Development and validation of a prognostic risk score system for COVID-19 inpatients: A multi-center retrospective study in China

Ye Yuan (yye@hust.edu.cn)
Huazhong University of Science and Technology
Chuan Sun
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Xiuchan Tang
2School of Mechanical Science and Engineering, Huazhong University of Science and Technology
Cheng Cheng
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Laurent Mombaerts
Luxembourg Centre for System Biomedicine, University of Luxembourg
Maolin Wang
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Tao Hu
.Department of Emergency, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology
Chenyu Sun
AMITA Health Saint Joseph Hospital Chicago
Yuqi Guo
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Xiuting Li
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Hui Xu
Department of Anesthesiology, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology
Tongxin Ren
Huazhong University of Science and Technology-Wuxi Research Institute
Yang Xiao
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Yaru Xiao
Department of Emergency, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology

Hongling Zhu
Division of Cardiology, Department of Internal Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology
Chuming Chen
Department of Biomedical Engineering, Southern University of Science and Technology
Yingxia Liu
Department of Biomedical Engineering, Southern University of Science and Technology
Zhichao Liang
.Department of Infectious Diseases, Shenzhen Key Laboratory of Pathogen and Immunity, State Key Discipline of Infectious Disease, The Third People's Hospital of Shenzhen, Second Hospital Affiliated to Southern University of Science and Technology
Zhiguo Cao
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Hai-Tao Zhang
School of Artificial Intelligence and Automation, Huazhong University of Science and Technology
Ioannis Ch. Paschaldis
Department of Electrical and Computer Engineering, Division of Systems Engineering, and Department of Biomedical Engineering, Boston University
Quanying Liu
.Department of Infectious Diseases, Shenzhen Key Laboratory of Pathogen and Immunity, State Key Discipline of Infectious Disease, The Third People's Hospital of Shenzhen, Second Hospital Affiliated to Southern University of Science and Technology
Jorge Goncalves
Department of Plant Sciences, University of Cambridge
Qiang Zhong
Department of Emergency, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology
Li Yan
Department of Emergency, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology

---

Research Article

Keywords: COVID-19, Risk Score, Mortality Prediction

Posted Date: July 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-41151/v1

Version of Record: A version of this preprint was published at Engineering on November 1st, 2020. See the published version at https://doi.org/10.1016/j.eng.2020.10.013.

Abstract

Coronavirus Disease 2019 (COVID-19) has become a world-wide pandemic. Hospitalized patients of COVID-19 suffer from a high mortality rate, motivating the development of convenient and practical methods for clinicians to promptly identify high-risk patients. Here we developed a risk score using clinical data from 1,479 inpatients admitted to Tongji Hospital, Wuhan, China (development cohort) and externally validated with data from two other centers: 141 inpatients from Jinyintan Hospital in Wuhan (validation cohort 1) and 432 inpatients from the Third People's Hospital Shenzhen (validation cohort 2). The risk score is based on three biomarkers readily available in routine blood samples and can be easily translated into a probability of death. The risk score can predict the mortality of individual patients more than 12 days in advance with more than 90% accuracy across all cohorts. Moreover, the Kaplan-Meier score shows that patients upon admission can clearly be differenciated into low, medium or high risk, with an AUC score of 0.9551. In summary, a simple risk score was validated to predict death in patients infected with COVID-19 and was validated in independent cohorts.

Introduction

The outbreak of Coronavirus Disease 2019 (COVID–19), a disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV–2), began in early December, 20191,2. As of June 24, 2020, more than 8 million individuals have been confirmed to be COVID–19 positive globally, with an overall mortality rate of more than 5%3. Among these patients, some developed pneumonia, and even progressed into severe acute respiratory failure (ARDS) rapidly with a very poor prognosis and even higher mortality4,5. In addition to pneumonia and ARDS, SARS-CoV–2 also leads to damage to other organs and systems, such as large-vessel strokes6. In a retrospective cohort study from China, 26% of the hospitalized patients required intensive care unit (ICU) care7. In New York, among 2,634 patients who were discharged or died, 14.2% were treated in the ICU, and 12.2% received invasive mechanical ventilation8. In Italy, among critically ill patients, almost all of them required respiratory support, and nearly nine in ten of these critically ill patients needed endotracheal intubation9. Despite all these efforts, the mortality remained high7–9. In the process of caring for COVID–19 patients, particularly the critically ill, healthcare providers are subjected to a deluge of lab results for an increasing number of hospitalized patients. It is arduous to identify the most important information for decision making, especially in urgent or emergent situations. It is therefore imperative to identify risk factors and parameters to build an accurate prognostic model for early intervention and management.

