

Establishment of a clinical nomogram model to predict the progress to severe COVID-19

Changli Tu (✉ tuchangli2008@126.com)

the fifth affiliated hospital of sun Yat-sen university <https://orcid.org/0000-0002-3478-6465>

Guojie Wang

Fifth Affiliated Hospital of Sun Yat-sen University

Cuiyan Tan

Fifth Affiliated Hospital of Sun Yat-sen University

Meizhu Chen

Fifth Affiliated Hospital of Sun Yat-sen University

Hu Peng

Fifth Affiliated Hospital of Sun Yat-sen University

Ying Wang

Fifth Affiliated Hospital of Sun Yat-sen University

Yingjian Liang

Fifth Affiliated Hospital of Sun Yat-sen University

Yiyang Huang

Fifth Affiliated Hospital of Sun Yat-sen University

Zhenguo Wang

Fifth Affiliated Hospital of Sun Yat-sen University

Jian Wu

Fifth Affiliated Hospital of Sun Yat-sen University

Kongqiu Wang

Fifth Affiliated Hospital of Sun Yat-sen University

Qinhuan Huang

Fifth Affiliated Hospital of Sun Yat-sen University

Jin Huang

Fifth Affiliated Hospital of Sun Yat-sen University

Xiaobin Zheng

Fifth Affiliated Hospital of Sun Yat-sen University

Qiuyue Chen

Fifth Affiliated Hospital of Sun Yat-sen University

Yayuan Geng

huiying medical technology Co.

Na Guo

huiying medical technology Co.

Xiaorong Zhou

Fifth Affiliated Hospital of Sun Yat-sen University

Xinran Liu

Fifth Affiliated Hospital of Sun Yat-sen University

Jing Liu

Fifth Affiliated Hospital of Sun Yat-sen University

Hong Shan

Fifth Affiliated Hospital of Sun Yat-sen University

Research

Keywords: clinical nomogram model , severe corona virus disease 2019 (COVID-19), novel coronavirus pneumonia, logistic regression

Posted Date: March 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17574/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background:

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is the leading cause of a public health emergency in the world, accompanying with high mortality in severe corona virus disease 2019(COVID-19), thereby early detection and stopping the progress to severe COVID-19 is important. Our aim is to establish a clinical nomogram model to calculate and predict the progress to severe COVID-19 timely and efficiently.

Methods:

In this study, 65 patients with COVID-19 had been included retrospectively in the Fifth Affiliated Hospital of Sun Yat-sen University from January 17, to February 11, 2020. Patients were randomly assigned to train dataset (n=51 with 15 progressing to severe COVID-19) and test dataset (n=14 with 4 progressing to severe COVID-19). Lasso algorithm was applied to filter the most classification relevant clinical factors. Based on selected factors, logistic regression model was fit to predict the severe from mild/common. Meanwhile in nomogram sensitivity, specificity, AUC (Area under Curve), and calibration curve were depicted and calculated by R language, to evaluate the prediction performance to severe COVID-19.

Results:

High ratio of sever COVID-19 patients (26.5%) had been found in our retrospective study, and 84% of these cases progress to severe or critical after 5 days from their first clinical examination. In these 65 patients with COVID-19, 77 clinical characteristics in first examination were collected and analyzed, and 37 ones had been found different between non-severe and severe COVID-19. But when all these factors were analyzed in establishment of prediction model, six factors are crucial for predicting progress of severe COVID-19 via Lasso algorithm. Based on these six factors, including increased fibrinogen, hyponatremia, decreased PaO₂, multiple lung lobes involved, down-regulated CD3⁺T-lymphocyte and fever, a logistic regression model was fit to discriminate severe and common COVID-19 patients. The sensitivity, specificity and AUC were 0.93, 0.86, 0.96 in the train dataset and 0.9, 1.0, 1.0 in test dataset respectively. Nomogram-predicted probability was more consistent with actual probability by R language.

Conclusions:

In summary, an efficient and reliable clinical nomogram model had been established, which indicate increased fibrinogen, hyponatremia, decreased PaO₂, multiple lung lobes involved, down-regulated CD3⁺T-lymphocyte and fever at the first clinical examination, could predict progress of patients to severe COVID-19.

Introduction

Since December 2019, a number of pneumonia cases infected with the novel coronavirus have been detected in Wuhan, Hubei province, China. The Coronavirus Study Group (CSG) of the International Committee on Taxonomy of Viruses formally considered this virus as a sister of severe acute respiratory syndrome coronaviruses (SARS-COV), the species Severe acute respiratory syndrome-related coronavirus and designates it as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in February 07, 2020^[1].According to World Health Organization(WHO), the disease caused by Novel Corona Virus(SARS-CoV-2) is now officially called Coronavirus Disease(COVID-19).By 24:00 on February 21, 2020, China had confirmed 76,288 cases and 2345 deaths, with a mortality rate of 3.07%^[2].With the rapid spread of the disease, we have therefore made the assessment that COVID-19 can be charcaterized as a pandemic by WHO on March 11, 2020.

Studies have shown that ^[3]among patients diagnosed with COVID-19, the mortality rates of severe novel coronavirus pneumonia (NCP), common pneumonia and non-pneumonia are respectively 5.88%, 0.12% and 0%. Our objective is to pick up key clinical and imaging characteristics of NCP patients during their first visit and establish a clinical nomogram to predict the progress to severity, and stop the progress early, thereby decline the mortality of COVID-19 patients.

Methods

Study design and patient population

This study was a retrospective single-center study, which included 83 confirmed COVID-19 adult patients (more than 18 years old) admitted to the Fifth Affiliated Hospital of Sun Yat-sen University from January 17, 2020 to February 11, 2020. A confirmed case of COVID-19 was defined as positive for novel coronavirus (SARS-CoV-2) nucleic acid by real-time fluorescent RT-PCR. NCP was diagnosed and clinically classified according to the new coronavirus pneumonia diagnosis and treatment plan (trial version 6) [4] drafted by the National Health Committee of the People's Republic of China. The clinical classifications of NCP are as follows: (1) mild, with mild symptoms and no obvious signs of pneumonia in imaging; (2) common, with fever, respiratory tract symptoms, and obvious imaging signs indicating pneumonia; (3) severe, with one of the following: a) respiratory rate ≥ 30 beats / min; b) mean oxygen saturation of resting state $\leq 93\%$; c) arterial blood oxygen partial pressure / oxygen concentration ≤ 300 mmHg (1 mmHg = 0.133 kPa);

d) pulmonary imaging showed a significant increase in $> 50\%$ within 24 ~ 48 h; (4) critical, with one of the following conditions: a) respiratory failure and requiring mechanical ventilation; b) Shock; c) ICU admission due to combined organ failure. All participants were followed until the end of disease course. We selected patients who had finished the first lung CT examination and admission within 7 days from onset, and completed all required laboratory, questionnaire and medical examination. This retrospective observational study was approved by the Research Ethics Committee of The Fifth Affiliated Hospital of Sun Yat-sen University (approval series number K153-1) and the need for informed consent was waived, considering the retrospective study design.

