

A New Risk Model Based on the Machine Learning Approach for Prediction of Mortality in the Respiratory Intensive Care Unit

Peng Yan

Chinese PLA General Hospital <https://orcid.org/0000-0002-2398-6189>

Siwan Huang

Ping An Healthcare Technology

Tiange Chen

Ping An Healthcare Technology

Xiang Li

Ping An Healthcare Technology

Guoxin Mo

Chinese PLA General Hospital

Huan Wu

Chinese PLA General Hospital

Jianqiao Xu

Chinese PLA General Hospital

Guotong Xie

Ping An Healthcare Technology

Lixin Xie (✉ xielx301@126.com)

Chinese PLA General Hospital

Research article

Keywords: critical care, RICU, prediction model, mortality

Posted Date: October 30th, 2019

DOI: <https://doi.org/10.21203/rs.2.16574/v1>

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

[Read Full License](#)

Abstract

Background and objectives : Intensive care unit (ICU) resources are inadequate for the large population in China, so it is essential for physicians to evaluate the condition of patients at admission. In this study, our objective was to construct a machine learning risk prediction model for mortality in respiratory intensive care units (RICUs).

Methods: This study involved 817 patients who made 1,063 visits and who were admitted to the RICU from January 1, 2012, to December 31, 2017. Potential predictors such as demographic information, laboratory results, vital signs and clinical characteristics were considered. Among the 1,063 visits, the RICU mortality rate was 13.5%. We constructed eXtreme Gradient Boosting (XGBoost) models and compared predictive performances with random forest models, logistic regression models and clinical scores such as Acute Physiology and Chronic Health Evaluation II (APACHE II) and the sequential organ failure assessment (SOFA) system.

Results : For this dataset, XGBoost models achieved the best performance with the area under the receiver operating characteristics curve (AUROC) of 0.923 (95% CI: 0.889 – 0.957) in the test set, which was significantly greater than APACHE II (0.811, 95% CI: 0.778 – 0.844) and SOFA (0.805, 95% CI: 0.770 – 0.840). The Hosmer-Lemeshow statistic was 12.667 with a P-value of 0.124, which indicated a good calibration of our predictive model in the test set. The final model contained variables that were previously known to be associated with mortality, but it also included some features absent from the clinical scores. The mean N-terminal pro-B-type natriuretic peptide (NT-proBNP) of survivors was significantly lower than that of the non-survival group (2066.43 pg/mL vs. 8232.81 pg/mL; $p < 0.001$).

Conclusions : Our results showed that the XGBoost model could be a suitable model for predicting RICU mortality with easy-to-collect variables at admission and help intensivists improve clinical decision-making for RICU patients. We found that NT-proBNP can be a good indicator of poor prognosis.

Background

Intensive care unit (ICU) costs account for a large portion of increasing health care expenditures, and ICU resources are also inadequate to accommodate large populations, especially in China.^{1, 2} To allocate resources and reduce high costs, physicians have made many efforts to ensure that limited resources go to patients who can benefit most from them. Therefore, the evaluation of patients' clinical condition at ICU admission, which is a risk assessment of mortality, has received much attention.

To determine the illness severity of patients, many clinical scores have been developed. Acute physiology and chronic health evaluation II (APACHE II) scores³ can stratify the risk of mortality depending on clinical characteristics collected on the first ICU day involving twelve acute physiological variables, age and chronic health information. A sequential organ failure assessment (SOFA) system⁴ is used to quantify the degree of six organ failures to evaluate severity. However, the accuracy of both scores for predicting mortality is limited. Moreover, the care conditions and illness severity of ICU patients in

developing countries are different from those in developed countries, so there is a need to construct a risk model that can be applied to patients in developing countries.

In recent years, with the development of machine learning, many studies have deployed multivariate prediction models such as eXtreme Gradient Boosting (XGBoost) models, random forest models and logistic regression models for the research of healthcare problems.^{5, 6} This study aimed to develop a new risk prediction model based on a machine learning method to predict the mortality of patients on their first day in the ICU with easy-to-obtain variables.

Methods

Data Source

The data was obtained from the hospital information system of the respiratory intensive care unit (RICU) in the Chinese People's Liberation Army General Hospital (PLAGH). PLAGH has 125 clinical, medical and technological departments, 4000 patient beds, and an annual volume of more than 3.8 million outpatient visits, over 110000 admissions and more than 65000 operations. Every year, an average of nearly 200 patients are admitted to the RICU in PLAGH.