Artificial intelligence (AI) technologies have had a surprising effectiveness in the medical domain, with a performance exceeding that of humans, especially for many image classification tasks10–12. Several AI-based researches have been conducted and shown promising results in addressing the challenges to control and predict COVID–19 spread and death toll13–19. Interpretable AI-based models (e.g., tree models) can enhance the confidence of medical professionals by helping them understand machine

decisions. Inspired by the interpretability properties of decision-trees, our previous work19 successfully identified three laboratory features from common blood tests that can accurately predict the mortality of patients with COVID−19. It has been demonstrated that particular laboratory features, including lymphopenia, lactate dehydrogenase (LDH), inflammatory markers (e.g., C-reactive protein [CRP], ferritin), D-dimer (>1 mcg/mL), prothrombin time (PT), troponin, and creatine phosphokinase (CPK), are associated with poor outcomes7,20,21. Older age has also been shown to be associated with increased mortality8,22–24.

In this study, we built an AI model that can generate real-time risk scores and help identify patients with a higher risk of mortality before becoming critically ill, and allowing prompt early intervention. In addition, our scores allow clinicians to monitor the disease progression and adjust therapies accordingly.

Results

Patients' characteristics and outcomes

A total of 1,479 COVID−19 patients were eligible and their relevant clinical information was collected and analyzed. Clinical characteristics, epidemiological history, symptoms on set, outcomes and results of lab tests were all included (see Table 1). The median age was 62, and no significant differences between genders were found (male 50.9% vs. female 49.1%). The majority of patients (71.9%) were local residents of Wuhan. In addition, 8.3 % of the 1,479 patients were familial clusters and 3.9% had an history of close contact. Notably, 21.6% patients had no known history of close contact or exposure, indicating the existence of other untraced transmission routes. COVID−19 patients exhibited variable clinical symptoms: 72.5% of patients manifested a fever, followed by symptoms of respiratory infection, such as cough (35.7%), shortness of breath (9.5%), fatigue(5.5%). In addition, gastrointestinal and neurological symptoms existed. Patients had complained of more than one symptom at a time. Tongji Hospital received a large number of severe and critical patients and, as a consequence, the mortality rate was high at 17.4% at earlier stage. In contrast, the Third People's Hospital in Shenzhen had only 4 deaths out of a total of 432 patients. Hence, most of the focus of the paper is on Tongji and Jinyintan Hospitals. Figures for the Third People's Hospital Shenzhen are mostly in Supplementary Information.

Model development and performance

The risk of mortality for individual patients was predicted with the following simple and explainable model from Logistic Regression (LR) developed in the Methods Section:

Risk score = 0.00850×LDH+0.0204×hs-CRP−0.150×Lymphocytes(%)−2.30, Probability of death = $\boldsymbol{\sigma}$(Risk Score),

where *LDH, hs-CRP,* and *Lymphocytes* are input predictors of the LR model and $\boldsymbol{\sigma}$ is the sigmoid function, i.e.,

$$\sigma(x)=\frac{1}{13e^{5x}}.$$

To simplify the use of the model in a clinical setting, Supplementary Tables S1 to S4 include a lookup tables for quickly computing the risk score and the probability of death for a patient. Because different patients had different admission dates and various lengths of stay, the predictive performance was evaluated backwards in time, i.e. as a function of the number of days between the blood sample and the eventual outcome (death or discharge). Its predictability is illustrated in Figure 1 and Supplementary Figure S4. The model achieved more than 95% (90%, 98%) cumulative AUC value for 20 days in advance for Tongji hospital (Jinyintan Hospital, the Third People's Hospital Shenzhen).

Figure 2 and Supplementary Figure 6 plot the distributions of scores for survived and deceased patients and the probabilies of death, using measurements that were taken within 10 days of the outcome. The risk score clearly separates blood samples of survived and deceased patients in all datasets, including both external validation datasets that were not used in model development. From a particular blood sample, a physician can easily calculate the probability of death; the higher the score, the higher this probability and risk for a patient.

Probability of death as a function of the risk score. The model (red curve) follows almost perfectly the probability of death (blue) calculated directly from the data.

Validation of the risk score

Next, risk score can be used to categorize patients to different risk groups upon admission, shown in the Kaplan-Meier Survival Curve (Figure 3). We applied the risk scores of patients at admission, and classified patients to three groups according to their scores: low-risk group (65.6%), intermediate-risk group (5.9%), and high-risk group (28.5%). It was observed that, in the development cohort, the 30-day mortality rates for low-, intermediate- and high- risk groups were 1.8%, 12.5% and 53.7%, respectively, showing a significant difference in the mortality rate. In the external validation cohort 2, the 30-day mortality rates for low-, intermediate- and high- risk groups were shown in Supplementary Figure S3. These results demonstrated that, the risk score could also be used to predict the mortality for individual patients as early as at patients' admission.

