

Original Paper

Noninvasive Real-Time Mortality Prediction in Intensive Care Units Based on Gradient Boosting Method: Model Development and Validation Study

Huizhen Jiang^{1*}, BSc; Longxiang Su^{2*}, MD; Hao Wang², MD; Dongkai Li¹, MD; Congpu Zhao¹, BSc; Na Hong³, PhD; Yun Long², MD; Weiguo Zhu¹, MD

¹Department of Information Center, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China

²Department of Critical Care Medicine, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China

³Digital Health China Technologies Co., Ltd, Beijing, China

*these authors contributed equally

Corresponding Author:

Weiguo Zhu, MD

Department of Information Center, State Key Laboratory of Complex Severe and Rare Diseases

Peking Union Medical College Hospital

Chinese Academy of Medical Science and Peking Union Medical College

1 Shuaifuyuan, Dongcheng District

Beijing, 100730

China

Phone: 86 01069154149

Email: Zhuwg@pumch.cn

Abstract

Background: Monitoring critically ill patients in intensive care units (ICUs) in real time is vitally important. Although scoring systems are most often used in risk prediction of mortality, they are usually not highly precise, and the clinical data are often simply weighted. This method is inefficient and time-consuming in the clinical setting.

Objective: The objective of this study was to integrate all medical data and noninvasively predict the real-time mortality of ICU patients using a gradient boosting method. Specifically, our goal was to predict mortality using a noninvasive method to minimize the discomfort to patients.

Methods: In this study, we established five models to predict mortality in real time based on different features. According to the monitoring, laboratory, and scoring data, we constructed the feature engineering. The five real-time mortality prediction models were RMM (based on monitoring features), RMA (based on monitoring features and the Acute Physiology and Chronic Health Evaluation [APACHE]), RMS (based on monitoring features and Sequential Organ Failure Assessment [SOFA]), RMML (based on monitoring and laboratory features), and RM (based on all monitoring, laboratory, and scoring features). All models were built using LightGBM and tested with XGBoost. We then compared the performance of all models, with particular focus on the noninvasive method, the RMM model.

Results: After extensive experiments, the area under the curve of the RMM model was 0.8264, which was superior to that of the RMA and RMS models. Therefore, predicting mortality using the noninvasive method was both efficient and practical, as it eliminated the need for extra physical interventions on patients, such as the drawing of blood. In addition, we explored the top nine features relevant to real-time mortality prediction: invasive mean blood pressure, heart rate, invasive systolic blood pressure, oxygen concentration, oxygen saturation, balance of input and output, total input, invasive diastolic blood pressure, and noninvasive mean blood pressure. These nine features should be given more focus in routine clinical practice.

Conclusions: The results of this study may be helpful in real-time mortality prediction in patients in the ICU, especially the noninvasive method. It is efficient and favorable to patients, which offers a strong practical significance.

(*JMIR Med Inform* 2021;9(3):e23888) doi: [10.2196/23888](https://doi.org/10.2196/23888)

KEYWORDS

real time; mortality prediction; intensive care unit; noninvasive

Introduction

Patients in intensive care units (ICUs) are usually suffering from the most severe and complicated diseases. Thus, they require more intensive care and hospital resources [1]. Research shows that the cost of ICUs accounts for 22% of total hospital costs [2]. The cost of doctors and nurses in the ICU is also a massive burden. Therefore, hospitals usually use scoring systems to help assess patients' risks and then place more efforts on improving the patient care and management. Scoring systems such as the Acute Physiology and Chronic Health Evaluation (APACHE) [3] systems II, III, and IV; the Simplified Acute Physiology Score II (SAPS II) [4]; and the Sequential Organ Failure Assessment (SOFA) score [5] are commonly used to estimate the illness severity of patients in the ICU [6,7]. However, the scoring systems cannot reflect the condition of patients in real time, and clinical staff must spend plenty of time calculating the scores to make decisions. Further, the scores alone are insufficient for the needs of the clinical staff. Johnson and Mark [8] found that the gradient boosting method outperformed the scoring systems on predicting mortality, which provided inspiration for our study.

Meanwhile, the severity and mortality of ICU patients can be specifically assessed in real time using machine learning methods. This would allow doctors and nurses to prepare lifesaving interventions ahead of time and provide families with more time to make decisions [9]. Hence, precisely predicting the mortality of ICU patients is significant. Machine learning technology has significantly changed lives in many aspects in recent years, even in the health care field [10,11]. Usually, the shortest time period for predicting mortality is 24 h [12,13], which is not sufficient for the ICU staff to obtain the real-time condition of patients. Kim et al [14] presented a deep learning method to predict the mortality of patients 6 h to 60 h prior to death, where the time period was a little longer than the real time. With regard to machine learning techniques, the ensemble and neural network models demonstrate better performance in predicting mortality [2]. Brand et al [15] proposed a deep learning method to predict mortality based only on heart rate, respiratory rate, and blood pressure, which had an accuracy of 76.3%, but its performance was not as good as that of other

methods. Besides, the neural network model cannot interpret the gap between the input and the output. Further, it is vulnerable to attack when the training set is inadvertently being modified [16].

In this study, we established a real-time mortality prediction model based on clinical data where we explored a noninvasive method to predict mortality by only monitoring features. Because frequent laboratory examination can cause physical trauma to patients whose bodies are already weak, using a model that can show general performance and is noninvasive is clinically meaningful.

