

## ARTICLE OPEN



# A predictive score for progression of COVID-19 in hospitalized persons: a cohort study

Jingbo Xu<sup>1,10</sup>, Weida Wang<sup>2,10</sup>, Honghui Ye<sup>1,10</sup>, Wenzheng Pang<sup>1,10</sup>, Pengfei Pang<sup>3,10</sup>, Meiwen Tang<sup>1</sup>, Feng Xie<sup>1</sup>, Zhitao Li<sup>1</sup>, Bixiang Li<sup>1</sup>, Anqi Liang<sup>4</sup>, Juan Zhuang<sup>1</sup>, Jing Yang<sup>1</sup>, Chunyu Zhang<sup>1</sup>, Jiangnan Ren<sup>5</sup>, Lin Tian<sup>6</sup>, Zhonghe Li<sup>7</sup>, Jinyu Xia<sup>8</sup>, Robert P. Gale<sup>9</sup>, Hong Shan<sup>3</sup> and Yang Liang<sup>2</sup>

Accurate prediction of the risk of progression of coronavirus disease (COVID-19) is needed at the time of hospitalization. Logistic regression analyses are used to interrogate clinical and laboratory co-variables from every hospital admission from an area of 2 million people with sporadic cases. From a total of 98 subjects, 3 were severe COVID-19 on admission. From the remaining subjects, 24 developed severe/critical symptoms. The predictive model includes four co-variables: age (>60 years; odds ratio [OR] = 12 [2.3, 62]); blood oxygen saturation (<97%; OR = 10.4 [2.04, 53]); C-reactive protein (>5.75 mg/L; OR = 9.3 [1.5, 58]); and prothrombin time (>12.3 s; OR = 6.7 [1.1, 41]). Cutoff value is two factors, and the sensitivity and specificity are 96% and 78% respectively. The area under the receiver-operator characteristic curve is 0.937. This model is suitable in predicting which unselected newly hospitalized persons are at-risk to develop severe/critical COVID-19.

npj Primary Care Respiratory Medicine (2021)31:33; <https://doi.org/10.1038/s41533-021-00244-w>

## INTRODUCTION

Approximately 10 percent of persons with SARS-CoV-2-infection are hospitalized because they develop severe/critical coronavirus disease (COVID-19)<sup>1–3</sup>. According to the interim guidance of the Centers for Disease Control and Prevention (CDC)<sup>5</sup>, in areas with sustained community-level outbreaks<sup>4</sup>, typically urban, most hospital admissions are for persons with severe/critical and/or comorbidities. This selection bias reflects limited acute care resources but is not representative of the global epidemiology of either SARS-CoV-2-infection or of COVID-19. Consequently, most prognostic and predictive scores using admission co-variables are for death from COVID-19 in persons with severe/critical COVID-19 rather than disease progression in persons with less severe disease<sup>6,7</sup>. For example, in a report from Italy of 1591 subjects, it was mentioned that all subjects were admitted to an intensive care unit (ICU), 99 percent of evaluable subjects required respiratory support, and the case fatality rate for ICU subjects was 26 percent<sup>8</sup>.

Understandably, most reports of large series of persons with COVID-19 are from urban centers where SARS-CoV-2-infection is an epidemic. However, as the numbers of confirmed cases and endemic areas increase every day, it is with high to certain probability that more and more areas currently with sporadic or clustered cases will eventually become areas with sustained community-level outbreaks and would involve hospitalizations for less severe disease under current containments. This is what we have seen in the historical influenza epidemics according to CDC's pandemic interval framework<sup>9</sup>. Under these circumstances, a precise and convenient triage strategy would be especially important in allocating health care capacity. To address this issue,

we studied outcomes of 98 consecutive subjects with COVID-19 in a region of 2 million persons where most admissions were for persons with mild or moderate COVID-19. We were able to use these data to develop a predictive model of the risk of progression to severe/critical COVID-19. These data may help physicians prioritize use of medical resources accordingly.

## RESULTS

From January 17 to February 13, 2020, 98 patients with COVID-19 in Zhuhai were admitted to the hospital (Table 1); 46 subjects were male. Their median age was 47 years (interquartile range [IQR], 34–62; range, 10 months to 80 years). From the total number of patients, 77 had traveled to an epidemic area and 18 had contact with a SARS-CoV-2-infected person; 45 subjects had comorbidities on admission, including hypertension ( $N = 17$ ), diabetes ( $N = 7$ ), cancer ( $N = 5$ ), tuberculosis ( $N = 2$ ), and chronic kidney disease ( $N = 2$ ). The median duration from symptoms onset to admission was 3 days (IQR, 1.0–5.3). On admission, 13 subjects were classified as having mild disease, 79 were classified as moderate, and 3 were classified as patients with severe COVID-19. None of these subjects were critical. (3 subjects were not classified on admission).