Clinical Data

Clinical data were extracted from electronic medical records, including basic demographic data, symptoms, vital signs, clinical classification, interval from onset to admission, complications and the SARS-CoV-2 RNA clearance time which was defined as the time of after two consecutive negative real-time PCR results with 1-day interval. Laboratory evaluations included total blood count, T-cell subgroup classification, coagulation function, liver and kidney function, electrolyte, blood gas analysis, lactate dehydrogenase(LDH), creatine kinase(CK), creatine kinase-MB(CK-MB), Hydroxybutyrate Dehydrogenase(α -HBDH), C-reactive protein (CRP), procalcitonin(PCT), cardiac troponin I (cTnI), brain natriuretic peptide(BNP), D-dimer, random blood sugar, and Hepatitis B surface antigen (HBsAg). The first lung CT examination performed within 7 days after onset (uCT760, United Imaging; Shanghai, China) were included, too. Each patient was placed in the supine position and scanned in the inspiratory phase. The thickness and interval were set as 1 mm. Lung CT images were screened independently by two imaging physicians with more than 5 years of experience in Picture Archiving and Communication Systems (PACS) workstation and a consensus were reached. For any dispute, another imaging physician with more than 10 years of experience in diagnosis was involved and made the final decision. The evaluation details includes: ☐ground-glass opacity(GGO); ☐consolidation; ☐stripes; ☐Number of affected lung (ranging from 0–2), location and number of affected lobe (right upper lobe, right middle lobe, right lower lobe, left upper lobe and left lower lobe, ranging 0–5), respectively; ☐Vascular enlargement^[5]; ☐bronchial abnormalities (bronchiectasis, distortion, wall thickening, etc); ☐nodules; ☐mediastinal lymph node enlargement, with the short axis of lymph node greater than 10 mm as the standard; ☐pleural effusion; ☐other abnormalities (such as emphysema, fibrosis, calcification, etc.)

Statistical analysis

The clinical data were shown as the mean \pm standard with normal distribution (Kolmogorov-Smirnov test, $P > 0.05$), and as the median (minimum, maximum) with non-normal distribution. Difference between non-severe(mild/common) and severe(severe/critical) groups was analyzed with either Independent Samples T Test or Mann–Whitney U test depending on the distribution of the data (Table 1). Statistical analysis was performed using SPSS software (Version 26.0, IBM). A P value < 0.05 was considered significant.

Table 1
Demographics and clinical characteristics of patients infected with SARS-CoV-2

	all patients (n = 65)	Severe group (n = 19)	Non-severe group (n = 46)	p value
Characteristic				
Age,years	44(19,80)	58(32,80)	38(19,75)	0.001
Gender				0.317
Male	28(43.1%)	10(52.6%)	18(39.1%)	
Female	37(56.9%)	9(47.4%)	28(60.9%)	
Signs and symptoms				
Fever	46(70.8%)	18(94.7%)	28(60.9%)	0.006
Cough	26(40%)	6(31.6%)	20(43.5%)	0.373
Fatigue	1(1.5%)	1(5.3%)	0(0)	0.292
Nasal discharge	2(3.1%)	0(0)	2(4.3%)	1.000
Pain (headache/pharyngodynia/muscle ache, etc)	14(21.5%)	5(26.3%)	9(19.6%)	0.529
Other clinical symptoms	6(9.2%)	1(5.3%)	5(10.9%)	0.662
Body T,°C	37.2(36.0,39.2)	37.9(36.6,39.2)	36.9(36.0,38.5)	< 0.001
R,/min	20(12,22)	20(14,22)	20(12,22)	0.898
P, /min	85(61,111)	84(62,110)	85(61,111)	0.948
SBP,mmHg	125(96,193)	130(108,174)	123(96,193)	0.502
DBP,mmHg	82(62,111)	85 (62,97)	82(64,111)	0.849
BMI,Kg/m2	23.7(15.6,33.0)	24.2(20.1,33.0)	23.5(15.6,31.0)	0.287
underlying diseases(N/Y)	23(35.4%)	10(52.6%)	13(28.3%)	0.062
circulatory basic diseases (hypertension, coronary heart disease)	12(18.5%)	7(36.8%)	5(10.9%)	0.030
respiratory basic diseases (chronic bronchitis, lung cancer,etc)	3(4.6%)	2(10.5%)	1(2.2%)	0.202
endocrine system basic diseases (diabetes)	8(12.3%)	5(26.3%)	3(6.5%)	0.041
Other systemic diseases (fractures, cerebral infarction, etc.)	10(15.4%)	3(15.8%)	7(15.2%)	1.000
Laboratory findings				
Nucleic acid continued positive for 7 days (+, -)	55(84.6%)	17(89.5%)	38(82.6%)	0.71
White blood cell count,×10 ⁹ /L	4.47(1.91,12.52)	3.76(2.63,6.52)	4.58(1.91,12.52)	0.039
Lymphocyte count,×10 ⁹ /L	1.55 ± 0.67	1.20 ± 0.38	1.70 ± 0.71	< 0.001
CRP,mg/L	4.86(0.01,128.31)	29.41(0.80,115.14)	2.67(0.01,128.31)	< 0.001
D-dimer,ng/ml	98.5(6.0,1539.0)	105.0(75.0,433.0)	91.0(6.0,1539.0)	0.07
Tbil,umol/L	7.67(2.90,29.29)	8.34(2.93,22.16)	7.16(2.9,29.29)	0.453

	all patients (n = 65)	Severe group (n = 19)	Non-severe group (n = 46)	p value
Albumin,g/L	39.43 ± 3.52	36.92 ± 3.26	40.46 ± 3.10	< 0.001
LDH,U/L	160(109,294)	186(148,294)	149(106,280)	< 0.001
potassium,mmol/L	3.70 ± 0.36	3.64 ± 0.42	3.70 ± 0.34	0.969
PaO2(air breathing),mmHg	97.57 ± 16.06	85.32 ± 11.41	102.64 ± 15.00	< 0.001
PaCO2(air breathing),mmHg	39.41 ± 3.56	37.59 ± 3.94	40.16 ± 3.14	0.007
Lac lactic acid,mmol/L	1.5(0.6,13.1)	1.4(0.6,2.7)	1.5(0.6,13.1)	0.811
CD3(+) T lymphocytes,/ul	984 ± 450	643 ± 320	1125 ± 422	< 0.001
CD3(+)CD4(+) T lymphocytes,/ul	521(83,1565)	377(83,893)	582(228,1565)	0.001
CD3(+)CD8(+) T lymphocytes,/ul	340.95 ± 172.82	214.16 ± 123.19	393.33 ± 163.75	< 0.001
Neutrophil count,x109/L	2.58(0.65,9.14)	2.49(1.18,4.31)	2.58(0.65,9.14)	0.834
Monocyte count,x109/L	0.5(0.23,1.74)	0.46(0.23,0.91)	0.51(0.24,1.74)	0.521
NLR	1.78(0.4,9.38)	2.16(0.74,9.38)	1.70(0.40,5.68)	0.025
Hemoglobin,g/L	137.72 ± 18.51	133.11 ± 22.23	139.63 ± 16.64	0.199
Platelet count,x109/L	182(81,393)	140(81,199)	193(87,193)	< 0.001
PCT(<0.5ug/L)	8(12.3%)	7(36.8%)	1(2.2%)	< 0.001
PT,S	12.1 (10.4,16.5)	12.6(10.4,16.5)	11.9(10.6,15.4)	0.093
APTT,S	31.7(13.8,322.0)	33.0(20.3,38.4)	31.2(13.8,322.0)	0.116
FIB,g/L	3.09 ± 0.79	3.64 ± 0.66	2.86 ± 0.73	< 0.001
ALT,U/L	15(5.6,80)	17.6(5.6,80)	14.6(6.4,68.1)	0.246
AST,U/L	20.7(10.4,60.0)	22.9(11.8,60.0)	17.9(10.4,39.2)	0.013
Dbil,umol/L	3.0(1.2,7.9)	3.4(1.4,7.9)	2.80(1.2,6.0)	0.147
Ibil,umol/L	4.20(1.51,15.29)	4.13(1.52,15.29)	4.21(1.51,13.58)	0.697
TP, g/L	69.03 ± 5.34	66.64 ± 4.81	70.01 ± 5.28	0.019
Globin, g/L	29.6(20.6,151.0)	30.2(23.4,37.2)	29.4(20.6,151.0)	0.846
a-HBDH,U/L	124.5(84.0,244.0)	141.5(117.0,213.0)	119.5(84.0,244.0)	0.002
CK,U/L	68(36,935)	84(47,183)	63(36,935)	0.068
CK-MB,U/L	13.4(5.7,48.5)	14.2(9.1,48.5)	12.5(5.7,30.4)	0.189
BUN,mmol/L	3.8(1.7,56.9)	4.0(2.6,56.9)	3.5(1.7,9.7)	0.007
Crea,μmol/L	58.7(33.1,798.0)	63.5(44.2,361.0)	56.5(33.1,798.0)	0.306
Serum Sodium, mmol/L	139.32 ± 2.70	137.11 ± 2.05	140.24 ± 2.40	< 0.001

