Retrospective Study

Clinical diagnosis of severe COVID-19: A derivation and validation of a prediction rule

Ming Tang, Xia-Xia Yu, Jia Huang, Jun-Ling Gao, Fu-Lan Cen, Qi Xiao, Shou-Zhi Fu, Yang Yang, Bo Xiong, Yong-Jun Pan, Ying-Xia Liu, Yong-Wen Feng, Jin-Xiu Li, Yong Liu

ORCID numbers: Ming Tang 0000-0002-1238-0755; Xia-Xia Yu 0000-0003-4811-5125; Jia Huang 0000-0002-7874-1672; Jun-Ling Gao 0000-0001-9174-1709; Fu-Lan Cen 0000-0003-0475-059X; Qi Xiao 0000-0002-6758-1941; Shou-Zhi Fu 0000-0001-7221-4815; Yang Yang 0000-0001-8490-7387; Bo Xiong 0000-0002-1286-3597; Yong-Jun Pan 0000-0003-1387-7162; Ying-Xia Liu 0000-0002-5059-6290; Yong-Wen Feng 0000-0003-4485-3534; Jin-Xiu Li 0000-0002-5959-2042; Yong Liu 0000-0002-5208-888X.

Author contributions: Huang J, Cen FL, Xiao Q, Fu SZ, Yang Y, Xiong B and Pan YJ contributed data curation; Tang M, Liu Y and Yu XX contributed formal analysis; Liu YX, Feng YW and Li JX contributed investigation; Tang M, Liu Y and Yu XX contributed methodology; Liu Y, Feng YW and Li JX contributed project administration; Cen FL, Liu YX and Feng YW contributed resources; Yu XX contributed software; Li JX contributed supervision; Gao JL contributed validation; Tang M wrote the original draft; all authors have read and approved the final manuscript.

Institutional review board statement: This study was approved by the Institutional...
Abstract

BACKGROUND

The widespread coronavirus disease 2019 (COVID-19) has led to high morbidity and mortality. Therefore, early risk identification of critically ill patients remains crucial.

AIM

To develop predictive rules at the time of admission to identify COVID-19 patients who might require intensive care unit (ICU) care.

METHODS

This retrospective study included a total of 361 patients with confirmed COVID-19 by reverse transcription-polymerase chain reaction between January 19, 2020, and March 14, 2020 in Shenzhen Third People’s Hospital. Multivariate logistic regression was applied to develop the predictive model. The performance of the predictive model was externally validated and evaluated based on a dataset involving 126 patients from the Wuhan Asia General Hospital between December 2019 and March 2020, by area under the receiver operating curve (AUROC), goodness-of-fit and the performance matrix including the sensitivity, specificity, and precision. A nomogram was also used to visualize the model.

RESULTS

Among the patients in the derivation and validation datasets, 38 and 9 participants (10.5% and 2.54%, respectively) developed severe COVID-19, respectively. In univariate analysis, 21 parameters such as age, sex (male), smoker, body mass index (BMI), time from onset to admission (> 5 d), asthenia, dry cough, expectoration, shortness of breath, asthenia, and Rox index < 18 (pulse oxygen saturation, SpO2)/FiO2 × respiratory rate, RR) showed positive correlations with severe COVID-19. In multivariate logistic regression analysis, only six parameters including BMI [odds ratio (OR) 3.939; 95% confidence interval (CI): 1.409-11.015; P = 0.017], fever [OR (6.794; 95% CI: 1.401-32.951; P = 0.016], Charlson index [OR (2.917; 95% CI: 1.279-6.654; P = 0.011)], PaO2/FiO2 ratio [OR (17.570; 95% CI: 1.117-276.383; P = 0.041], and neutrophil/lymphocyte ratio [OR (3.574; 95% CI: 1.048-12.191; P = 0.042] were found to be independent predictors of COVID-19. These factors were found to be significant risk factors for severe patients confirmed with COVID-19. The AUROC was 0.941 (95% CI: 0.901-0.981) and 0.936 (95% CI: 0.886-0.987) in both datasets. The calibration properties were good.

CONCLUSION

The proposed predictive model had great potential in severity prediction of COVID-19 in the ICU. It assisted the ICU clinicians in making timely decisions for the target population.

Key Words: COVID-19; Communicable diseases; Clinical decision rules; Prognosis; Nomograms

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: This study established a risk-prediction model to estimate the prognosis of patients with coronavirus disease 2019 (COVID-19) for clinicians to more objectively calculate the severity of an individual patient and optimize the subsequent treatment. The model with the aforementioned six predictors could be used in clinical practice to identify high-risk dialysis patients for more investigations and interventions. This study provided a simple form of a probability prediction model to identify patients with severe COVID-19.

Citation: Tang M, Yu XX, Huang J, Gao JL, Cen FL, Xiao Q, Fu SZ, Yang Y, Xiong B, Pan YJ, Liu YX, Feng YW, Li JX, Liu Y. Clinical diagnosis of severe COVID-19: A derivation and validation of a prediction rule. World J Clin Cases 2021; 9(13): 2994-3007
INTRODUCTION

In December 2019, an outbreak of coronavirus disease 2019 (COVID-19) occurred in Wuhan, China[1,2]. Up to April 18, 2020, there were 83251 cases reported in China[3] and 1442312 internationally[4]. The estimated case-fatality rate associated with COVID-19 in Wuhan might be as high as 5.25% [95% confidence interval (CI): 4.98-5.51][5,6].

