


# A composite risk model predicts disease progression in early stages of COVID-19: A propensity score-matched cohort study

Jianjun Xu<sup>1,\*</sup>, Yang Gao<sup>1,\*</sup>, Shaobo Hu<sup>1,2,\*</sup>, Suzhen Li<sup>3</sup>, Weimin Wang<sup>1</sup>, Yuzhe Wu<sup>1</sup>, Zhe Su<sup>1</sup>, Xing Zhou<sup>1</sup>, Xiang Cheng<sup>1</sup> and Qichang Zheng<sup>1</sup> 

## Abstract

**Background:** Recently, studies on COVID-19 have focused on the epidemiology of the disease and clinical characteristics of patients, as well as on the risk factors associated with mortality during hospitalization in critical COVID-19 cases. However, few research has been performed on the prediction of disease progression in particular group of patients in the early stages of COVID-19.

**Methods:** The study included 338 patients with COVID-19 treated at two hospitals in Wuhan, China, from December 2019 to March 2020. Predictors of the progression of COVID-19 from mild to severe stages were selected by the logistic regression analysis.

**Results:** COVID-19 progression to severe and critical stages was confirmed in 78 (23.1%) patients. The average value of the neutrophil-to-lymphocyte ratio (NLR) was higher in patients in the disease progression group than in the improvement group. Multivariable logistic regression analysis revealed that elevated NLR, LDH and IL-10 were independent predictors of disease progression. The optimal cut-off value of NLR was 3.75. The values of the area under the curve, reflecting the accuracy of predicting COVID-19 progression by NLR was 0.739 (95%CI: 0.605–0.804). The risk model based on NLR, LDH and IL-10 had the highest area under the ROC curve.

**Conclusions:** The performed analysis demonstrates that high concentrations of NLR, LDH and IL-10 were independent risk factors for predicting disease progression in patients at the early stage of COVID-19. The risk model combined with NLR, LDH and IL-10 improved the accuracy of the prediction of disease progression in patients in the early stages of COVID-19.

## Keywords

COVID-19, SARS-CoV-2, neutrophil-to-lymphocyte ratio, systemic immune-inflammation index, propensity score matching, predictor, risk model

Accepted: 17th March 2021

## Introduction

The novel coronavirus, named by the World Health Organization (WHO) as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was first identified in the city of Wuhan, China, in December 2019.<sup>1</sup>

<sup>1</sup>Department of Hepatobiliary Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

<sup>2</sup>Department of General Surgery, The People's Hospital of Honghu City, Honghu, China

<sup>3</sup>Department of Gastroenterology, Wuhan Asia General Hospital, Wuhan, China

\*Co-first author.

## Corresponding author:

Qichang Zheng, Department of Hepatobiliary Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, No.1277, Liberation Avenue, Jiangnan District, Wuhan 430022, China.  
Email: zqc\_wh@yahoo.com

Given the rapid outbreak of the coronavirus disease 2019 (COVID-19) caused by SARS-CoV-2 in China and its rapid spread of COVID-19 pandemic worldwide, global concerns have emerged.<sup>2</sup> The extent and burden of COVID-19 are particularly evident in countries lacking adequate medical resources. As of 24 April 2020, 84,312 patients have been diagnosed with COVID-19 in China, and 4642 (5.5%) of them died.

The analysis of 44,672 confirmed cases of COVID-19 in China revealed that the majority of SARS-CoV-2 infections were mild or general type, while 889 (1.2%) were asymptomatic, 13.8% were severe and 4.7% were critical cases.<sup>3</sup> The mortality rate in critical COVID-19 cases is approximately 49%.<sup>3</sup> Severe patients usually develop dyspnoea or hypoxaemia one week after the onset of the disease. Critical COVID-19 patients develop rapidly acute respiratory distress syndrome, septic shock, refractory metabolic acidosis, coagulation disorders and multi-organ failure.<sup>4</sup> In the early days of COVID-19 spread in Wuhan, patients with severe or critical COVID-19 were admitted to a hospital, while patients with mild or general COVID-19 were advised to be isolated at home. With increased awareness of the spread and progression of COVID-19, efforts have been made throughout China to admit and treat all confirmed or suspected COVID-19 patients. It has been reported that patients in Guangzhou, China, had a better prognosis than in Wuhan. Patients in Wuhan were more likely to be admitted to an ICU and had higher mortality.<sup>5</sup> This difference highlights the importance of appropriate care and treatment in early mild cases, especially in patients at high risk of disease progression. Early detection of risk factors for progression to severe illness can help to provide adequate supportive care and treatment, decreasing the number of patients that develop the severe condition, reducing mortality and alleviating the shortages of medical resources.

The neutrophil-to-lymphocyte ratio (NLR) and systemic immune-inflammation index (SII), both easily calculated from a routine blood test, are indicators of inflammation and immune response.<sup>2,6</sup> Elevated NLR represents a risk factor for mortality not only from infectious diseases, but also from malignancies, intracerebral haemorrhage and dermatomyositis.<sup>7-9</sup> Similarly, increased SII has a prognostic value in a variety of malignancies, including hepatocellular carcinoma, breast cancer and oesophageal squamous cell carcinoma.<sup>10-12</sup> It has been documented that severe and critical COVID-19 cases tend to have higher neutrophil counts and lower lymphocyte counts.<sup>1,13</sup> Whether NLR or SII could be an independent predictor of COVID-19 progression in the early (mild or moderate) stages of the disease remains to be determined.

Thus far, studies on COVID-19 have focused on the epidemiology of the disease and clinical characteristics of patients,<sup>14-19</sup> as well as on the risk factors associated with mortality during hospitalization in critical COVID-19 cases. However, no research has been performed on the prediction of progression in patients in the early stages of the disease. To address this lack of data, the present research was designed to determine whether baseline values of NLR and SII in patients with mild or moderate COVID-19 can be used as reliable predictors of the progression of the disease in its early stages.

