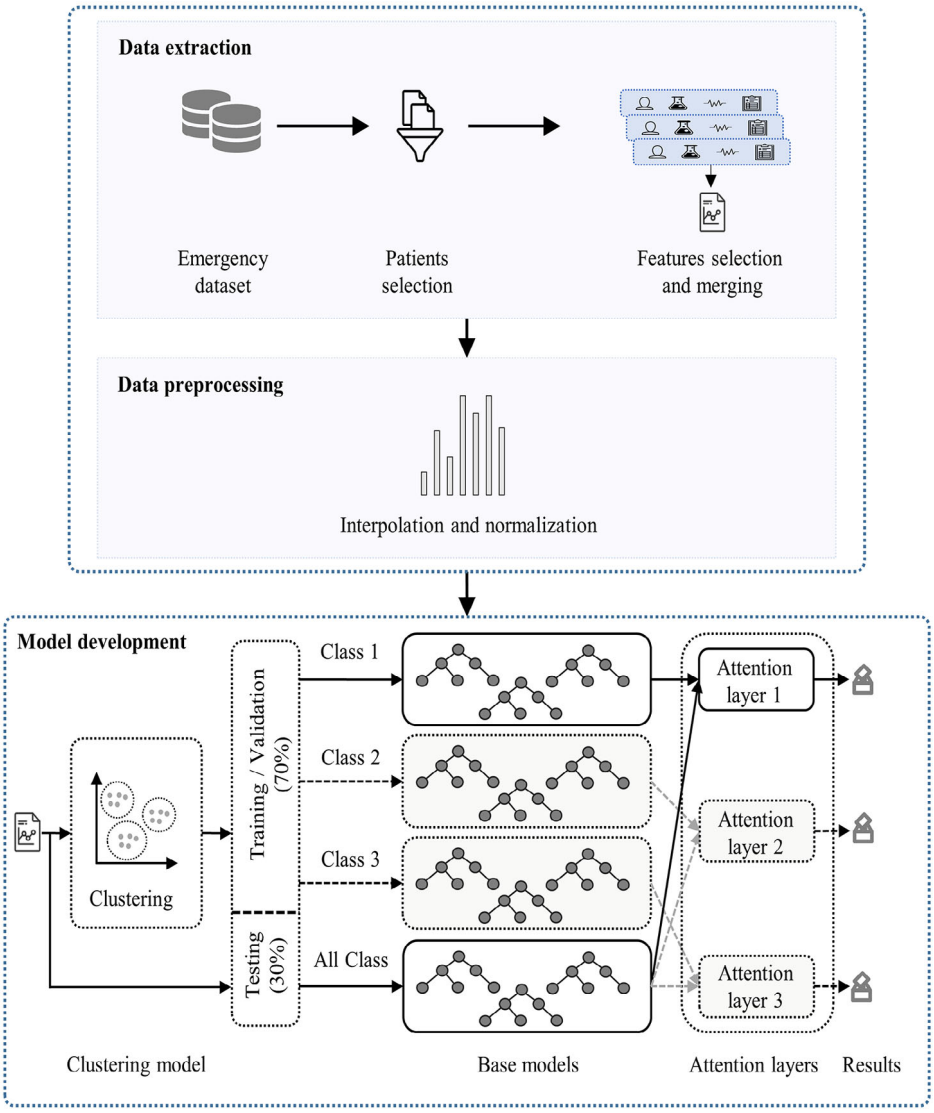

FIGURE 2 Flowchart of the model: data extraction, data preprocessing, and model development.

collected from the first test results of patients after admission to the ED. The intubated patients were intubated before hospitalization or during ED retention.

3.2 | Model implementation

After the above model framework was established, the clinical dataset was randomly divided according to the training:testing ratio of 7:3. Ten-fold cross-validation was executed in training subset to find the optimal model parameters. The experiments were repeated until most suitable set of model parameters were achieved. For the cluster model, the cluster number *k* was determined by the physicians' experience and set as 3. SHapley Additive exPlanations (SHAP) is a game theoretic approach that calculates the importance of features for any ML model. The results show the contribution of different features to the prediction model (Figure 3). According to SHAP analysis, white blood cell (WBC) count, international normalized ratio (INR), respiratory rate (RR), and pH are identified as the top four influential factors in predicting intubation. Five features related to infection and three other features crucial for intubation prediction were selected for the cluster features: WBC, neutrophil absolute value, neutrophil percentage, lymphocyte absolute value, lymphocyte percentage, INR, RR, and pH. The patients were divided into three subgroups.

3.3 | Comparison of clustering groups

After clustering, patients were divided into 1–3 groups. principal component analysis (PCA) dimension reduction was used to visually display the distribution of patients according to cluster labels, as shown in Figure 4. The characteristics of patients in the intubation and non-intubation groups with different subgroups are shown in the Supporting Information S3. *t*-Tests were performed between the intubation and

TABLE 1 Study cohort characteristics.