Methods

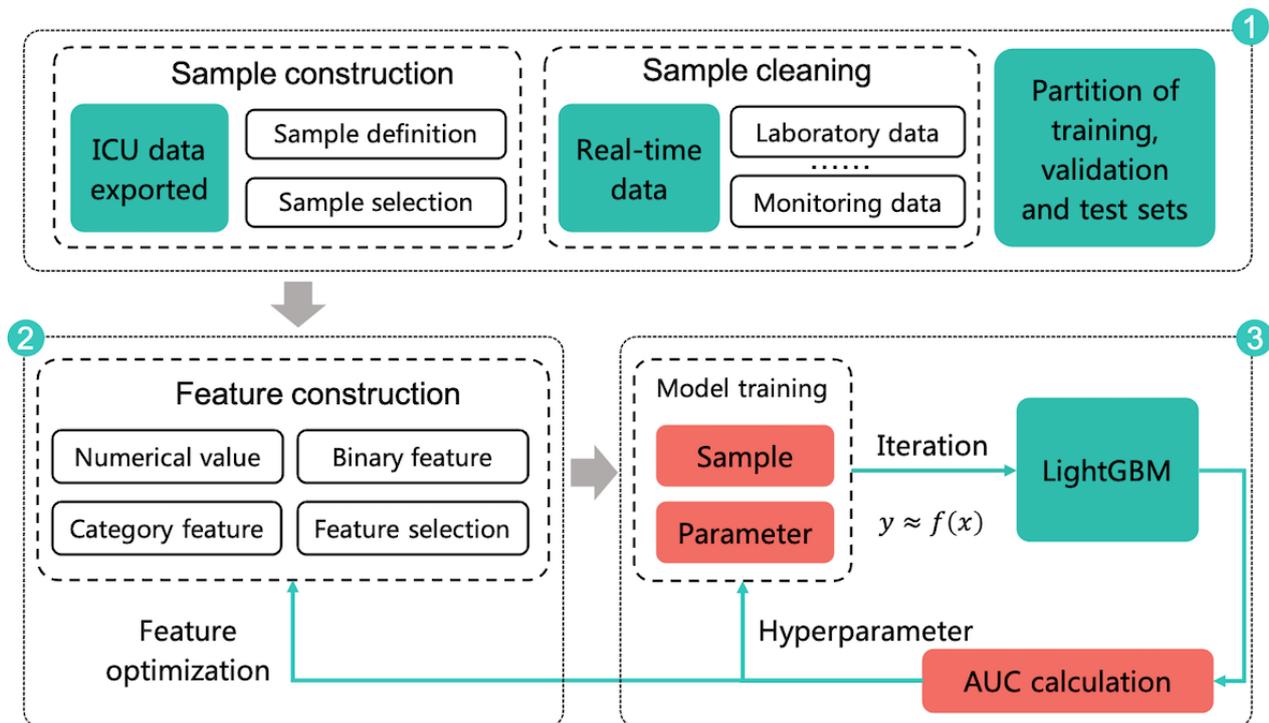
Data Sources

We used the ICU data from Peking Union Medical College Hospital from 2013 to 2018. A total of 13,649 patients were investigated in our experiments with the privacy information filtered out. We mined features from three types of data: real-time monitoring, laboratory, and scoring data. The main features from the monitoring data are listed in [Multimedia Appendix 1](#).

Prediction Model

In this study, we constructed the real-time mortality prediction model based on the clinical data. The data of the patients in the ICU were updated all the time, and the model could predict each patient's mortality once the data were updated; the model could predict the mortality after 2 h at any time if the data were not updated during the 2 h period. Therefore, it is a real-time prediction model. The modeling process involved three steps, which are shown in [Figure 1](#). First, we constructed and cleaned the sample data according to the clinical data. Second, after dividing the data into training and test sets, we normalized all types of data as features. Third, we used the LightGBM method to train the data and optimized the model by adjusting the parameters. LightGBM and XGBoost are both gradient boosting decision tree methods, and LightGBM has good performance and high training efficiency [17]. In this paper, we also compared the performance of the LightGBM and XGBoost methods.

Figure 1. Modeling process, including sample construction and cleaning, feature engineering, model training, and optimization. AUC: area under the curve; ICU: intensive care unit.



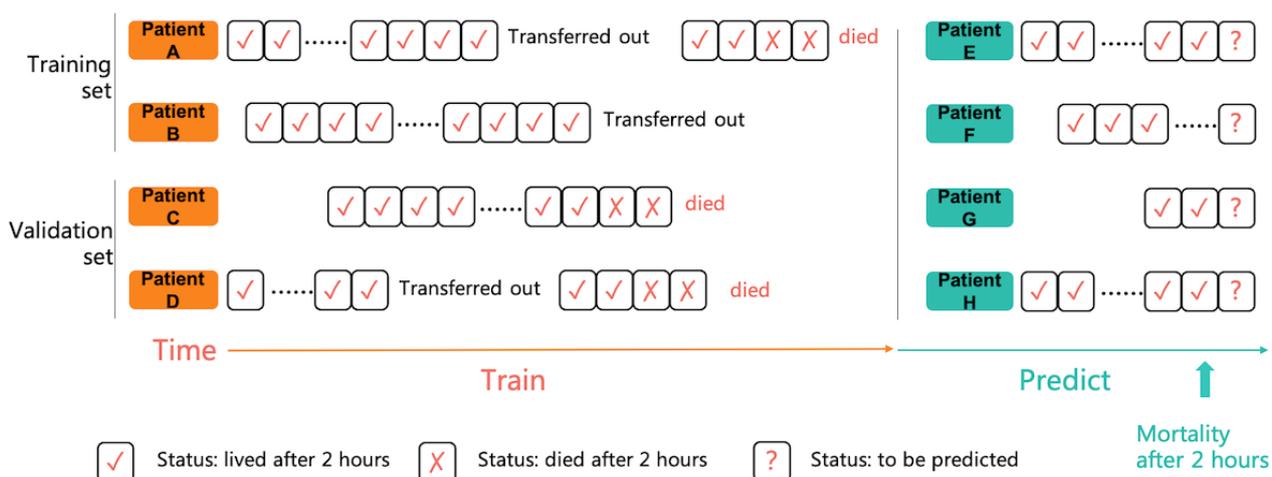
Sample Construction

Usually, patients in ICUs are weak and at high risk. Therefore, focusing on the real-time condition of an ICU patient by the clinical staff is meaningful. In this research, we predicted the mortality of a patient after 2 h based on the clinical data.

Figure 2 shows how we constructed the data by the hour. One record of a patient was captured in each hour, and each patient may have a sample sequence based on the timeline. As a result, there might be several samples for one patient. For example,

patient A had been admitted to the ICU twice, and patients B and C had each been admitted to the ICU once. There were 2 samples for patient A. During patient A's second stay in the ICU, the third box contained a cross, representing a status of "died after 2 h," so he/she died 2 h after he/she was admitted into the ICU because one box meant one sample in 1 h. The data of patients A and B were the training data, and the data of patients C and D were the validation data. Samples of patients E, F, G, and H were the test data set. Then, the samples were constructed according to the process shown in Figure 2. For the 13,649 patients, we constructed 1,172,652 samples in all.

Figure 2. Feature engineering process. Each square represents a 1 h record in the intensive care unit (ie, one sample). The symbols in the squares indicate the status of the patient.



Feature Engineering

Feature engineering is the key process in machine learning. The modeling performance depends on the feature engineering quality to a large extent.

In this study, two data types existed: numerical and categorical data from the monitoring, laboratory, and scoring data. For the numerical data, we directly considered the numerical value as the feature, such as the heart rate and temperature. The categorical data included gender and positive or negative status. We used the LabelEncoder method [18] to normalize these categorical data. LabelEncoder is a method that converts text data into multinumeric values. It can convert two-class and multiclass features. For example, the positive and negative states were represented by 0 and 1, respectively. We left the missing value blank to ensure the authenticity of the data.