Most of the patients diagnosed with COVID-19 had mild disease[5,6]. For these patients, general isolation and symptomatic treatment are available, and intensive care unit (ICU) care is needed to reduce the number of deaths and alleviate the shortage of medical resources unless their condition worsened rapidly. In such a case, risk stratification of patients by severity of the illness at an early stage is necessary.

Predictive models or scoring systems have been developed to quantify the impact of risk factors. Complex statistical models were introduced during the model development step, and these could be used by clinicians for decision-making. Also, the risk of severity in patients could be easily estimated and monitored over time, thus motivating them to modify their behavior. Several scoring systems that targeted the severity prediction have already been developed. However, most of these proposed models did not include a description of the study population, and the calibration of predictions is rarely assessed[7]. Liang et al[8] developed a risk score based on the characteristics of COVID-19 patients at the time of admission to the hospital based on a nationwide cohort in China. However, this risk scoring system consisted of 10 items, which were relatively complex and time-consuming. Besides, hemoptysis occurred rarely. Moreover, this system was not compared with the other traditional acute respiratory distress syndrome (ARDS) scoring systems, and the calibration of this system has not been evaluated.

Hence, in this study, a total of 361 patients diagnosed with COVID-19 in the Shenzhen Third People’s Hospital were included and a rule-based on risk-prediction model for the severity of COVID-19 patients was developed. Moreover, the developed predicted model was also externally validated, and its performance was carefully calibrated.

MATERIALS AND METHODS

Patient recruitment and study design

This multicenter retrospective cohort study included 487 adult patients with COVID-19 (aged ≥ 18 years) as confirmed by the reverse transcription-polymerase chain reaction (RT-PCR) between January 19, 2020, and March 14, 2020, in Shenzhen Third People’s Hospital and the Wuhan Asia General Hospital. Ethical approval was obtained from the institutional review board (study number: RC2020-102) of these two hospitals. The exclusion criteria were as follows: Patients aged less than 18 years, who have undergone surgeries recently, or had a cerebral hemorrhage, myocardial infarction, and other diseases. Also patients with incomplete medical data or who refused to participate in this study were excluded. All patients who presented with COVID-19 between December 2019 and March 2020 to the Wuhan Asia General Hospital were amalgamated to form an external validation dataset. The study design is shown in Figure 1.

Variables

Independent variables included sociodemographic factors, clinical symptoms, comorbidity, travel and contact history, and laboratory tests. The outcome variables were whether patients were defined as severe or critical. Patients were defined as severe if they met any of the following criteria[9]: (1) Shortness of breath, more than 30 times/min; (2) Oxygen saturation at rest of less than 93%; and (3) Arterial partial pressure of oxygen/fraction of inspired oxygen (FiO2) ≤ 300 mmHg. Cases were defined as critical if they met any of the following criteria: (1) Respiratory failure and requiring mechanical ventilation; (2) Shock; and (3) With any other organ failure that required ICU care.
Data analysis and predictive rule development

Logistic regression was applied to identify the independent factors that were associated with critical COVID-19 patients. The following parameters were included for analysis: Age ≥ 60 years, sex, body mass index (BMI), illness onset to admission, dry cough, upper respiratory tract symptoms (any sign of stuffy nose or sore throat), productive sputum, fever, shortness of breath, fatigue, and Charlson index. Laboratory results (presence or otherwise in the first 3 d), including thrombocytopenia platelet count, PaO$_2$/FiO$_2$ (P/F) ratio, neutrophil/lymphocyte ratio (NLR), and alanine aminotransferase/aspartate aminotransferase (ALT/AST) ratio, were also identified for each patient. Stepwise multivariate regression was used to select adjusted independent predictors ($P < 0.05$ was considered significant).

The predicted probabilities provided by the proposed predictive model were used to identify the patients’ risk of illness severity\cite{10}. To ease the use of the model and retain the effects of each variable, the fitted regression coefficients were rescaled by dividing their values by the smallest fitted coefficient and rounded to the nearest integer. For each binary risk factor, the optimal condition was defined as the absence of the factor. Zero points were assigned to patients who were in good condition under the corresponding risk factor. The sum of these scores represented the overall risk of the patient.

The 10-fold cross-validations were included to internally validate the performance of the newly proposed predictive model. The data were randomly split into 10 approximately equal sizes, in which 9 were used to develop the model and one was used for internal validation. Furthermore, the newly developed predictive model was externally validated by the data collected from the Wuhan Asia General Hospital.

Validation sample and model performance

In addition to area under the receiver operating curve (AUROC), the predictive performance of the proposed model was measured by its sensitivity, specificity, and precision. The goodness-of-fit test was applied by comparing the observed and predicted events of COVID-19 using the risk group deciles by the Hosmer-Lemeshow $\chi^2$ test.

RESULTS

Patient demographics

A total of 421 patients were admitted to the Shenzhen Third People’s Hospital during the study period. Sixty patients were excluded due to incomplete medical records at

\begin{figure}
\centering
\includegraphics[width=\textwidth]{flowchart.png}
\caption{Flowchart of the study selection process.}
\end{figure}
the time of the study, and their age was less than 18 years. A population of 361 patients was included in this analysis, and randomly separated into 4 groups. Three groups were used as the training set (262 patients), while the remaining 1 group was used as an internal testing set (99 patients). In the training set, 224 (62.0%) patients had mild and 38 (10.5%) had severe disease. The age of patients with mild disease was 48.5 [interquartile range (IQR), 36.0-59.0] years; 54 (24.1%) were seniors (≥ 60 years); and 106 (47.3%) were males. However, the median age of patients with severe disease was 64 (IQR, 57.5-66.0) years; 26 (68.4%) were seniors; and 29 (76.3%) were males. Similarly, in the testing set, 90 (25.8%) patients had mild and 9 (2.5%) patients had severe disease. The median age of patients with mild disease was 43.0 (IQR, 34.0-54.8) years; 14 (15.6%) were seniors (≥ 60 years); and 38 (42.2%) were males. However, the median age of patients with severe disease was 64.0 (IQR, 62.0-69.0) years; 8 (88.9%) were seniors; and 6 (66.7%) were males. More details of these patients are shown in Table 1.

