



A Novel Prediction Model for Long-Term SARS-CoV-2 RNA Shedding in Non-Severe Adult Hospitalized Patients with COVID-19: A Retrospective Cohort Study

Chen Lu Huang · Ling Fei · WeiXia Li · Wei Xu · Xu Dong Xie ·

Qiang Li · Liang Chen

Received: January 22, 2021 / Accepted: March 17, 2021 / Published online: March 31, 2021
© The Author(s) 2021

ABSTRACT

Introduction: Due to the lack of clear direction (evidence) on the duration of viral shedding and thus potential for transmission, this retrospective study aimed to come up with a prediction model of prolonged coronavirus disease-19 (COVID-19) transmission or infection-spreading potential.

Methods: A total of 1211 non-severe patients with COVID-19 were retrospectively enrolled. Multivariate Cox regression was performed to identify the risk factors associated with long-term SARS-CoV-2 RNA shedding, and a prediction model was established.

Results: In the training set, 796 patients were divided into the long-term (> 21 days) group ($n = 116$, 14.6%) and the short-term (≤ 21 days) group ($n = 680$, 85.4%) based on their viral shedding duration. Multivariate analysis identified that age > 50 years, comorbidity, CD4-positive T-lymphocytes count ($CD4 + T$ cell) ≤ 410 cells/ul, C-reactive protein

(CRP) > 10 mg/L, and the corticosteroid use were independent risk factors for long-term SARS-CoV-2 RNA shedding. Incorporating the five risk factors, a prediction model, named as the CCCCA score, was established, and its area under the receiver operator characteristic curve (AUROC) was 0.87 in the training set and 0.83 in the validation set, respectively. In the validation set, using a cut-off of 8 points, we found sensitivity, specificity, positive predictive value, and negative predictive value of 51.7%, 92.2%, 33.3%, and 96.2%, respectively. Long-term SARS-CoV-2 RNA shedding increased from 14/370 (3.8%) in patients with CCCCA < 8 points to 15/45 (33.3%) in patients with CCCCA ≥ 8 points.

Conclusion: Using the CCCCA score, clinicians can identify patients with long-term SARS-CoV-2 RNA shedding.

Keywords: 2019 novel coronavirus disease; Prediction model; Risk factors; Severe acute respiratory syndrome coronavirus 2; Viral shedding

Chen Lu Huang, Ling Fei and WeiXia Li contributed equally.

C. L. Huang · L. Fei · W. Li · W. Xu ·
X. D. Xie · Q. Li (✉) · L. Chen (✉)
Department of Liver Diseases, Shanghai Public
Health Clinical Center, Fudan University, Shanghai
201508, China
e-mail: liqiang66601@163.comL. Chen
e-mail: chenliang@shphc.org.cn

Key Summary Points

Age, comorbidity, CD4+ T cell count, CRP, and corticosteroid use were independent risk factors associated with prolonged SARS-CoV-2 RNA shedding.

A prediction model was established, called the CCCCA score, which scores from 5 (age \leq 50 years, absence of comorbidity, CD4+ T cell $>$ 410 cells/ul, CRP \leq 10 mg/L, no corticosteroid use) to 14 points (age $>$ 50 years, presence of comorbidity, CD4+ T cell \leq 410 cells/ul, CRP $>$ 10 mg/L, corticosteroid use).

Using the CCCCA model, clinicians can identify patients with prolonged SARS-CoV-2 RNA shedding, which is of importance for determining an appropriate period of isolation and planning strategies to reduce COVID-19 transmission.

DIGITAL FEATURES

This article is published with digital features, including a summary slide, to facilitate understanding of the article. To view digital features for this article go to <https://doi.org/10.6084/m9.figshare.14222525>.

INTRODUCTION

In December 2019, a novel coronavirus, called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), caused a global outbreak of coronavirus disease-19 (COVID-19) [1]. Phylogenetic analysis of the complete viral genome revealed that SARS-CoV-2 was most closely related (89.1% similarity) to severe acute respiratory syndrome coronavirus (SARS-CoV) that is the infectious agent responsible for the outbreak of severe acute respiratory syndrome (SARS) in 2003 [2]. At present, the COVID-19

pandemic has become a global public health issue. As of 12 January 2021, the overall upward trend of COVID-19 is continuing, which brings the cumulative numbers to over 88 million reported cases and over 1.9 million deaths globally since the start of the pandemic [3].