Model Training

In this research, we needed to predict the real-time condition of a patient 2 h after each moment. Actually, this process was a binary classification problem (ie, life or death). LightGBM is a gradient boosting method that is superior in dealing with the binary classification problem and has high efficiency and performance, especially in dealing with structured data. The 1,172,652 samples were randomly divided into three parts. One-third of the samples was set as the training set, one-third was set as the validation set, and the rest was set as the test set.

The area under the curve (AUC) was used to evaluate the model's performance.

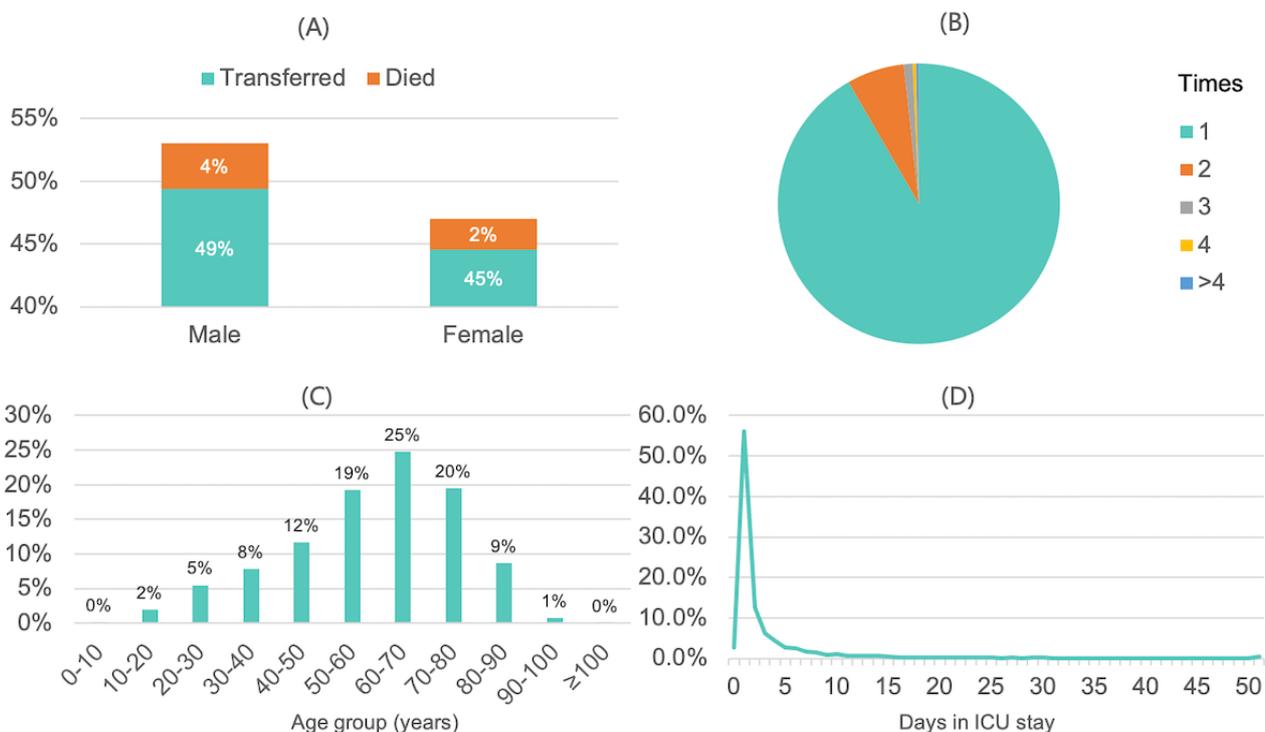
Based on different features, we constructed five real-time mortality prediction models:

- RMM: based on monitoring features;
- RMA: based on monitoring features and APACHE;
- RMS: based on monitoring features and SOFA;
- RMML: based on monitoring and laboratory features; and
- RM: based on all monitoring, laboratory, and scoring features.

Results

In presenting the results of our study, we will focus on the results of the models in the test set. Figure 3 shows the distributions and proportions of patients in the ICU in the data set. Figure 3A shows that male patients in the ICU outnumbered female patients irrespective of whether they were transferred out or died. Figure 3B shows that more than 12,510 patients were transferred into the ICU only once, and 898 patients stayed in the ICU twice. Patients between the ages of 50 and 80 years accounted for 8700 of the total number of patients, as shown in Figure 3C. In addition, patients between the ages of 60 and 70 years represented the largest group, accounting for one-quarter of the total number of patients. Figure 3D shows the length of stay of patients in the ICU in a single visit; we can observe that most patients stayed in the ICU for fewer than 5 days.

Figure 3. Distribution of the data set. (A) Proportion of patients that were transferred out of the intensive care unit (ICU) or died, according to gender. (B) Distribution of the number of times patients transferred into the ICU. (C) Age group distribution of ICU patients. (D) Length of stay of patients in the ICU.



First, we evaluated the influence of the scoring systems through extensive experiments; the results are shown in Figures 4 and 5. Figure 4 shows that the RMM model outperformed the RMA and RMS models. Overall, all three models showed an upward

trend with the increase in tree number and became stable after the tree number reached 200. The RMS model demonstrated better performance than the RMA model. Therefore, the SOFA scoring system was more valuable than the APACHE scoring

system in predicting mortality. Compared with the RMA and RMS models, the RMM model was superior and demonstrated the best performance (AUC 0.8264) when the tree number was 299.

Figure 4. Performance of the RMM, RMA, and RMS models with parameter variation. Each point on the line represents one experiment. AUC: area under the curve.