Characteristics	Intubation group ($n = 2289$)	Non-intubation group ($n = 12,300$)	p-Value
Demographics			
Male, *n* (%)	1547 (67.6)	7932 (64.5)	0.004*
Age (years), mean (SD)	65.2 (16.5)	61.6 (25.1)	0.001*
Weight (kg), mean (SD)	62.9 (12.8)	63.8 (14.0)	0.152
Chief complaints, *n* (%)			
Injuries	501 (21.9)	2528 (20.6)	0.148
Digestive system diseases	111 (4.9)	1439 (11.7)	0.000*
Respiratory diseases	435 (19.0)	1563 (12.7)	0.000*
Neurological diseases	719 (31.4)	3303 (26.9)	0.000*
Cardiovascular diseases	314 (13.7)	1506 (12.2)	0.050
Intoxication	35 (1.5)	238 (1.9)	0.188
Comorbidities, *n* (%)			
Hypertension	544 (23.8)	3073 (25.0)	0.215
Diabetes	261 (11.4)	1445 (11.7)	0.637
Cancer	158 (6.9)	1084 (8.6)	0.255
Vital signs, mean (SD)			
Respiratory rate (breaths/min)	17.9 (6.6)	15.7 (4.0)	0.000*
Heart rate (breaths/min)	96.7 (26.2)	90.2 (22.7)	0.000*
Systolic blood pressure (mmHg)	140.2 (38.3)	136.1 (30.7)	0.000*
Diastolic blood pressure (mmHg)	79.6 (22.5)	78.1 (17.4)	0.000*
Temperature (°C)	36.9 (1.8)	36.9 (1.0)	0.000*
Oxygen saturation (%)	98.1 (16.3)	97.8 (7.5)	0.000*
Laboratory tests, mean (SD)			
pH	7.52 (1.19)	7.44 (0.53)	0.000*
Base excess (mmol/L)	−4.12 (7.59)	−1.18 (5.15)	0.000*
Total carbon dioxide (mmol/L)	24.0 (9.9)	25.7 (9.0)	0.000*
Blood oxygen (vol%)	17.4 (4.7)	16.9 (3.9)	0.000*
Partial pressure of oxygen (mmHg)	133.2 (79.9)	115.8 (54.4)	0.000*
Partial pressure of carbon dioxide (mmHg)	43.2 (20.3)	37.6 (11.0)	0.000*
Blood urea nitrogen (mmHg)	9.40 (8.21)	7.83 (6.29)	0.000*
White blood cell (10^9/L)	14.0 (13.9)	10.4 (7.3)	0.000*
Neutrophil absolute value (10^9/L)	11.1 (6.5)	8.3 (5.3)	0.000*
Neutrophil percentage (%)	80.4 (14.5)	77.6 (13.0)	0.025*
Lymphocyte absolute value (10^9/L)	1.67 (2.31)	1.40 (3.64)	0.000*
Lymphocyte percentage (%)	13.6 (12.8)	15.8 (11.2)	0.001*
Platelet (10^9/L)	201 (105)	207 (91)	0.000*
Na^+ (mmol/L)	217.3 (76.4)	212.4 (73.6)	0.019*
Ca^+ (mmol/L)	2.29 (0.54)	2.34 (0.49)	0.410
K^+ (mmol/L)	6.17 (2.86)	5.98 (2.47)	0.000*
Cl^- (mmol/L)	163.6 (59.9)	160.4 (57.9)	0.238
HCO_3^- (mmol/L)	21.3 (7.2)	22.6 (4.9)	0.000*
Hemoglobin (g/L)	125 (31)	127 (29)	0.000*
Creatine (μmol/L)	123.1 (150.2)	118.2 (442.9)	0.770
Lactate (mmol/L)	3.98 (4.11)	2.13 (2.08)	0.000*

(Continues)

TABLE 1 (Continued)

Characteristics	Intubation group ($n = 2289$)	Non-intubation group ($n = 12,300$)	p-Value
Total bilirubin (μmol/L)	21.3 (31.5)	19.7 (29.6)	0.022*
Prothrombin time (s)	16.0 (7.3)	14.0 (3.6)	0.000*
Prothrombin time activity (%)	81.6 (24.6)	93.7 (21.0)	0.000*
International normalized ratio	1.32 (0.96)	1.11 (0.46)	0.000*

Abbreviation: SD, standard deviation.
*$p < 0.05$ indicates a significant difference.