	all patients (n = 65)	Severe group (n = 19)	Non-severe group (n = 46)	p value
Serum Chlorine, mmol/L	100.97 ± 3.16	99.34 ± 3.01	101.64 ± 3.01	0.007
random Serum glucose, mmol/L	5.3(4.1,18.5)	6.3(4.9,16.4)	5.0(4.1,18.5)	< 0.001
HBsAg(+)	6(9.2%)	3(15.8%)	3(6.5%)	0.347
PH(arterial blood)	7.39 ± 0.03	7.40 ± 0.04	7.38 ± 0.03	0.026
BNP(> 500 pg/ml)	13(20%)	9(47.4%)	4(8.7%)	0.001
cTnl(<0.3ug/L)	3(4.6%)	2(10.5%)	1(2.2%)	0.202
CT imaging features				
one or both lungs (0–2)	2(0,2)	2(1,2)	1(0,2)	0.001
number of involved lung lobes (0–5)	2(0,5)	5(1,5)	1(0,5)	< 0.001
right upper lobe	28(43.1%)	15(78.9%)	13(28.3%)	< 0.001
right middle lobe	20(30.8%)	13(68.4%)	7(15.2%)	< 0.001
right lower lobe	38(58.5%)	16(84.2%)	22(47.8%)	0.007
left upper lobe	29(44.6%)	16(84.2%)	13(28.3%)	< 0.001
left lower lobe	37(56.9%)	16(84.2%)	21(45.7%)	0.004
Ground-glass opacity (GGO)	44(67.7%)	16(84.2%)	28(60.9%)	0.067
consolidation	12(18.5%)	7(36.8%)	5(10.9%)	0.030
Vascular enlargement	35(53.8%)	15(78.9%)	20(43.5%)	0.009
stripes	6(9.2%)	4(21.1%)	2(4.3%)	0.055
nodules	14(21.5%)	4(21.1%)	10(21.7%)	1.000
bronchial abnormalities	5(7.7%)	3(15.8%)	2(4.3%)	0.144
mediastinal lymph node enlargement	3(4.6%)	3(15.8%)	0(0)	0.022
pleural effusion	1(1.5%)	1(5.3%)	0(0)	0.292
Others (emphysema/tuberculosis, etc)	20(30.8%)	9(47.4%)	11(23.9%)	0.062
T: Temperature; R: Respiration Rate; P: Pulse ; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; BMI: Body Mass Index; CRP: C-reactive protein; TBil: total bilirubin; LDH: lactic dehydrogenase; PaO2: arterial partial pressure of oxygen; PaCO2: partial pressure of carbon dioxide in artery; NLR: Neutrophil -to-Lymphocyte Ratio; PCT: procalcitonin; PT: Prothrombin time; APTT: activated partial thromboplastin time; FIB: fibrinogen; ALT: Alanine transaminase; AST: glutamic oxalacetic transaminase; Dbil: direct bilirubin; Ibil: indirect bilirubin; TP: total protein; α-HBDH: α-hydroxybutyrate dehydrogenase; CK: creatine kinase; CK-MB: creatine kinase-MB; BNN: urea nitrogen; Crea: Creatine kinase; BNP: brain natriuretic peptide; cTnl: cardiac troponin I.				

Patients were randomly assigned to train dataset (n = 51 with 15 progressing to severe NCP) and test dataset (n = 14 with 4 progressing to severe NCP) depended on 8:2 ratio. Standardization was carried out on continuous factors to control the mean value as 0 and the standard variance as 1. Lasso algorithm was applied to filter the most classification relevant clinical factors. Based on

selected factors, logistic regression model was fit to predict the severe from mild. 10-fold cross validation was applied to find the optimal lambda value in lasso. In this study, 1 standard error of minimum lambda, was deemed as the optimal lambda. R version 3.6.2 was used to set up lasso algorithm and rms package was utilized to fit logistic regression model and prediction nomogram. Meanwhile in nomogram sensitivity, specificity, AUC (Area under Curve), and calibration curve were depicted and calculated by R language, to evaluate the prediction performance to severe NCP.

Results

1. High ratio of sever/critical COVID-19 patients

When we retrospective analyzed the data of 83 COVID-19 adult patients, we found that there were 22 severe/critical cases (26.5%), among which there were 18 patients (84%) progress to severe or critical after their first examination. Because we aimed to establish a nomogram prediction model in early stage of COVID-19, we excluded the patients with longer interval more than 7 days between the first lung CT examination or admission and onset (Fig.1). At last, 65 COVID-19 patients were included in our study, including 46 mild/common cases and 19 severe/critical cases (Fig.1). Among the 19 severe/critical cases, the median time from first examination to severe exacerbation was 5 days.

2. The original clinical characteristics in COVID-19 patients

In these 65 COVID-19 patients in our hospital, 77 clinical characteristics were collected and analyzed (Table1). The average age of COVID-19 patients were 44 years, and severe/critical cases were older than mild/common ones. But there was no difference in gender ratio between. The most common symptom in COVID-19 patients was fever (70.8%), followed by cough (40%) and pain (21.5%). Both the frequency of fever occurrence and body temperature in severe/critical cases were higher than mild/common ones. In consideration of underlying diseases, the circulatory and endocrine system basic diseases were more common in severe/critical cases, maybe due to their age. Otherwise, some difference in laboratory findings had been found between severe/critical and mild/common groups ($P < 0.05$, table1.), including WBC count, LYM count, CRP, ALB, LDH, PaO₂ (air breathing), PaCO₂ (air breathing), CD3 (+)T-lymphocytes count, CD3 (+)CD4 (+)T-lymphocytes count, CD3 (+)CD8 (+)T-lymphocytes count, NLR, PLT count, PCT, FIB, AST, TP, α -HBDH, BUN, serum sodium, serum chloride, random blood sugar, PH value(arterial blood), and increased BNP. As for imaging features in chest CT of COVID-19 patients, more bilateral lung and lung lobes were involved, higher frequency of consolidation and mediastinal lymph node enlargement had been shown in the severe/critical cases, comparing with mild/common ones ($P < 0.05$, Table1.). And the most common imaging sign in COVID-19 patients was ground-glass opacity (Fig2), happening in 67.79% of all patients and 84.2% of severe cases.