**Independent risk factor evaluation**

In univariate analysis, 21 parameters such as age, sex (male), smoker, BMI, time from onset to admission (> 5 d), asthenia, dry cough, expectoration, shortness of breath, asthenia, and Rox index < 18 (pulse oxygen saturation, \( \text{SpO}_2 \)/(FiO\_2 \times \text{respiratory rate}, \text{RR}) showed positive correlations with severe COVID-19.

In multivariate logistic regression analysis, only six parameters including BMI [odds ratio (OR) 3.939; 95%CI: 1.409-11.015; \( P = 0.009 \)], time from onset to admission (≥ 5 d) (OR 7.107; 95%CI: 1.449-34.849; \( P = 0.016 \)), fever (OR 6.794; 95%CI: 1.401-32.951; \( P = 0.017 \)), Charlson index (OR 2.917; 95%CI: 1.279-6.654; \( P = 0.011 \)), \( P/F \) ratio (OR 17.570; 95%CI: 1.117-276.383; \( P = 0.041 \)), and NLR (OR 3.574; 95%CI: 1.048-12.191; \( P = 0.042 \)) were found to be independent predictors for COVID-19 (Table 2).

**Final model**

Based on the results of multivariable modeling, a final prediction rule was developed using the categorical versions of the retained variables. The clinical prediction rule points were assigned after standardization to the lowest regression coefficient. The nomogram was established based on the features selected by logistic regression, which were used for predicting severe COVID-19 (Figure 2). For easy application, the nomogram was further simplified by point scores. The following point scores were used to measure the magnitude of the association of each predictor with COVID-19: BMI, 1 point if more than 24 and 2 for 27; time from onset to admission (≥ 5 d), 2 points; fever (≥ 37.3°C), 2 points; Charlson index, 1 point; \( P/F \) score ≤ 300 mmHg, 6 points; and NLR, 1 point if more than 2 and 2 if more than 4 (Table 3). The maximum total risk score was 17, and this was most likely to be severe for any given COVID-19 patient. The cutoff score was selected using the Youden index\(^{[11]}\).

**Model calibration**

The calibration curves showed that the predicted rates agreed with that of the actual results observed in the cohort. According to the Hosmer-Lemeshow \( \chi^2 \) test, the adequacy of this prediction model was shown to be good (\( P = 0.901 \)) (Figure 3).

**Performance of the predictive model and comparison with CURB-65**

The predicted probability of severe COVID-19 was calculated as follows: LOGIT (\( p \)) = 3.939 × BMI + 7.107 × Time from onset to admission + 6.794 × Fever + 2.917 × Charlson index + 17.547 × \( P/F \) + 3.574 × NL. Proposed score rules were also generated (Table 4).

The model showed high potential for predicting severe COVID-19 [area under the curve (AUC) 0.941 (95%CI: 0.901-0.981) for internal validation and 0.936 (0.886-0.987) for external validation]. The precision, sensitivity, and specificity for internal validation were 0.270 (0.234-0.307), 0.930 (0.896-0.964), and 0.622 (0.602-0.641), respectively. The values for external validation were 0.413 (0.261-0.563), 0.927 (0.873-0.980), and 0.804 (0.728-0.880), respectively (Table 5). The proposed predictive model performed better when compared with the CURB-65 models.

Using the 10-fold cross-validation, the proposed model was further compared with the CURB-65 models. The results showed that the proposed model had higher AUC (0.941), and sensitivity and specificity when compared with that of the CURB-65 models (AUC 0.704) (Table 5).
### Table 1: Patient demographics and clinical characteristics in the training and internal validation sets

|                      | Training sets                                      |                      | Internal testing sets                                 |                      |
|----------------------|----------------------------------------------------|----------------------|-------------------------------------------------------|----------------------|
|                      | Mild, n = 224 (62.0%)                              | Severe, n = 38 (10.5%) | Mild, n = 90 (25.8%)                                  | Severe, n = 9 (2.5%)  |
| Age (yr)             | 48.5 (56.0-59.0)                                   | 64 (57.5-66.0)       | 43.0 (34.0-54.8)                                     | 64.0 (62.0-69.0)      |
| ≥ 60 yr              | 54 (24.1)                                          | 26 (68.4)            | 0.000                                                | 14 (15.6)             |
|                      |                                                    |                      |                                                       | 8 (88.9)              | 0.000                |
| Sex                  | 0.002                                              |                      |                                                       |                      | 0.159                |
| Male                 | 106 (47.3)                                         | 29 (76.3)            | 38 (42.2)                                            | 6 (66.7)              |
| Female               | 118 (52.7)                                         | 9 (23.7)             | 52 (57.8)                                            | 3 (33.3)              |
| BMI > 24             | 57 (25.4)                                          | 11 (28.9)            | 0.000                                                | 22 (24.4)             |
|                      |                                                    |                      |                                                       | 4 (44.4)              | 0.024                |
| BMI > 27             | 21 (9.4)                                           | 14 (36.8)            | 5 (5.6)                                              | 2 (22.2)              |
| Onset to hospitalization (> 5 d) | 62 (27.7)                  | 25 (65.8)            | 0.000                                                | 34 (37.8)             |
|                      |                                                    |                      |                                                       | 5 (55.6)              | 0.000                |
| Symptoms             |                                                    |                      |                                                       |                      |
| Fever > 37.3°C       | 172 (76.8)                                         | 29 (76.3)            | 0.000                                                | 78 (86.7)             |
| Fever > 39°C         | 8 (3.6)                                            | 7 (18.4)             | 1 (1.1)                                              | 2 (22.2)              |
| Charlson index > 0   | 72 (32.2)                                          | 33 (86.8)            | 0.000                                                | 39 (43.3)             |
|                      |                                                    |                      |                                                       | 8 (88.9)              | 0.000                |
| P/F                  | 13 (5.8)                                           | 16 (42.1)            | 0.000                                                | 3 (3.3)               |
|                      |                                                    |                      |                                                       | 3 (33.3)              | 0.006                |
| NLR > 2              | 80 (35.7)                                          | 15 (39.5)            | 0.000                                                | 39 (43.3)             |
|                      |                                                    |                      |                                                       | 3 (33.3)              | 0.009                |
| NLR > 4              | 24 (10.7)                                          | 17 (44.7)            | 7 (7.8)                                              | 4 (44.4)              |
| Score ≥ 7            | 31 (13.8)                                          | 35 (92.1)            | 0.000                                                | 13 (13.1)             |
|                      |                                                    |                      |                                                       | 9 (100)               | 0.000                |