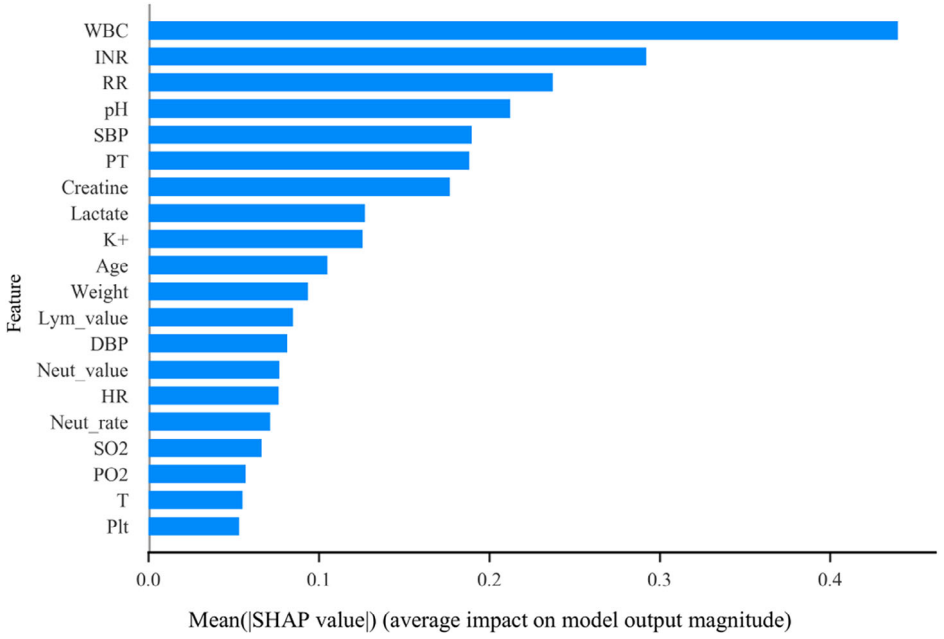

FIGURE 3 SHapley Additive exPlanations (SHAP) analysis of eXtreme Gradient Boosting (XGBoost) model. Feature importance ranking of the model. The *x*-axis is the mean SHAP value and the *y*-axis is the feature.

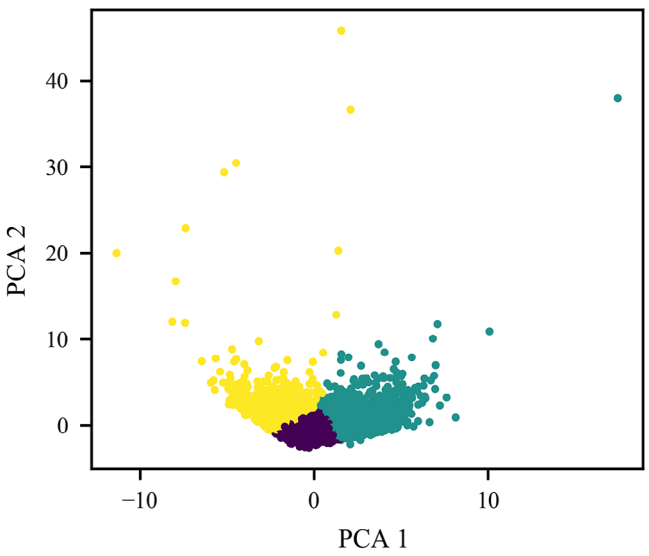

FIGURE 4 Visual distribution of patient cluster. The *x*-axis is the PCA 1 and the *y*-axis is the PCA 2. Dots represent sample points, and different colors represent different clusters.

non-intubation groups. Table 1 summarizes the demographic characteristics; those used for clustering differ significantly among the three groups. The intubation rate was increased from groups 1 to 3. As shown in Figure 5, both RR and INR showed an increasing trend between the intubation and the non-intubation groups. The pH among the three groups did not show an increasing or decreasing trend, but the average pH of groups 1 and 2 was within the normal range, and the pH of group 3 was the lowest (7.34), suggesting that the risk of decompensated acidosis was maximal. For infection-related indicators, WBC, neutrophil absolute value, and neutrophil percentage showed an increasing trend, while the percentage of lymphocytes showed a decreasing trend.

The significant difference test between the three groups is shown in Table 2.

3.4 | Performance of the model

Table 3 shows the performance comparison of all base models for different subgroups of patients. XGBoost has the best model performance