Comparison with other standard scores

The score from the proposed model was compared with the scores of other well-used models reported previously, such as qSOFA, CURB 65 and CRB 65 in both development and external validation cohorts. The minimal requirement for different scores is 829 patients with available measurements in the development cohort. As shown in Figure 3, the AUC for the scores of our model, CRB 65, CURB 65 and qSOFA were 0.9551, 0.7393, 0.8130, and 0.7480,

respectively. The ROC and AUC for the external validation dataset are shown in Supplementary Figures S1 and S2. It can be seen that the proposed score system is better than standard score systems for predicting the outcome of patients with COVID–19.

Discussion

The proportion of critical or fatal cases is quite high among hospitalized COVID–19 patients8,35. Although the mortality rate is only 1.4–2.3% based on large-scale epidemiological studies5, about one in three to four hospitalized patients were admitted to ICU4,8,35,36, and 71% to 97.3% of the critically ill patients eventually needed respiratory support8,35–37, while 15% of the ICU patients required extracorporeal membrane oxygenation (ECMO)4. Despite many practices, including respiratory support, different medication regimens, and even ECMO, the case-fatality rate for these critically ill patients was still very high4,8,35–37. Retrospective studies have suggested that the onset of dyspnea was relatively late (median 6.5 days after symptoms onset), but the progression to ARDS could be swift thereafter (median 2.5 days after onset of dyspnea) among patients who developed critical illness36–38. In addition, the high mortality rate in Wuhan in the early outbreak, and in some other areas around the world, exceeded the capacity of local medical resources. These suggest that it is critical to promptly identify patients who are likely to have poor prognosis and higher risk of becoming critically ill.

Although COVID–19 is a multifaceted disease with uncertainty surrounding effective treatments and wide variation in clinical course and prognosis, multiple laboratory features, including lymphopenia, LDH, inflammatory markers, D-dimer, PT, troponin, and CPK, are associated with poor prognosis27,28,39. Our study demonstrates that risk of death among patients with COVID–19 is predictable using a risk score computed from only three common clinical blood samples: LDH, hs-CRP and Lymphocyte (%). As shown in Supplementary Figure S5, these three predictors provided a good separation between survived and deceased patients, in blood samples taken within 10 of patients' outcome. Front-line clinicians can monitor the disease progression of a patient by applying the proposed risk score to available blood samples. This provides monitoring and screening out high risk patients in real time and as laboratory data become available. Overall, the model serves an accurate indicator for early detection and intervention to reduce the mortality rate, and can potentially monitor the progression of the disease to effectively review and adjust clinical management by healthcare providers.

The significance of our work is five-fold. First, the model may identify high- risk patients early enough and provide them with alternative therapies such as using appropriate respiratory support and other treatments as soon as possible. Second, the model is not based on thresholds and instead provides a continuous probability of death. Thresholds are useful on extreme values, but can be misleading when risk scores are near the thresholds. Instead, probabilities of outcomes provide a level of confidence in the prediction. Third, it provides a simple formula to precisely and quickly quantify the risk of death from just three features of a blood sample. Fourth, the three key features can be conveniently collected at any hospital, even in areas where healthcare resources are limited. The features are objective and quantitative, and therefore avoid any bias of subjective clinical judgements. Last but not least, our

research has been constructed using Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)40 guidance with internal and external validation datasets from multiple centers, and the validation of our model has been confirmed by two cohorts of patients from different hospitals. There are, however, several limitations of the model. First, the patients in the development cohort were from Tongji Hospital, and most were severe and critical. Hence, the cohort may not accurately represent patients with asymptomatic or mild or moderate cases of COVID–19 and the samples could have selection bias. Second, we did not model the effects of different therapies since treatments were not controlled and varied from patient to patient. Finally, this study provided evidence that the risk score could help clinicians determine early intervention for patients with COVID–19 in three Chinese hopitals. We require further investigation and validation involving other hospitals and countries. In particular, it is possible that different hospitals have distinct laboratory, therapies and discharging protocols and that these may affect blood samples and, as a consequence, the interpretation of the risk score.

In conclusion, a simple prognostic risk score system was developed based on a logistic regression classifier to predict death risk for COVID–19 patients, and was validated wth independent cohorts from multiple centers. This risk score system may help healthcare providers to promptly identify patients with poor prognosis and initiate appropriate intervention early to improve the prognosis.

Methods

Study design and support:

The study was approved by the Tongji Hospital Ethics Committee.

Data Resources:

Two separate cohorts of COVID–19 patients were used for model development and validation. The electronic medical records of 1,479 hospitalized COVID–19 cases admitted to Tongji Hospital in Wuhan, China, from January 10th to March 8th, 2020, were used to train the model. On the other hand, electronic medical records of 141 inpatients from Jinyintan Hospital in Wuhan, and 432 inpatients from the Third People's Hospital Shenzhen were used to validate the model. Epidemiological, demographic, clinical, laboratory, medications, nursing record, and outcome data from electronic medical records were extracted. Data monitoring and recording were performed in the same way for both cohorts. The clinical outcomes were followed up to March 8th, 2020, as shown in Table 1.