During hospitalization, four subjects received mechanical ventilation; ten subjects in the moderate severity cohort were administered corticosteroids. Further, 17 subjects (15 in the moderate cohort) were administered chloroquine, 12 (11 in the moderate cohort) subjects were administered lopinavir/ritonavir (LPV/r), and 13 subjects were administered intravenous immunoglobulin.

<sup>1</sup>Department of Hematology, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>2</sup>Department of Hematologic Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, Guangdong, China. <sup>3</sup>Department of Interventional Therapy, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>4</sup>Department of Rheumatology, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>5</sup>Department of Gastroenterology, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>6</sup>Department of Pharmacy, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>7</sup>Department of Nephrology, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>8</sup>Department of Infectious Diseases, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong, China. <sup>9</sup>Haematology Research Centre, Department of Immunology and Inflammation, Imperial College London, London, UK. <sup>10</sup>These authors contributed equally: Jingbo Xu, Weida Wang, Honghui Ye, Wenzheng Pang, Pengfei Pang. ✉email: xujingbo@mail.sysu.edu.cn; shanhong@mail.sysu.edu.cn; liangyang@sysucc.org.cn

**Table 1.** Baseline characteristics of patients with COVID-19 on admission by severity at triage and severity during hospitalization.

Parameters	Entire cohort <sup>a</sup> (N = 98)	No. (%)							
		Severity at triage				Severity during hospitalization			
		Mild (n = 13)	Moderate (n = 79)	Severe (n = 3)	N/A (n = 3)	P-value	Mild/Moderate (n = 74)	Severe/Critical (n = 24)	P-value
Age, years (median, IQR)	46.5 (34.3–62.0)	34.0 (29.0–42.5)	50.5 (36.0–60.5)	78.0 (65.0–80.0)	37.0 (6.0–63.0)	0.004	39.0 (31.3–56.3)	63.0 (46.0–66.5)	<0.001
Sex						0.55			0.06
Female	53 (54.1)	5 (38.5)	45 (57.0)	1 (33.3)	2 (66.7)		44 (59.5)	9 (37.5)	
Male	45 (45.9)	8 (61.5)	34 (43.0)	2 (66.7)	1 (33.3)		30 (30.5)	15 (62.5)	
Epidemiological history (N = 92)			n = 74	n = 2		0.04	n = 73	n = 19	0.06
Travel to epidemic area before onset	71 (77.2)	7 (53.8)	60 (81.1)	1 (50.0)	3 (100)		59 (80.8)	12 (63.2)	
Known sick contacts	17 (18.5)	6 (46.2)	11 (14.9)	1 (50.0)	0 (0)		14 (19.2)	3 (15.8)	
Unknown exposure	4 (4.3)	0 (0)	3 (4.1)	0 (0)	0 (0)		0(0)	4 (21.1)	
Comorbidities	44 (44.9)	3 (23.1)	37 (46.8)	2 (66.7)	2 (66.7)	0.27	30 (40.5)	14 (58.3)	0.13
Hypertension	17 (17.3)	1 (7.7)	15 (19.0)	1 (33.3)	0 (0)	0.54	9 (12.2)	8 (33.3)	0.05
Diabetes Mellitus	7 (7.1)	0 (0)	6 (7.6)	1 (33.3)	0 (0)	0.28	2 (2.7)	5 (20.8)	0.05
Malignancies	5 (5.1)	0 (0)	3 (3.8)	1 (33.3)	1 (33.3)	0.03	4 (5.4)	1 (4.2)	0.83
Tuberculosis	2 (2.0)	0 (0)	2 (2.5)	0 (0)	0 (0)	1.00	0 (0)	2 (8.3)	0.16
Chronic kidney diseases	2 (2.0)	1 (7.7)	1 (1.3)	0 (0)	0 (0)	0.35	1 (1.4)	1 (4.2)	0.42
Stroke	1 (1.0)	0 (0)	1 (1.3)	0 (0)	0 (0)	1.00	1 (1.4)	0 (0)	0.57
Coronary heart disease	1 (1.0)	0 (0)	1 (1.3)	0 (0)	0 (0)	1.00	1 (1.4)	0 (0)	0.57
Symptoms									
Paucisymptomatic	15 (15.3)	1 (7.7)	12 (15.2)	0 (0)	2 (66.7)	0.13	14 (18.9)	1 (4.2)	0.02
Fever	58 (59.2)	9 (69.2)	48 (60.8)	1 (33.3)	0 (0)	0.13	39 (52.7)	19 (79.2)	0.01
T <sub>max</sub> on admission (°C, IQR)	38.0 (37.5–38.3)	38.0 (37.7–38.8)	38.0 (37.5–38.3)	37.6	N/A	0.36	37.8 (37.5–38.3)	38.0 (37.8–38.3)	0.50
Cough	17 (17.3)	2 (15.4)	14 (17.7)	1 (33.3)	0 (0)	0.83	14 (18.9)	3 (12.5)	0.48
Sore throat	5 (5.1)	0 (0)	5 (6.3)	0 (0)	0 (0)	1.00	5 (6.8)	0 (0)	0.02
Diarrhea	4 (4.1)	0 (0)	2 (2.5)	1 (33.3)	1 (33.3)	0.04	2 (2.7)	2 (8.3)	0.36
SpO <sub>2</sub> on admission, % IQR (N = 97) <sup>b</sup>	98.0 (96.9–98.8)	98.0 (97.1–98.8)	98.2 (96.9–98.8)	94.0 (91.0–94.5)	98.0 (98.0–99.9)	0.04	98.3 (97.7–98.9)	96.1 (94.6–97.9)	0.03
Days to admission, IQR <sup>c</sup>	3.0 (1.0–5.3)	2 (1–5)	3 (1–5)	6 (5–12)	N/A	0.43	3 (1–6)	3 (1–4)	0.58