This hospital information system consists of clinical and demographic information, laboratory results, mortality and lengths of RICU stays for patients admitted to the RICU until the time of discharge. The data were retrospectively collected and then analysed.

From January 1, 2012, to December 31, 2017, a total of 1,350 patients who made 1,643 visits to the RICU were identified. Patients were excluded if their length of RICU stay was less than 24 hours or if they had missing values in important features. After exclusion, 817 patients with 1,063 visits remained in this study (Fig. 1). We considered each visit as an independent incidence.

Data description

The entire dataset of 1,063 visits was randomly split into train and test datasets at a 7:3 ratio. A total of 726 visits with 101 deaths in the RICU (13.9%) were placed in the train set, while 337 visits with 43 deaths in the RICU (12.8%) were placed in the test set.

Outcome and candidate predictors

We defined our outcome as recorded death in the RICU. All features were derived from original datasets. After data cleaning and processing, 107 features with missing values less than 30% were considered (details in Table S1.).

The features included demographic information at the time of RICU admission, such as age, sex and marital status, laboratory results, vital signs and nursing assessment, such as pressure ulcer risk. These

candidate variables were defined as the worst value within 24 hours after RICU admission. Clinical characteristics such as infection within 24 hours after RICU admission were also considered.

Statistical analysis

We first used standard descriptive statistics to describe each feature between the survival group and the non-survival group at RICU discharge. For continuous variables, the mean value and standard deviation were calculated, and differences between groups were assessed by univariate analysis with the Mann-Whitney U test. For categorical variables, group size and proportion were used, and comparisons between groups were examined by the chi-square test.

Then, we implemented the eXtreme Gradient Boosting⁷ (XGBoost) method to predict the mortality of patients admitted to the RICU. The XGBoost method uses the gradient boosting decision tree algorithm: new models are generated to correct the errors of previous fitting. After training, it combines all the models together to achieve better predictive performance. We reached our final model through a hyper-parameter selection process using the k-fold cross validation area under the curve ($k = 5$). We started with default parameters, tuned the maximum depth of a tree and minimum sum of the instance weight needed in a child and tested various numbers of other parameters to avoid overfitting on the training set.

The final XGBoost model used a 'gbtree' booster with a learning objective of logistic regression for binary classification. The model deployed a maximum depth of 5, a minimum sum of instance weights of 20 and a gamma value of 5. The learning rate was set to 0.02 and boosting iterations were set to 560. The subsample ratio of the training instance and of columns when constructing each tree were all set to 0.6. The other parameters were set to their default values.

We compared XGBoost models with random forest models⁸ and logistic regression⁹ models. For random forest models, we first tried to determine the maximum depth of a tree with one tree model trained on the data. Then, we trained with different numbers of trees until the out-of-bag error did not decrease further. Finally, we achieved our model of 100 estimators with a depth of 4. For logistic regression models, we trained models with different regularizations. Our final model used the L2 penalty. As these models cannot handle missing data as automatically as XGBoost models, we generated multiple imputations for incomplete multivariate data.

After the final model was constructed, we employed the Shapley Additive exPlanations (SHAP) method¹⁰ to better interpret the non-linear relationship between variables and outcome. SHAP assigns a value to each feature for each prediction by computing a weighted average of differences between models with one feature included and withheld for all possible feature subsets. With the higher SHAP value, the attribution of this feature to the risk prediction becomes larger. Therefore, we could explain the output of our final model from the SHAP values.

The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUROC) were used to measure the predictive performance of the models. The area under the ROC curve of predictive models

and clinical scores were compared by the DeLong test.¹¹ Calibration of the risk model was assessed by the quintile plot of observed risk versus expected risk and the Hosmer-Lemeshow goodness-of-fit test.¹²

All statistical analyses were performed on Python 3.6.4 and R 3.5.1 (packages: mice, pROC, ResourceSelection). A two-tailed p value < 0.05 was considered statistically significant.

Results

Table 1 summarizes the selected demographics, clinical characteristics, laboratory results and vital signs of the study cohort. Supplementary Table S1 compares all the features between the survival group and the non-survival group. The rate of RICU mortality was 13.5%. Among all RICU patients, 73.2% were male and 88.7% were married. The average age in the survival group was 68.48 years, while the average age in the non-survival group was 80.43 years. The mean N-terminal pro-B-type natriuretic peptide (NT-proBNP) of survivors was significantly lower than that of the non-survival group (2066.43 pg/mL vs. 8232.81 pg/mL; $p < 0.001$). The mean length of RICU stay was 20.32 days.