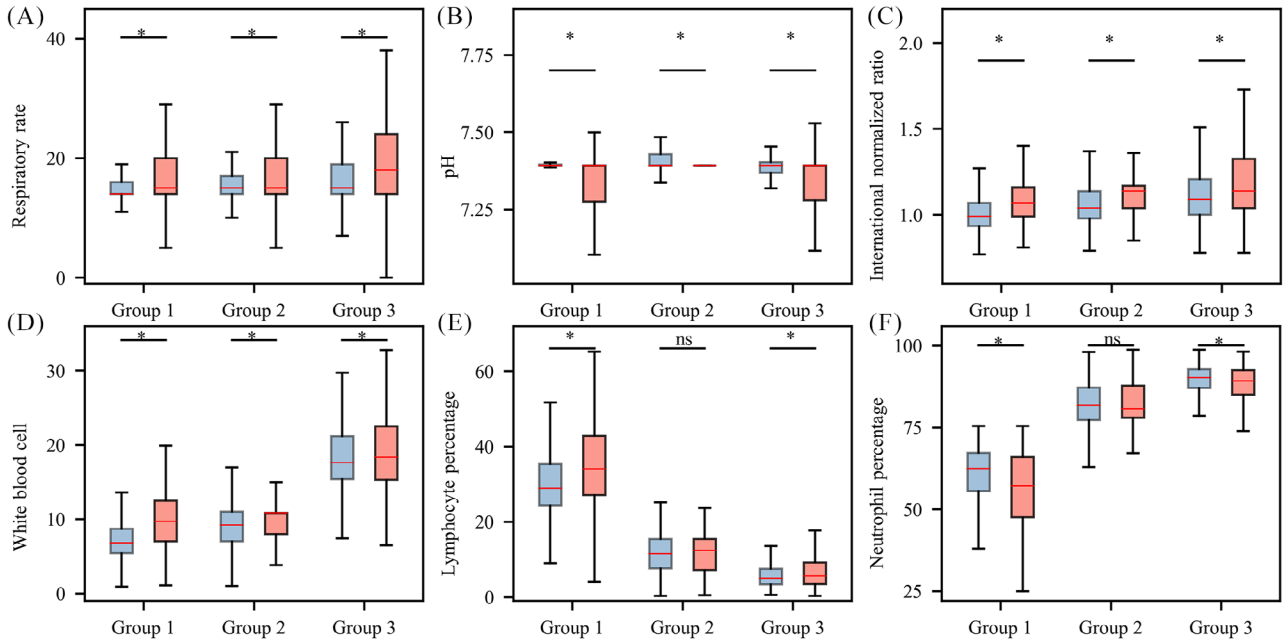

FIGURE 5 Distribution of clustering features. (A) Respiratory rate, (B) pH, (C) international normalized ratio, (D) white blood cell, (E) lymphocyte percentage, and (F) neutrophil percentage. The *x*-axis represents different clustering groups, the *y*-axis represents different characteristic values. The blue represents the intubation group and the orange represents the non-intubation group. The asterisk (*) means that there is a significant difference between the intubation group and non-intubation group, ns means that there is no significant difference.

TABLE 2 Characteristics of different subgroups.

Characteristics	Group 1	Group 2	Group 3	*p*-Value
Age (years), mean (SD)	58.6 (18.7)	64.0 (27.3)	60.8 (18.1)	0.000*
Weight (kg), mean (SD)	64.4 (14.7)	63.4 (13.5)	63.6 (13.4)	0.000*
Respiratory rate (breaths/min), mean (SD)	15.5 (3.6)	15.7 (4.3)	17.7 (6.2)	0.000*
pH, mean (SD)	7.38 (0.10)	7.42 (0.06)	7.34 (0.13)	0.000*
White blood cell (10^9/L), mean (SD)	7.8 (4.1)	9.0 (3.0)	19.8 (13.3)	0.000*
Neutrophil absolute value (10^9/L), mean (SD)	4.7 (2.1)	7.6 (2.7)	17.2 (5.9)	0.000*
Neutrophil percentage (%), mean (SD)	59.9 (9.4)	82.5 (6.8)	89.0 (6.3)	0.000*
Lymphocyte absolute value (10^9/L), mean (SD)	2.5 (1.9)	1.0 (0.5)	1.3 (1.7)	0.000*
Lymphocyte percentage (%), mean (SD)	31.4 (9.0)	11.3 (5.1)	6.2 (4.4)	0.000*
International normalized ratio, mean (SD)	1.06 (0.33)	1.12 (0.35)	1.30 (1.05)	0.000*
Intubation percentage	10.2%	14.0%	27.9%	0.000*
Number of patients	3537	8327	2725	–

Abbreviation: SD, standard deviation.
*$p < 0.05$ indicates a significant difference.

in all patient groups, with an AUROC of 0.8706, 0.8396, 0.7532, and 0.8353. Among all XGBoost-based models, model 1 has the highest AUROC (0.8706), while model 3 has the lowest value (0.7532). For the model trained and tested with all the data, the AUROC was lower than that of models 1 and 2 but higher than that of model 3, with an AUROC of 0.8353.

Table 4 shows the model performance after adding the attention mechanism. Compared to the base models without an increased attention mechanism, AUROC increased by 0.0025, 0.0049, and 0.0076, respectively. Among the three patient groups, the AUROC for group 1 had the highest value of 0.8731, while that of group 3 was the lowest at 0.7608. Also, SEN and SPE reached their maximum values in group 1, 0.8347 and 0.7849, respectively.

3.5 | Feature importance analysis

The SHAP of different subgroups is shown in Figure 6. INR, pH, and creatine were the top three features in terms of feature importance for patients in group 1. For patients in group 2, the top three features

TABLE 3 Performance comparison of all base models.