COVID-19 coronavirus disease 2019, IQR interquartile range, N/A not available, SpO<sub>2</sub> saturation of peripheral oxygen, T<sub>max</sub> peak temperature.

<sup>a</sup>Three patients did not have severity stratification at triage.

<sup>b</sup>Missing data in 1 patient in Moderate group.

<sup>c</sup>Duration from symptom onset to admission in symptomatic patients, from probable exposure to admission in asymptomatic patients.

At the time of final follow-up (median 55 days from admission; IQR, 52–58; range, 37–79 days), the highest COVID-19 severity scores were mild in 8 subjects, moderate in 66, severe in 19, and critical in 5. Severity grade shifts are summarized in Table 2. Among 92 mild and moderate patients on admission, 21 (22.8%) progressed during hospitalization, including 3 with critical illness. The progression rates were 15.4% and 24.1% in the mild and moderate groups respectively. The median duration of hospitalization or interval to death was 18 days, which was not significantly different among the severity cohorts. The median duration of virus shedding was 8 days (IQR: 4–10 days; range, 1–19 days), and was similar among the severity cohorts. Detailed laboratory results of patients with COVID-19 on admission by severity is illustrated in Table 3.

In subjects with mild or moderate severity disease on admission (N = 95), we interrogated co-variables associated with risk of progression to severe/critical disease. Some binary co-variables were excluded because of low sensitivity and specificity. Continuous variables were tested in receiver-operator characteristic (ROC) curves to identify cutoff values (Fig. 1 and Supplementary Fig. 1), transformed into categorical variables, and entered in multivariate backward stepwise logistic regression analysis with clinical co-variables significantly associated with the risk of progression (as shown in Tables 1 and 3). Several duplicates and co-linear covariates were excluded such as International Normalized Ratio, CD4 and CD8-positive cell concentrations. C-reactive protein (CRP) > 5.75 mg/L, prothrombin time (PT) > 12.3 s, age > 60 years, and blood oxygen saturation (SpO<sub>2</sub>) < 97% correlated with a severer

disease characteristic, showing a clinical value in categorizing patients with a higher risk of progression to severe/critical diseases. Co-variables, odds ratios and 95% confidence intervals are shown in Fig. 2. The score of each point was defined as the relative weights assigned according to the regression coefficient of each categorical co-variate, namely 1 point for each. The area under the ROC curve (AUROC) of the score was 0.937. The cutoff value in dividing subjects into high- and low-risk groups with the potential risk for progressing to severe/critical cases was 2, with a sensitivity of 96% and specificity of 78%. The hazard ratio of progression to severe/critical COVID-19 in subjects with a score ≥ 2 was 42 (11–164) compared with subjects whose score was < 2. 59 percent (43, 73%) subjects with a score ≥ 2 developed severe/critical COVID-19 compared with 2 percent (0.3, 9.0%, P < 0.001) of subjects with a score < 2.

## DISCUSSION

In our study of 98 consecutive, unselected subjects with SARS-CoV-2-infection and COVID-19 (including all hospitalized persons) in an area of 2 million people with sporadic or clustered cases, we identified four admission co-variables that were significantly associated with progression to severe/critical disease. We used these co-variables to develop a predictive score that identified subjects with a 40-fold increased risk of progression of COVID-19 to a severe/critical stage with a sensitivity of 96% and a specificity

**Table 2.** Overview of therapeutic interventions and clinical outcomes for patients with COVID-19 by severity at triage.