BMI: Body mass index; P/F: PaO$_2$/FiO$_2$; NLR: Neutrophil/lymphocyte ratio.

### DISCUSSION

With the ongoing COVID-19 pandemic and a severe shortage of ICU resources, it is critical for first-line clinicians to promptly predict the prognosis of patients with COVID-19 and optimize the undermined medical sources. As COVID-19 is a new pandemic and its clinical features are not yet fully illustrated, clinicians are more prone to estimate the severity of individual patients through their subjective experience. To solve this critical problem, our study established a risk-prediction model to estimate the prognosis of patients with severe COVID-19, in order that clinicians can more objectively calculate the severity of an individual patient and optimize the subsequent treatment. This risk-prediction model was based on the data obtained from 361 patients with severe COVID-19, in which 262 were in the derivation set, 99 in the internal validation set, and 126 in the external validation set. The model demonstrated high accuracy in predicting the severity of COVID-19 with an AUC of 0.948 when compared with other models.

Among dozens of clinical parameters affecting the prognosis of patients with COVID-19, many indices were shown to be significantly correlated with the severity of the disease. However, only six clinical and laboratory characteristics were identified as independent factors that could distinguish patients with severe COVID-19 from those with mild and moderate COVID-19. Older age (> 60 years) alone could not predict the prognosis independently, whereas comorbidity together with older age led to a worse prognosis.

The present model demonstrated that the P/F ratio was the highest risk factor for predicting the worst prognosis. Patients with lower oxygenation index at the time of admission had 16 times higher risk of developing severe COVID-19. Previously, the guidelines have recommended early identification of patients with severe pneumonia based on a lower oxygenation index\[^9,12\]. As the clinical classification is primarily based on the oxygenation index, it is not surprising that the oxygenation index is predictive for the development of severe COVID-19. The P/F ratio showed a close relationship with the Qsp/Qt index, and a low P/F ratio indicates pulmonary parenchymal injury caused by COVID virus infection or a subsequent immunological response\[^13,14\].
Another important finding of the present study was the benefit of timely admission after the onset of COVID-19 symptoms, which was associated with a six times higher risk of COVID-19 infection. This finding was in agreement with the China Ministry of Health Guidelines\[12\] which advocate early hospitalization, but different from the other practices in many other countries which did not mention early admission at all or emphasize home quarantine\[15,16\]. Patients with mild and moderate COVID-19 were encouraged to stay at home to prevent the surge of COVID-19 and hospital collapse until they were critically ill\[17-20\]. The present study indicated that early admission could not only help avoid local transmission but also reduce the incidence of severity and criticality of COVID-19.

Patients with fever (> 37.3°C) are more likely to progress to severe COVID-19 when compared with those without fever. Fever generally implies an ongoing immunological response to severe acute respiratory syndrome CoV-2 (SARS-CoV-2) virus. Higher body temperature might indicate a stronger response due to severe pathogenic infection or overreaction to the SARS-CoV-2 virus in some individual cases. As indicated by the research in China and Europe, the infection storm induced by cytokines and other immune molecules might cause deterioration of clinical symptoms and pulmonary damage. Unfortunately, the exact role of immunological response in the etiology of SARS-CoV-2 virus, whose pathological features are sometimes unpredictable and novel to clinicians, is currently unclear. Although SARS-CoV-2 and SARS are both caused by coronavirus, the implication of fever in patients with COVID-19 was not in concordance with that in patients with SARS. Patients who did not have fever during admission were more likely to die from SARS (44% vs 7%, \(P = 0.045\)\[21\]). The physiological mechanism still remains unclear, and further studies are warranted to elucidate this mechanism. Nevertheless, the aforementioned finding is important as fever can easily be screened and assist in the early identification of severe COVID-19.