Group	Base model	AUROC	AUPRC	SEN	SPE	ACC
1	LR	0.7501	0.5087	0.6514	0.8489	0.8286
	ANN	0.8617	0.4566	0.8165	0.7523	0.7589
	RF	0.8486	0.4387	0.8165	0.7450	0.7524
	XGBoost	0.8706	0.4799	0.8073	0.7856	0.7881
2	LR	0.7085	0.5111	0.6809	0.7360	0.7283
	ANN	0.7763	0.3680	0.7208	0.7193	0.7194
	RF	0.8276	0.5445	0.7379	0.7337	0.7343
	XGBoost	0.8396	0.6009	0.7721	0.7514	0.7543
3	LR	0.6603	0.5788	0.5614	0.7593	0.7041
	ANN	0.7124	0.4697	0.6667	0.6661	0.6663
	RF	0.7358	0.5014	0.7412	0.6424	0.6699
	XGBoost	0.7532	0.5277	0.7149	0.6915	0.6980
All	LR	0.7148	0.5303	0.6556	0.7740	0.7554
	ANN	0.7943	0.4325	0.7340	0.7204	0.7225
	RF	0.8264	0.5221	0.7653	0.7228	0.7293
	XGBoost	0.8353	0.5416	0.7762	0.7277	0.7353

Abbreviations: ACC, accuracy; ANN, artificial neural network; AUPRC, area under the precision recall curve; AUROC, area under the receiver operating characteristic curve; LR, logistic regression; RF, random forest; SEN, sensitivity; SPE, specificity; XGBoost, eXtreme Gradient Boosting.

TABLE 4 Performance of the model.

Group	AUROC	AUPRC	SEN	SPE	ACC
1	0.8731	0.4992	0.8349	0.7849	0.7900
2	0.8445	0.5995	0.7606	0.7612	0.7611
3	0.7608	0.5718	0.7368	0.7069	0.7152

Abbreviations: ACC, accuracy; AUPRC, area under the precision recall curve; AUROC, area under the receiver operating characteristic curve; SEN, sensitivity; SPE, specificity.

of feature importance were INR, lymphocyte absolute value, and pH. For patients in group 3, systolic blood pressure (SBP), RR, and creatine contributed maximally to intubation prediction. In groups 1 and 2, the greater the INR, the greater the risk of intubation. Also, RR, SBP, and lymphocyte absolute value were also positively correlated with the risk of intubation, while pH and creatine were negatively associated with the risk of intubation.

4 | LIMITATIONS

Despite the valuable insights gained from this study, it is important to acknowledge its limitations. First, as a single-center retrospective study, the findings may not be generalizable to other settings or populations. Additionally, some patient data loss could have potentially biased the results and affected data quality. While efforts were made to replace missing data with average values, this approach may not accurately reflect individual patient characteristics.

Furthermore, the output of our model is only "intubation" or "non-intubation," it is ultimately up to clinicians to determine when intubation should occur based on their clinical judgment and assessment of each patient's unique situation. Finally, the study included a large amount of patient data during the COVID-19 pandemic, especially in the early stages of the pandemic. At this stage, in order to prevent the rapid deterioration of the patient's condition, clinicians are more inclined to early intubation. Therefore, in the post-epidemic era, the effectiveness and applicability of the model may need to be reassessed and adjusted. Overall, while this study offers valuable insights into factors associated with intubation outcomes in critically ill patients, its limitations highlight the need for continued research in this area. By addressing these limitations and building upon existing knowledge, we can improve our understanding of critical care management and ultimately enhance patient outcomes. In future studies, we aim to strengthen the predictive ability of the model by introducing initial diagnosis in the ED.

5 | DISCUSSION

Based on data derived from 15,924 patients, we developed a model with an AUROC of 0.84, SEN of 0.78, and SPE of 0.73 to predict the need for intubation in critically ill patients in the ED. The predictive