Model Performance

The XGBoost model achieved the best performance, with the highest area under the ROC curve of 0.923 (95% CI: 0.889–0.957) in the test set, followed by the random forest model and logistic regression models. The area under the ROC curve of all models can be found in Fig. 2. The XGBoost model greatly outperformed the random forest models ($p = 0.018$), logistic regression models ($p < 0.001$) and clinical scores ($p < 0.001$). The calibration of the XGBoost model is illustrated in Fig. 3. The Hosmer-Lemeshow statistic of the model was 12.67 ($p = 0.124$) in the test set, indicating no evidence of poor fit.

Feature Importance and Interpretations

Feature importance was calculated by the amount of improvement in accuracy after a feature was introduced onto a branch. This indicated how valuable each feature was in the processing of prediction model construction. The ten most important features derived from the final XGBoost model are presented in Fig. 4. We applied the SHAP method to better understand the non-linear relationship between those variables and the outcome in Fig. S1-S10. NT-proBNP contributed the most to the model. When its value was larger than approximately 1200 pg/mL, the risk of death started to increase rapidly (Fig. S1). The age at RICU admission also had great influences, and it had two clear changing points at approximately 80 and 85 (Fig. S2). Urea (Fig. S3) ranked as the third most important feature, followed by lactic acid (Fig. S4), blood glucose (Fig. S5), and respiratory rate (Fig. S7). Their relationships with the outcome all shared an S shape: increasing slowly initially, then becoming steeper in the middle and flattening off again at the end. The pressure ulcer risk score (Fig. S8) and red blood cell count (Fig. S10) shared an inverse S-shaped curve. When variables such as red blood cells in urine (Fig. S6) and myoglobin (Fig. S9) were below the normal value, the risk of death increased sharply.

Discussion

We trained three machine learning approach-based models and assessed their performance with each other and clinical scores using the area under the ROC curve. Among them, the XGBoost model had the best predictive power with an area under the ROC curve of 0.923 (95% CI: 0.889–0.957) in the test set; this model was significantly greater than other models and had higher scores ($p < 0.05$). Moreover, XGBoost also showed good calibration by the Hosmer-Lemeshow goodness-of-fit test (12.67; $p = 0.124$). Therefore, our study established a new mortality prediction model that can be applied to RICU patients.

Many studies have attempted to develop risk models for predicting mortality in the ICU. Most of the studies focused on time-series data. Ghassemi M et al.¹³ developed time-varying models with a combination of latent topic features and baseline features that reported an AUC of 0.85. Another study¹⁴ in a German tertiary care centre showed a real-time prediction for ICU mortality after cardiothoracic surgery. However, some studies^{13,15} concluded that data collected within 24 hours after admission contributed the most to the predictive power of the model. Furthermore, due to inadequate ICU resources, many clinical decisions should be made within a limited time after admission to avoid treatment delays. Our predictive model was based on baseline features, which were all easy to collect, to help intensivists assess the clinical conditions of patients in the first 24 hours after they are admitted to the ICU.

We also identified several important features. As expected, baseline age and respiratory rate were important in the model and were also included in clinical scores like APACHE II and SOFA. Our results showed that patients aged 80 years and over had a higher ICU death risk (Fig. S2), a conclusion that was supported by the study of Bagshaw S M, et al.¹⁶ Similar to our findings (Fig. S4-S5), previous studies have identified that high blood glucose¹⁷ and lactic acid level¹⁸ play an important role in mortality prediction. NT-proBNP, the feature that most influenced the model, has been widely recognized as an excellent diagnostic and prognostic marker in heart failure patients.¹⁹ Many studies have investigated its usefulness in ICU patients. One study²⁰ proposed that a single measurement of NT-proBNP at admission might be a potential prognostic marker in unselected ICU patients, and survivors had significantly lower NT-proBNP than non-survivors, which was consistent with our result (Fig. S1). As the third most important variable, a high urea level (Fig. S3)²¹ has been reported to be an independent predictive value for the 12-month mortality prediction of elderly admitted to the ICU, combined with the presence of acute renal failure, the need for mechanical ventilation, a low Glasgow Coma Score (GCS) and age. Pressure ulcer risk score classifies the risk for pressure ulcers: a score under 14 indicates an increased risk of pressure ulcer development.²² In addition, patients who developed pressure ulcers were more likely to die during their hospital stay,²³ which was consistent with our results (Fig. S8).