Although a negative SARS-CoV-2 ribonucleic acid (RNA) test is not required in all countries for non-immunocompromised individuals, the main importance of SARS-CoV-2 RNA is the public health implications for planning strategies to reduce transmission. In a previous study, long-term virus shedding was documented in patients with mild symptoms and in asymptomatic patients, and this should be considered when making decisions regarding clinical and public health, and strategies for the prevention and control of SARS-CoV-2 infection [4]. In addition, the duration of viral shedding is often considered in determining an appropriate period of isolation, as it is often used as a marker of infectivity. In China, one of the release criteria for hospitalized patients with COVID-19 is SARS-CoV-2 RNA shedding cessation, which is defined as the occurrence of two consecutive negative results of nasopharyngeal or throat swab specimens with at least 24-h intervals. However, patients with prolonged SARS-CoV-2 RNA shedding presented difficulties in terms of their safe discharge from isolation treatment centers. SARS-CoV-2 RNA shedding from respiratory tract specimens has been found to persist for up to 63 days after symptom onset among patients with COVID-19 [5, 6].

A number of papers have demonstrated prolonged viral shedding in severe illness. In the study by Liu et al., patients with severe COVID-19 tended to have a high viral load and a long virus-shedding period [7]. Zheng et al. also found that the virus persisted longer with higher load and peaked later in the respiratory tissue of patients with severe COVID-19 [8]. Li et al. identified nine (50%) asymptomatic patients and nine (50%) patients with mild symptoms of COVID-19 at admission, and found that the median duration of virus shedding was 11.5, 28, and 31 days for pre-symptomatic, asymptomatic, and mildly symptomatic patients, respectively [4]. In conclusion, there have been several studies

evaluating risk factors for prolonged shedding of SARS-CoV-2 RNA. However, specifically in non-severe but not asymptomatic patients, the data are scarce and no prediction score has so far been provided. Therefore, we aimed to clarify the high-risk factors associated with long-term SARS-CoV-2 RNA shedding, and to establish an accurate prediction model for the prediction of long-term SARS-CoV-2 RNA shedding.

METHODS

Study Population

A total of 1332 consecutive COVID-19 patients with age ≥ 18 years old, who were admitted to Shanghai Public Health Clinical Center, between January 20, 2020 and November 30, 2020, were included in our study. Shanghai Public Health Clinical Center, located in Shanghai, China, is a tertiary teaching hospital, and the only designated hospital for adult patients with COVID-19 (≥ 18 years old). Exclusion criteria: (1) patients presenting with severe COVID-19 ($n = 38$); (2) patients who were still being treated up to November 30, 2020 ($n = 83$). Finally, 1211 non-severe adult hospitalized patients with COVID-19 were enrolled. The 796 patients admitted to hospital between January 20 and August 31, 2020 constituted the training set, and 415 patients admitted to hospital between September 1 and November 30, 2020 constituted the validation set. The flow chart of the study population is shown in Fig. 1.

The study was approved by the Clinical Research Ethics Committee of the Shanghai Public Health Clinical Center (No. 2020-S219-01). Although this is a retrospective study, at hospital admission, all patients were provided verbal consent of the possibility of their clinical data being used for a medical study. The procedures followed were in accordance with the ethical standards of the Helsinki Declaration (1964, amended most recently in 2008) of the World Medical Association.

Diagnostic Criteria

Patients with COVID-19 were confirmed according to the positive results of SARS-CoV-2 RNA tests in nasopharyngeal or throat swab specimens using the real-time reverse transcription polymerase chain reaction (RT-PCR) method, which is the gold standard for detecting SARS-CoV-2 infection based on the World Health Organization interim guidance [9]. Patients with severe COVID-19 were diagnosed according to at least one of the following standards [10]: (1) respiratory distress, respiratory rates ≥ 30 breaths/min; (2) resting oxygen saturation $\leq 93\%$; (3) oxygenation index ≤ 300 mmHg; (4) requirement of mechanical ventilation; (5) shock; and (6) combined with other organ failures. Patients with non-severe COVID-19 were defined as patients without any of the above six standards. SARS-CoV-2 RNA shedding cessation was defined as the occurrence of two consecutive RT-PCR negative results of nasopharyngeal or throat swab specimens with at least 24-h intervals. SARS-CoV-2 RNA shedding duration was defined as the time from illness onset (symptom onset for symptomatic patients, first positive RT-PCR tests for asymptomatic patients) to SARS-CoV-2 RNA shedding cessation. At present, there are no unified standards for the cut-off value of prolonged SARS-CoV-2 RNA shedding. Based on previous studies [4, 8, 11–13], which reported that the median duration of viral shedding was 12–21 days, in this study, prolonged or long-term SARS-CoV-2 RNA shedding was defined as SARS-CoV-2 RNA shedding duration of more than 21 days. Patients were divided into two groups: the long-term SARS-CoV-2 RNA shedding group (> 21 days), and the short-term SARS-CoV-2 RNA shedding group (≤ 21 days).