	No. (%)					P-value
	Entire cohort (n = 98)	Severity at triage				
		Mild (n = 13)	Moderate (n = 79)	Severe (n = 3)	N/A (n = 3)	
<b>Therapeutic approach</b>						
Mechanical ventilation	4 (4.1)	0 (0)	2 (2.5)	2 (66.7)	0 (0)	0.01
Glucocorticoid	10 (10.2)	0 (0)	10 (12.7)	0 (0)	0 (0)	0.66
Chloroquine	17 (17.3)	1 (7.7)	15 (19.0)	0 (0)	1 (33.3)	0.54
LPV/r	12 (12.2)	1 (7.7)	11 (13.9)	0 (0)	0 (0)	1.00
IVIg	13 (13.3)	2 (15.4)	9 (11.4)	1 (33.3)	1 (33.3)	0.18
<b>Final severity stratification during hospitalization</b>						
Mild	8 (8.2)	6 (46.2)	1 (1.3)	0 (0)	1 (33.3)	<0.001
Moderate	66 (67.3)	5 (38.5)	59 (74.7)	0 (0)	2 (66.7)	
Severe	19 (19.4)	2 (15.4)	16 (20.3)	1 (33.3)	0 (0)	
Critical	5 (5.1)	0 (0)	3 (3.8)	2 (66.7)	0 (0)	
<b>Outcomes</b>						
Length of stay in hospital, days, IQR	18.0 (14.8–25.3)	20.0 (16.0–23.5)	18.0 (15.0–26.0)	19.0 (9.0–23.0)	13.0 (11.0–19.0)	0.54
Died in hospital	1 (1.0)	0 (0)	0 (0)	1 (33.0)	0 (0)	0.06
Discharged from hospital	96 (98.0)	13 (100)	78 (98.7)	2 (66.7)	3 (100)	0.16
Days to negativity, IQR <sup>a</sup>	8.0 (4.0–10.0)	9 (3–12)	8 (5–10)	4 (4–9)	N/A	0.64

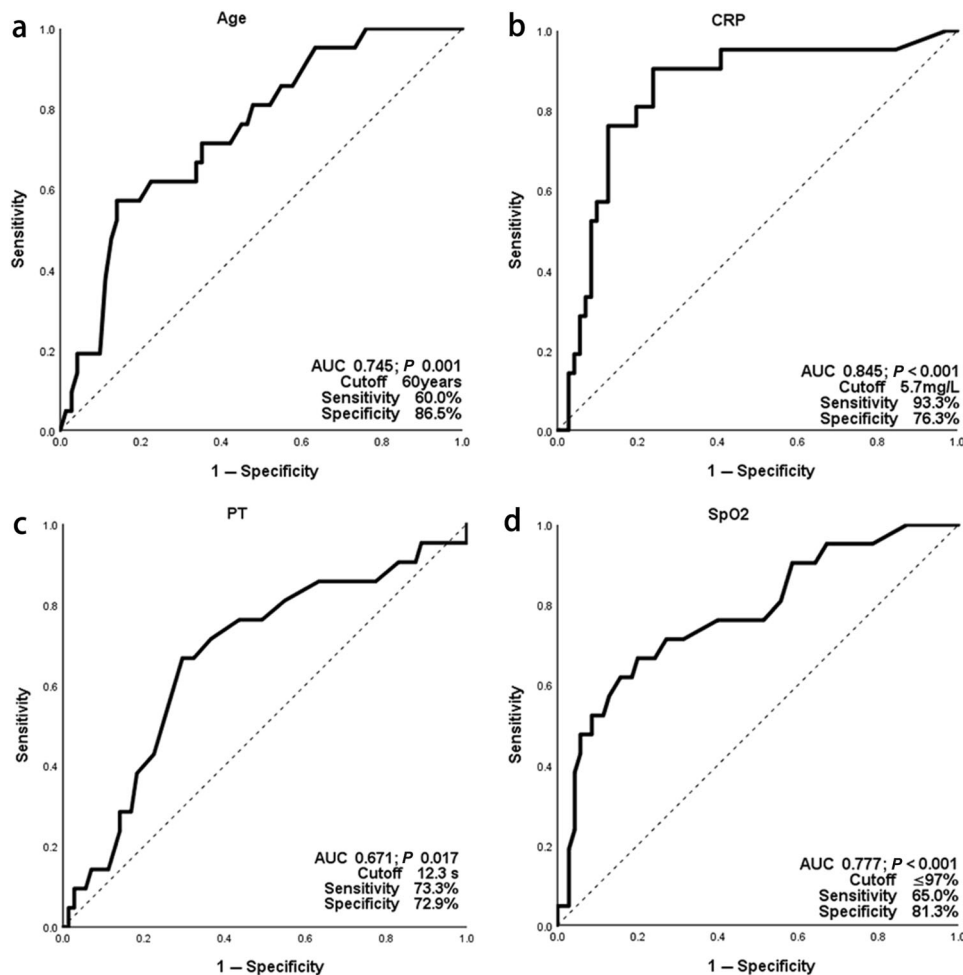
COVID-19 coronavirus disease 2019, N/A not available, LPV/r lopinavir/ritonavir, IVIG intravenous immunoglobulin, IQR interquartile range.