The proposed model suggested that the prognosis was partially determined by the patients’ general condition, such as the Charlson index and BMI. The Charlson index has been used to predict the long-term prognosis of patients with comorbidities. It has a specific scoring system that can also include age when predicting the lethality, and compare with other indices of comorbid disease\[22\]. This might give Charlson index a better predictive power, as advancing age is usually associated with chronological deterioration of physical functions. In contrast, the proposed model suggested that

### Table 2 Demographic clinical characteristics of external validation sets

|                      | Mild, \(n = 110\) (%) | Severe, \(n = 16\) | \(P\) value |
|----------------------|-----------------------|--------------------|-------------|
| **Age (yr)**         | 49.0 (35.0-57.8)      | 70.0 (64.0-78.0)   |             |
| \(> 60\)             | 27 (24.6)             | 12 (75.0)          | 0.000       |
| **Sex**              |                       |                    | 0.007       |
| Male                 | 51 (46.4)             | 13 (81.3)          |             |
| Female               | 59 (53.6)             | 3 (18.7)           |             |
| **BMI > 24**         | 30 (27.3)             | 8 (50.0)           | 0.021       |
| \(> 27\)             | 15 (13.6)             | 4 (25.0)           |             |
| **Onset to hospitalization (> 5 d)** | 37 (33.6) | 15 (94.0) | 0.000       |
| **Symptoms**         |                       |                    |             |
| Fever > 37.3°C       | 87 (79.1)             | 12 (75.0)          | 0.025       |
| Fever > 39°C         | 4 (3.6)               | 3 (18.8)           |             |
| Charlson index > 0   | 54 (49.1)             | 15 (93.8)          | 0.000       |
| P/F                  | 1 (0.9)               | 6 (37.5)           | 0.000       |
| NLR > 2              | 44 (40.0)             | 2 (12.5)           | 0.000       |
| NLR > 4              | 6 (5.5)               | 14 (87.5)          |             |
| Score ≥ 7            | 23 (20.9)             | 15 (93.8)          | 0.000       |

BMI: Body mass index; P/F: PaO\(_2\)/FiO\(_2\); NLR: Neutrophil/lymphocyte ratio.
Table 3 Risk factors of severe patients with coronavirus disease 2019

| Variable                        | Univariate analysis | Multivariate analysis |
|---------------------------------|---------------------|-----------------------|
| Age ≥ 60 yr                     | 6.821 (3.224-14.431) | -                     |
| Sex                             | 3.587 (1.624-7.924)  | -                     |
| BMI                             | 2.701 (1.727-4.224)  | 3.939 (1.409-11.015)  |
| Illness onset at admission      | 5.025 (2.418-10.441) | 7.107 (1.449-34.849)  |
| Smoker                          | 4.060 (1.912-8.621)  | -                     |
| Dry cough                       | 3.395 (1.604-7.184)  | -                     |
| Productive sputum               | 2.612 (1.284-5.312)  | -                     |
| Fever                           | 4.615 (1.948-10.935) | 6.794 (1.401-32.951)  |
| Shortness of breath             | 3.312 (1.161-9.448)  | -                     |
| Fatigue                         | 3.450 (1.661-7.164)  | -                     |
| Charlson index                  | 2.371 (1.813-3.100)  | 2.917 (1.279-6.654)   |
| PaO2/FIO2                       | 11.804 (5.028-27.714) | 17.570 (1.117-276.832) |
| Platelet count                  | 0.423 (0.180-0.992)  | -                     |
| NLR                             | 3.766 (2.271-6.245)  | 3.574 (1.048-12.191)  |
| D-Dimer                         | 4.385 (2.102-9.148)  | -                     |
| ESR                             | 11.025 (2.583-47.048) | -                     |
| hs-CRP                          | 4.314 (2.359-7.887)  | -                     |
| AST                             | 6.332 (2.753-14.563) | -                     |
| Cr                              | 16.258 (3.994-66.188) | -                     |
| IL-6                            | 2.346 (1.224-4.494)  | -                     |
| PCT                             | 3.961 (1.062-14.772) | -                     |
| ROX                             | 0.237 (0.111-0.505)  | -                     |

BMI: Body mass index; NLR: Neutrophil/lymphocyte ratio; ESR: Erythrocyte sedimentation rate; CRP: C-reactive protein; AST: Aspartate aminotransferase; IL: Interleukin; PCT: Procalcitonin.

advanced age (> 60 years) alone was not enough to independently predict the prognosis of COVID-19, but the existing comorbidity might better predict the development of severe COVID-19 according to the data in the present study. This finding was consistent with that of previous reports which showed that comorbidity is an important independent predictor of mortality in COVID-19[6,23-25]. However, the study did not define the role of comorbidity quantitatively. Our study findings are novel in evaluating the relationship of comorbidity in patients with severe COVID-19 with Charlson index.

BMI (> 24) was associated with a 2.94 times higher risk of severe COVID-19. Higher BMI increases the risk of a variety of diseases, including cardiovascular diseases, diabetes, and other metabolism disorders, subsequently increasing the mortality rate, according to several large-cohort studies[26,27]. This was in agreement with that of a previous study, which reported on the role of obesity in predicting patients’ COVID-19 severity[28]. A retrospective study proved that obesity was a risk factor for the severity of infections caused by other respiratory viruses including influenza[29] and Middle East respiratory syndrome coronavirus[30].

According to the proposed model, D-dimer had poor predictive ability, which contradicted previous findings[29]. However, prospective studies are needed to further explore the link between abnormal D-dimer levels and the severity and mortality of patients with COVID-19[9,12].

The present study demonstrated that the NLR indicated a 2.57 times higher risk of the development of severe COVID-19. The NLR acts as a biomarker for assessing the severity of bacterial infections and the prognosis of patients with pneumonia and
Table 4 The proposed score rules for clinical diagnosis of severe coronavirus disease 2019

| Variables                        | Description | Score |
|----------------------------------|-------------|-------|
| BMI                              | ≥ 24        | 1     |
|                                 | ≥ 27        | 2     |
| Time from onset to admission     | ≥ 5 d       | 2     |
| Fever                            | ≥ 37.3 °C   | 2     |
|                                 | ≥ 39 °C     | 4     |
| Charlson index                   |             | 1     |
| P/F score                        | < 300       | 6     |
| NLR                              | ≥ 2         | 1     |
|                                 | ≥ 4         | 2     |

BMI: Body mass index; P/F: PaO$_2$/FiO$_2$; NLR: Neutrophil/lymphocyte ratio.