Data Collection

The demographic characteristics, epidemiologic histories, comorbidity, laboratory parameters, and clinical outcomes were extracted from the electronic medical records. Comorbidity included at least one of the following: hypertension,

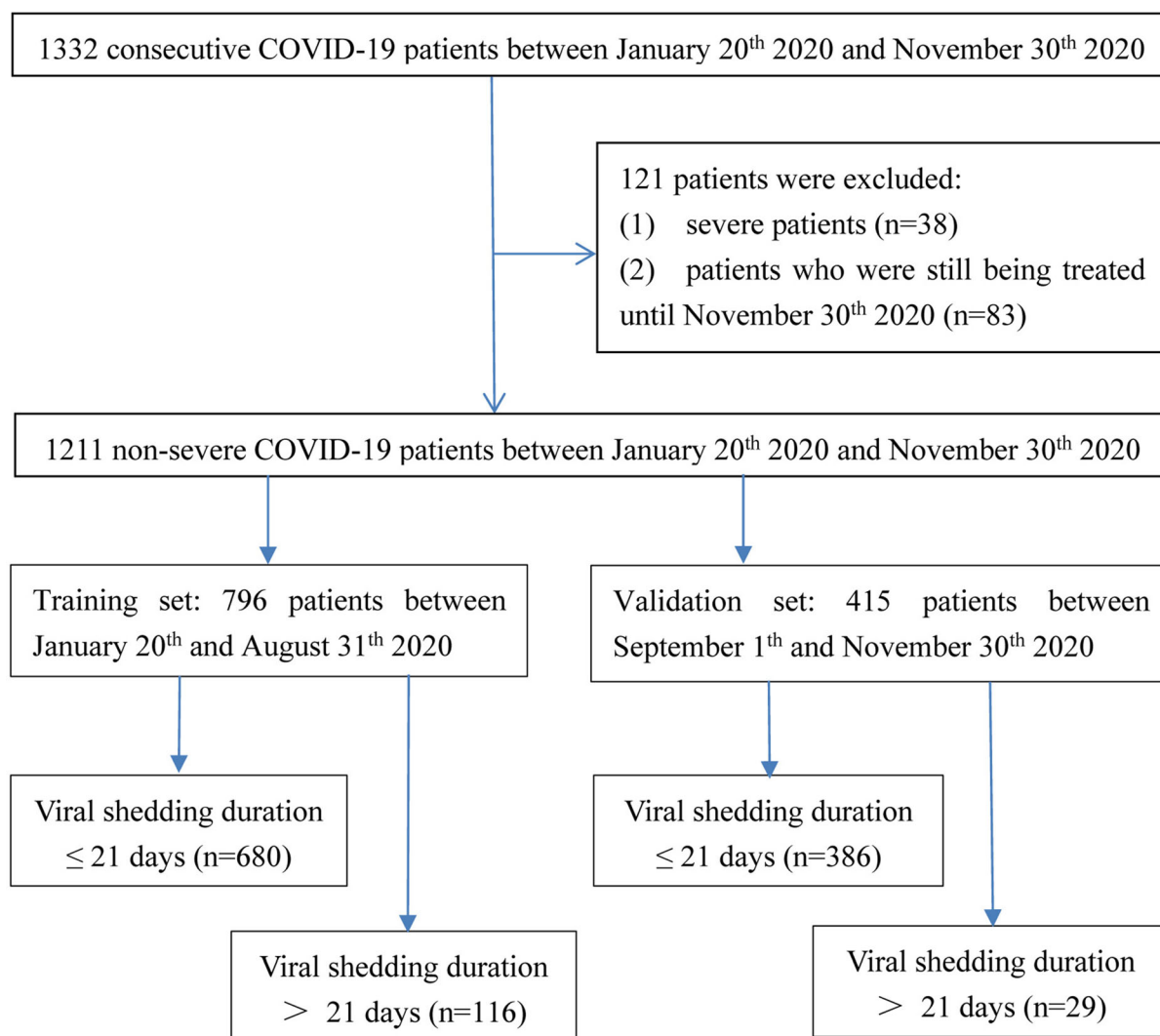


Fig. 1 Flow diagram of the study population. Finally, 1211 non-severe patients with COVID-19 were enrolled. The 796 patients admitted to hospital between January 20th 2020 and August 31st 2020 constituted the training

set, and 415 patients admitted to hospital between September 1st 2020 and November 30st 2020 constituted the validation set

diabetes, chronic cardiovascular disease, chronic liver disease, chronic lung disease, chronic kidney disease, acquired immunodeficiency syndrome, malignancy, and other immunocompromised factors including corticosteroid use and autoimmune diseases. For patients in this study, the vital signs were monitored daily; laboratory parameters were examined every 3–7 days, as were chest CT scans. The laboratory parameters including white blood cell (WBC) count, lymphocyte count, CD4-positive T-lymphocytes count

(CD4 + T cell count), C-reactive protein (CRP), lactate dehydrogenase (LDH), D-dimer, and erythrocyte sedimentation rate (ESR) were collected both at baseline and at all points during hospitalization.