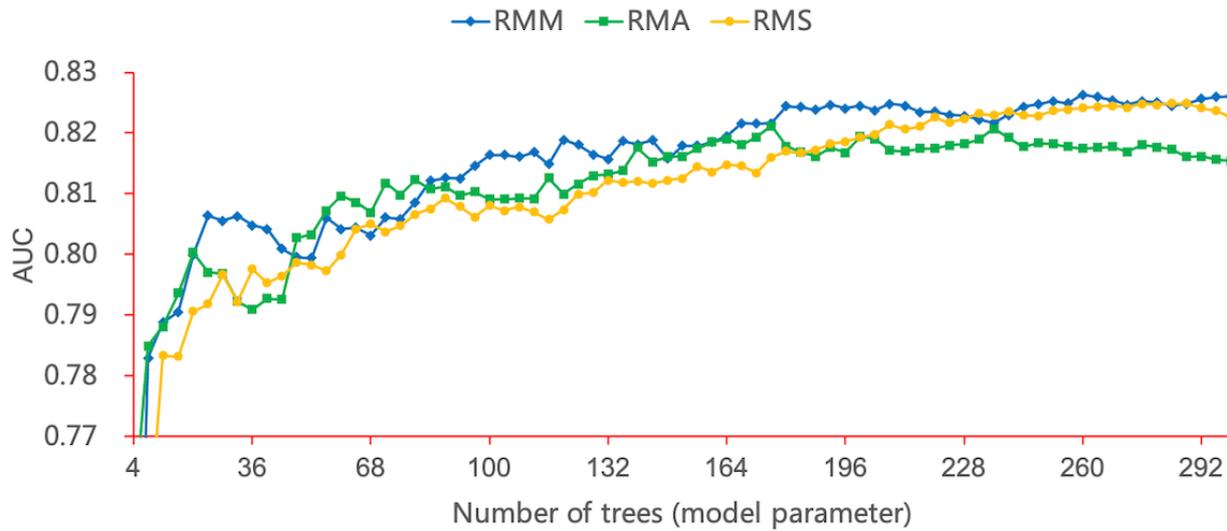
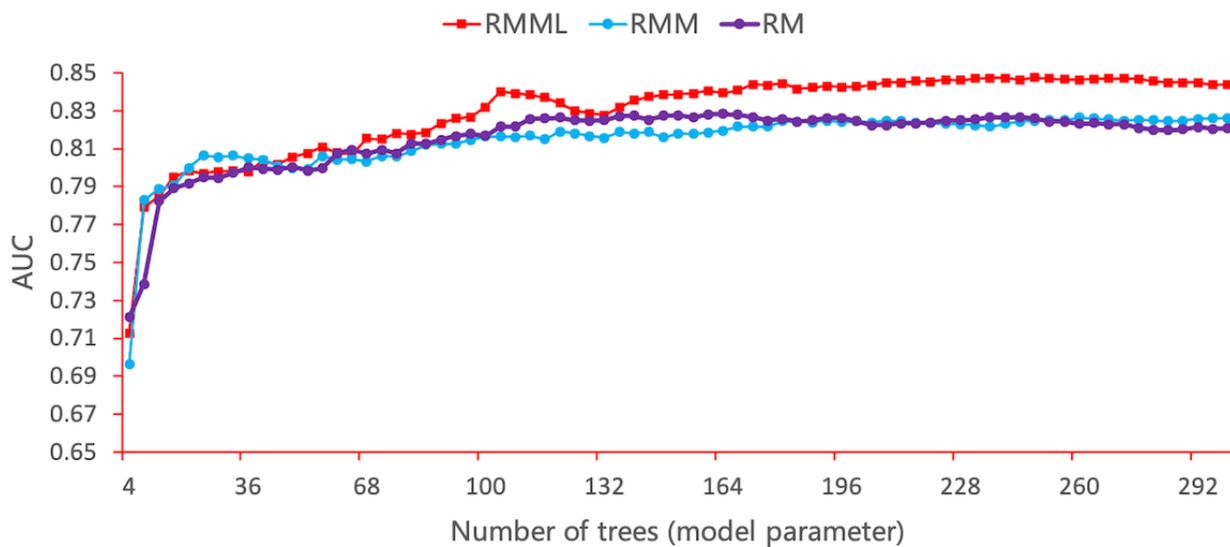


Figure 5 shows the results of the experiments that we conducted on the RMML, RMM, and RM models to compare their performance in terms of the monitoring, laboratory, and scoring features. The RMML model exhibited the best performance based on the monitoring and laboratory features than the other two models. When the tree number was 234, RMML obtained the best AUC (0.8476). In the RM model, monitoring, laboratory, and all scoring features were considered. The RM model exhibited worse performance than the RMM model.

Figure 5. Performance of the RMML, RMM, and RM models with parameter variation. Each point on the line represents one experiment. AUC: area under the curve.



In addition, we repeated the experiments above using XGBoost. The AUCs of XGBoost were relatively lower than those of LightGBM, as shown in Table 1. The best performance with XGBoost was 0.8452 on the RMML model and 0.8154 on the RMM model. As well, we showed the RMML and RMM models using LightGBM and XGBoost on the validation set, and the results are shown in Table 1. Therefore, LightGBM outperformed XGBoost in these experiments.

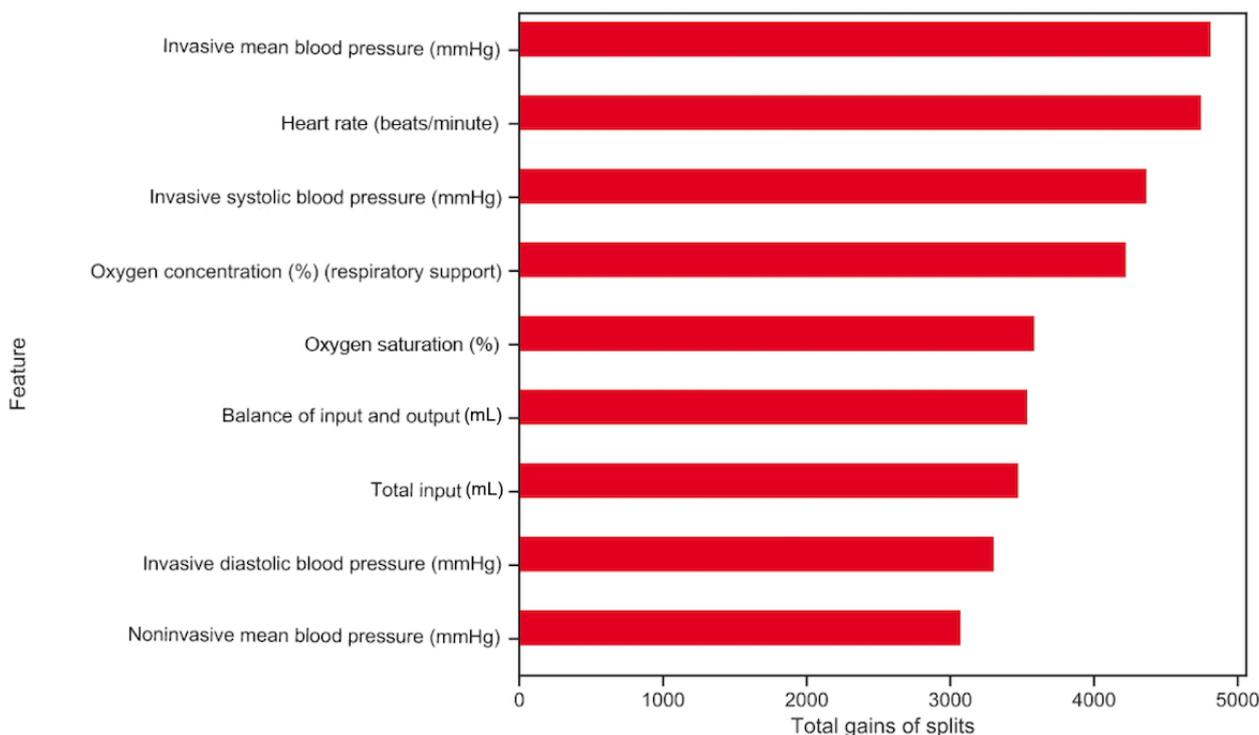
Table 1. Performance of the RMML and RMM models using LightGBM and XGBoost.