<sup>a</sup>Duration from admission to the date that patients tested a second PCR negativity in throat or nasal swab specimen.

**Table 3.** Laboratory results of patients with COVID-19 on admission by severity during hospitalization.

Parameters (median, IQR)	Entire cohort (n = 98)	Severity during hospitalization		P-value
		Mild/Moderate (n = 74)	Severe/Critical (n = 24)	
<b>Hematologic</b>				
White-cell count (x10 <sup>9</sup> /L)	5.0 (4.0–6.5)	5.0 (4.2–6.4)	5.0 (3.6–6.6)	0.71
Lymphocyte count (x10 <sup>9</sup> /L)	1.6 (1.1–2.1)	1.7 (1.3–2.2)	1.1 (0.6–1.5)	0.001
Hemoglobin (g/L)	136.5 (125.8–148.0)	137.0 (126.8–146.0)	135.0 (119.3–149.8)	0.42
Platelet (x10 <sup>9</sup> /L)	194.5 (164.5–251.0)	209.5 (178.0–272.0)	167 (133.5–197.5)	<0.001
CD4 + T cell count (/mm <sup>3</sup> )	547 (404–751)	597 (467–841)	432 (220–547)	0.004
CD8 + T cell count (/mm <sup>3</sup> )	337 (236–492)	402.5 (293–524.8)	216 (122–289)	<0.001
<b>Liver function</b>				
Alanine aminotransferase (U/L)	15.5 (11.1–28.0)	15.0 (10.0–25.8)	21.7 (13.8–42.1)	0.05
Aspartate aminotransferase (U/L)	20.4 (15.3–29.8)	19.0 (14.4–25.3)	27.3 (18.5–38.9)	0.01
Albumin (g/L)	38.7 (36.0–41.5)	39.0 (36.3–42.3)	36.7 (35.5–41.1)	0.53
<b>Coagulation</b>				
Prothrombin time (s)	12.0 (11.5–12.6)	11.8 (11.4–12.5)	12.5 (12.0–13.1)	0.01
International normalized ratio	1.05 (1.01–1.11)	1.04 (1.00–1.10)	1.09 (1.05–1.14)	0.03
Activated partial thromboplastin time (s)	31.6 (29.1–33.5)	31.6 (28.8–33.1)	32.1 (29.7–34.2)	0.81
Fibrinogen (g/L)	3.03 (2.56–3.72)	2.87 (2.51–3.41)	3.77 (2.88–4.29)	0.001
Thrombin time (s)	13.8 (13.0–15.1)	14.2 (13.3–15.1)	13.3 (12.5–14.1)	0.35
<b>Blood chemistry</b>				
Sodium (mmol/L)	140.0 (138.0–142.0)	139.0 (137.0–141.5)	142.0 (139.3–143.0)	0.02
Blood glucose (mmol/L)	5.34 (4.90–6.48)	5.34 (4.90–6.79)	5.36 (4.89–6.16)	0.69
<b>Kidney function</b>				
Blood urea nitrogen (μmol/L)	3.70 (2.90–4.30)	3.50 (2.70–4.20)	3.90 (3.65–4.55)	0.30
Creatinine (mmol/L)	59.6 (47.8–71.6)	57.0 (47.0–70.8)	67.1 (53.2–75.2)	0.84
<b>Inflammatory</b>				
Lactate dehydrogenase (U/L)	164.0 (142.5–201.5)	158.0 (135.0–196.0)	180.5 (156.0–215.8)	0.02
Hydroxybutyric dehydrogenase (U/L)	129.0 (108.0–155.0)	128.0 (107.0–155.0)	141.5 (119.3–175.8)	0.06
Lactate (mmol/L)	1.50 (1.10–1.95)	1.50 (1.20–1.90)	1.05 (0.80–2.20)	0.61
C-reactive protein (mg/L)	4.10 (0.56–17.60)	1.16 (0.35–5.60)	26.70 (12.35–44.25)	0.002

COVID-19 coronavirus disease 2019, IQR interquartile range.



**Fig. 1** Receiver operating characteristic curves to identify cut-offs of co-variables on admission in patients with mild/moderate SARS-CoV-2-infection. These four co-variables were entering the final model and shown here, other co-variables were shown in Supplementary Fig. 1. **a** Age; **b** CRP; **c** PT; **d** SpO<sub>2</sub>. CRP c-reactive protein, PT prothrombin time, SpO<sub>2</sub> blood oxygen saturation.

of 78%. The AUROC was 0.937. Our model is simple to use based on readily available co-variables.