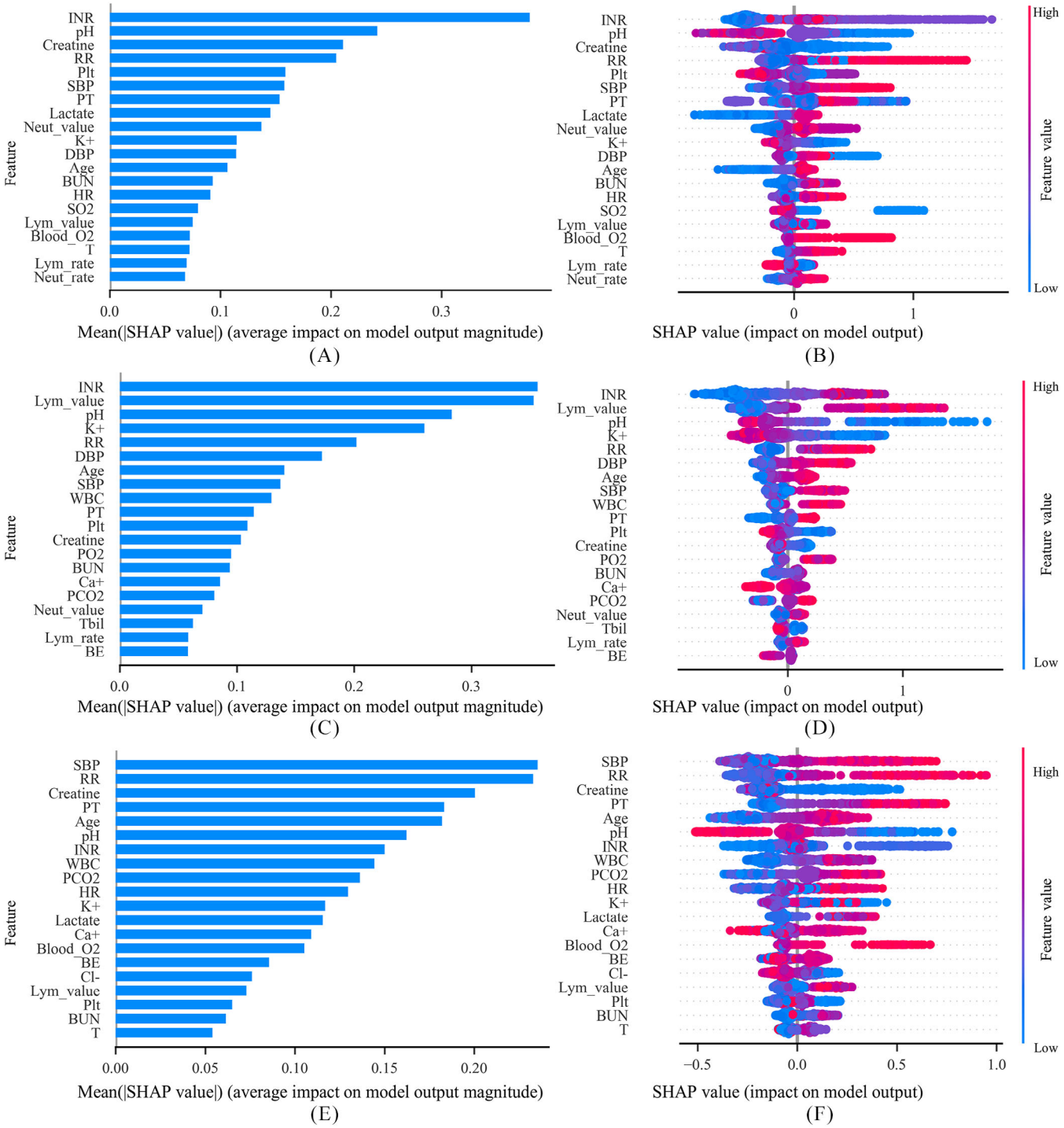

FIGURE 6 SHapley Additive exPlanations (SHAP) analysis of base model. (A) Mean |SHAP value| ranking of base model 1. (B) SHAP value ranking of base model 1. (C) Mean |SHAP value| ranking of base model 2. (D) SHAP value ranking of base model 2. (E) Mean |SHAP value| ranking of base model 3. (F) SHAP value ranking of base model 3. In (B), (D), and (F), the x-axis is SHAP value, and the y-axis is feature importance ranking. A dot represents a sample, and the color indicates the value of that feature.

model utilizes only bedside parameters that are routinely available at the time of ED admission, and may be employed in clinical settings to alert physicians about patients who are at an elevated risk of requiring intubation during their ED stay, without imposing any additional workload on medical or nursing staff.

The ML models vary according to the scenarios. This study explored the performance of different ML models in predicting tracheal intubation in the ED.[24] The XGBoost model exhibits the highest AUROC and superior predictive performance in both subgroups and overall, surpassing other models. Therefore, we employed XGBoost to

construct a multifactorial prediction model for tracheal intubation, which can serve as a clinical decision-making tool to advance precision medicine.

Many predictive intubation models are based on data from ICU, such as MIMIC-IV, eICU, HiRID, ANZICS, and PIC.[25–27] Venturini et al. collected data from 3425 ICU patients and proposed a new ML method to predict the timing of intubation within 5 days of ICU admission based on the concept of cure–survival model.[28] Using MIMIC-III and eICU databases, Daniela et al. developed and validated a predictive model for intubation in children between 24 h and 7 days following pediatric ICU admission.[29] These models exhibited good predictive performance, but were trained in more homogenous critically ill populations than we have studied here and focused on longer-term outcomes less relevant to ED decision making.[30,31] To our knowledge, our model is the first intubation prediction model constructed using ED data. Different from ICU, we found that WBC, INR, RR, and pH are the most important risk factors for tracheal intubation in ED, which provided new insights for clinical practice.

The current model incorporates multiple features, including age, RR, oxygen saturation, and others. Through analyzing the correlation between these features and endotracheal intubation, a reliable predictive model was established. Furthermore, cross-validation and other evaluation metrics were employed to assess the performance of the model while optimization and adjustments were conducted. At present, it is necessary to input the detection indicators in the order of the model interface, that is, the features, when the model is trained, and the model gives the prediction results. In our future work, we will consider encapsulating the model to simplify the difficulty of practical operation and facilitate clinical use. Future studies would improve and expand upon this model. For instance, incorporating additional features and data such as patient medical history and laboratory test results enhance the prediction ACC. Furthermore, integrating the model with other medical devices and systems can enable real-time monitoring and prediction of patients' endotracheal intubation needs.