Model and method	Area under the curve	
	Test set	Validation set
RMML		
LightGBM	0.8476	0.8483
XGBoost	0.8452	0.8466
RMM		
LightGBM	0.8264	0.8269
XGBoost	0.8154	0.8167

Because the RMML model demonstrated the best performance, we analyzed the relevant features that predicted mortality in that model. Figure 6 shows the top nine features relevant to the mortality prediction. The “gain” of the feature splitting implies the importance of the feature in the model, which was computed during the model training. Thus, the bigger the gains of the feature, the more important the feature was in the model. It was shown that invasive mean blood pressure was the most important feature related to mortality prediction. Among the top nine

features, heart rate, invasive systolic blood pressure, oxygen concentration, oxygen saturation, balance of input and output, total input, invasive diastolic blood pressure, and noninvasive mean blood pressure were all vital sign features in the monitoring. “Balance of input and output” was the difference between input and output data, while “total input” was the input data only. They all demonstrated a relatively strong correlation with the mortality prediction.

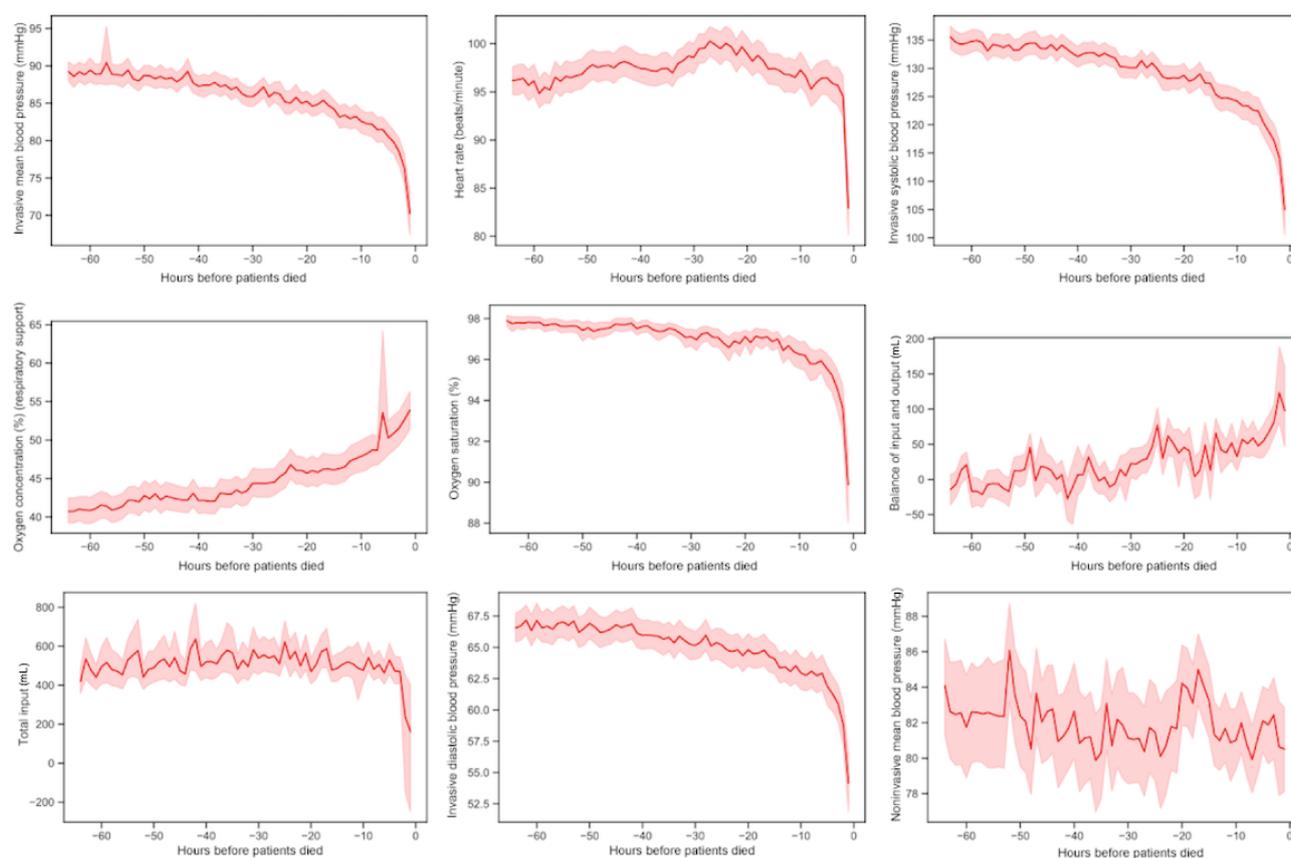
Figure 6. Top nine features relevant to mortality prediction. The horizontal bar represents the gain of each feature in the model; a bigger gain means more relevance and importance in the mortality prediction.



In addition, we exploited the variation in each of the top nine features with time during the last 64 h before a patient died. Figure 7 shows that all nine features showed an obvious trend with the time variation. “Oxygen concentration” and “balance of input and output” exhibited an upward trend during the final 64 h before the patient died. The other seven features all

decreased with time during the final 64 h before the patient died. For example, invasive diastolic blood pressure exhibited a downward trend and sharply declined in the last 5 h. Similarly, the top eight features all rapidly changed in the last 5 h. The “noninvasive mean blood pressure” exhibited dithering but an overall decreasing trend.

Figure 7. Variation in the top nine features relevant to mortality prediction with time. The abscissa represents 64 h before the patients died, and the ordinate represents the value of the feature.



Discussion

In this paper, we used clinical data to predict the real-time mortality of ICU patients. Several models were established based on different features. Extensive experiments showed that the models that used the machine learning method were superior to the scoring systems. More importantly, they can be employed to predict real-time mortality in a noninvasive manner.