## Methods

### Study design and participants

The present investigation was designed as a retrospective study involving a total of 476 COVID-19 patients admitted consecutively to the Union Hospital of Huazhong University of Science and Technology and the Wuhan Asia General Hospital, from December 2019 to March 2020. The clinical outcomes, discharge from hospital or death in hospital were recorded up to 31 March 2020. Both hospitals are located in Wuhan, Hubei Province, and were designated by the Chinese government as hospitals responsible for the treatment of COVID-19 patients. The disease was diagnosed according to the WHO guidelines.

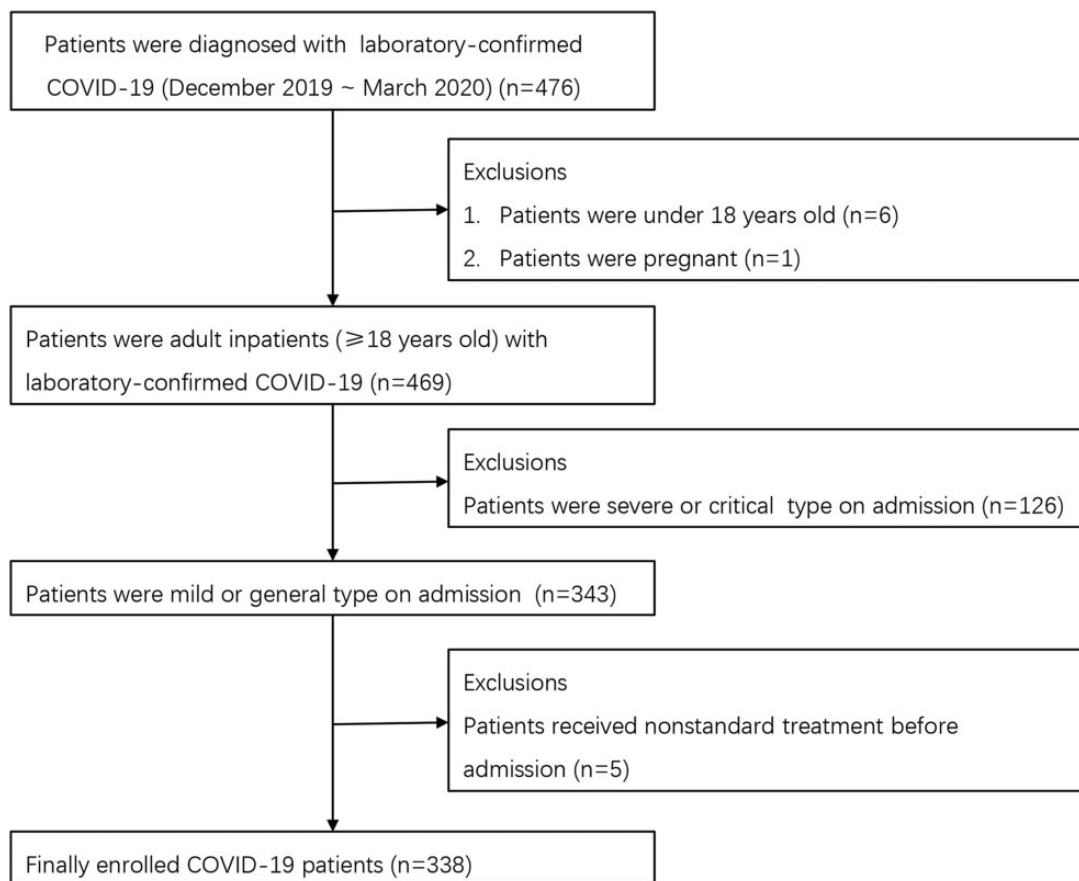
Exclusion criteria were as follows: (1) patients that were pregnant or under 18 years old; (2) patients with severe or critical COVID-19 at admission; (3) patients receiving nonstandard treatment before the admission (Figure 1). Based on the exclusion criteria, 338 COVID-19 patients were included in the final analyses.

All patients admitted to the hospital with mild or moderate COVID-19 were divided into the improvement group and the progression group. The improvement group included patients who recovered after admission and were discharged from the hospital. The progression group included patients who after the admission progressed to severe or critical condition, or died in the hospital, regardless whether the final outcome was discharged or death.

The protocol of the study was approved by the Research Ethics Committee of the Tongji Medical College. The data used in the study were anonymized, and the requirement for informed consent was waived by the Committee.

### Data collection

Complete clinical data for all COVID-19 patients were collected from the medical records of the patients. They included demographic, clinical, laboratory, imaging,



**Figure 1.** Flow diagram of patient disposition: A total of 338 COVID-19 patients were enrolled in this study.

treatment and outcome information. The laboratory data and CT scan of the lung were collected corresponded to the results of the first test performed upon admission. Prior to the analysis, the patient information was de-identified and anonymized. The NLR was calculated from the results of neutrophil and lymphocyte counts. SII was calculated according to the results of a routine blood test, according to the formula:  $SII = N \times P/L$ , where N, P and L represents the count of neutrophils, platelets and lymphocytes, respectively.<sup>6</sup>

### Laboratory examination

Laboratory confirmation of the infection of SARS-CoV-2 was conducted by local CDC in accordance with the Chinese CDC protocol. Pharyngeal swab samples were collected from all patients and the samples were stored in a viral-transport medium for laboratory testing. Real-time RT-PCR was performed to exclude infection with other respiratory viruses including influenza A virus, coxsackie virus, influenza B virus, respiratory syncytial virus, enterovirus and parainfluenza

virus. Automatic biochemical analyser, AU5800 (Beckman Coulter, USA), was used for measuring the concentrations of ALT, AST, albumin and creatinine. Liquid Assayed Multiquel was performed in QC procedures. Automated blood analyser, XE-2100 (Sysmex, Japan), was used for measuring the count of white blood cells, neutrophils, lymphocytes and platelets. The lymphocyte test kit (Beckman Coulter Inc., FL, USA) was used for lymphocyte subset analysis. Plasma cytokines (IL10, IL6, IL4, IL2, IFN- $\gamma$  and TNF- $\alpha$ ) were detected using the human Th1/2 cytokine kit II (BD Ltd, Franklin lakes, NJ, USA). All laboratory tests were conducted in accordance with the product manual.