3. Selection of significant predicting factor, establishment of a clinical nomogram model to predict the progress to severe COVID-19

After analyzed above 77 clinical characteristics in first examination in 65 COVID-19 patients, 37 ones had been found significant different between **mild/common and severe/critical** cases. But not all these 37 clinical characteristics could be as predictors. 10-fold cross validation lasso algorithm was applied to filter the most classification relevant clinical factors, according to the optimal lambda=0.151(Fig3). Based on this lambda, 6 related factors, including Temperature, Na⁺, FIB, number of affected lung lobes, CD3(+) T lymphocytes and PaO₂(air breathing), had been selected. The Lasso Coefficients for these 6 selected factors in this models were follow: Elevated body temperature(37.7°C or 98.6°F) increased fibrinogen(3.52 g/L), Serum Sodium (139.5 mmol/L), decreased PaO₂ (83.5 mm Hg), multiple lung lobes involved (4 lobes), down-regulated CD3(+)T-lymphocyte (851/ul), which indicated contribution to the model and the prediction strength to severe COVID-19 (Fig4.). After establishment of clinical nomogram prediction model with above 6 factors, the efficiency and performance of this model had been tested with logistic regression analysis. We found the high sensitivity (0.93), specificity (0.86) and AUC (0.96) in train dataset, and sensitivity (0.90), specificity (1.0) and AUC (1.0) in test dataset respectively (Fig5.). Otherwise, calibration curve of this model was depicted and calculated by R language, nomogram-predicted probability is more consistent with actual probability (Fig6), which indicated the prediction model to severe COVID-19 was efficient and reliable.

Discussion

The main manifestations of COVID-19 patients were fever, weakness and dry cough. Most of the severe patients developed dyspnea and/or hypoxemia one week after the onset of the disease, and the severe patients rapidly proceeded to acute respiratory distress

syndrome, septic shock, etc. In the face of the new emergence of SARS-CoV-2, there is no targeted antiviral drug [2]. The severity of the COVID-19 is directly related to the prognosis. In the cohort of 41 COVID-19 patients, 22 (55%) cases developed severe dyspnea, 13 (32%) cases were admitted to the intensive care unit and of which 6 patients died [6]. It can be concluded that the case fatality rate in this cohort is about 14.6%. From our experience in diagnosis and treatment, early screening of severe patients and early intervention (such as early immune supportive therapy, high-flow oxygen therapy, etc.) contributes to the control of the disease and may reduce mortality. The illness was mild in 16 of 19 severe cases (84%) at admission. If the progress could be predicted and the intervention can be carried out in advance at the stage of disease, the prognosis would be improved.

In this retrospective study, consulting the previous literature [3][7][8], we included 77 clinical risk factors related to the clinical classification of COVID-19, of which 37 had statistically significant differences between non-severe and severe patients ($P < 0.05$). At the same time, these 77 clinical parameters were screened by Lasso algorithm, and a new predictive model for predicting the clinical outcome of COVID-19 was established and verified. According to this model, COVID-19 patients were divided into non-severe group and severe group. In this model, fibrinogen, hyponatremia, decreased PaO₂(air breathing), multiple lung lobes involved, down-regulated CD3(+) T-lymphocyte and fever are closely related to COVID-19 severe illness. The prediction efficiency is high and ideal by testing with logistic regression.

We speculate that it will contribute to identify severe COVID-19 patients by these 6 clinical parameters at this early stage.

At present, PaO₂ and pulmonary imaging have been widely used to predict the severity of pulmonary diseases and the survival of patients [4][9]. The lower PaO₂(air breathing), the more severe patient's condition and the worse prognosis are shown. Imaging changes in Chest CT in NCP are rapid. High-resolution CT (HRCT) of the chest is critical for early detection, evaluation of disease severity and follow-up of patients with NCP [10]. Based on NCP guideline (trial version 6) [4], the diagnosis of severe NCP had include that Chest CT showed a significant increase in $> 50\%$ within 24 ~ 48 h. but the judgement on the baseline Chest CT, is also very important to predict the possibility of progress to severe cases. This study found that COVID-19 patients with severe prognosis had more lung lobes and bilateral lungs involvement in the early stage of the disease, and the incidence of consolidation in severe patients was higher than that in non-severe patients, suggesting that the pulmonary exudation may progress faster in severe patients. Chest CT findings, especially the number of involved lung lobes, contribute to the prediction of COVID-19 clinical outcome.

Studies have shown [11] that SARS-CoV-2 and SARS-CoV have good sequence homology and may share a common ancestor with SARS/ SARS-like coronaviruses. The levels of CD3(+), CD4(+) and CD8(+) T-lymphocyte subsets in peripheral blood of SARS patients decrease significantly. Possible reasons for the SARS-associated lymphopenia may be direct attack and destruction to lymphocytes by SARS-CoV, or immune-mediated lymphocyte destruction, bone marrow or thymus suppression and apoptosis. T lymphocytes migrate to the lungs directly or by cytokine-mediated approach to phagocytose virus. But a lot of inflammatory factors are released accompanying T lymphocytes destruction, which may cause inflammatory storm and progress to severe COVID-19 [12][13].

In this study, the CD3 (+) T lymphocyte count in patients with severe COVID-19 was significantly lower than that in mild patients, which was confirmed as important predictor in our clinical nomogram prediction model. But subgroup of CD3 (+)CD4(+) T lymphocyte and CD3(+)CD8(+) T lymphocytes count had been filtrated by Lasso algorithm, so the function and meaning of subgroup of T lymphocytes need to be studied further. Meanwhile, if T lymphocytes subgroup testing is not available, lymphocyte count and the neutrophil-to-lymphocyte ratio (NLR) maybe also provide some reference to predict of severe, if the change is great. For example in the study of Huang C, NLR was identified as the independent risk factor for severe illness in patients with COVID-19, and $NLR \geq 3.13$ facilitated severe cases [7].

In this study, the serum sodium in the severe group was lower than that in non-severe group. Sodium metabolism is regulated by many factors, such as renin-angiotensin-aldosterone system, vasopressin, atrial natriuretic peptide. Studies have shown [9] that SARS-CoV-2 S protein has a strong binding affinity with human ACE2. ACE2 has another important function to participate in renin-angiotensin system, which is very important in regulating electrolyte balance. Certainly, the related mechanism of hyponatremia in patients with COVID-19 remains to be further studied. But the higher Lasso coefficients in this prediction model indicated that correcting hyponatremia maybe is important to stop progress to severe stage.