The diagnosis of COVID–19 patients were based on the following diagnostic criteria from the National Health and Health Commission of China25: 1) SARS- CoV–2 nucleic acid positive in respiratory or blood samples detected by RT- PCR; 2) high homology between virus sequence detected in respiratory or blood samples and the known sequence of SARS-CoV–2.

Development of an AI-based risk score system

A LR classifier was applied to train the model to fit the outcome from three predictors, including LDH, high-sensitivity-CRP (hs-CRP) and Lymphocyte (%), which were automatically chosen in previous study19. These factors have also been frequently observed as key risk factors for COVID–19 patients26–28. All patients' measurements collected within 10 days to their definite outcomes were used for model training. The output classes were defined as the outcome of the patient, either death or survival, after ICU time. The LR model aims at predicting the risk groups of hospitalized patients (low-, intermediate-, and high-risk) according to different levels of their risk scores: 0–30 defined as low-risk, 30–50 as intermediate-risk and 50–100 as high-risk, in the development cohort and validation cohort.

Performance assessment and comparison

Our score system was benchmarked against several state-of-the-art models developed using other machine learning approaches and standard metrics were used to quantify the performance of different models. The area under the curve (AUC) of the receiver operating characteristic (ROC) curve at one specific day was used to evaluate model effectiveness. Further, the associated cumulative AUC score29 was also introduced as a time-dependent measure to evaluate the risk of mortality for individual patients, computed backwards in time from the day of discharge or death. The performance of our system was also compared with those of other standard models, such as qSOFA (quick Sequential [Sepsis-related] Organ Failure Assessment), CURB 65 and CRB 65 in both development and validation cohorts30–34.

Declarations

Acknowledgments Funding: L. Y. is supported by special fund for novel coronavirus pneumonia from the department of science and technology of Hubei province (2020FCA035). **Author Contributions**: Y. Y. conceived the study; L. Y., Q. Z. collected data; Y. Y., C. S., and X. T. discovered the model; C. S., L. Y., H. Z., Y. X., L. M., H. X., Y. P., J. G., and Y. Y. drafted the manuscript; all authors provided critical review of the manuscript and approved the final draft for publication. **Competing interests**: The authors declare no competing interests.**Data and code availability**: The code implementation is available from Y. Y.

The need for consent that the participate/patient consented to participate and/or publish was waived by the approving ethics committee.

References

1. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A Novel Coronavirus Emerging in China - Key Questions for Impact Assessment. The New England journal of medicine 2020;382:692-4.

2. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. Lancet 2020;395:470-3.
3. Coronavirus disease (COVID-2019) Situation Report 155, WHO, 23 June
4. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020;395:497-506.
5. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. The New England journal of medicine 2020;382:1708-20.
6. Oxley TJ, Mocco J, Majidi S, et al. Large-Vessel Stroke as a Presenting Feature of Covid-19 in the Young. The New England journal of medicine
7. Zhou F, Yu T, Du R, et Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054-62.
8. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. Jama
9. Grasselli G, Zangrillo A, Zanella A, et al. Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy. Jama 2020;323:1574-81.
10. Dunnmon JA, Yi D, Langlotz CP, Re C, Rubin DL, Lungren Assessment of Convolutional Neural Networks for Automated Classification of Chest Radiographs. Radiology 2019;290:537-44.
11. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med 2019;25:24-9.
12. Zhu H, Cheng C, Yin H, et al. Automatic Multi-Label ECG Diagnosis of Impulse or Conduction Abnormalities in Patients with Deep Learning Algorithm: A Cohort Study. Lancet Digital Health
13. Ienca M, Vayena On the responsible use of digital data to tackle the COVID- 19 pandemic. Nat Med 2020;26:463-4.
14. Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine
15. Santosh KC. AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data. J Med Syst 2020;44:93.
16. Ayyoubzadeh SM, Ayyoubzadeh SM. Predicting COVID-19 Incidence Through Analysis of Google Trends Data in Iran: Data Mining and Deep Learning Pilot Study. JMIR public health and surveillance 2020;6:e18828.
17. Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health J Thorac Dis 2020;12:165-74.
18. Liu F, Zhang Q, Huang C, et CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients. Theranostics 2020;10:5613-22.
19. Yan L, Zhang H-T, Goncalves J, et An interpretable mortality prediction model for COVID-19 patients. Nature Machine Intelligence 2020. https://doi.org/10.1038/s42256-020-0180-7