### Study definitions

The classification of the severity of COVID-19 was based on the “Diagnosis and Treatment Protocol for COVID-19 (Trial Version 7)” (4). The clinical classifications of illness severity of COVID-19 were as follows: (1) mild type, with mild clinical symptoms and the absence of signs of pneumonia on imaging; (2)

moderate type, with fever, respiratory tract symptoms and signs of pneumonia were identified on imaging; (3) severe type, characterized by one of the following: (a) respiratory distress, respiratory rate  $\geq 30$  breaths/min; (b) mean oxygen saturation  $\leq 93\%$  in the resting state; (c) oxygenation index  $\leq 300$  mmHg; and (4) critical type, characterized by one of the following: (a) shock; (b) respiratory failure requiring mechanical ventilation; (c) organ failure requiring ICU admission. Fever was defined as an axillary temperature of at least  $37.3^{\circ}\text{C}$ .<sup>20</sup> The duration of viral shedding was defined as the time from the onset of illness to the second negative nucleic acid test.<sup>4</sup>

### Statistical analysis

Summary statistics of the demographic data, clinical characteristics, laboratory results and radiographic findings were expressed as median and the interquartile range (IQR) for continuous variables and as frequencies and proportions for categorical variables. If the variance in the improvement group and the progression group was the same, continuous variables were compared by the Student's *t*-test; otherwise, the Welch's *t*-test was used. Categorical variables were compared by the  $\chi^2$  test or Fisher's exact test.

The relationship between the NLR treated as a continuous variable, and the progression of COVID-19 was examined first; subsequently, the relationship was evaluated considering NLR as a categorical variable according to the best threshold value. These relationships were examined using univariate and multivariate logistic regression analyses, and odds ratio (OR) and 95% confidence interval (CI) were calculated. After univariate logistic regression analysis using the backward stepwise procedure (LR), only the variables with a *P*-value of less than 0.1 were considered for multivariate analysis to identify predictors of progression of COVID-19. To avoid overfitting in multivariate analysis, six of those variables were selected on the basis of clinical constraints and previous studies. Previous investigations have documented that older age was associated with poor prognosis in COVID-19 patients.<sup>20</sup> Recent studies indicated that patients with severe COVID-19 had higher C-reactive protein (CRP) and lactate dehydrogenase (LDH) concentrations than patients with non-severe COVID-19.<sup>13</sup> Moreover, in comparison with patients with mild COVID-19, the concentrations of IL-10 and IFN $\gamma$  in patients with severe COVID-19 were significantly increased in the early stage of the disease, and most detected cytokines peaked in the serum three to six days after the onset.<sup>21,22</sup> Therefore, NLR, CRP, LDH, interleukin-10 (IL-10) and interferon  $\gamma$  (IFN $\gamma$ ) were selected as the

six variables for the multivariable logistic regression analysis.

To further validate the association between elevated NLR and disease progression, propensity score matching (PSM) was used to eliminate confounding bias.<sup>23</sup> All COVID-19 patients were divided into two groups, the high NLR ( $\geq 3.75$ ) group and the low NLR ( $< 3.75$ ) group, based on the best threshold value of NLR predicting disease progression. The propensity score was calculated using the logistic regression model in which age, gender, hypertension, coronary heart disease, diabetes, chronic obstructive pulmonary disease, fever and cough were considered. The matching was performed using a 1:1 ratio. By this approach, patients in the high NLR group were matched with patients in the low NLR group having the closest propensity score. Patients selected by PSM were enrolled in a new cohort and subjected to further analysis of the association between NLR elevation and disease progression. The best threshold value of NLR was calculated according to the Youden index that was the sum of sensitivity and specificity minus 1. The two-sided *P*-value of less than 0.05 was considered to indicate a statistically significant difference. SPSS 24.0 software was used to perform PSM and for statistical analysis.

## Results

### Patient characteristics

A total of 338 COVID-19 patients, in which the diagnosis was confirmed by the detection of SARS-CoV-2 RNA, were enrolled (Figure 1). Their demographic, clinical, laboratory, radiographic and outcomes data are listed in Table 1. Six patients died during hospitalization, and 332 were discharged. The median age of the patients was 57.5 years (IQR 35.0–68.0), and the majority were female (56.8%). Comorbidities were present in 154 (45.6%) patients, with hypertension being the most common comorbidity, followed by diabetes and coronary heart disease. The most frequent symptoms on admission were fever, followed by cough, fatigue, chest distress and myalgia. In routine blood tests, the median values of neutrophils, lymphocytes, platelet, SII and NLR were  $3.05 (\times 10^9/\text{L})$ ,  $1.30 (\times 10^9/\text{L})$ ,  $205.50 (\times 10^9/\text{L})$ ,  $503.43 (\times 10^9/\text{L})$ , 2.31, respectively.

### Relationship between clinical characteristics and COVID-19 progression

Seventy-eight (23.1%) COVID-19 patients developed disease progression after the admission, and six (7.7%) of them died while hospitalized (Table 1). Table 1 also indicates that patients in the progression