In addition to its functions of coagulation and hemostasis, fibrinogen is also an acute-phase protein that participates in a series of pathophysiological processes in vivo, such as inflammatory tissue damage and repair, especially under acute stress. With the

aggravation of lung injury in ARDS patients, coagulation activity further decreased, and fibrinolytic activity further increased, and fibrinogen level gradually increased with the aggravation of the disease^[14]. Elevation of peripheral blood fibrinogen in SARS patients may play an important role in development and progress of the disease and its treatment^[15]. This data shows that the blood fibrinogen of patients with severe COVID-19 was higher than non-severe patients. The CT imaging grouping of pulmonary imaging abnormalities indicated that the fibrinogen level was increased in those with extensive lung tissue lesions, indicating the progression of lung injury, while the fibrinogen exudation was more significant in those with severe lung disease. Increased fibrinogen can be used as an important diagnostic and prognostic indicator for the severity and rapid progression of pulmonary lesions in COVID-19 patients.

Although high fever is not common and typical in COVID-19 patients, fever had been found in 70.8% of all patients. In this study, the average temperature in the first physical examination of patients in severe group was higher than that in non-severe group, which indicated we should pay more attention to patients with high fever and monitor closely to avoid progress to severe stage.

Conclusions

In summary, an efficient and reliable clinical nomogram model had been established, which indicate increased fibrinogen, hyponatremia, decreased PaO₂, multiple lung lobes involved, down-regulated CD3(+) T-lymphocyte and high fever at the first clinical examination, could predict progress of patients to severe COVID-19. This prediction model is very helpful for us to early assess, regulate treatment and stop the progress to severe COVID-19. Certainly, practical application of the model needs to be verified by expanding the sample size of several other research centers and by prospective predictive research.

Declarations

Acknowledgements:

The authors would like to thank the study participants who donated their samples for the advancement of scientific knowledge.

Author contributions:

Conceived and designed and experiments: Changli Tu, Guojie Wang, Jing Liu, Hong Shan. Collection of samples: Changli Tu, Guojie Wang, Cuiyan Tan, Meizhu Chen, Hu Peng, Yingjian Liang, Yiyi Huang, Zhenguo Wang, Jian Wu, Kongqiu Wang, Qinhuan Huang, Jin Huang, Xiaobin Zheng, Xiaorong Zhou, Xinran Liu. Performed the experiments: Changli Tu, Guojie Wang, Cuiyan Tan.

Data analysis: Changli Tu, Guojie Wang, Ying Wang, Qiuyue Chen, Yayuan Geng, Na Guo, Xiaorong Zhou. Contributed reagents/materials/analysis tools: Yayuan Geng, Na Guo. Wrote the paper: Changli Tu, Guojie Wang. Critical review and approval: Jing Liu, Hong Shan.

Conflict of interest

The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Ethics approval and consent to participate

This retrospective observational study was approved by the Research Ethics Committee of The Fifth Affiliated Hospital of Sun Yat-sen University (approval series number K153-1) and the need for informed consent was waived, considering the retrospective study design.

Consent for publication

Not applicable.

Availability of supporting data

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

Funding

Not applicable.

References

[1] Alexander E. Gorbalenya, Susan C. Baker, Ralph S. Baric, et al. Severe acute respiratory syndrome-related coronavirus: The species and its viruses—a statement of the Coronavirus Study Group. www.biorxiv.org/content/10.1101/2020.02.07.937862v1.

Date last updated: February 11 2020. Date last accessed: February 11 2020.

- [2]National Health Commission of the People's Republic of China. The latest news on Health Emergency Office. 2020. www.nhc.gov.cn/xcs/yqtb/202002/543cc508978a48d2b9322bdc83daa6fd.shtml.Date last updated: February26 2020. Date last accessed: February212020.
- [3]YangYang, Qing-Bin Lu, Ming-Jin Liu, et al. Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China. *atmedRxiv preprint* doi: <https://doi.org/10.1101/2020.02.10.20021675>.
- [4]National Health Commission of the People's Republic of China. Diagnosis and treatment of pneumonia infected with novel coronavirus. 2020. www.nhc.gov.cn/zyygj/s7653p/202002/8334a8326dd94d329df351d7da8aefc2.shtml.Date last updated: February19 2020. Date last accessed: February19 2020.
- [5]Zhao W, Zhong Z, Xie X, et al.Relation Between Chest CT Findings and Clinical Conditions of Coronavirus Disease (COVID-19) Pneumonia: A Multicenter Study.*AJR Am J Roentgenol*. 2020 Mar 3:1-6. doi: 10.2214/AJR.20.22976.
- [6] Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020 Feb 15;395(10223):497-506.doi: 10.1016/S0140-6736(20)30183-5.
- [7]Jingyuan Liu, Yao Liu, Pan Xiang, et al. Neutrophil-to-Lymphocyte Ratio Predicts Severe Illness Patients with 2019 Novel Coronavirus in the Early Stage.*at medRxiv preprint* doi: <https://doi.org/10.1101/2020.02.10.20021584>.
- [8]Wei-jie Guan, Zheng-yi Ni, Yu Hu 3,etal.Clinical characteristics of 2019 novel coronavirus infection in China.*at medRxiv preprint* doi: <https://doi.org/10.1101/2020.02.06.20020974>.
- [9]Cao MS, Sheng J, Wang TZ, et al. Acute exacerbation of idiopathic pulmonary fibrosis: usual interstitial pneumonitis vs. possibleusual interstitial pneumonitis pattern. *Chin Med J (Engl)*. 2019 Sep 20;132(18):2177-2184. doi: 10.1097/CM9.0000000000000422.
- [10] Pan Y, Guan H, Zhou S, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia(2019-nCov): a study of 63 patients in Wuhan, china. *Eur Radiol*. 2020 Feb 13.doi: 10.1007/s00330-020-06731-x.
- [11] Xu X, Chen P, Wang J,et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci China Life Sci*. 2020 Mar;63(3):457-460. doi: 10.1007/s11427-020-1637-5.
- [12] Lu Y1, Gong EC, Zhang QY,et al. Expression of SARS-CoV in various types of cells in lung tissues. *Beijing Da Xue Xue Bao Yi Xue Ban*. 2005 Oct 18;37(5):453-7.
- [13] National Research Project For SARS Beijing Group Beijing 100020 China.Dynamic changes of T-lymphocytes and immunoglobulins in patients with severe acute respiratory syndrome. *Zhonghua Yi Xue Za Zhi*. 2003 Jun 25;83(12):1014-7.
- [14]Idell S. Endothelium and disordered fibrin turnover in the injured lung: newly recognized pathways. *Crit Care Med*. 2002 May;30(5 Suppl):S274-80.DOI: 10.1097/00003246-200205001-00017
- [15] Gou CY, Li XH, Liang LC, et al. Changes and significance of peripheral blood fibrinogen level in patients with severe acuterespiratory syndrome.*Zhonghua Shi Yan He Lin Chuang Bing Du Xue Za Zhi*. 2007 Sep;21(3):258-60.