20. Wu C, Chen X, Cai Y, Xia Ja, Song Y. Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. JAMA Internal Medicine
21. Shi S, Qin M, Shen B, et al. Association of Cardiac Injury With Mortality in Hospitalized Patients With COVID-19 in Wuhan, China. JAMA cardiology
22. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. Jama
23. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. Jama
24. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. MMWR Morb Mortal Wkly Rep 2020;69:343-6.
25. Commision CNH. Diagnosis and treatment of novel coronavirus pneumonia in China (trial version 7)
26. Ruan Q, Yang K, Wang W, Jiang L, Song Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med 2020:1-3.
27. Mo P, Xing Y, Xiao Y, et al. Clinical characteristics of refractory COVID-19 pneumonia in Wuhan, China. Clin Infect Dis
28. Livesey A, Garty F, Shipman AR, Shipman KE. Lactate dehydrogenase in dermatology practice. Clin Exp Dermatol
29. Lambert J, Chevret S. Summary measure of discrimination in survival models based on cumulative/dynamic time-dependent ROC curves. Stat Methods Med Res 2016;25:2088-102.
30. Singer M, Deutschman CS, Seymour CW, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). Jama 2016;315:801-10.
31. Chen YX, Wang JY, Guo SB. Use of CRB-65 and quick Sepsis-related Organ Failure Assessment to predict site of care and mortality in pneumonia patients in the emergency department: a retrospective study. Crit Care 2016;20:167.
32. Ranzani OT, Prina E, Menéndez R, et al. New Sepsis Definition (Sepsis-3) and Community-acquired Pneumonia Mortality. A Validation and Clinical Decision- Making Study. Am J Respir Crit Care Med 2017;196:1287-97.
33. Ferreira M, Blin T, Collercandy N, et al. Critically ill SARS-CoV-2-infected patients are not stratified as sepsis by the Annals of intensive care 2020;10:43.
34. Santana AR, Amorim FF, Soares FB, Godoy LDS, Almeida… Comparison of CURB-65 and CRB-65 as predictors of death in community-acquired pneumonia in adults admitted to an ICU. Critical Care 2013;17.
35. Myers LC, Parodi SM, Escobar GJ, Liu VX. Characteristics of Hospitalized Adults With COVID-19 in an Integrated Health Care System in Jama 2020.

36. Wang D HB, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, Jama the Journal of the American Medical Association 2020.
37. Xu Z, Shi L, Wang Y, et al. Pathological findings of COVID-19 associated with acute respiratory distress syndrome. The Lancet Respiratory medicine 2020;8:420-2.
38. Yang X, Yu Y, Xu J, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. The Lancet Respiratory medicine
39. Ruan Q, Yang K, Wang W, Jiang L, Song Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China. Intensive Care Med 2020:1-3.
40. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. Br J Surg 2015;102:148-58.

Tables

Table 1 Clinical features of the studied patients.

Characteristics	Overall
Age, years, Median(Q1, Q3)	62.00 (48.50,70.00)
Sex ,n (%)	
Male	753(50.9)
Female	726(49.1)
Epidemiological history， n(%)	
Wuhan residents	1063(71.9)
Contact with confirm or suspected patients	57(3.9)
Familial cluster	123(8.3)
Health worker	8(0.5)
Contact with HUANAN SEAFOOD MARKET	7(0.5)
Undefined contact history	320(21.6)
Symptoms onset, n(%)	
Myalgia or arthralgia	11(0.7)
Fatigue	82(5.5)
Diarrhea	46(3.1)
Abdominal pain	4(0.3)
Headache	4(0.3)
Chest pain	7(0.5)
Sore throat	12(0.8)
Shortness of breath	141(9.5)
Coma	1(0.1)
Fever	1072(72.5)

Cough	528(35.7)
Palpitation	3(0.2)
Asymptomatic	43(2.9)
Outcomes, n(%)	
Survival rate	1222(82.6)
Mortality rate	257(17.4)
Lab test	
Lactate dehydrogenase, Median(Q1, Q3),U/L	209.00 (176.00,289.50)
Lymphocytes, Median(Q1, Q3),%	24.65 (15.00,32.20)
High sensitive C-reactive protein, Median(Q1, Q3),mg/L	3.60 (1.10,27.50)
Leukocytes, Median(Q1, Q3), ×109/L	5.84 (4.72,7.87)
Eosinophils, Median(Q1, Q3), ×109/L	0.08 (0.02,0.14)
Basophils, Median(Q1, Q3), ×109/L	0.02 (0.01,0.03)
Neutrophils, Median(Q1, Q3), ×109/L	3.64 (2.66,5.51)
Lymphocytes, Median(Q1, Q3), ×109/L	1.34 (0.88,1.76)
Monocytes, Median(Q1, Q3), ×109/L	0.48 (0.36,0.61)
Erythrocytes, Median(Q1, Q3), × 1012/L	4.02 (3.61,4.44)
Thrombocytes, Median(Q1, Q3), ×109/L	213.00 (159.00,275.75)
Alanine aminotransferase, Median(Q1, Q3), U/L	24.00 (15.00,39.00)
Aspartate transaminase, Median(Q1, Q3),U/L	22.00 (17.00,32.00)
Albumin, Median(Q1, Q3), g/L	36.10 (32.10,39.20)
Total bilirubin, Median(Q1, Q3), μmol/L	8.60 (6.40,12.40)