**Table 1.** Baseline characteristics of COVID-19 patients (*n* = 338).

Covariate	Total ( <i>n</i> = 338)	Progressed group ( <i>n</i> = 78)	Improved group ( <i>n</i> = 260)	<i>P</i> value
Demographics and clinical characteristics				
Age, years	57.5 (35–68)	62.5 (41–70)	56.0 (34–67)	0.024
Gender				0.548
Female	192 (56.8)	42 (53.8)	150 (57.7)	
Male	146 (43.2)	36 (46.2)	110 (42.3)	
Comorbidity	154 (45.6)	38 (48.7)	116 (44.6)	0.523
Hypertension	90 (26.6)	24 (30.8)	66 (25.4)	0.345
Coronary heart disease	32 (9.5)	6 (7.7)	26 (10.0)	0.541
Diabetes	36 (10.7)	8 (10.3)	28 (10.8)	0.898
Chronic obstructive lung disease	14 (4.1)	4 (5.1)	10 (3.8)	0.618
Other	82 (24.3)	20 (25.6)	62 (23.8)	0.746
Fever (temperature $\geq 37.3^{\circ}\text{C}$ )	258 (76.3)	72 (92.3)	186 (71.5)	<0.001
Sore-throat	68 (20.1)	14 (17.9)	54 (20.8)	0.586
Nasal congestion	24 (7.1)	6 (7.7)	18 (6.9)	0.817
Cough	208 (61.5)	58 (74.4)	150 (57.7)	0.008
Chest distress	132 (39.1)	50 (64.1)	82 (31.5)	<0.001
Fatigue	180 (53.3)	48 (61.5)	132 (50.8)	0.095
Myalgia	118 (34.9)	36 (46.2)	82 (31.5)	0.018
Headache	70 (20.7)	20 (25.6)	50 (19.2)	0.220
Nausea or vomiting	28 (8.3)	16 (20.5)	12 (4.6)	<0.001
Diarrhoea	94 (27.8)	28 (35.9)	66 (25.4)	0.069
Palpitate	38 (11.2)	22 (28.2)	16 (6.2)	<0.001
Chest pain	14 (4.1)	2 (2.6)	12 (4.6)	0.425
Laboratory findings on admission				
White blood cell count, $\times 10^9/\text{L}$	5.14 (3.90–6.66)	4.90 (3.76–6.87)	5.17 (4.11–6.56)	0.431
<4, $\times 10^9/\text{L}$	85 (25.1)	23 (29.5)	62 (23.8)	0.314
>10, $\times 10^9/\text{L}$	10 (3.0)	2 (2.6)	8 (3.1)	0.815
Neutrophil count, $\times 10^9/\text{L}$	3.05 (2.23–4.28)	3.34 (2.52–5.12)	2.92 (2.17–4.19)	0.014
Lymphocyte count, $\times 10^9/\text{L}$	1.30 (0.97–1.63)	0.93 (0.68–1.21)	1.40 (1.09–1.80)	<0.001
NLR	2.31 (1.52–3.82)	3.90 (2.27–5.56)	2.08 (1.43–3.27)	<0.001
Platelet count, $\times 10^9/\text{L}$	205.50 (161.75–269.75)	198.50 (157.50–272.50)	206.00 (163.00–268.75)	0.656
SII, $\times 10^9/\text{L}$	503.43 (289.26–878.25)	807.86 (409.93–1478.95)	453.16 (261.10–772.36)	<0.001
Haemoglobin, g/ dL	12.70 (11.50–13.60)	12.80 (11.10–13.60)	12.70 (11.60–13.60)	0.348
ALT, U/L	23 (16.75–36)	27 (20–38)	21 (16–33)	0.530
AST, U/L	23 (19–33)	29 (23–39)	21.50 (17–29)	0.003
Albumin, g/L	37.30 (33.08–40.80)	35.40 (28.90–40.20)	37.75 (33.70–41.00)	<0.001
Creatinine, $\mu\text{mol/L}$	64.70 (56.20–78.13)	70.00 (58.30–82.60)	63.50 (54.40–77.40)	0.904
BUN, mmol/L	3.93 (3.15–4.93)	3.90 (3.00–5.29)	3.94 (3.21–4.75)	0.898
Creatine kinase, U/L	47 (28.25–73.25)	66 (42.00–109)	44 (1.60–66)	<0.001
LDH, U/L	213 (173.75–261.25)	243 (216–332)	193 (160–245)	<0.001
CRP, mg/L	28.69 (8.39–30.38)	28.69 (10.50–57.30)	28.69 (7.81–28.69)	<0.001
ESR, mm/h	35 (14–51)	35 (25–60)	27 (12–45)	0.002
IL-2, pg/mL	2.70 (2.53–2.81)	2.70 (2.50–2.74)	2.70 (2.53–2.81)	0.829
IL-4, pg/mL	2.14 (1.92–2.31)	2.14 (2.10–2.26)	2.14 (1.90–2.31)	0.385
IL-6, pg/mL	13.74 (4.97–16.71)	16.71 (7.93–16.71)	12.28 (4.44–16.71)	0.062

(continued)

Table 1. Continued

Covariate	Total (n = 338)	Progressed group (n = 78)	Improved group (n = 260)	P value
IL-10, pg/mL	4.20 (3.44–4.33)	4.20 (4.20–5.14)	4.11 (3.34–4.21)	<0.001
TNF $\alpha$ , pg/mL	3.32 (2.21–4.27)	2.48 (2.16–4.27)	3.41 (2.22–4.27)	0.173
IFN $\gamma$ , pg/mL	2.34 (1.98–2.44)	2.34 (2.21–2.50)	2.34 (1.96–2.43)	0.001
CD3, $\times 10^9$ per L	0.92 (0.69–1.22)	0.66 (0.49–0.86)	1.01 (0.77–1.32)	<0.001
CD4, $\times 10^9$ per L	0.55 (0.42–0.74)	0.42 (0.30–0.52)	0.60 (0.47–0.87)	<0.001
CD8, $\times 10^9$ per L	0.29 (0.21–0.42)	0.24 (0.13–0.31)	0.32 (0.22–0.45)	<0.001
CD4/CD8	1.85 (1.39–2.59)	1.75 (1.26–2.48)	1.85 (1.40–2.61)	0.751
Duration of viral shedding after COVID-19 onset, days	20 (17–26)	20.50 (18–25)	20 (17–26.75)	0.912
Imaging features				
Consolidation	172 (50.9)	58 (74.4)	114 (43.8)	<0.001
Ground-glass opacity	251 (74.3)	61 (78.2)	190 (73.1)	0.364
Bilateral patchy shadowing	261 (77.2)	66 (84.6)	195 (75.0)	0.076
Outcomes				
Death	6/338 (1.8)	6/78 (7.7)	0/260 (0)	NS
Discharge	332/338 (98.2)	72/78 (92.3)	260/260 (100)	

NLR: neutrophil-to-lymphocyte ratio; SII: systemic immune-inflammation index; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BUN: blood urea nitrogen; LDH: lactate dehydrogenase; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; IL: interleukin; TNF $\alpha$ : tumor necrosis factor  $\alpha$ ; IFN $\gamma$ : interferon  $\gamma$ .