Figures

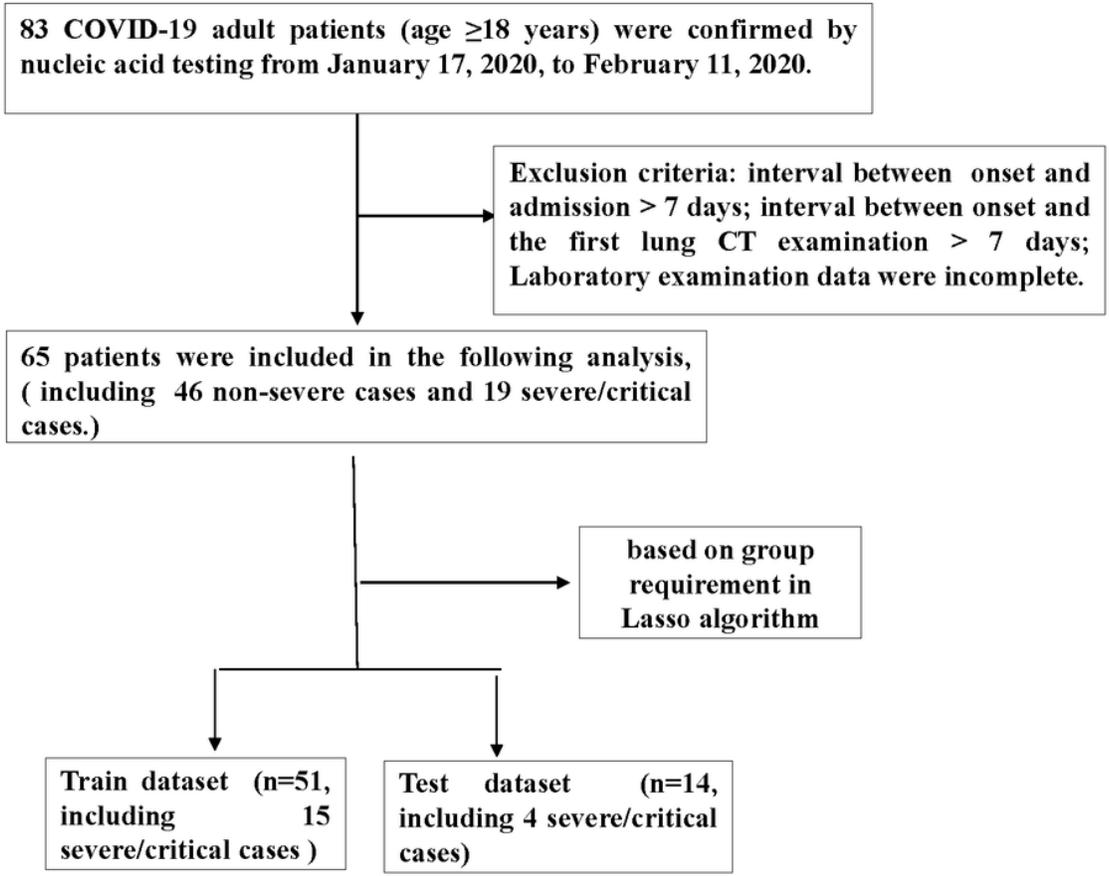


Fig. 1: Screening flow chart of COVID-19 patients.

Figure 1

Screening flow chart of COVID-19 patients.

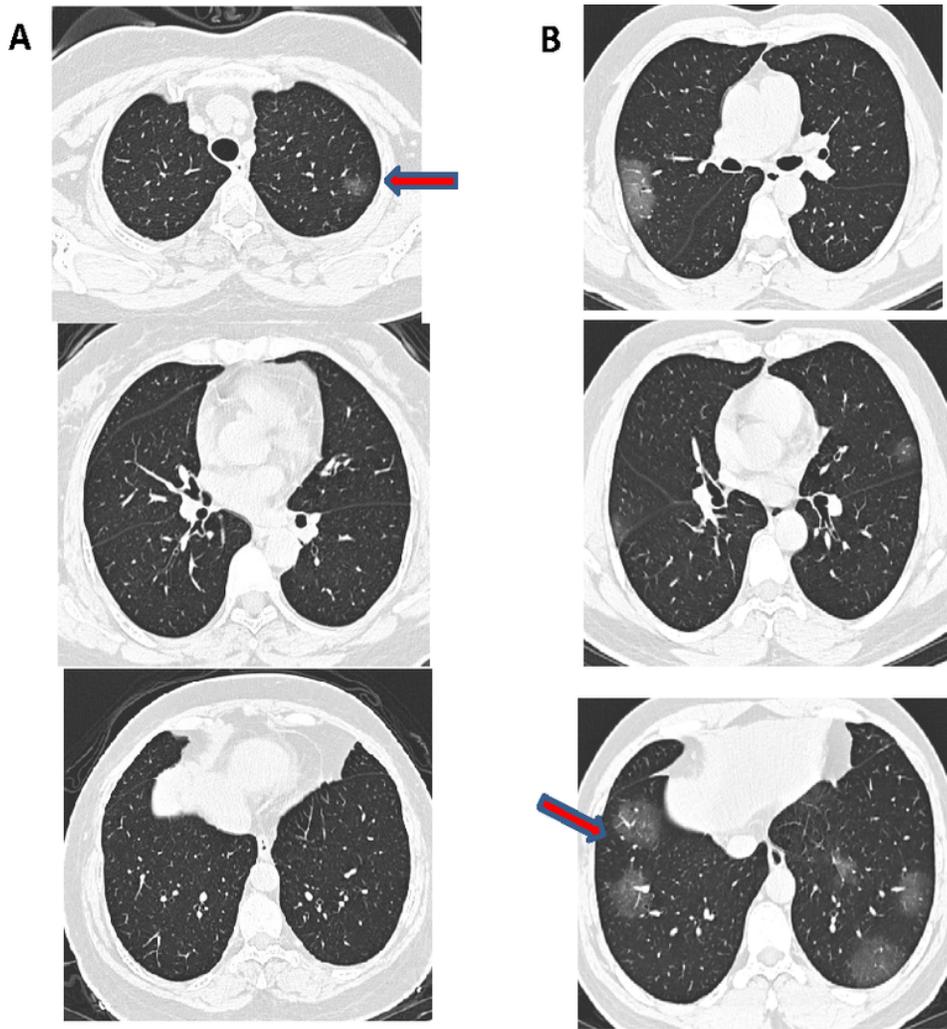


Fig.2: Typical ground glass opacity change in Chest CT of COVID-19 patients at the early stage. A: single small GGO in left upper lobe in a common COVID-19 patient; B: GGO in multiple lobes in both lungs in a COVID-19 patients who progress to severe later.

Figure 2

Typical ground glass opacity change in Chest CT of COVID-19 patients at the early stage. A: single small GGO in left upper lobe in a common COVID-19 patient; B: GGO in multiple lobes in both lungs in a COVID-19 patients who progress to severe later.