There are several reports of prognostic and predictive scores of outcomes of COVID-19, although most studies have important biases and are not representative of real-world experience with the SARS-CoV-2 pandemic and with COVID-19<sup>10</sup>. For example, in regions where authorities imposed home isolation and social distancing, most persons with mild/moderate COVID-19 were not hospitalized<sup>11</sup>. On the other hand, in urban regions with large numbers of cases of COVID-19 and limited intensive care resources such as mechanical ventilation, most hospital admissions were for persons with severe/critical COVID-19 alone<sup>12–14</sup>. There were also obvious selection biases as to why and where people were hospitalized in these studies, and hospitalized persons from a center or a few centers were unlikely representatives of the distribution of cases of COVID-19 in a region, especially a region with sporadic or clustered cases of SARS-CoV-2-infection<sup>15,16</sup>. In several studies, there was censoring of subjects still in-hospital, which biased the interpretations of results accordingly<sup>12,17</sup>. In contrast, we were able to identify every patient of COVID-19 in our area, all of whom were actually hospitalized. We also conducted a complete follow-up of all the subjects. These biases are obvious when we consider the 1 death in our study versus an average of 10–20 percent in other studies<sup>14,18</sup>. Our subjects were more likely to be similar to a typical non-epidemic

setting of exposure to SARS-CoV-2; therefore, our prognostic score is more likely to be widely useful<sup>1,19,20</sup>.

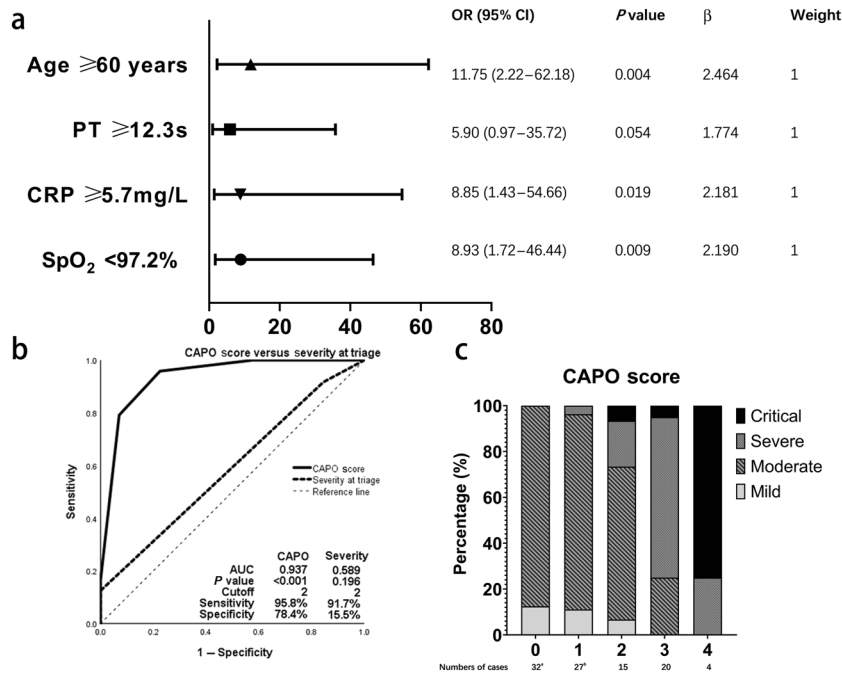
Similar to other studies, we found that age, CRP, and SpO<sub>2</sub> on admission correlated with outcomes<sup>12,21–23</sup>. However, our study differed from other predicting tools<sup>12,17,21,22</sup>, wherein we identified a new risk factor, PT. Previous studies indicate that coagulation disorders are common in patients with severe COVID-19<sup>24,25</sup> and are associated with an increased risk of acute respiratory distress syndrome<sup>26</sup>. We suggested early monitoring of PT to predict the likelihood of progression.

In conclusion, our study had limitations including its retrospective design, relatively few subjects, and no validation cohort. We also lacked detailed data on post-admission interventions. However, other than oxygen supplementation and mechanical ventilation, none of the other interventions proved effective<sup>27</sup>. Also, our aim was to predict outcomes from admission to better allocate medical resources. Our score is easily implemented and should assist physicians to identify persons with COVID-19 on admission at the greatest risk to develop severe/critical disease.

## METHODS

### Subjects

This retrospective observational study was conducted at the Fifth Affiliated Hospital of Sun Yat-sen University, the largest tertiary academic hospital in Zhuhai. The Institutional Review Board of the Fifth Affiliated Hospital of



**Fig. 2** Development and performance of predicting score of progression risk in patients with mild/moderate SARS-CoV-2-infection. **a** Multivariate analysis associated with progression to severe/critical illness. **b** Characteristic curves of predicting score and severity at triage for prediction of progression. **c** Distributions of final severity stratification by predicting score on admission. CRP c-reactive protein, PT prothrombin time, SpO<sub>2</sub> blood oxygen saturation, OR odds ratio, CAPO CRP, Age, PT, SpO<sub>2</sub>.