Note: Data are median (IQR), n (%), or n/N (%). P values were calculated by t-test, Welch's t-test,  $\chi^2$  test, or Fisher's exact test, as appropriate.

group were 6.5 years older than in the improvement group (62.5 vs. 56.0 years,  $P=0.024$ ). There was no significant difference in the ratio of males to females between the two groups ( $P=0.548$ ). Compared with patients in the improvement group, patients in the progression group had a higher rate of developing a fever, chest distress, myalgia, nausea or vomiting and palpitation (all  $P<0.05$ ). They also had a higher count of neutrophils (3.34 vs.  $2.92 \times 10^9/L$ ,  $P=0.014$ ), lower count of lymphocytes (0.93 vs.  $1.40 \times 10^9/L$ ,  $P<0.001$ ) and higher NLR (3.90 vs. 2.08,  $P<0.001$ ) and SII ( $807.86$  vs.  $453.16 \times 10^9/L$ ,  $P<0.001$ ) values. In addition, they had higher concentrations of creatine kinase, LDH, CRP, ESR, IL-10 and IFN $\gamma$  (all  $P<0.05$ ) on admission. The incidence of lung consolidation detected by imaging was higher in the progression group (74.4% vs. 43.8%,  $P<0.001$ ).

### Risk factors associated with COVID-19 progression

To identify the predictors of COVID-19 progression, multiple clinical parameters were evaluated by logistic regression analysis (Table 2). In univariable analysis, the odds of disease progression were higher in patients with fever or cough (Table 2). Age, elevated neutrophil count, reduced lymphocyte count, elevated NLR, SII, creatine kinase, LDH, CRP, ESR, IL-10 and IFN $\gamma$  were also associated with disease progression.

To avoid overfitting in multivariate analysis, age, NLR, CRP, LDH, IL-10 and IFN $\gamma$  were selected for multivariate regression analysis, as justified in the Methods section. Multivariable logistic regression analysis documented that increased NLR, LDH and IL-10 were independent predictors of disease progression in COVID-19 patients with ORs of 1.183 (95% CI: 1.045–1.339), 1.007 (95% CI: 1.003–1.010) and 1.344 (95% CI: 1.080–1.672), respectively.

A risk model was established based on NLR, LDH and IL-10, and receiver operating characteristic (ROC) curves were performed to assess the value of risk model and other single parameters (Figure 2). The values of the area under the ROC curve (AUROC), reflecting the accuracy of predicting COVID-19 progression, were 0.739 (95% CI: 0.605–0.804) for NLR, superior to other single parameters (Table S1). Moreover, the risk model based on NLR, LDH and IL-10 had the highest AUROC (Table S1).

The ROC curve analysis indicated that the optimal cut-off values for predicting the progression of COVID-19 were 3.75 for NLR, 213.5 for LDH, 4.15 for IL-10, respectively. Table 3 presents sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) (95% CI) of risk model and other single parameters. The Yoden index of risk model was superior to other single parameters (Table 3).

**Table 2.** Risk factors associated with COVID-19 progression.

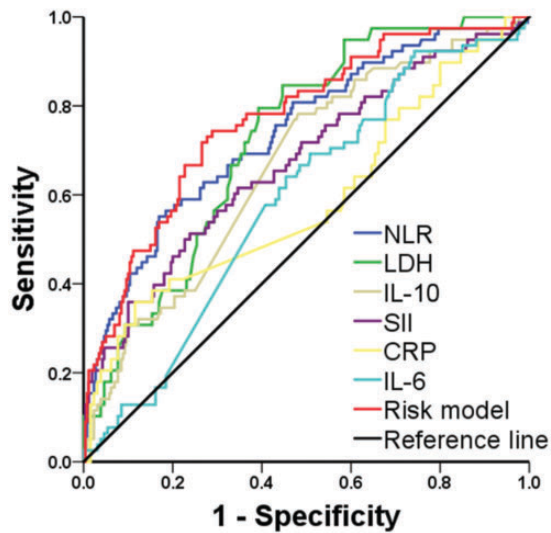
Covariate	Univariable OR (95% CI)	P value	Multivariable OR (95% CI)	P value
<b>Demographics and clinical characteristics</b>				
Age, years	1.017 (1.002–1.033)	0.025		
Gender	0.856 (0.514–1.423)	0.548		
Hypertension	1.306 (0.749–2.278)	0.346		
Coronary heart disease	0.750 (0.297–1.894)	0.543		
Diabetes	0.947 (0.413–2.172)	0.898		
Chronic obstructive lung disease	1.351 (0.412–4.434)	0.619		
Fever (temperature $\geq 37.3^{\circ}\text{C}$ )	4.774 (1.990–11.456)	<0.001		
Cough	2.127 (1.209–3.740)	0.009		
<b>Laboratory findings on admission</b>				
White blood cell count, $\times 10^9/\text{L}$	0.946 (0.825–1.085)	0.430		
<4, $\times 10^9/\text{L}$	1.335 (0.760–2.348)	0.315		
>10, $\times 10^9/\text{L}$	0.829(0.172–3.987)	0.815		
Neutrophil count, $\times 10^9/\text{L}$	1.205 (1.037–1.400)	0.015		
Lymphocyte count, $\times 10^9/\text{L}$	0.072 (0.033–0.157)	<0.001		
NLR	1.299 (1.153–1.464)	<0.001	1.183 (1.045–1.339)	0.008
Platelet count, $\times 10^9$ per L	0.999 (0.996–1.002)	0.656		
SII, $\times 10^9/\text{L}$	1.001 (1.000–1.001)	<0.001		
Haemoglobin, g/dL	0.992 (0.976–1.008)	0.348		
ALT, U/L	1.003 (0.993–1.013)	0.530		
AST, U/L	1.019 (1.005–1.034)	0.006		
Albumin, g/L	0.902 (0.857–0.950)	<0.001		
Creatinine, $\mu\text{mol/L}$	1.000 (0.992–1.007)	0.904		
BUN, mmol/L	1.008 (0.887–1.146)	0.898		
Creatine kinase, U/L	1.009 (1.005–1.014)	<0.001		
LDH, U/L	1.009 (1.005–1.012)	<0.001	1.007 (1.003–1.010)	<0.001
CRP, mg/L	1.016 (1.007–1.025)	0.001		
ESR, mm/h	1.016 (1.006–1.026)	0.002		
IL-6, pg/mL	1.010 (0.999–1.021)	0.077		
IL-10, pg/mL	1.487 (1.229–1.799)	<0.001	1.344 (1.080–1.672)	0.008
TNF $\alpha$ , pg/mL	0.932 (0.841–1.033)	0.180		
IFN $\gamma$ , pg/mL	1.699 (1.130–2.556)	0.011		
CD3, $\times 10^9$ per L	0.034 (0.012–0.092)	<0.001		
CD4, $\times 10^9$ per L	0.008 (0.002–0.039)	<0.001		
CD8, $\times 10^9$ per L	0.006 (0.001–0.050)	<0.001		
CD4/CD8	1.022 (0.892–1.172)	0.751		