Our model introduced some novel predictors of mortality risk for RICU patients, such as red blood cells in the urine, myoglobin and red blood cell count. For example, myoglobin has long been evaluated as an early marker of myocardial infarction diagnostics and is associated with the mortality of patients after cardiac surgery.²⁴ However, the association of ICU mortality with myoglobin has not been discussed in previous work. In future studies, we may further examine the association between these novel variables and ICU mortality. Although relationships between several top influential features and outcome have been

briefly discussed, our primary objective was to build a predictive model that combined all the features but not to identify individual risk predictors.

There are some limitations we would like to acknowledge. Our study was conducted at a single centre retrospectively. The external validations from separate institutions would be necessary to test the generalized predictive power of the model for future application. Moreover, our model was only based on the population of the RICU. Further study on the population from mixed types of ICU and a comparison between different types of ICU would be advised.

Conclusion

We developed a new predictive model of RICU mortality based on a machine learning approach. Our final model, with its superior predictive performance and good calibration, would help intensivists make better clinical decisions to identify patients who are at the most risk.

Abbreviations

ICU: intensive care unit; RICU: respiratory intensive care unit; SOFA: sequential organ failure assessment; APACHE II: acute physiology and chronic health evaluation II; PLAGH: Chinese People's Liberation Army General Hospital; RF: random forest; LR: logistic regression; XGBoost: extreme gradient boosting; ROC: receiving operating characteristic; AUROC: area under an ROC curve; GCS: Glasgow Coma Score; SHAP: shapley additive explanations; NT-proBNP: N-terminal pro-B-type natriuretic peptide.

Declarations

Acknowledgements

The authors thank the Medical Big Data Central for providing the previous medical record data.

Author's Contributions

PY, SH, TC, XL, GXM, HW, JQX: Collected and analysed data. PY, SH: Reviewed the data and wrote the manuscript. GTX, LXX, provided critical revision of the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding

The study was supported by the China National Key Research Program (2018ZX09201013), and the China National Geriatric Clinical Center Program (NCRCG-PLAGH-2017003), which had no role in the study design, data collection, analysis, interpretation, or manuscript preparation.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This retrospective study used previously documented medical records, so it was exempted from clinical ethics. This study was a retrospective data analysis and we did not registered. The need for informed consent was waived.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Declarations

This study was conducted in accordance with the amended Declaration of Helsinki.

Table

Table 1 Selected demographics, clinical characteristics, laboratory results and vital signs of RICU patients. Data are summarized as the mean \pm standard deviation or n (%).

Characteristics	Total (N=1063)	Survivors (N=919)	Non-survivors (N=144)	P value
Demographic variables				
Age, years	70.1 ± 17.47	68.48 ± 17.67	80.43 ± 11.76	<0.001
Male	778 (73.18)	668 (72.68)	110 (76.38)	0.406
Married	943 (88.71)	817 (88.90)	126 (87.5)	0.725
Vital signs				
Body height, cm	167.69 ± 7.62	167.69 ± 7.63	167.67 ± 7.57	0.980
Body weight, kg	64.69 ± 13.88	64.81 ± 13.94	63.9 ± 13.44	0.657
Respiratory rate, /min	23.1 ± 6.77	22.75 ± 6.76	25.28 ± 6.45	<0.001
Heart rate, /min	95.67 ± 25.3	93.29 ± 23.7	110.85 ± 29.65	<0.001
Total fluid output, mL/day	1735.97 ± 865.89	1769.19 ± 840.31	1555.6 ± 977.07	0.003
Total fluid intake, mL/day	2385.89 ± 806.52	2329.16 ± 791.12	2694.87 ± 822.31	<0.001
Mean arterial pressure, mmHg	85.23 ± 20.85	86.54 ± 19.8	77.61 ± 24.91	<0.001
Saturation of peripheral oxygen (SpO ₂), %	95.92 ± 6.62	95.97 ± 6.36	95.67 ± 7.8	0.035
Serum laboratory tests				
Creatinine, mol/L	92.35 ± 100.94	86.66 ± 99.52	127.23 ± 102.95	<0.001
Red blood cell count, 10 ¹² /L	3.65 ± 0.81	3.72 ± 0.79	3.19 ± 0.78	<0.001
Troponin T, ng/mL	0.09 ± 0.32	0.07 ± 0.2	0.25 ± 0.68	<0.001
Glucose, mmol/L	7.15 ± 2.83	6.94 ± 2.71	8.49 ± 3.19	<0.001
Direct bilirubin, mol/L	6.55 ± 16.48	5.6 ± 11.29	12.4 ± 33.57	<0.001
Myoglobin, ng/mL	133.78 ± 313.8	108.82 ± 270.3	280.88 ± 473.62	<0.001
Prothrombin time, s	15.4 ± 4.19	15.02 ± 3.18	17.72 ± 7.54	<0.001
Lactate dehydrogenase, U/L	292.57 ± 345.64	279.32 ± 351.84	374.22 ± 292.73	<0.001
NT-proBNP, pg/mL	2948.23 ± 6187.95	2066.43 ± 4648.29	8232.81 ± 10311.47	<0.001
Urea, mmol/L	8.57 ± 7.03	7.66 ± 5.91	14.11 ± 10.16	<0.001
C-reactive protein, mg/dL	6.05 ± 6.99	5.62 ± 6.97	8.54 ± 6.6	<0.001
Urinalysis				
Red blood cells in urine, /L	379.21 ± 1719.98	356.17 ± 1763.88	535.05 ± 1383	<0.001
Arterial blood gas test				