Serum creatinine, Median(Q1, Q3), μmol/L	69.00 (57.00,85.00)
Blood urine nitrogen, Median(Q1, Q3),mmol/L	4.50 (3.54,6.00)
Sodium, Median(Q1, Q3), mmol/L	140.40 (138.40,142.20)
Chlorine, Median(Q1, Q3), mmol/L	101.90 (99.70,104.00)
Potassium, Median(Q1, Q3), mmol/L	4.34 (4.01,4.69)

Figures

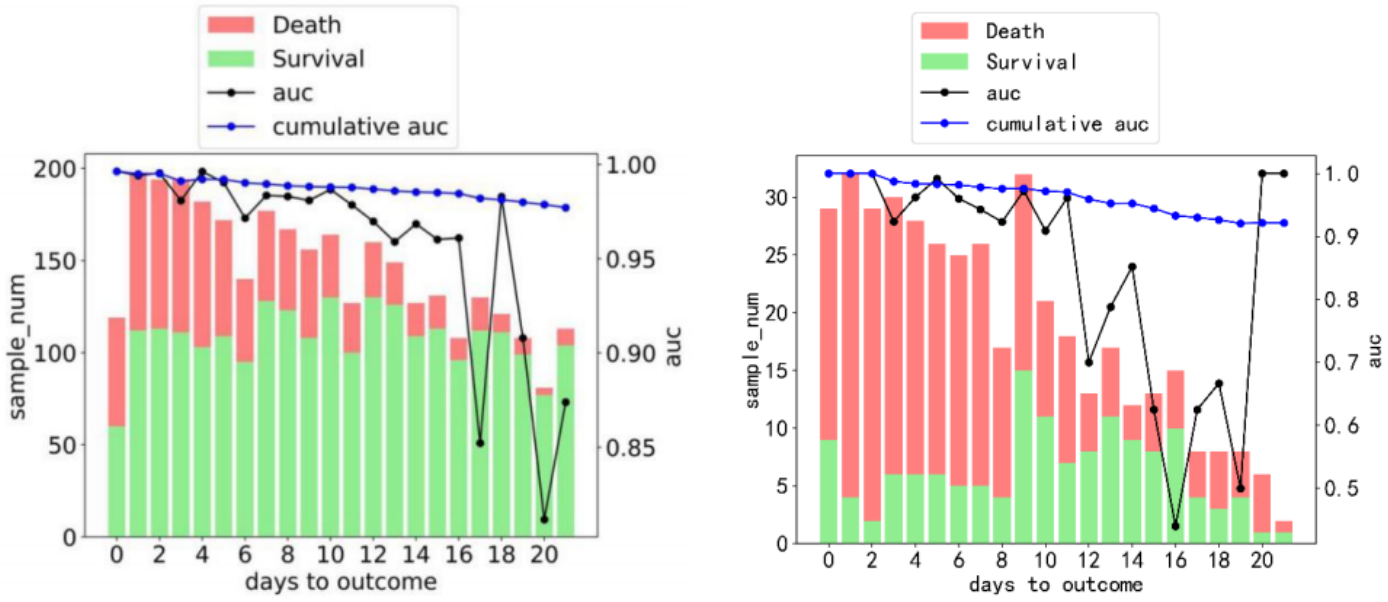

Figure 1

The performance of the proposed model (AUC Score and cumulative AUC Score29) as a function of the number of days until the outcome for all patients in the development cohort (Tongji Hospital, left) and external validation cohort 1 (Jinyintan Hospital, right).