Table 5 Comparison of predictive performance between the proposed method and CURB-65

|                          | Proposed model | CURB-65 |
|--------------------------|----------------|---------|
| Internal validation      |                |         |
| AUROC                    | 0.941 (0.901-0.981) | 0.704 (0.564-0.863) |
| Precision                | 0.270 (0.234-0.307)  | 1.000   |
| Sensitivity              | 0.930 (0.896-0.964)  | 0.156 (0.085-0.228) |
| Specificity              | 0.622 (0.602-0.641)  | 1.000   |
| External validation      |                |         |
| AUROC                    | 0.936 (0.886-0.987)  | 0.654 (0.518-0.791) |
| Precision                | 0.413 (0.261-0.565)  | 0.481 (0.254-0.708) |
| Sensitivity              | 0.927 (0.873-0.980)  | 0.292 (0.186-0.399) |
| Specificity              | 0.804 (0.728-0.880)  | 0.954 (0.925-0.983) |

AUROC: Area under the receiver operating curve.

tumors$^{[31,32]}$. An increase in the NLR indicates a poor clinical prognosis$^{[33]}$. These results indicate that COVID-19 might affect a series of immune responses, including changes in T lymphocytes, serving as an important factor for predicting clinical deterioration. In addition, patients with severe illness might be associated with bacterial infection due to deteriorated immune function. The results of the present study were consistent with those of previous findings. For example, the study by Wang et al$^{[23]}$ followed up the laboratory test results of patients with COVID-19 and found that 5 patients had an increased neutrophil count and reduced lymphocyte count before their death$^{[23]}$. Liu et al$^{[34]}$ suggested that patients aged ≥ 50 years and a NLR ≥ 3.13 were at high risk and should be promptly transferred to the ICU for invasive respiratory support$^{[34]}$.

The model with the aforementioned six predictors could be used in clinical practice to identify high-risk dialysis patients. A simple form of a probability prediction model was put forward to identify patients with severe COVID-19. The cutoff values of the probabilities that discriminated between patients with mild and severe COVID-19 are provided in Table 3. According to the Youden index, a cutoff value of ($P = 0.439$) yielded a sensitivity and a specificity of 0.950 and 0.259, respectively, which was optimal to identify patients with severe COVID-19 when compared with other potential models.

Several models have previously been developed for predicting the prognosis of pneumonia. CURB-65 is a relatively widely accepted model$^{[35]}$. The CURB-65 score is comprised of 5 separate elements: Confusion, uremia, respiratory rate, blood pressure, and age ≥ 65 years. This was initially developed by the British Thoracic Society in 1987.
as the CRB criteria\cite{}, and was later modified to CURB-65 and validated by Lim et al\cite{37} in an international, multicenter, derivation/validation study. This score was evaluated for predicting mortality if the case was identified as having severe pneumonia. On the contrary, CURB-65 is not exclusively designed for COVID-19 patients. Thus, some features in CURB-65 were invaluable for stratifying COVID-19. For example, confusion, urea level > 7 mmol/L, and low blood pressure were rarely observed at the time of admission, contributing little to the identification of severe COVID-19. The risk score developed by Liang et al\cite{8} based on the characteristics of patients with COVID-19 at the time of admission consisted of 10 items, which is relatively complex and time-consuming. Besides, hemoptysis rarely occurs. Moreover, our model did not compare with the other traditional ARDS score systems, and the calibration of this system was not evaluated. Based on the clinical data, the proposed model calculated not only the factors of comorbidity information and BMI, but also special factors associated with COVID-19, such as the NLR, showing significantly improved performance.

The proposed predictive model, including the aforementioned 6 factors, assists in predicting the prognosis with high accuracy. It was promising in clinical decision-making to identify patients in the early stage with severe COVID-19. The model had high accuracy when tested using internal and external data from different places. However, the fitness of the model in other ethnicities has not yet been validated and is subjective to adjustment according to local data.

However, the present study had several limitations that undermine the prediction of the proposed prognostic model. The study included two of the largest hospitals in China, yet the model might not be generalizable to patients with COVID-19 in other

---

**Figure 2** Nomogram. BMI: Body mass index; PF: PaO$_2$/FiO$_2$; NL: Neutrophil/lymphocyte.

**Figure 3** Calibration and discrimination plot of the coronavirus disease 2019 critical prediction model.
hospitals. This was the largest COVID-19 study so far. Still, the findings were limited by the small sample size and unequal distribution of COVID-19 patients between the hospitals.

**CONCLUSION**

In summary, the proposed risk-prediction model for COVID-19 in this study showed significant association with the presence of obesity, delayed admission, fever, existing comorbidities, hypoxemia, and higher NLR. The model accurately predicted the severity of patients with COVID-19. This could practically help clinicians to perform timely interventions and control measures to prevent overcrowding or delayed diagnosis of severe COVID-19 in hospital settings. The predictive utility of the model warrants further validation and improvement. Other prospective studies can assist in developing a better model based on the present findings.

**ARTICLE HIGHLIGHTS**

**Research background**

The outbreak of coronavirus disease 2019 (COVID-19) led to high mortality and the intensive care unit (ICU) is needed to reduce the number of deaths. This causes a shortage of medical resources, especially in ICU settings.