Sun Yat-sen University approved the study protocol and classified the study minimal-risk and voided requirement for informed consent. COVID-19 containment (including active detection and quarantine of close contacts and travelers from epidemic areas [Hubei province]) was led by the Chinese government<sup>28</sup>. All consecutive subjects with confirmed SARS-CoV-2-infection, determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR; Novel Coronavirus [2019-nCoV][Real Time Multiplex RT-PCR Kit [Z-RR-0479-02-25, Shanghai ZJ Liferiver Bio-tech Co., Ltd. Shanghai, China)], between 17 January and 13 February 2020 were enrolled in the study. Subjects received therapy in accordance with the interim guidelines of the National Health Commission, China<sup>29</sup>. Lopinavir/ritonavir, chloroquine, and arbidol were administered to some subjects. Data lock was 21 March, 2020.

## Data

Demographic data, clinical symptoms, and laboratory results were collected and extracted from the hospital electronic medical records. Two investigators independently coordinated and integrated the data with discordances adjudicated by reviewing original records. Subject identifiers were deleted, thereby creating an anonymized dataset. Laboratory assessments included complete blood count with differential, liver and kidney function tests, coagulation tests, and C-reactive protein and lymphocyte subsets.

## Definitions

Criteria for the diagnosis of COVID-19 followed the interim guidelines of the National Health Commission, China<sup>29</sup>. A confirmed case was based on the exposure history, which included exposure to suspected cluster outbreaks, clinical manifestations (fever and/or respiratory symptoms), chest computed tomography imaging, and results of qRT-PCR for SARS-CoV-2 and anti-SARS-CoV-2 IgM and IgG antibodies using enzyme-linked immunosorbent assay. Classification of COVID-19 severity was based on the interim or 7th edition guidelines of the National Health Commission<sup>29</sup>. Severity was stratified on admission and revised based on disease progression during hospitalization. Outcomes were evaluated at the date of last follow-up, discharge or death, and by whether the subject required continued hospitalization. The highest severity during hospitalization was designed as the primary outcome to develop the predictive model for the likelihood of progression.

## Statistics

Descriptive statistics were used to summarize demographic data. Results were reported as medians and IQRs, means with standard deviations, or counts and frequencies. Continuous variables were compared using the t-test or one-way analysis of variance (ANOVA). Categorical dependent parameters were compared using the chi-square test and Fisher's exact test. Cut-off values were identified following Youden index of ROC curve. All tests were two-sided, and a *P*-value < 0.05 was considered significant. The Statistical Package for Social Sciences (SPSS) 25.0 software (IBM, Armonk, NY, USA), VassarStats (Vassarstats.net), and the GraphPad Prism (version 8.2.1) (GraphPad Software, San Diego, CA, USA) were used for statistical analyses and illustrations.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 6 January 2021; Accepted: 19 April 2021;

Published online: 03 June 2021

## REFERENCES

- Verity, R. et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020).
- Wu, Z. & McGoogan, J. M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* <https://doi.org/10.1001/jama.2020.2648> (2020).
- Garg, S. et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019-COVID-NET, 14 States, March 1–30, 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 458–464 (2020).
- World Health Organization Regional Office for Europe. About Pandemic Phases. <https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/data-and-statistics/pandemic-influenza/about-pandemic-phases> (WHO, 2020).