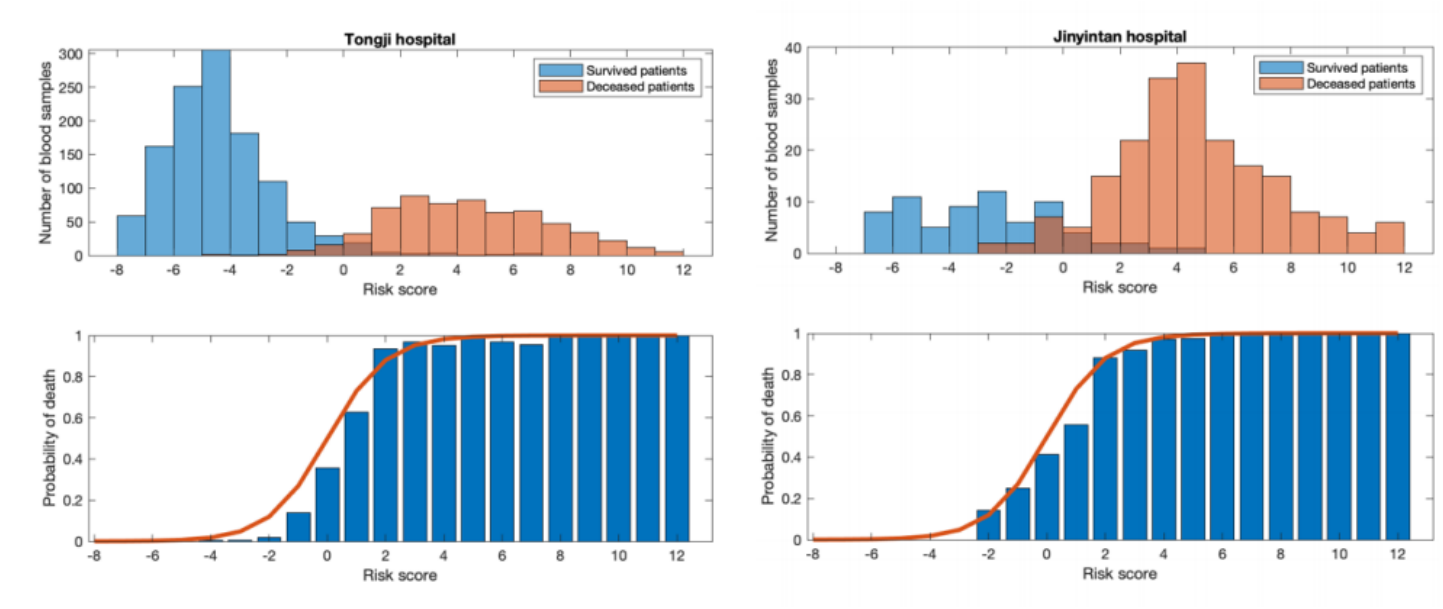

Figure 2

(Top) Distributions of scores for survived and deceased patients for Tongji hospital (left) and the Third People's Hospital Shenzhen (right), from blood samples taken within 10 days from patients' outcome. (Bottom) Probability of death as a function of the risk score. The model (red curve) follows almost perfectly the probability of death (blue) calculated directly from the data.

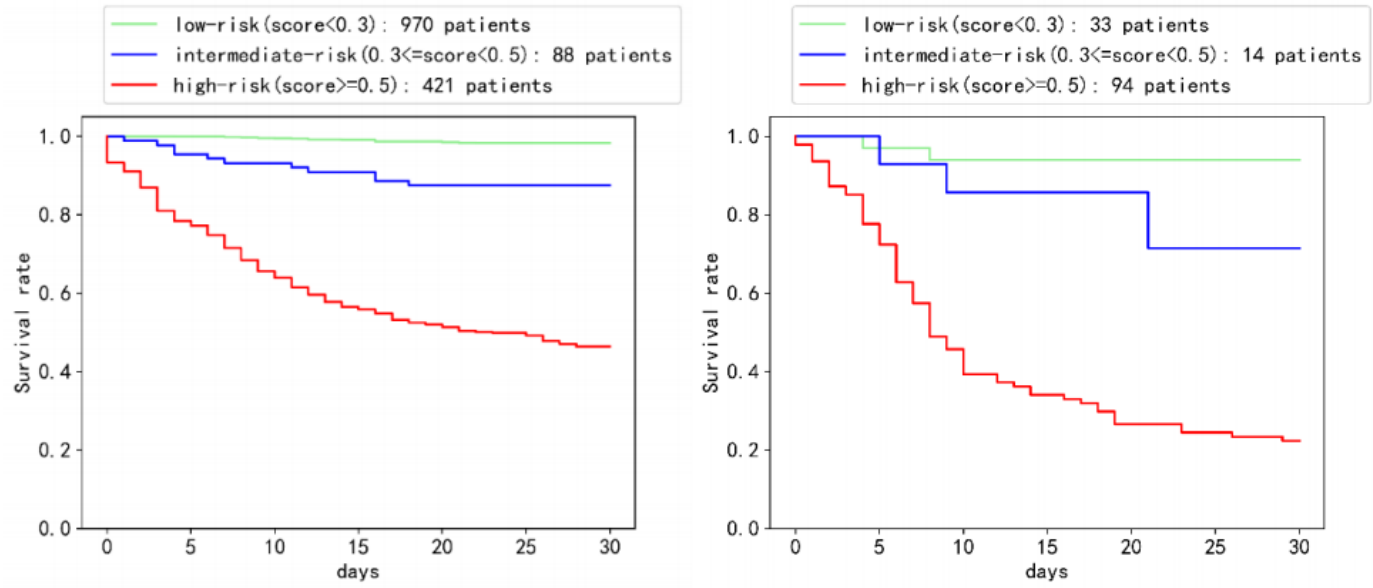

Figure 3

Kaplan-Meier Survival Curve for the development cohort and external validation cohort 1. Left: in the development cohort, the 30-day mortality rates for low-, intermediate- and high- risk groups were 1.8%, 12.5% and 53.7%, respectively, showing a significant difference in the mortality rate. Right: in the external

validation cohort 1, 23.4% of the patients were in the low- risk group, 9.9% of the patients were in the intermediate-risk group, and 66.7% of the patients were in the high-risk group.