Constant care of ICU patients is necessary against their life-threatening conditions. Intensive care is based on more financial support and more professional hospital staff [19,20]. The US health care spending was approximately 17% of the gross domestic product (GDP) in 2011 and may reach 26% of the GDP by 2035 [21]. In addition to the cost, the mortality rate in ICUs cannot be ignored. Studies show that ICUs have the highest mortality rate of all hospital units (16.2% [22] and 22.4% [23]). Therefore, helping predict patient mortality in ICUs is significant, as it could save time for nurses and doctors by more efficiently measuring the risk of ICU patients. It would be better if there was less trauma to patients in the clinical process.

Commonly, hospital staff use scoring systems to help predict the severity status of ICU patients. Most of these scoring systems calculate the scores based on the worst values during the first 24 h after ICU admission [24]. The SAPS score only uses the data in the first hour after ICU admission, which are more robust because the missing data have a lesser effect on specificity [25]. Saleh et al [24] compared APACHE II and III, SAPS II, and

SOFA and showed that APACHE II and III demonstrated better performance than the others. However, Yap et al [26] verified that the National Early Warning Score demonstrated the best performance for predicting the severity status of patients with emphysematous pyelonephritis patients. Tan et al [27] explored the ability of the scoring systems to predict sepsis mortality in the short term (less than 30 days in the hospital) and long term (more than 30 days). They discovered that the sensitivity and specificity were similar in both factors, whereas geographical region had a significant effect on the short-term mortality prediction. Therefore, the scoring systems can show different performance on different diseases and under different situations [28,29]. In addition, Nielsen et al [30] compared the APACHE II and SAPS II with the aggregation of the APACHE II and SAPS II, and the aggregation of APACHE II and SAPS II outperformed each single model. Similarly, Fei et al [31] presented the use of the fibrin degradation product level and APACHE II scores in parallel to improve the prediction performance.

Machine learning technologies have been increasingly used in the health care field because of their excellent performance [20,32]. Further, machine learning models have been confirmed to perform better in predicting the severity status of ICU patients than the scoring systems. Henry et al [33] proposed a supervised learning model to predict the risk of patients getting septic shock, and machine learning was found to have higher sensitivity and specificity than the scoring systems. An ensemble machine learning model was investigated by Pirracchio et al

[2], and the results showed better performance for the machine learning model than for the common scoring systems. In recent years, many studies have focused on using the neural network model to predict mortality [34,35]. Most of the experiments demonstrated that the neural network model outperformed the other models. Norrie [36] innovatively proposed a prespecified library of models and established an optimum model. However, the neural network model is difficult to explain in terms of the black box principle [32,37], which is not clear for high recursion [38]. Using the inherently interpretable models is vital and important in health care because decisions in health care involve high stakes [39]. Awad et al [20] demonstrated that the decision tree model is interpretable and better than the neural network model in predicting mortality. Similarly, Blanco-Justicia et al [40] used a depth-limited decision tree model to avoid the black box problem. In reality, the gradient boosting methods usually perform better than the deep learning method on structured data, especially on a small data set.

The limitation of this study is that the data were obtained from one hospital only. The structure and quality of data may vary in different hospitals. In the future, we would try to improve our model based on multicenter data.

In the present study, we constructed the real-time mortality prediction model based on the monitoring, laboratory, and scoring data. Compared with the RMM, RMA, RMS, and RM models, the RMML model demonstrated the best performance. Moreover, we found that the invasive mean blood pressure, heart rate, and invasive systolic blood pressure were the top three features relevant to the mortality prediction. In addition, the RMM model performed better than the RMA and RMS models. Therefore, noninvasively predicting real-time mortality would be meaningful. Not only can the results of our research provide support for decision making by clinical staff, but our method is also better for patients because the real-time mortality prediction is noninvasive.

Acknowledgments

This study was supported by the National Key Research & Development Program of China (project 2018YFC0116905) and the CAMS Innovation Fund for Medical Sciences (CIFMS; project 2016-I2 M-2-004).

Authors' Contributions

HJ, LS, YL, and WZ contributed to the study concept and design. DL, HW, and CZ contributed to the acquisition of the data set, and HJ created the figures. LS, HJ, and NH consulted on the analyses. All authors interpreted the results, contributed to the manuscript, and approved the final draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The main monitoring features.

[\[DOCX File , 17 KB-Multimedia Appendix 1\]](#)

References

1. Angus DC, Black N. Improving care of the critically ill: institutional and health-care system approaches. *Lancet* 2004 Apr 17;363(9417):1314-1320 [FREE Full text] [doi: [10.1016/S0140-6736\(04\)16007-8](https://doi.org/10.1016/S0140-6736(04)16007-8)] [Medline: [15094279](https://pubmed.ncbi.nlm.nih.gov/15094279/)]
2. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. *Lancet Respir Med* 2015 Jan;3(1):42-52 [FREE Full text] [doi: [10.1016/S2213-2600\(14\)70239-5](https://doi.org/10.1016/S2213-2600(14)70239-5)] [Medline: [25466337](https://pubmed.ncbi.nlm.nih.gov/25466337/)]
3. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. *Crit Care Med* 2006 May;34(5):1297-1310. [doi: [10.1097/01.CCM.0000215112.84523.F0](https://doi.org/10.1097/01.CCM.0000215112.84523.F0)] [Medline: [16540951](https://pubmed.ncbi.nlm.nih.gov/16540951/)]
4. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;270(24):2957-2963. [doi: [10.1001/jama.270.24.2957](https://doi.org/10.1001/jama.270.24.2957)] [Medline: [8254858](https://pubmed.ncbi.nlm.nih.gov/8254858/)]
5. Vincent J, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med* 1998 Nov;26(11):1793-1800. [doi: [10.1097/00003246-199811000-00016](https://doi.org/10.1097/00003246-199811000-00016)] [Medline: [9824069](https://pubmed.ncbi.nlm.nih.gov/9824069/)]
6. Rothwell P. Physiological scoring systems and audit. *Lancet* 1993 Jul 31;342(8866):306. [Medline: [8101332](https://pubmed.ncbi.nlm.nih.gov/8101332/)]
7. Platon L, Amigues L, Ceballos P, Fegueur N, Daubin D, Besnard N, et al. A reappraisal of ICU and long-term outcome of allogeneic hematopoietic stem cell transplantation patients and reassessment of prognosis factors: results of a 5-year cohort study (2009-2013). *Bone Marrow Transplant* 2016 Feb;51(2):256-261. [doi: [10.1038/bmt.2015.269](https://doi.org/10.1038/bmt.2015.269)] [Medline: [26569092](https://pubmed.ncbi.nlm.nih.gov/26569092/)]
8. Johnson AEW, Mark RG. Real-time mortality prediction in the Intensive Care Unit//AMIA Annual Symposium Proceedings. 2018 Apr 16 Presented at: AMIA Annu Symp Proc. 2017; 2017; Washington p. 994-1003.