In clinical practice, clinicians should finally choose whether to perform early intubation according to the results of the model, their own clinical judgment and the patient's condition changes. Although this study screened important risk factors for endotracheal intubation in the ED, it should be noted that no single feature can be used alone for high-precision intubation prediction. In our model, multiple features are combined for prediction. At the same time, according to the results of cluster analysis, clinicians should allocate more energy and medical resources to patients in group 3 to reduce mortality. SHAP analysis revealed inconsistencies in feature importance across different subgroups. Clinicians should focus on distinct features for patients with different groups, such as coagulation function and pH for groups 1 and 2, and blood pressure and respiratory function for group 3.

Overall, our study highlights the potential benefits of using predictive modeling in emergency medicine. By leveraging data analytics and artificial intelligence techniques, we can improve clinical decision making and ultimately enhance patient care.

AUTHOR CONTRIBUTIONS

Hongbo Ding, Xue Feng, and Qi Yang designed the study, extracted and analyzed the data, and wrote the paper as the co-first author. Siyi Zhu and Mei Zhao contributed to the analysis of the results in a statistical aspect. Yichang Yang, Xiaozhen Ji, Yangbo Kang, Jiashen Shen, and Shanxiang Xu assisted in the support of clinical knowledge and reviewed the paper. Gangmin Ning and Yongan Xu were in-charge of the overall direction of the study as the corresponding author. All authors gave final approval of the paper for submission.

ACKNOWLEDGMENTS

The authors would like to thank the patients who participated in this study. This work was supported by the National Natural Science Foundation of China (grants 81571916, 81372079, and 81201478), the Key Research and Development (R&D) Program of Zhejiang Province (grants 2022C03G2013326 and 2024C3186), the Zhejiang Province Public Welfare Foundation Project (grant LGF22H150009), and the Major Project of National-Zhejiang Provincial Administration of Traditional Chinese Medicine (grant GZY-ZJ-KJ-24030) .

CONFLICT OF INTEREST STATEMENT

The authors declare they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this published article and its additional files.

ETHICS STATEMENT

This study was reviewed and approved by the Institutional Review Board (project no. 2022-274) of Second Affiliated Hospital Zhejiang University School of Medicine. Informed consent is not required for this study and all patient data are de-identified. All procedures in this research were performed in accordance with the Declaration of Helsinki and relevant regulations.

ORCID

Hongbo Ding MM https://orcid.org/0009-0005-3051-9537
Yongan Xu PhD, MD https://orcid.org/0000-0002-8547-0014

REFERENCES

1. Wong IMJ, Ferguson ND, Urner M. Invasive mechanical ventilation. *Intensive Care Med*. 2023;49(6):669-672. doi:10.1007/s00134-023-07079-8
2. Walter JM, Corbridge TC, Singer BD. Invasive mechanical ventilation. *South Med J*. 2018;111(12):746-753. doi:10.14423/SMJ.0000000000000905
3. Semler MW, Casey JD, Lloyd BD, et al. Oxygen-saturation targets for critically ill adults receiving mechanical ventilation. *N Engl J Med*. 2022;387(19):1759-1769. doi:10.1056/NEJMoa2208415
4. Bayram B, Sanci E. Invasive mechanical ventilation in the emergency department. *Turk J Emerg Med*. 2019;19(2):43-52. doi:10.1016/j.tjem.2019.03.001
5. Pham T, Brochard LJ, Slutsky AS. Mechanical ventilation: state of the art. *Mayo Clin Proc*. 2017;92(9):1382-1400. doi:10.1016/j.mayocp.2017.05.004