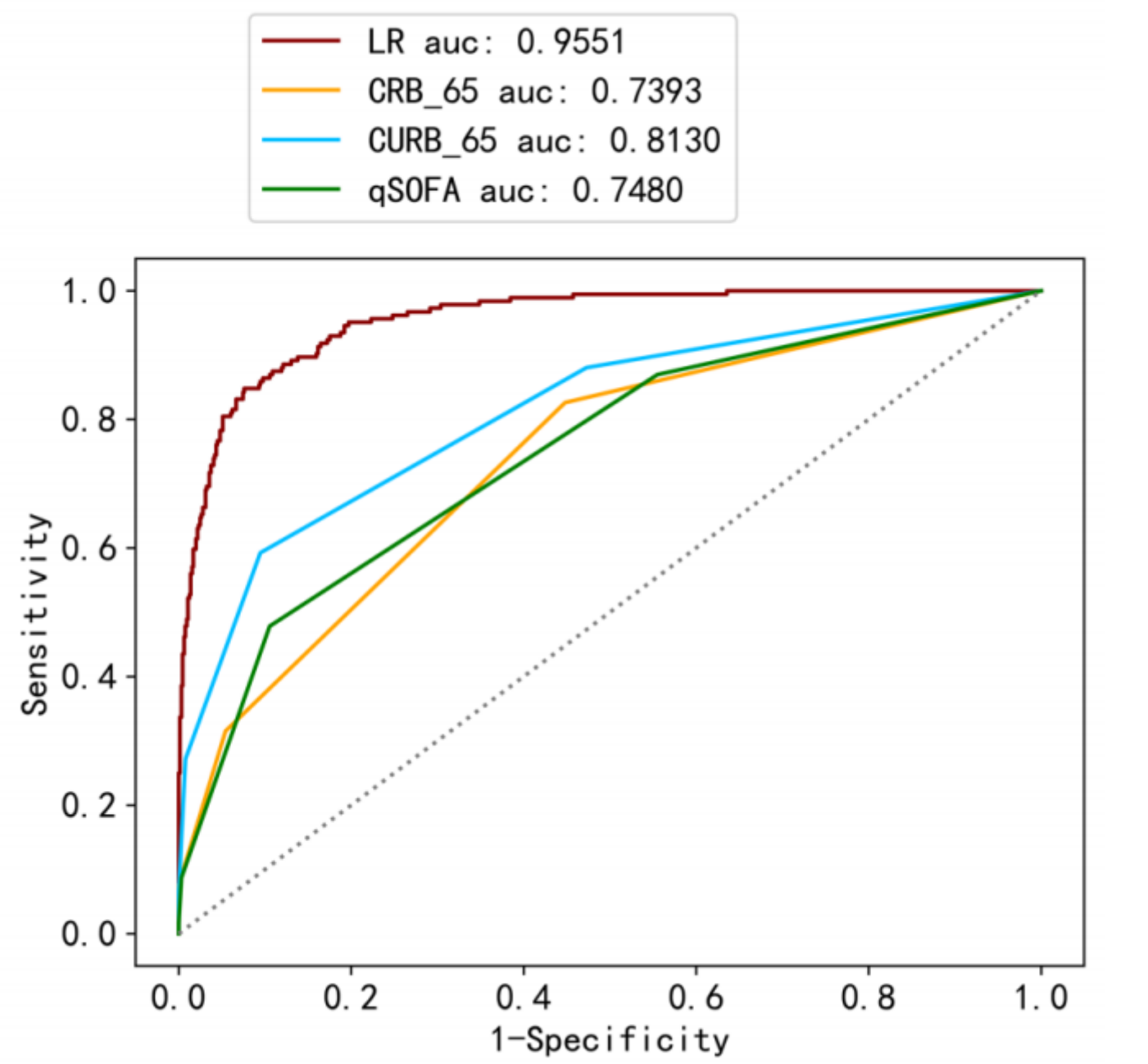

Figure 4

Comparative analysis of ROC of different scoring systems in the 829 patients from the development cohort who had available measurements at admission (minimal requirement for different scores) shows that the proposed model has a larger AUC than other models reported previously.

Supplementary Files

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

- SupplementalMaterial.pdf