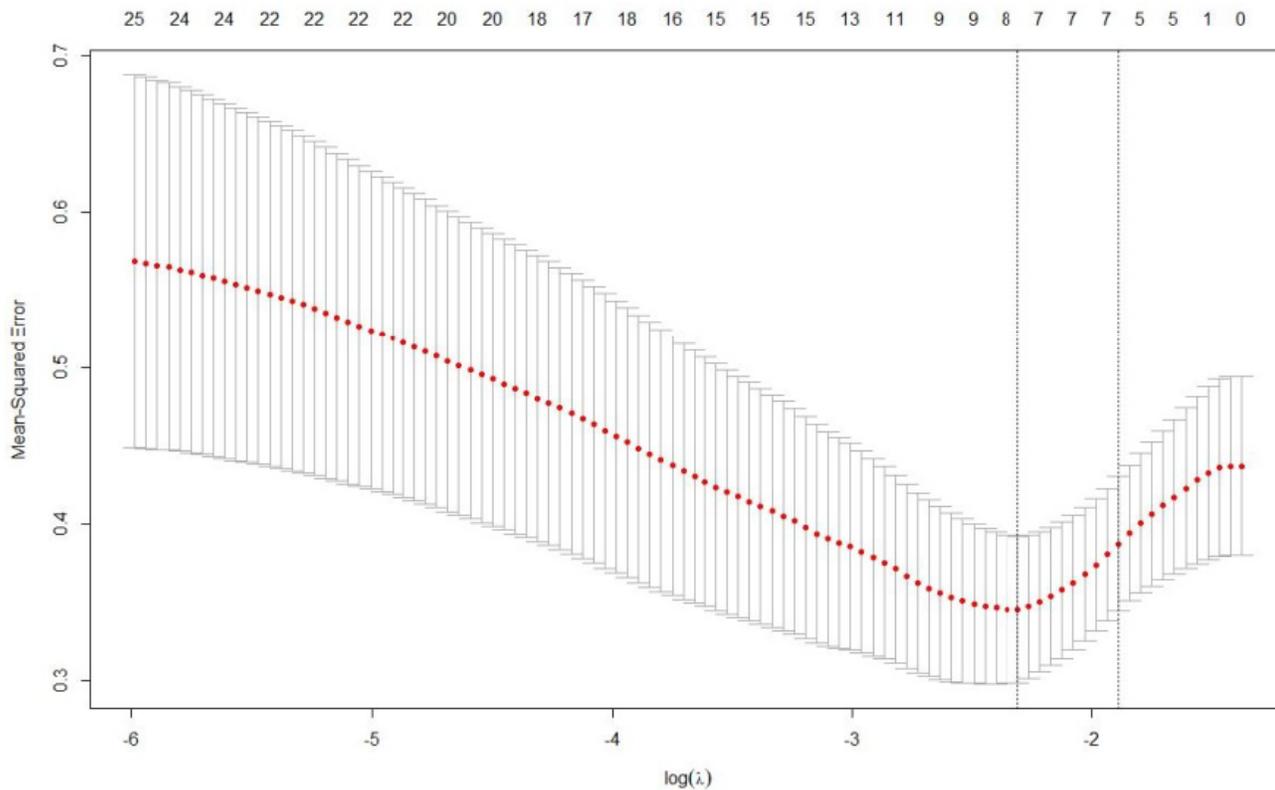


Fig.3: Testing and selection of a set of factors most relevant to predict the progress of severe COVID-19 via Lasso algorithm

Lasso algorithm was applied to filter the most classification relevant clinical factors, according to the optimal $\lambda=0.151$.

Figure 3

Testing and selection of a set of factors most relevant to predict the progress of severe COVID-19 via Lasso algorithm Lasso algorithm was applied to filter the most classification relevant clinical factors, according to the optimal $\lambda=0.151$.

Lasso Coefficients

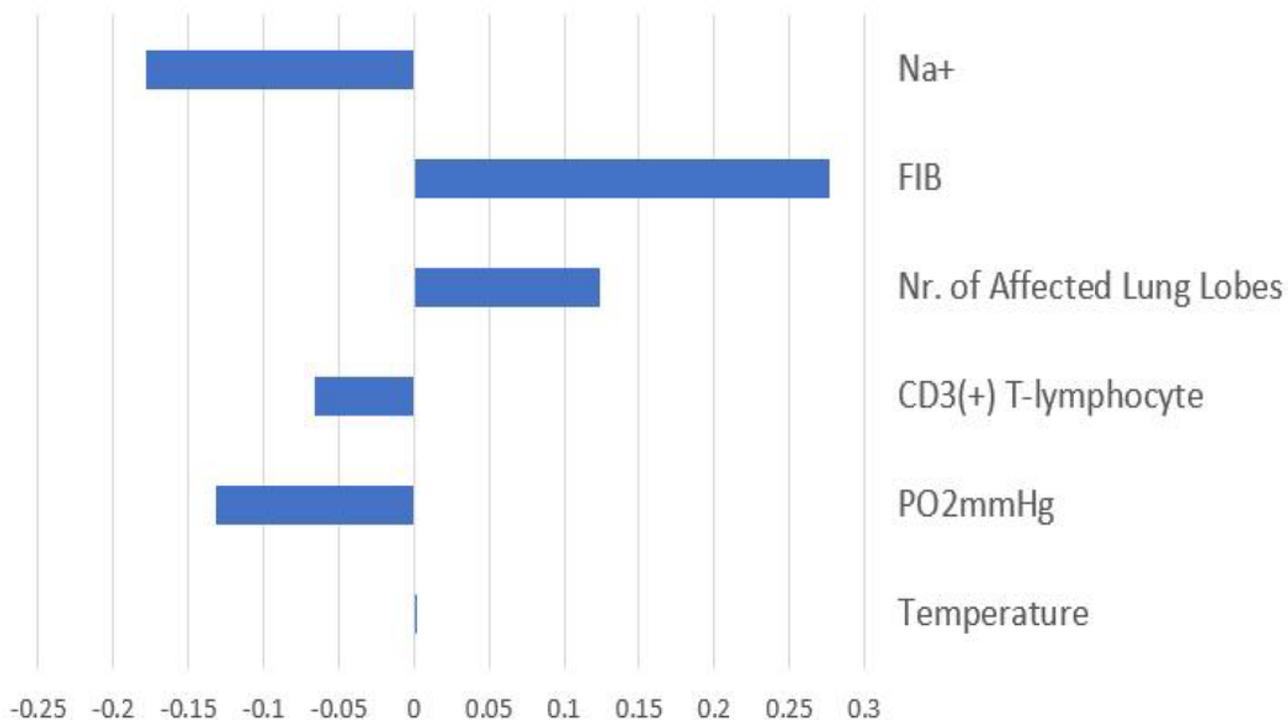


Fig.4:Lasso Coefficients for each selected factors in the Lasso models.
Bar chart lists the selected factors on the y-axis with corresponding coefficients on the x-axis. Generally, the higher the absolute value of the coefficient, the greater contribution to the model.

Figure 4

Lasso Coefficients for each selected factors in the Lasso models. Bar chart lists the selected factors on the y-axis with corresponding coefficients on the x-axis. Generally, the higher the absolute value of the coefficient, the greater contribution to the model.