6. Gumustop S, Gallo-Bernal S, McPeake F, Briggs D, Gee MS, Pianykh OS. Predicting health crises from early warning signs in patient medical records. *Sci Rep*. 2022;12(1):19267. doi:10.1038/s41598-022-23900-8
7. Mudumbai SC, Barr J, Scott J, et al. Invasive mechanical ventilation in California over 2000–2009: implications for emergency medicine. *West J Emerg Med*. 2015;16(5):696-706. doi:10.5811/westjem.2015.6.25736
8. Bolourani S, Brenner M, Wang P, et al. A machine learning prediction model of respiratory failure within 48 hours of patient admission for COVID-19: model development and validation. *J Med Internet Res*. 2021;23(2):e24246. doi:10.2196/24246
9. Weingart SD. Managing initial mechanical ventilation in the emergency department. *Ann Emerg Med*. 2016;68(5):614-617. doi:10.1016/j.annemergmed.2016.04.059
10. Stefan MS, Nathanson BH, Higgins TL, et al. Comparative effectiveness of noninvasive and invasive ventilation in critically ill patients with acute exacerbation of chronic obstructive pulmonary disease. *Crit Care Med*. 2015;43(7):1386-1394. doi:10.1097/CCM.0000000000000945
11. Mohr NM, Wessman BT, Bassin B, et al. Boarding of critically ill patients in the emergency department. *Crit Care Med*. 2020;48(8):1180-1187. doi:10.1097/CCM.0000000000004385
12. Duan J, Wang S, Liu P, et al. Early prediction of noninvasive ventilation failure in COPD patients: derivation, internal validation, and external validation of a simple risk score. *Ann Intensive Care*. 2019;9(1):108. doi:10.1186/s13613-019-0585-9
13. Carrillo A, Gonzalez-Diaz G, Ferrer M, et al. Non-invasive ventilation in community-acquired pneumonia and severe acute respiratory failure. *Intensive Care Med*. 2012;38(3):458-466. doi:10.1007/s00134-012-2475-6
14. Wang Z, Zhang L, Huang T, et al. Developing an explainable machine learning model to predict the mechanical ventilation duration of patients with ARDS in intensive care units. *Heart Lung*. 2022;58:74-81. doi:10.1016/j.hrtlng.2022.11.005
15. Goecks J, Jalili V, Heiser LM, Gray JW. How machine learning will transform biomedicine. *Cell*. 2020;181(1):92-101. doi:10.1016/j.cell.2020.03.022
16. Jing Y, Qin P, Fan X, et al. Deep learning-assisted gait parameter assessment for neurodegenerative diseases: model development and validation. *J Med Internet Res*. 2023;25:e46427. doi:10.2196/46427
17. Pladet LCA, Barten JMM, Vernooij LM, et al. Prognostic models for mortality risk in patients requiring ECMO. *Intensive Care Med*. 2023;49(2):131-141. doi:10.1007/s00134-022-06947-z
18. Naqvi SAA, Tennankore K, Vinson A, Roy PC, Abidi SSR. Predicting kidney graft survival using machine learning methods: prediction model development and feature significance analysis study. *J Med Internet Res*. 2021;23(8):e26843. doi:10.2196/26843
19. Shehab M, Abualigah L, Shambour Q, et al. Machine learning in medical applications: a review of state-of-the-art methods. *Comput Biol Med*. 2022;145:105458. doi:10.1016/j.compbiomed.2022.105458
20. Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. *Nat Rev Mol Cell Biol*. 2022;23(1):40-55. doi:10.1038/s41580-021-00407-0
21. Siu BMK, Kwak GH, Ling L, Hui P. Predicting the need for intubation in the first 24 h after critical care admission using machine learning approaches. *Sci Rep*. 2020;10(1):20931. doi:10.1038/s41598-020-77893-3
22. Arvind V, Kim JS, Cho BH, Geng E, Cho SK. Development of a machine learning algorithm to predict intubation among hospitalized patients with COVID-19. *J Crit Care*. 2021;62:25-30. doi:10.1016/j.jcrc.2020.10.033
23. Wong AI, Cheung PC, Kamaleswaran R, Martin GS, Holder AL. Machine learning methods to predict acute respiratory failure and acute respiratory distress syndrome. *Front Big Data*. 2020;3:579774. doi:10.3389/fdata.2020.579774
24. Jiang M, Pan CQ, Li J, Xu LG, Li CL. Explainable machine learning model for predicting furosemide responsiveness in patients with oliguric acute kidney injury. *Ren Fail*. 2023;45(1):2151468. doi:10.1080/0886022X.2022.2151468
25. Duan J, Chen L, Liu X, et al. An updated HACOR score for predicting the failure of noninvasive ventilation: a multicenter prospective observational study. *Crit Care*. 2022;26(1):196. doi:10.1186/s13054-022-04060-7
26. Hyland SL, Faltys M, Huser M, et al. Early prediction of circulatory failure in the intensive care unit using machine learning. *Nat Med*. 2020;26(3):364-373. doi:10.1038/s41591-020-0789-4
27. Wang Z, Zhang L, Huang T, et al. Developing an explainable machine learning model to predict the mechanical ventilation duration of patients with ARDS in intensive care units. *Heart Lung*. 2023;58:74-81. doi:10.1016/j.hrtlng.2022.11.005
28. Venturini M, Van Keilegom I, De Corte W, Vens C. Predicting time-to-intubation after critical care admission using machine learning and cured fraction information. *Artif Intell Med*. 2024;150:102817. doi:10.1016/j.artmed.2024.102817
29. Chanci D, Grunwell JR, Rafiei A, et al. Development and validation of a model for endotracheal intubation and mechanical ventilation prediction in PICU patients. *Pediatr Crit Care Med*. 2024;25(3):212-221. doi:10.1097/PCC.0000000000003410
30. Burdick H, Lam C, Mataraso S, et al. Prediction of respiratory decompensation in Covid-19 patients using machine learning: the READY trial. *Comput Biol Med*. 2020;124:103949. doi:10.1016/j.compbiomed.2020.103949
31. Le Terrier C, Suh N, Wozniak H, et al. Delayed intubation is associated with mortality in patients with severe COVID-19: a single-centre observational study in Switzerland. *Anaesth Crit Care Pain Med*. 2022;41(4):101092. doi:10.1016/j.accpm.2022.101092

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ding H, Feng X, Yang Q, et al. A risk prediction model for efficient intubation in the emergency department: A 4-year single-center retrospective analysis. *JACEP Open*. 2024;5:e13190. https://doi.org/10.1002/emp2.13190