OR: odds ratio. NLR: neutrophil-to-lymphocyte ratio; SII: systemic immune-inflammation index; ALT: alanine aminotransferase; AST: aspartate aminotransferase; BUN: blood urea nitrogen; LDH: lactate dehydrogenase; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; IL: interleukin; TNF $\alpha$ : tumor necrosis factor  $\alpha$ ; IFN $\gamma$ : interferon  $\gamma$ .

### Risk factors associated with COVID-19 progression analysed after propensity score matching

To further validate the association between elevated NLR and disease progression, PSM was used to reduce confounding bias. The patients were divided into two groups based on the best threshold value of

NLR predicting disease progression: the high NLR ( $\geq 3.75$ ) group and the low NLR ( $< 3.75$ ) group. The high NLR group had a higher incidence of COVID-19 progression than the low NLR group ( $P < 0.001$ ). Seventy-four propensity score-matched pairs of COVID-19 patients were selected and enrolled in a new cohort. However, a statistically significant



**Figure 2.** ROC curve: Predicting COVID-19 progression with risk model and other single parameters.

difference in COVID-19 progression continued to be present between the high and low NLR groups ( $P < 0.001$ ) (Table 4). After univariate analysis, the variables with  $P < 0.1$ , including age, NLR, LDH, CRP, IL-10 and  $\text{IFN}\gamma$ , were selected for multivariate regression analysis using the backward stepwise procedure (LR). This analysis confirmed that  $\text{NLR} \geq 3.75$  was still an independent predictor of disease progression in COVID-19 patients (Table S2).

## Discussion

The present study has demonstrated that NLR, measured at the early stage of the disease, is a significant independent predictor for the progression of COVID-19 to severe and critical stages. Moreover, the finding that  $\text{NLR} \geq 3.75$  is an independent predictor of disease progression was also confirmed in the propensity score-matched cohort. The PPV of NLR is actually quite poor. Moreover, the risk model combined with NLR, LDH and IL-10 enlarged the accuracy of the prediction of COVID-19 progression, with AUROC of 0.77 (95% CI: 0.71–0.83). The relevance of early adequate treatment to prevent mild or moderate cases from developing into severe ones is well-recognized since the treatment of critical COVID-19 patients requires not only significant medical resources but, most importantly, results in a high mortality rate.<sup>20</sup> When thinking about COVID, there is a spectrum of asymptomatic to very severe illness and the precise point on the journey is very time dependent. So getting a cohort who are in a similar clinical state on the journey to more severe illness is difficult, especially when the criteria for

hospitalization depends on signs and symptoms. So the rate of disease progression will also affect this. One person might have a slow disease progression and present late with mild symptoms, another might have rapid disease progression and present earlier with mild or moderate symptoms.

Although previous studies reported several models for predicting disease progression to a severe or critical state,<sup>24–34</sup> both mild and severe COVID-19 patients were included in these papers dealing with COVID-19 prediction of disease progression to very severe illness.<sup>24–34</sup> It seems unreasonable to apply a risk model based on the analysis of the parameters of severe COVID-19 patients to mild COVID-19 patients. Consequently, patients admitted to the hospital with mild or moderate COVID-19 were included in this study, and patients with severe or critical COVID-19 at admission were excluded. All patients admitted to the hospital with mild or moderate COVID-19 were divided into the improvement group and the progression group. The laboratory data and CT scan of the lung were collected corresponded to the results of the first test performed upon admission. This current analysis provides a new evidence-based risk model for early identification of mild or moderate COVID-19 patients at high risk of COVID-19 progression.

Previous studies have documented that older age is a significant independent predictor of mortality in MERS and SARS.<sup>35,36</sup> The current investigation has found that older age is also associated with increased odds of in-hospital death in COVID-19 patients.<sup>37</sup> However, age is not an independent predictor of the progression of COVID-19 to severe or critical stages. COVID-19 is commonly susceptible in the population, and there is a risk of progression to severe disease.<sup>4</sup> A possible explanation of this apparent discrepancy might be that older patients with severe or critical COVID-19 have a higher risk of death. Since patients with severe or critical COVID-19 on admission were not included in the study, the fatality rate among the analysed population was lower than that documented in previous reports.<sup>37</sup>

In agreement with previous research,<sup>4</sup> the present analysis has shown that an increased concentration of LDH at the early stage of COVID-19 was also a significant independent predictor for disease progression. Moreover, COVID-19 patients in the progression group had higher concentrations of IL-10 and IL-6 at the early stage of the disease, although differences in IL-6 between the groups were not statistically significant. However, IL-10 at the early stage was also a significant independent predictor for COVID-19 progression. The identification of elevated concentrations of inflammatory factors in COVID-19 patients is consistent with previous studies.<sup>38</sup> The current work



**Table 3.** Sensitivity, specificity and negative and positive predictive values of a NLR of 3.75 and LDH of 213.5 U/L in the prediction of COVID-19 progression.