Lactic acid, mmol/L	1.78 ± 1.07	1.68 ± 0.94	2.36 ± 1.48	<0.001
Oxygenation index	263.77 ± 107.2	272.67 ± 106.47	211.54 ± 96.32	<0.001
Nursing assessments				
Glasgow Coma Score (GCS)	12.89 ± 3.51	13.16 ± 3.3	11.1 ± 4.27	<0.001
Norton score for pressure ulcer risk	13.73 ± 3.07	14.05 ± 3.02	11.72 ± 2.58	<0.001

NT-proBNP = N-terminal pro-B-type natriuretic peptide.

References

- Halpern, N. A., & Pastores, S. M. (2010). Critical care medicine in the United States 2000–2005: an analysis of bed numbers, occupancy rates, payer mix, and costs.. *Critical Care Medicine*, 38(1), 65–71.
- Du, B., Xi, X., Chen, D., & Peng, J. (2010). Clinical review: Critical care medicine in mainland China. *Critical Care*, 14(1), 206–206.
- Knaus, W. A., Draper, E. A., Wagner, D. P., & Zimmerman, J. E. (1985). APACHE II: a severity of disease classification system.. *Critical Care Medicine*, 13(10), 818–829.
- Vincent, J. L., De Mendonca, A., Cantraine, F., Moreno, R., Takala, J., Suter, P.M., ... & Blecher, S. (1998). Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: Results of a multicenter, prospective study. *Critical Care Medicine*, 26(11), 1793–1800.
- Zheng, T., Xie, W., Xu, L., He, X., Zhang, Y., You, M., ... & Chen, Y. (2017). A machine learning-based framework to identify type 2 diabetes through electronic health records. *International Journal of Medical Informatics*, 120–127.
- Van Rosendael, A. R., Maliakal, G., Kolli, K. K., Beecy, A., Alaref, S. J., Dwivedi, A., ... & Min, J. K. (2018). Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry. *Journal of Cardiovascular Computed Tomography*, 12(3), 204–209.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *knowledge discovery and data mining*, 785–794.
- Breiman L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the royal statistical society series b-methodological*, 58(1), 267–288.