9. Shickel B, Loftus TJ, Ozrazgat-Baslanti L, Ebadi T, Bihorac A, Rashidi P. DeepSOFA: A Real-Time Continuous Acuity Score Framework using Deep Learning. ArXiv e-prints 2018 [[FREE Full text](#)]
10. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. *JAMA* 2018 Apr 03;319(13):1317-1318. [doi: [10.1001/jama.2017.18391](https://doi.org/10.1001/jama.2017.18391)] [Medline: [29532063](https://pubmed.ncbi.nlm.nih.gov/29532063/)]
11. Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, et al. Integrating machine learning and multiscale modeling-perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. *NPJ Digit Med* 2019;2:115 [[FREE Full text](#)] [doi: [10.1038/s41746-019-0193-y](https://doi.org/10.1038/s41746-019-0193-y)] [Medline: [31799423](https://pubmed.ncbi.nlm.nih.gov/31799423/)]
12. Harutyunyan H, Khachatryan H, Kale DC, Ver Steeg G, Galstyan A. Multitask learning and benchmarking with clinical time series data. *Sci Data* 2019 Jun 17;6(1):96 [[FREE Full text](#)] [doi: [10.1038/s41597-019-0103-9](https://doi.org/10.1038/s41597-019-0103-9)] [Medline: [31209213](https://pubmed.ncbi.nlm.nih.gov/31209213/)]
13. Kim MJ, Kim YH, Sol IS, Kim SY, Kim JD, Kim HY, et al. Serum anion gap at admission as a predictor of mortality in the pediatric intensive care unit. *Sci Rep* 2017 May 03;7(1):1456 [[FREE Full text](#)] [doi: [10.1038/s41598-017-01681-9](https://doi.org/10.1038/s41598-017-01681-9)] [Medline: [28469150](https://pubmed.ncbi.nlm.nih.gov/28469150/)]
14. Kim SY, Kim S, Cho J, Kim YS, Sol IS, Sung Y, et al. A deep learning model for real-time mortality prediction in critically ill children. *Crit Care* 2019 Aug 14;23(1):279 [[FREE Full text](#)] [doi: [10.1186/s13054-019-2561-z](https://doi.org/10.1186/s13054-019-2561-z)] [Medline: [31412949](https://pubmed.ncbi.nlm.nih.gov/31412949/)]
15. Brand L, Patel A, Singh I, Brand C. Real Time Mortality Risk Prediction: A Convolutional Neural Network Approach. In: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018). 2018 Presented at: Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies; January 19-21, 2018; Funchal, Madeira, Portugal p. 463-470. [doi: [10.5220/0006596204630470](https://doi.org/10.5220/0006596204630470)]
16. Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z. Berkay Celik, Ananthram Swami. Practical Black-Box Attacks against Machine Learning. In: McDaniel, I. Goodfellow, S. Jha, Z. B. Celik, A. Swami, Practical black-box attacks against machine learning, pp. 506?. New York: Association for Computing Machinery; 2017 Apr Presented at: ASIA CCS '17: ACM Asia Conference on Computer and Communications Security; April, 2017; Abu Dhabi United Arab Emirates p. 506-519. [doi: [10.1145/3052973.3053009](https://doi.org/10.1145/3052973.3053009)]
17. Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree//Advances in neural information processing systems. In: Advances in neural information processing systems. Red Hook, NY: Curran Associates Inc; 2017 Presented at: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing System; December 2017; Long Beach, CA, USA p. 3149-3157.
18. Bisong E. Introduction to Scikit-learn. In: Building Machine Learning and Deep Learning Models on Google Cloud Platform. Berkeley, CA: Apress; 2019:215-229.
19. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet* 2011 Aug 27;378(9793):815-825. [doi: [10.1016/S0140-6736\(11\)60814-3](https://doi.org/10.1016/S0140-6736(11)60814-3)] [Medline: [21872750](https://pubmed.ncbi.nlm.nih.gov/21872750/)]
20. Awad A, Bader-El-Den M, McNicholas J. Patient length of stay and mortality prediction: A survey. *Health Serv Manage Res* 2017 May;30(2):105-120. [doi: [10.1177/0951484817696212](https://doi.org/10.1177/0951484817696212)] [Medline: [28539083](https://pubmed.ncbi.nlm.nih.gov/28539083/)]
21. Baicker K, Goldman D. Patient cost-sharing and healthcare spending growth. *J Econ Perspect* 2011;25(2):47-68. [doi: [10.1257/jep.25.2.47](https://doi.org/10.1257/jep.25.2.47)] [Medline: [21595325](https://pubmed.ncbi.nlm.nih.gov/21595325/)]
22. Vincent J, Marshall JC, Namendys-Silva SA, François B, Martin-Loeches I, Lipman J, ICON investigators. Assessment of the worldwide burden of critical illness: the intensive care over nations (ICON) audit. *Lancet Respir Med* 2014 May;2(5):380-386. [doi: [10.1016/S2213-2600\(14\)70061-X](https://doi.org/10.1016/S2213-2600(14)70061-X)] [Medline: [24740011](https://pubmed.ncbi.nlm.nih.gov/24740011/)]
23. Machado FR, Cavalcanti AB, Bozza FA, Ferreira EM, Angotti Carrara FS, Sousa JL, SPREAD Investigators, Latin American Sepsis Institute Network. The epidemiology of sepsis in Brazilian intensive care units (the Sepsis PREvalence Assessment Database, SPREAD): an observational study. *Lancet Infect Dis* 2017 Nov;17(11):1180-1189. [doi: [10.1016/S1473-3099\(17\)30322-5](https://doi.org/10.1016/S1473-3099(17)30322-5)] [Medline: [28826588](https://pubmed.ncbi.nlm.nih.gov/28826588/)]
24. Saleh A, Ahmed M, Sultan I, Abdel-lateif A. Comparison of the mortality prediction of different ICU scoring systems (APACHE II and III, SAPS II, and SOFA) in a single-center ICU subpopulation with acute respiratory distress syndrome. *Egypt J Chest Dis Tuberc* 2015 Oct;64(4):843-848. [doi: [10.1016/j.ejcdt.2015.05.012](https://doi.org/10.1016/j.ejcdt.2015.05.012)]
25. Nagrebetsky A, Bittner EA. Missing Data and ICU Mortality Prediction: Gone But Not to Be Forgotten. *Crit Care Med* 2017 Dec;45(12):2108-2109. [doi: [10.1097/CCM.0000000000002780](https://doi.org/10.1097/CCM.0000000000002780)] [Medline: [29148991](https://pubmed.ncbi.nlm.nih.gov/29148991/)]
26. Yap X, Ng C, Hsu K, Chien C, Goh ZNL, Li C, et al. Predicting need for intensive care unit admission in adult emphysematous pyelonephritis patients at emergency departments: comparison of five scoring systems. *Sci Rep* 2019 Nov 12;9(1):16618 [[FREE Full text](#)] [doi: [10.1038/s41598-019-52989-7](https://doi.org/10.1038/s41598-019-52989-7)] [Medline: [31719593](https://pubmed.ncbi.nlm.nih.gov/31719593/)]
27. Tan TL, Tang YJ, Ching LJ, Abdullah N, Neoh H. Scientific reports 2018 Nov 12;8(1):16698 [[FREE Full text](#)] [doi: [10.1038/s41598-018-35144-6](https://doi.org/10.1038/s41598-018-35144-6)] [Medline: [30420768](https://pubmed.ncbi.nlm.nih.gov/30420768/)]
28. Lee H, Yoon S, Oh S, Shin J, Kim J, Jung C, et al. Comparison of APACHE IV with APACHE II, SAPS 3, MELD, MELD-Na, and CTP scores in predicting mortality after liver transplantation. *Sci Rep* 2017 Sep 07;7(1):10884 [[FREE Full text](#)] [doi: [10.1038/s41598-017-07797-2](https://doi.org/10.1038/s41598-017-07797-2)] [Medline: [28883401](https://pubmed.ncbi.nlm.nih.gov/28883401/)]