Variable	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)
NLR	0.55 (0.43–0.66)	0.83 (0.77–0.87)	0.86 (0.81–0.90)	0.49 (0.39–0.60)
LDH	0.79 (0.69–0.87)	0.61 (0.55–0.67)	0.91 (0.85–0.94)	0.38 (0.30–0.46)
IL-10	0.78 (0.67–0.86)	0.52 (0.46–0.58)	0.89 (0.82–0.93)	0.33 (0.26–0.40)
Risk model	0.74 (0.63–0.82)	0.71 (0.65–0.76)	0.90 (0.85–0.94)	0.44 (0.35–0.52)

NLR: neutrophil-to-lymphocyte ratio; LDH: lactate dehydrogenase; IL: interleukin; NPV: negative predictive value; PPV: positive predictive value.

**Table 4.** Comparison of clinical parameters between the low and high NLR groups.

Covariate	Unmatched cohort			Propensity score-matched cohort		
	NLR < 3.75 (n = 251)	NLR ≥ 3.75 (n = 87)	P value	NLR < 3.75 (n = 74)	NLR ≥ 3.75 (n = 74)	P value
Age	53 (35–66)	67 (42–75)	<0.001	63.5 (39–67)	62.5 (39.3–70)	0.741
Gender			0.002			0.742
Female	155 (61.8)	37 (42.5)		36 (48.6)	34 (45.9)	
Male	96 (38.2)	50 (57.5)		38 (51.4)	40 (54.1)	
Hypertension	61 (24.3)	29 (33.3)	0.101	22 (29.7)	25 (33.8)	0.596
Coronary heart disease	23 (9.2)	9 (10.3)	0.746	8 (10.8)	9 (12.2)	0.797
Diabetes	20 (8)	16 (18.4)	0.007	11 (14.9)	9 (12.2)	0.631
Chronic obstructive lung disease	10 (4)	4 (4.6)	0.804	4 (5.4)	2(2.7)	0.405
Fever	185 (73.7)	73 (83.9)	0.054	65 (87.8)	61 (82.4)	0.355
Cough	158 (62.2)	52 (59.8)	0.694	46 (62.2)	44 (59.5)	0.736
COVID-19 progression	35 (13.9)	43 (49.4)	<0.001	11 (14.9)	35 (47.3)	<0.001

Note: Data are median (IQR) or n (%). P values were calculated by t-test, Welch's t-test,  $\chi^2$  test, or Fisher's exact test, as appropriate. NLR: neutrophil-to-lymphocyte ratio.

documented that, except for IL-6, the serum concentration of cytokines peaked at three to six days after the onset of the disease in critically ill COVID-19 patients; these cytokines included IL-10, IL-2, IL-4, TNF $\alpha$  and IFN $\gamma$ .<sup>38</sup> In addition, concentrations of IL-10 and IL-6 continued to increase in patients with severe COVID-19, and with the concentrations of IL-6 beginning to decrease after 16 days.<sup>38</sup>

Some limitations of the present work should be acknowledged. First, this investigation was designed as a retrospective study, and this study was based on just 338 patients and relatively few cases of progressing COVID-19 were included in the analyses. Thus, large-scale multinational prospective cohort studies are necessary to confirm the results and strengthen the conclusions reached. Second, the proper cut-off value of single parameters is still to be verified by large-scale multinational prospective cohort studies. Then, mild or general cases were treated before they progressed to severe cases. As stated earlier, patients with mild

COVID-19 are at risk of progressing to the severe stage, and once they become severe, they have a high mortality rate. Therefore, it would be unreasonable not to treat the patients already admitted to the hospital. Although they were treated at different sites, the two designated hospitals followed the same guidelines for treating mild or moderate COVID-19 specified in the Diagnosis and Treatment Protocol for COVID-19 issued by the National Health Commission of China. Thus, the use of the data from two different centres did not affect the results of the study, since similar treatment strategies were employed.

In conclusion, the performed analysis demonstrates that high concentrations of NLR, LDH and IL-10 were independent risk factors for predicting disease progression in COVID-19 patients at the early stage of the disease. The values of the AUROC of NLR were superior to other single parameters. The risk model combined with NLR, LDH and IL-10 improved the accuracy of the prediction of disease progression in

patients in the early stages of COVID-19. Assessment of predictors might facilitate early identification of COVID-19 patients at high risk for disease progression and ensure timely administration of appropriate treatment to prevent mild cases from becoming severe.

### Acknowledgements

This work was supported by the Natural Science Foundation of Hubei Province (No. 2019CFB729), the National Natural Science Foundation of China (No. 81903173, No. 81874231).

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Natural Science Foundation of Hubei Province (No. 2019CFB729), the National Natural Science Foundation of China (No.81903173, No. 81874231).

### Ethical approval

This retrospective study was approved by the Tongji Medical College research ethics committee, Wuhan, China (NO: IORG0003571).

### Guarantor

Qichang Zheng.

### Contributorship

Jianjun Xu, Yang Gao, Shaobo Hu and Qichang Zheng researched literature and conceived the study. Suzhen Li, Weimin Wang, Yuzhe Wu, Zhe Su, Xing Zhou and Xiang Cheng were involved in protocol development, gaining ethical approval, patient recruitment and data analysis. Jianjun Xu, Shaobo Hu and Yang Gao wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

### ORCID iD

Qichang Zheng  <https://orcid.org/0000-0003-4570-0863>

### Supplemental material

Supplemental material for this article is available online.