10. Lundberg, S., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions. *neural information processing systems*, 4765–4774.
11. DeLong, E. R., DeLong, D. M., & Clarkepearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach.. *Biometrics*, 44(3), 837–845.
12. Hosmer D W, Lemeshow S. (2000). *Applied Logistic Regression*. New York, USA: John Wiley and Sons.
13. Ghassemi, M., Naumann, T., Doshivelez, F., Brimmer, N. J., Joshi, R., Rumshisky, A., & Szolovits, P. (2014). Unfolding physiological state: mortality modelling in intensive care units. *knowledge discovery and data mining*, 75–84.
14. Meyer, A., Zverinski, D., Pfahringer, B., Kempfert, J., Kuehne, T., Sundermann, S. H., ... & Eickhoff, C. (2018). Machine learning for real-time prediction of complications in critical care: a retrospective study. *The Lancet Respiratory Medicine*, 6(12), 905–914.
15. Hug, C. W., & Szolovits, P. (2009). ICU Acuity: Real-time Models versus Daily Models. *american medical informatics association annual symposium*, 260–264.
16. Bagshaw, S. M., Webb, S. A., Delaney, A., George, C., Pilcher, D., Hart, G. K., & Bellomo, R. (2009). Very old patients admitted to intensive care in Australia and New Zealand: a multi-centre cohort analysis.. *Critical Care*, 13(2), 1–14.
17. Falciglia, M., Freyberg, R. W., Almenoff, P. L., D'Alessio, D. A., & Render, M. L. (2009). Hyperglycemia-related mortality in critically ill patients varies with admission diagnosis. *Critical care medicine*, 37(12), 3001–3009.
18. Kruse, O., Grunnet, N., & Barfod, C. (2011). Blood lactate as a predictor for in-hospital mortality in patients admitted acutely to hospital: a systematic review. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, 19(1), 74–74.
19. De Lemos, J. A., Morrow, D. A., Bentley, J. H., Omland, T., Sabatine, M. S., McCabe, C. H., ... & Braunwald, E. (2001). The prognostic value of B-type natriuretic peptide in patients with acute coronary syndromes. *The New England Journal of Medicine*, 345(14), 1014–1021.
20. Meyer, B., Huelsmann, M., Wexberg, P., Karth, G. D., Berger, R., Moertl, D., ... & Heinz, G. (2007). N-terminal pro-B-type natriuretic peptide is an independent predictor of outcome in an unselected cohort of critically ill patients.. *Critical Care Medicine*, 35(10), 2268–2273.
21. Lown, D. J., Knott, J. C., Rechnitzer, T., & Macisaac, C. (2013). Predicting short-term and long-term mortality in elderly emergency patients admitted for intensive care.. *Critical Care and Resuscitation*, 15(1).
22. Norton D, McLaren R & Exton-Smith A. (1962). *An Investigation of Geriatric Nursing Problems in Hospital*. Churchill Livingstone, Edinburgh.
23. Becker, D., Tozo, T. C., Batista, S. S., Mattos, A. L., Silva, M. C., Rigon, S., ... & Duarte, P. A. (2017). Pressure ulcers in ICU patients: Incidence and clinical and epidemiological features: A multicenter study in southern Brazil. *Intensive and Critical Care Nursing*, 55–61.

24. Hofmann, D., Buettner, M., Rissner, F., Wahl, M., & Sakka, S. G. (2007). Prognostic value of serum myoglobin in patients after cardiac surgery. *Journal of Anesthesia*, 21(3), 304–310.

Figures

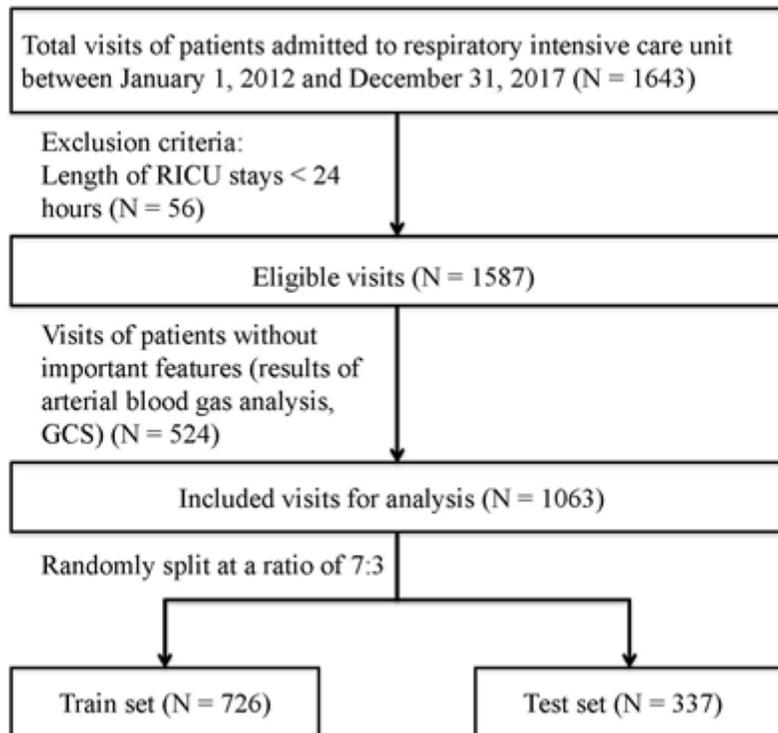


Figure 1

Flowchart of the enrolment procedures for the predictive model.