**Research motivation**

The goal of this study was to develop and externally validate prediction rules to risk stratify patients by severity at an early stage.

**Research objectives**

The study aimed to risk stratify patients by severity of the illness at an early stage to better microallocate limited medical resources.

**Research methods**

This multicenter retrospective cohort study included 487 adult patients with confirmed COVID-19 between January 19, 2020, and March 14, 2020, in Shenzhen Third People’s Hospital and the Wuhan Asia General Hospital. Independent variables included sociodemographic factors, clinical symptoms, comorbidity, travel and contact history, and laboratory tests. The outcome variables were whether patients were defined as severe or critical. Logistic regression was applied to identify the independent factors that were associated with critical COVID-19 patients. Stepwise multivariate regression was used to select adjusted independent predictors ($P < 0.05$ was considered significant). The 10-fold cross-validations were included to internally validate the performance of the newly proposed predictive model. The data were randomly split into 10 approximately equal sizes, in which 9 were used to develop the model and one was used for internal validation. Furthermore, the newly developed predictive model was externally validated by the data collected from the Wuhan Asia General Hospital. In addition to area under the receiver operating curve, the predictive performance of the proposed model was measured by its sensitivity, specificity, and precision. The goodness-of-fit test was applied by comparing the observed and predicted events of COVID-19 using the risk group deciles by the Hosmer-Lemeshow $\chi^2$ test.

**Research results**

The model with the aforementioned six predictors could be used in clinical practice to identify high-risk dialysis patients. A simple form of a probability prediction model was put forward to identify patients with severe COVID-19. The proposed predictive model, including the aforementioned 6 factors, assists in predicting the prognosis with high accuracy.

**Research conclusions**

The proposed risk-prediction model for COVID-19 in this study showed significant association with the presence of obesity, delayed admission, fever, existing comorbidities, hypoxemia, and higher neutrophil/lymphocyte ratio. The model
accurately predicted the severity of patients with COVID-19. This could practically help clinicians perform timely interventions and control measures to prevent overcrowding or delayed diagnosis of severe COVID-19 in hospital settings.

**Research perspectives**

The proposed predictive model had great potential in severity prediction of COVID-19 in the ICU. It assisted the ICU clinicians in making timely decisions for the target population.

**REFERENCES**

1. Liu H, Stratton CW, Tang YW. Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *J Med Virol* 2020; 92: 401-402 [PMID: 31950516 DOI: 10.1002/jmv.25678]

2. Hui DS, I Azhar E, Madani TA, Ntoumi F, Kock R, Dar O, Ippolito G, Mchugh TD, Memish ZA, Drosten C, Zumla A, Petersen E. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* 2020; 91: 264-266 [PMID: 31953166 DOI: 10.1016/j.ijid.2020.01.009]

3. Online dashboard for tracking the spread of the coronavirus outbreak in China. 2020 May 18 [cited 18 April 2020]. In: [Tencent Internet]. [about one screen]. Available from: https://news.qq.com/zt2020/page/feiyan.htm/

4. *Coronavirus Resource Center*. Coronavirus 2019-nCoV Global Cases by Johns Hopkins CSSE. 2020 April 18 [cited 18 April 2020]. In: Johns Hopkins CSSE [Internet]. [about one screen]. Available from: https://coronavirus.jhu.edu/map.html

5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395: 497-506 [PMID: 31986264 DOI: 10.1016/s0140-6736(20)30183-5]

6. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Hu Y, Qiu Y, Wang J, Liu Y, Wei Y, Xie J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020; 395: 507-513 [PMID: 32007143 DOI: 10.1016/s0140-6736(20)30211-7]

7. Wynants L, Van Calster B, Collins GS, Riley RD, Heinzé G, Schuit E, Bonten MJM, Damen JAA, Debray TPA, De Vos M, Dhiman P, Haller MC, Harhay MO, Henckaerts L, Kreuzberger N, Lohman A, Luijken K, Ma J, Andauer CL, Reitsma JB, Sergeant JC, Shi C, Skoetz N, Smits LJ, Snell KJE, Sperrin M, Spijkers R, Stereyerberg EW, Takada T, van Kuijk SMJ, van Royen FS, Wallisch C, Hooft L, Moons KGM, van Smeden M. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020; 369: m1328 [PMID: 32265220 DOI: 10.1136/bmj.m1328]

8. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, Li Y, Guan W, Sang L, Lu J, Xu Y, Chen G, Guo H, Guo J, Chen Z, Zhao Y, Li S, Zhang N, Zhong N, He J; China Medical Treatment Expert Group for COVID-19. Development and Validation of a Clinical Risk Score to Predict the Occurrence of Critical Illness in Hospitalized Patients With COVID-19. *JAMA Intern Med* 2020; 180: 1081-1089 [PMID: 32396163 DOI: 10.1001/jamainternmed.2020.2033]

9. Wei PF. Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7). *Zhonghua yixue zazhi (Engl)* 2020; 133: 1087-1095 [PMID: 32358325 DOI: 10.1097/CMA9.0000000000000819]

10. Fine MJ, Aubele TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, Coley CM, Marrie TJ, Kapoor WN. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med* 1997; 336: 243-250 [PMID: 8995086 DOI: 10.1056/NEJM199701233360402]

11. Ruoppi MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom J* 2008; 50: 419-430 [PMID: 18435502 DOI: 10.1002/bimj.200710415]

12. *World Health Organization*. Clinical management COVID-19. 2020 May 27 [cited 20 June 2020]. In: WHO office site [Internet]. [about 1 screen]. Available from: https://www.who.int/publications-detail-redirect临床management-of-covid-19