5. Centers for Disease Control and Prevention. *Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19)*, <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html#clinical-course> (Centers for Disease Control and Prevention, 2020).
6. Chen, R. et al. Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in China. *Chest* <https://doi.org/10.1016/j.chest.2020.04.010> (2020).
7. Bello-Chavolla, O. Y. et al. Predicting mortality due to SARS-CoV-2: a mechanistic score relating obesity and diabetes to COVID-19 outcomes in Mexico. *J. Clin. Endocrinol. Metab.* <https://doi.org/10.1210/clinem/dgaa346> (2020).
8. Grasselli, G. et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *JAMA* <https://doi.org/10.1001/jama.2020.5394> (2020).
9. Influenza Pandemic Framework Workgroup. Updated preparedness and response framework for influenza pandemics. *MMWR Morb. Mortal. Wkly. Rep.* 63(RR06), 1–9, <https://www.cdc.gov/flu/pandemic-resources/pdf/mmwr-rr6306.pdf> (2014).
10. Wynants, L. et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *BMJ* <https://doi.org/10.1136/bmj.m1328> (2020).
11. Gandhi, R. T., Lynch, J. B. & Del Rio, C. Mild or moderate covid-19. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMcp2009249> (2020).
12. Galloway, J. B. et al. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. *J. Infect.* <https://doi.org/10.1016/j.jinf.2020.05.064> (2020).
13. Di Bella, S. et al. Neck circumference as reliable predictor of mechanical ventilation support in adult inpatients with COVID-19: a multicentric prospective evaluation. *Diabetes Metab. Res. Rev.* e3354, <https://doi.org/10.1002/dmrr.3354> (2020).
14. Li, X. et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. *J. Allergy Clin. Immunol.* <https://doi.org/10.1016/j.jaci.2020.04.006> (2020).
15. Liang, W. H. et al. Clinical characteristics and outcomes of hospitalised patients with COVID-19 treated in Hubei (epicentre) and outside Hubei (non-epicentre): a nationwide analysis of China. *Eur. Respir. J.* **55**, <https://doi.org/10.1183/13993003.00562-2020> (2020).
16. CDC COVID-19 Response Team Geographic differences in COVID-19 cases, deaths, and incidence- United States, February 12–April 7, 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 465–471 (2020).
17. Gong, J. et al. A tool to early predict severe corona virus disease 2019 (COVID-19): a multicenter study using the risk Nomogram in Wuhan and Guangdong, China. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciaa443> (2020).
18. Liu, Y. et al. Neutrophil-to-lymphocyte ratio as an independent risk factor for mortality in hospitalized patients with COVID-19. *J. Infect.* <https://doi.org/10.1016/j.jinf.2020.04.002> (2020).
19. Wu, J. T. et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat. Med.* **26**, 506–510 (2020).
20. CDC Covid-Response Team. Severe outcomes among patients with coronavirus disease 2019 (COVID-19)-United States, February 12–March 16, 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 343–346 (2020).
21. Liang, W. et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients With COVID-19. *JAMA Intern. Med.* <https://doi.org/10.1001/jamainternmed.2020.2033> (2020).
22. Cheng, F. Y. et al. Using machine learning to predict ICU transfer in hospitalized COVID-19 patients. *J. Clin. Med.* **9**, <https://doi.org/10.3390/jcm9061668> (2020).
23. Herold, T. et al. Elevated levels of IL-6 and CRP predict the need for mechanical ventilation in COVID-19. *J. Allergy Clin. Immunol.* **146**, 128–136 (2020). e124.
24. Connors, J. M. & Levy, J. H. COVID-19 and its implications for thrombosis and anticoagulation. *Blood* **135**, 2033–2040 (2020).
25. Terpos, E. et al. Hematological findings and complications of COVID-19. *Am. J. Hematol.* <https://doi.org/10.1002/ajh.25829> (2020).
26. Wu, C. et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern. Med.* <https://doi.org/10.1001/jamainternmed.2020.0994> (2020).
27. COVID-19 Treatment Guidelines Panel. *Coronavirus Disease 2019 (COVID-19) Treatment Guidelines*. The National Institute of Allergy and Infectious Diseases-Rocky Mountain Laboratories (NIAID-RML) <https://www.covid19treatmentguidelines.nih.gov/> (2020).
28. Li, Z. et al. China CDC COVID-19 emergency response strategy team. Active case finding with case management: the key to tackling the COVID-19 pandemic. *Lancet* [https://doi.org/10.1016/S0140-6736\(20\)31278-2](https://doi.org/10.1016/S0140-6736(20)31278-2) (2020).
29. National Health Commission of the PRC. *Interim Guidance for Diagnosis and Treatment of Novel Coronavirus Pneumonia* (the 7th edn), <http://www.nhc.gov.cn/yzygj/s7653p/202003/46c9294a7dfe4cef80dc7f5912eb1989/files/ce3e6945832a438eaae415350a8ce964.pdf> (National Health Commission of the PRC, 2020).

## ACKNOWLEDGEMENTS

The authors acknowledge the expert work of Qie Yin from Hong Kong Baptist University (statistics) and Nan Li from Peking University Clinical Research Institute (epidemiology and statistics). Also, the authors especially thank the doctors, nurses, and other staffs in the the Fifth Affiliated Hospital of Sun Yat-sen University for their care for patients. This work was supported by the Program for Guangdong Introducing Innovative and Entrepreneurial Teams (2017ZT075096 to Y.L.) and the National Natural Science Foundation of China (Grant No. 81873428 and 81660682 to Y.L.). R.P.G. acknowledges support from the National Institute of Health Research (NIHR) Biomedical Research Centre funding scheme. Funding sources had input into the design, execution nor analyses of the data nor the decision to publish the results. Submission for publication was agreed by all authors all of whom had full access to of the data and take responsibility to submit the typescript for publication.

## AUTHOR CONTRIBUTIONS

Concept and design: J.X., Y.L., and H.S. Acquisition of data: H.Y., W.P., P.P., M.T., F.X., Z.L., B.L., A.L., J.Z., J.Y., C.Z., J.R., L.T., Z.L. and J.X. Drafting of the typescript: W.W. and R.P.G. Critical revision of the typescript: R.P.G., Y.L., J.X., and H.S. Statistical analysis: W.W., H.Y., W.P., and P.P. Supervision: Y.L., J.X. and H.S. Co-first authors: J.X., W.W., H.Y., W.P., P.P.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41533-021-00244-w>.

**Correspondence** and requests for materials should be addressed to J.X., H.S. or Y.L.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021