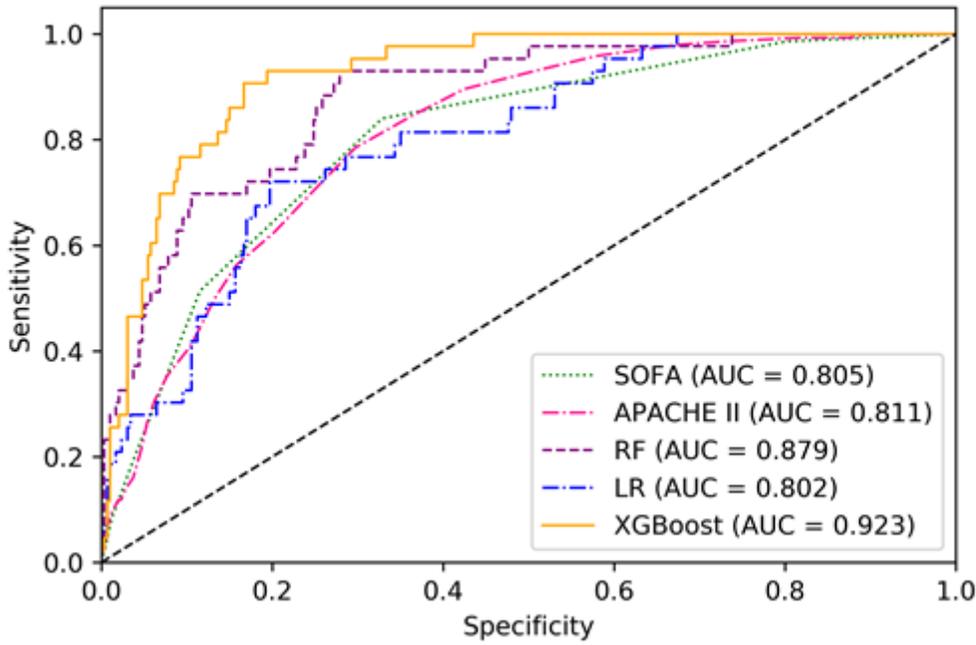


Figure 2

ROC curve of SOFA, APACHE II, RF, LR and XGBoost models. The XGBoost model performed significantly better than SOFA ($p < 0.001$), APACHE II ($p < 0.001$), RF ($p = 0.018$) and LR ($p < 0.001$) using the DeLong test. SOFA = sequential organ failure assessment; APACHE II = acute physiology and chronic health evaluation II; RF = random forest; LR = logistic regression; XGBoost = extreme gradient boosting.