29. Cui Y, Wang T, Bao J, Tian Z, Lin Z, Chen D. Comparison of Charlson's weighted index of comorbidities with the chronic health score for the prediction of mortality in septic patients. *Chin Med J (Engl)* 2014;127(14):2623-2627. [Medline: [25043078](#)]
30. Nielsen AB, Thorsen-Meyer H, Belling K, Nielsen AP, Thomas CE, Chmura PJ, et al. Survival prediction in intensive-care units based on aggregation of long-term disease history and acute physiology: a retrospective study of the Danish National Patient Registry and electronic patient records. *Lancet Digit Health* 2019 Jun;1(2):e78-e89 [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30024-X](#)] [Medline: [33323232](#)]
31. Fei A, Lin Q, Liu J, Wang F, Wang H, Pan S. The relationship between coagulation abnormality and mortality in ICU patients: a prospective, observational study. *Sci Rep* 2015 Mar 23;5:9391 [FREE Full text] [doi: [10.1038/srep09391](#)] [Medline: [25797521](#)]
32. Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al. Deep Learning for Health Informatics. *IEEE J Biomed Health Inform* 2016 Dec 29;21(1):4-21. [doi: [10.1109/JBHI.2016.2636665](#)] [Medline: [28055930](#)]
33. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. *Sci Transl Med* 2015 Aug 05;7(299):299ra122. [doi: [10.1126/scitranslmed.aab3719](#)] [Medline: [26246167](#)]
34. Shickel B, Loftus TJ, Adhikari L, Ozrazgat-Baslanti T, Bihorac A, Rashidi P. DeepSOFA: A Continuous Acuity Score for Critically Ill Patients using Clinically Interpretable Deep Learning. *Sci Rep* 2019 Feb 12;9(1):1879 [FREE Full text] [doi: [10.1038/s41598-019-38491-0](#)] [Medline: [30755689](#)]
35. Davoodi R, Moradi MH. Mortality prediction in intensive care units (ICUs) using a deep rule-based fuzzy classifier. *J Biomed Inform* 2018 Mar;79:48-59 [FREE Full text] [doi: [10.1016/j.jbi.2018.02.008](#)] [Medline: [29471111](#)]
36. Norrie J. Mortality prediction in ICU: a methodological advance. *Lancet Respir Med* 2015 Jan;3(1):5-6. [doi: [10.1016/S2213-2600\(14\)70268-1](#)] [Medline: [25466334](#)]
37. Zhengping Che, Sanjay Purushotham, Robinder Khemani, Yan Liu. Interpretable deep models for icu outcome prediction. In: *AMIA Annu Symp Proc*. 2017 Feb 10 Presented at: Purushotham, R. Khemani, Y. Liu, Interpretable deep models for icu outcome prediction, *Amia Annu Symp Proc* () 371?380; 2016; Chicago p. 371-380.
38. The Lancet Respiratory Medicine. Opening the black box of machine learning. *Lancet Respir Med* 2018 Nov;6(11):801. [doi: [10.1016/S2213-2600\(18\)30425-9](#)] [Medline: [30343029](#)]
39. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019 May 13;1(5):206-215. [doi: [10.1038/s42256-019-0048-x](#)]
40. Blanco-Justicia A, Domingo-Ferrer J, Martínez S, Sánchez D. Machine learning explainability via microaggregation and shallow decision trees. *Knowl-Based Syst* 2020 Apr;194:105532. [doi: [10.1016/j.knosys.2020.105532](#)]

Abbreviations

APACHE: Acute Physiology and Chronic Health Evaluation
AUC: area under the curve
GDP: gross domestic product
ICU: intensive care unit
SAPS II: Simplified Acute Physiology Score II
SOFA: Sequential Organ Failure Assessment

Edited by C Lovis; submitted 26.08.20; peer-reviewed by B Qian, P Karsmakers, X Li; comments to author 24.10.20; revised version received 17.12.20; accepted 25.01.21; published 25.03.21

Please cite as:

Jiang H, Su L, Wang H, Li D, Zhao C, Hong N, Long Y, Zhu W

Noninvasive Real-Time Mortality Prediction in Intensive Care Units Based on Gradient Boosting Method: Model Development and Validation Study

JMIR Med Inform 2021;9(3):e23888

URL: <https://medinform.jmir.org/2021/3/e23888>

doi: [10.2196/23888](#)

PMID: [33764311](#)

©Huizhen Jiang, Longxiang Su, Hao Wang, Dongkai Li, Congpu Zhao, Na Hong, Yun Long, Weiguo Zhu. Originally published in *JMIR Medical Informatics* (<http://medinform.jmir.org>), 25.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly

cited. The complete bibliographic information, a link to the original publication on <http://medinform.jmir.org/>, as well as this copyright and license information must be included.