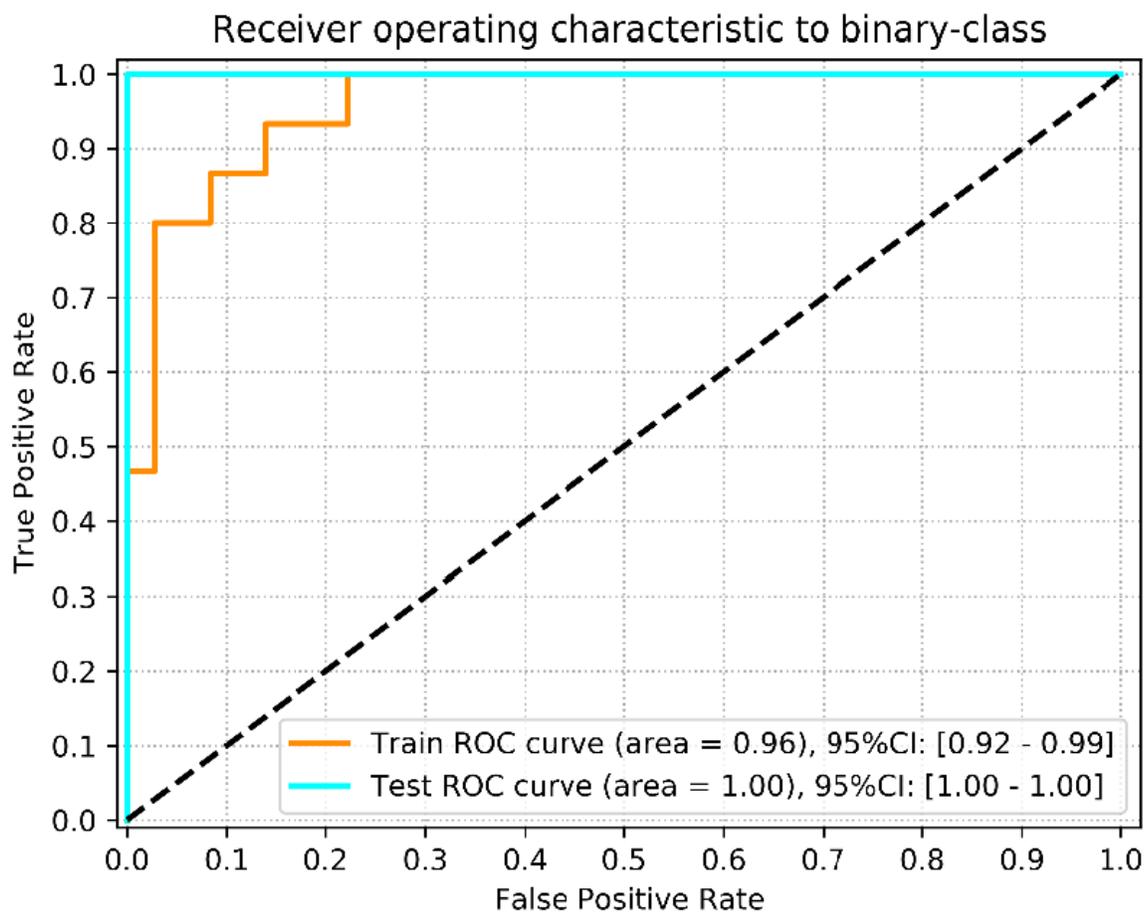


Fig:5: ROC Curve of prediction model based on selected factors.

ROC curves of Logistic regression model in train dataset (AUC, 0.96; 95% CI, 0.92-0.99; Sensitivity, 0.93; Specificity, 0.96) and test dataset (AUC, 1; 95% CI, 1-1; Sensitivity, 1; Specificity, 1) . ROC(Receiver Operating Characteristic) .

Figure 5

ROC Curve of prediction model based on selected factors. ROC curves of Logistic regression model in train dataset (AUC, 0.96; 95% CI, 0.92-0.99; Sensitivity, 0.93; Specificity, 0.96) and test dataset (AUC, 1; 95% CI, 1-1; Sensitivity, 1; Specificity, 1) . ROC(Receiver Operating Characteristic) .

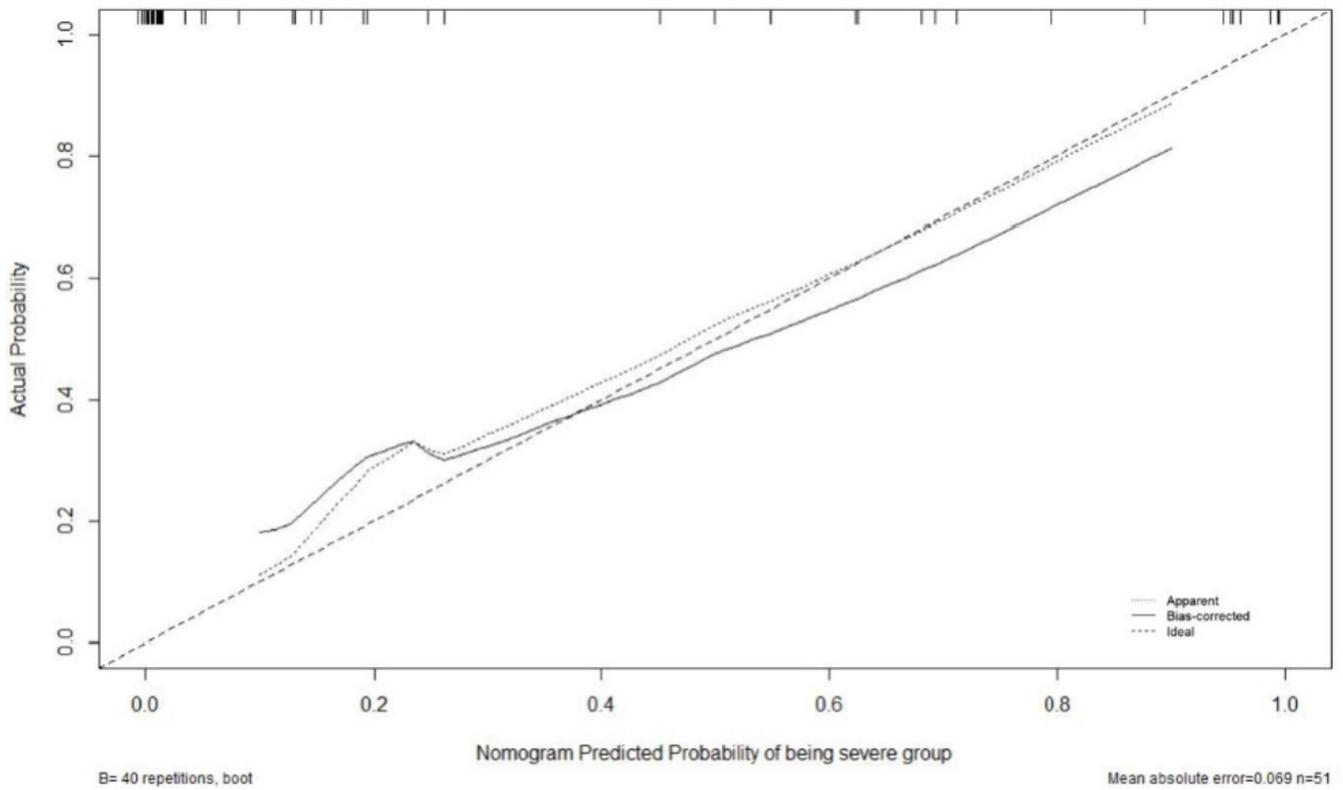


Fig.6: The calibration curve of clinical nomogram model to predict the progress to severe COVID-19 . Nomogram-predicted probability of progress to severe group is plotted on the x-axis; actual probability is plotted on the y-axis.

Figure 6

The calibration curve of clinical nomogram model to predict the progress to severe COVID-19 . Nomogram-predicted probability of progress to severe group is plotted on the x-axis; actual probability is plotted on the y-axis.