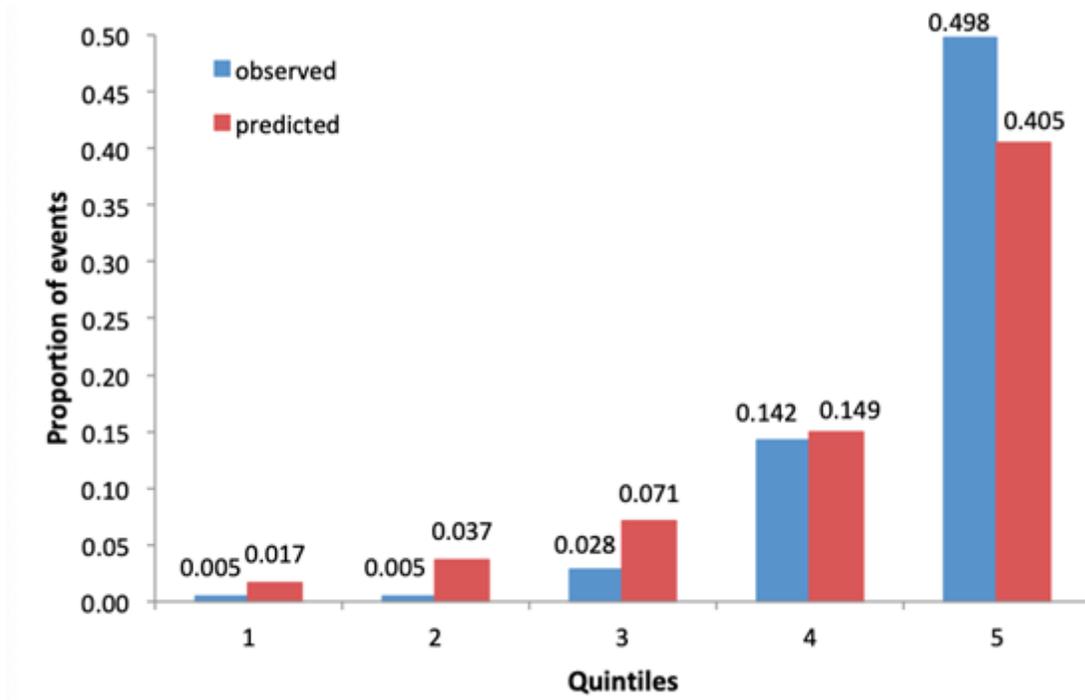


Figure 3

Calibration plot of the XGBoost model in the test set. Data were divided into quintiles by predicted risk. The blue bar represents the observed proportion of events, and the red bar represents the predicted risk. The Hosmer-Lemeshow statistic of the model was 12.67 with a p value of 0.124.

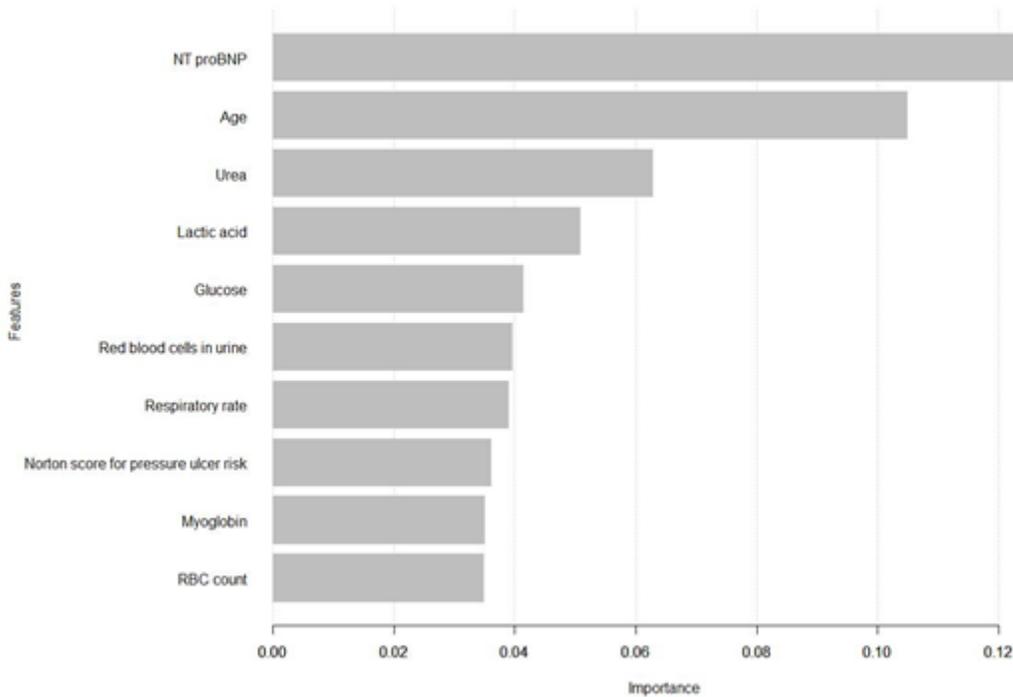


Figure 4

Importance of the top 10 influential variables in the XGBoost model. NT-proBNP contributed most to the model. As expected, age at RICU admission also had great influences. Urea ranked as the third most important feature, followed by lactic acid, blood glucose, red blood cells in urine, respiratory rate, pressure ulcer risk score, myoglobin and red blood cell count.

Supplementary Files

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

- [figures10.jpg](#)
- [Supplementaryfigurelegends.docx](#)
- [figures1.jpg](#)
- [figures9.jpg](#)
- [figures6.jpg](#)
- [figures2.jpg](#)
- [figures3.jpg](#)
- [tables1.docx](#)

- figures8.jpg
- figures4.jpg
- figures5.jpg
- figures7.jpg