13. Rehberg S, Maybauer MO, Enkhbaatar P, Maybauer DM, Yamamoto Y, Traber DL. Pathophysiology, management and treatment of smoke inhalation injury. *Expert Rev Respir Med* 2009; 3: 283-297 [PMID: 20161170 DOI: 10.1586/ers.09.21]

14. Horovitz JH, Carrico CJ, Shires GT. Pulmonary response to major injury. *Arch Surg* 1974; 108: 349-355 [PMID: 481333 DOI: 10.1001/archsurg.1974.01350270079014]

15. Bhimraj A, Morgan RL, Shumaker AH, Lavergne V, Baden L, Cheng VC, Edwards KM, Gandhi R, Muller WJ, O’Horo JC, Shoham S, Murad MH, Mustafa R, Sultan S, Falck-Ytter Y. Infectious Diseases Society of America Guidelines on the Treatment and Management of Patients with COVID-19. *Clin Infect Dis* 2020 [PMID: 32338708 DOI: 10.1093/cid/ciaa478]

16. Nicola M, O’Neill N, Sofrabi C, Khan M, Agha M, Agha R. Evidence based management guideline for the COVID-19 pandemic - Review article. *Int J Surg* 2020; 77: 206-216 [PMID: 32289472 DOI: 10.1016/j.ijsu.2020.04.001]
17 Junaidi I. Mild, moderate COVID-19 cases would be isolated in homes under new govt plan 2020. 2020 March 8 [cited 20 June 2020]. In: Dawn office site [Internet] [about 1 screen]. Available from: https://www.dawn.com/news/1539282

18 Shamsoddin E. A COVID-19 pandemic guideline in evidence-based medicine. Evid Based Dent 2020; 21: 71-73 [PMID: 32591667 DOI: 10.1038/s41432-020-0105-7]

19 Pascarella G, Strunia A, Piliego C, Bruno F, Del Buono R, Costa F, Scarlata S, Agrò FE. COVID-19 diagnosis and management: a comprehensive review. J Intern Med 2020; 288: 192-206 [PMID: 32348558 DOI: 10.1111/joim.13091]

20 World Health Organization. Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected. 2020 March 13 [cited 20 June 2020]. In: WHO office site [Internet] [about 2 screens]. Available from: https://www.who.int/docs/default-source/coronaviruse/clinical-management-of-novel-cov.pdf

21 Muller MP, Richardson SE, McGeer A, Dresser L, Raboud J, Mazzulli T, Loeb M, Louie M; Canadian SARS Research Network. Early diagnosis of SARS: lessons from the Toronto SARS outbreak. Eur J Clin Microbiol Infect Dis 2006; 25: 230-237 [PMID: 16586072 DOI: 10.1007/s10096-006-0127-x]

22 Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40: 373-383 [PMID: 3558716 DOI: 10.1016/0021-9681(87)90171-8]

23 Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. JAMA 2020; 323: 1061-1069 [PMID: 32031570 DOI: 10.1001/jama.2020.1585]

24 Yang X, Yu Y, Xu J, Shu H, Xia J, Liu H, Wu Y, Zhang L, Yu Z, Fang M, Yu T, Wang Y, Pan S, Zou X, Yuan S, Shang Y. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. Lancet Respir Med 2020; 8: 475-481 [PMID: 32105632 DOI: 10.1016/S2213-2600(20)30079-5]

25 Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020; 395: 1054-1062 [PMID: 32171076 DOI: 10.1016/S0140-6736(20)30566-3]

26 Kivimäki M, Kauorna E, Ferrie JE, Luukkonen R, Nyberg ST, Alfredsson L, Batter GD, Brunner EJ, Francss E, Goldberg M, Knutsson A, Koskenuo M, Nordin M, Oksanen T, Penti J, Rugulies R, Shipley MJ, Singh-Manoux A, Stepe0 A, Suominen SB, Theorell T, Vahtera J, Virtanen M, Westerholm P, Westerlund H, Zins M, Hamer M, Bell JA, Tabak AG, Jokela M. Overweight, obesity, and risk of cardiometabolic multimorbidity: pooled analysis of individual-level data for 120 813 adults from 16 cohort studies from the USA and Europe. Lancet Public Health 2017; 2: e277-e285 [PMID: 28626830 DOI: 10.1016/S2468-2667(17)30074-9]

27 Klatsky AL, Zhang J, Udaltsova N, Li Y, Tran HN. Body Mass Index and Mortality in a Very Large Cohort: Is It Really Healthier to Be Overweight? Perm J 2017; 21: 16-142 [PMID: 28678695 DOI: 10.7754/Clin.Lab.2018.180833]

28 Watkins J. Preventing a covid-19 pandemic. BMJ 2020; 368: m810 [PMID: 32111649 DOI: 10.1136/bmj.m810]

29 Moser JS, Galindo-Fraga A, Ortiz-Hernández AA, Gu W, Hunsberger S, Galán-Herrera JF, Guerrero ML, Ruiz-Palacios GM, Beigel JH; La Red ILI 002 Study Group. Underweight, overweight, and obesity as independent risk factors for hospitalization in adults and children from influenza and other viral respiratory infections. Influenza Other Respir Viruses 2019; 13: 3-9 [PMID: 30515985 DOI: 10.1016/j.invl.2016.07.001]
factors and outcome. The British Thoracic Society and the Public Health Laboratory Service. *Q J Med* 1987; 62: 195-220 [PMID: 3116595]

37 Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, Lewis SA, Macfarlane JT. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003; 58: 377-382 [PMID: 12728155 DOI: 10.1136/thorax.58.5.377]