### References

- Huang Y, Tu M, Wang S, et al. Clinical characteristics of laboratory confirmed positive cases of SARS-CoV-2 infection in Wuhan, China: a retrospective single center analysis. *Travel Med Infect Dis* 2020; 36: 101606.
- Liu Y, Du X, Chen J, et al. Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. *J Infect* 2020; 81: e6–e12.
- The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) – China, 2020. *China CDC Week* 2020; 2: 113–122.
- National Health Commission of the People's Republic of China. Chinese management guideline for COVID-19 (Trial Version 7), <http://www.nhc.gov.cn/zycjg/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml> (accessed 8 June 2020; in Chinese).
- Lei Z, Cao H, Jie Y, et al. A cross-sectional comparison of epidemiological and clinical features of patients with coronavirus disease (COVID-19) in Wuhan and outside Wuhan, China. *Travel Med Infect Dis* 2020; 35: 101664.
- Zhang K, Hua YQ, Wang D, et al. Systemic immune-inflammation index predicts prognosis of patients with advanced pancreatic cancer. *J Transl Med* 2019; 17: 30.
- Guthrie GJ, Charles KA, Roxburgh CS, et al. The systemic inflammation-based neutrophil-lymphocyte ratio: experience in patients with cancer. *Crit Rev Oncol Hematol* 2013; 88: 218–230.
- Giede-Jeppe A, Bobinger T, Gerner ST, et al. Neutrophil-to-lymphocyte ratio is an independent predictor for in-hospital mortality in spontaneous intracerebral hemorrhage. *Cerebrovasc Dis* 2017; 44: 26–34.
- Ha YJ, Hur J, Go DJ, et al. Baseline peripheral blood neutrophil-to-lymphocyte ratio could predict survival in patients with adult polymyositis and dermatomyositis: a retrospective observational study. *PLoS One* 2018; 13: e0190411.
- Zhao LY, Yang DD, Ma XK, et al. The prognostic value of aspartate aminotransferase to lymphocyte ratio and systemic immune-inflammation index for overall survival of hepatocellular carcinoma patients treated with palliative treatments. *J Cancer* 2019; 10: 2299–2311.
- Sun Y, Li W, Li AJ, et al. Increased systemic immune-inflammation index independently predicts poor survival for hormone receptor-negative, HER2-positive breast cancer patients. *Cancer Manag Res* 2019; 11: 3153–3162.
- Gao Y, Guo W, Cai S, et al. Systemic immune-inflammation index (SII) is useful to predict survival outcomes in patients with surgically resected esophageal squamous cell carcinoma. *J Cancer* 2019; 10: 3188–3196.
- Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020; 382: 1708–1720.
- Al-Tawfiq JA. Asymptomatic coronavirus infection: MERS-CoV and SARS-CoV-2 (COVID-19). *Travel Med Infect Dis* 2020; 35: 101608.
- Li X, Zeng W, Li X, et al. CT imaging changes of corona virus disease 2019(COVID-19): a multi-center study in Southwest China. *J Transl Med* 2020; 18: 154.
- Schlagenhauf P, Grobusch MP, Maier JD, et al. Repurposing antimetabolites and other drugs for COVID-19. *Travel Med Infect Dis* 2020; 34: 101658.
- Jiang XL, Zhang XL, Zhao XN, et al. Transmission potential of asymptomatic and paucisymptomatic SARS-CoV-2 infections: a three-family cluster study in China. *J Infect Dis* 2020; 221: 1948–1952.
- Nie X, Fan L, Mu G, et al. Epidemiological characteristics and incubation period of 7,015 confirmed cases with coronavirus disease 2019 outside Hubei Province in China. *J Infect Dis* 2020; 222: 26–33.
- Pan Y, Yu X, Du X, et al. Epidemiological and clinical characteristics of 26 asymptomatic SARS-CoV-2 carriers. *J Infect Dis* 2020; 221: 1940–1947.
- 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. *Lancet Respir Med* 2020; 8: 475–481.
- Moore BJB and June CH. Cytokine release syndrome in severe COVID-19. *Science* 2020; 368: 473–474.
- Zhao M. Cytokine storm and immunomodulatory therapy in COVID-19: role of chloroquine and anti-IL-6 monoclonal antibodies. *Int J Antimicrob Agents* 2020; 55: 105982.
- Kuss O, Blettner M and Borgermann J. Propensity score: an alternative method of analyzing treatment effects. *Dtsch Arztebl Int* 2016; 113: 597–603.
- Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020; 369: m1328.
- Yu H, Shao J, Guo Y, et al. Data-driven discovery of clinical routes for severity detection in covid-19 pediatric cases. *medRxiv* 2020; 03: 20032219.
- Zhou Y, Yang Z, Guo Y, et al. A new predictor of disease severity in patients with covid-19 in Wuhan, China. *medRxiv* 2020; 03: 20042119.
- Benchoufi M, Bokobza J, Anthony C, et al. Lung injury in patients with or suspected COVID-19: a comparison between lung ultrasound and chest CT-scanner severity assessments, an observational study. *medRxiv* 2020; 04: 20050997.
- Chassagnon G, Vakalopoulou M, Battistella E, et al. AI-driven quantification, staging and outcome prediction of COVID-19 pneumonia. *Med Image Anal* 2021; 67: 101860.
- Li Z, Zhong Z, Li Y, et al. From community-acquired pneumonia to COVID-19: a deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans. *Eur Radiol* 2020; 30: 6828–6837.

30. Lyu P, Liu X, Zhang R, et al. The performance of chest CT in evaluating the clinical severity of COVID-19 pneumonia: identifying critical cases based on CT characteristics. *Invest Radiol* 2020; 55: 412–421.
31. Wang C, Deng R, Gou L, et al. Preliminary study to identify severe from moderate cases of COVID-19 using NLR&RDW-SD combination parameter. *medRxiv* 2020; 04: 20058594.
32. Zhu Z, Cai T, Fan L, et al. Clinical value of immune-inflammatory parameters to assess the severity of coronavirus disease 2019. *Int J Infect Dis* 2020; 95: 332–339.
33. Ponti G, Maccaferri M, Ruini C, et al. Biomarkers associated with COVID-19 disease progression. *Crit Rev Clin Lab Sci* 2020; 57: 389–399.
34. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020; 323: 1061–1069.
35. Choi KW, Chau TN, Tsang O, et al. Outcomes and prognostic factors in 267 patients with severe acute respiratory syndrome in Hong Kong. *Ann Intern Med* 2003; 139: 715–723.
36. Hong KH, Choi JP, Hong SH, et al. Predictors of mortality in Middle East respiratory syndrome (MERS). *Thorax* 2018; 73: 286–289.
37. Zhou F, Yu T, Du R, et al. 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.
38. Liu J, Li S, Liu J, et al. Longitudinal characteristics of lymphocyte responses and cytokine profiles in the peripheral blood of SARS-CoV-2 infected patients. *EbioMedicine* 2020; 55: 102763.