Statistical Analysis

We used the Kolmogorov–Smirnov test to verify the normality of continuous variables. Normal distribution variables were presented using

mean and standard deviations, non-normal distribution continuous variables were presented using median and interquartile ranges (IQRs), and qualitative data were presented using frequency distribution. For the comparison of data between groups, we used the Student's *t* test for normal distribution variables, the non-parametric Mann–Whitney-test for non-normal distribution continuous variables, and the chi-squared test for qualitative variables. The significance of each variable was assessed by univariate and multivariate Cox analysis for investigating the independent risk factors of long-term SARS-CoV-2 RNA shedding. All the variables at a statistically significant level after multivariate Cox analysis were candidates for formulation of a scoring model, and their relevant points were determined by the hazard ratio (HR) values to reflect their weights of impact on the SARS-CoV-2 RNA shedding duration. The predictive performance of the scoring model was assessed using the receiver operating characteristic (ROC) curve and the area under the ROC curve (AUROC). The cut-off value was obtained by maximizing the Youden index (sensitivity + specificity-1). Diagnostic accuracy was evaluated by sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The statistical analyses were performed using the SPSS software, v.15.0 (SPSS; Chicago, IL, USA), MedCalc software, v.16.1 (MedCalc Software, Ostend, Belgium), and R software, v.3.6.1 (Foundation for Statistical Computing, Vienna, Austria). All significance tests were two-tailed, and $p < 0.05$ was considered as statistically significant.

RESULTS

Clinical Characteristics of Patients in the Training Set

Clinical characteristics of patients in the training set are shown in Table 1. Overall, 796 patients with COVID-19 were enrolled. The median age was 36 years (IQR 25–51), 456 patients (57.3%) were male, 203 patients (25.5%) were older than 50 years, 205 patients (25.8%) had obesity, 138 patients (17.3%) had a

smoking history, 134 patients (16.8%) had at least one comorbidity, and the median SARS-CoV-2 RNA shedding duration was 13 days (IQR 9–18). The 796 patients were divided into the long-term group ($n = 116$, 14.6%) and the short-term group ($n = 680$, 85.4%) based on their SARS-CoV-2 RNA shedding duration. Age, comorbidity, fever, WBC, lymphocyte count, CD4 + T cell count, CRP, and corticosteroid use were significantly different between the two groups (Table 1).

Independent Risk Factors Associated With Long-Term SARS-CoV-2 RNA Shedding

Cox regression analysis of prolonged viral RNA shedding duration is shown in Table 2. Multivariate Cox analysis identified that age > 50 years (HR 2.67, 95% CI 1.54–3.69, $p = 0.001$), comorbidity (HR 1.83, 95% CI 1.03–3.25, $p = 0.039$), CD4 + T cell count ≤ 410 cells/ul (HR 1.77, 95% CI 1.05–2.89, $p = 0.042$), CRP > 10 mg/L (HR 3.39, 95% CI 2.06–5.61, $p < 0.001$), and the corticosteroid use (HR 3.91, 95% CI 2.02–7.56, $p < 0.001$) were independent risk factors associated with long-term SARS-CoV-2 RNA shedding.

Construction of a Scoring Model

In order to facilitate clinical use and further assessment, a scoring model was established according to the results of multivariate Cox analysis, called the CCCCA score, which scores from 5 (age ≤ 50 years, absence of comorbidity, CD4 + T cell > 410 cells/ul, CRP ≤ 10 mg/L, no corticosteroid use) to 14 points (age > 50 years, presence of comorbidity, CD4 + T cell ≤ 410 cells/ul, CRP > 10 mg/L, corticosteroid use). Using Shanghai Public Health Clinical Center laboratory reference range standards, we chose the lower limit of normal (≤ 410 cells/ul) as cut-off for CD4 + T cell count, and the upper limit of normal (> 10 mg/L) as cut-off for CRP. The calculator of the CCCCA score is shown in Table 3. ROC analysis was used to assess the performance of the CCCCA score in the training set (Fig. 2a). For the prediction of long-term SARS-CoV-2 RNA shedding, the AUROC of the

Table 1 Clinical characteristics of patients in the training set

	All (<i>n</i> = 796)	Short term (<i>n</i> = 680)	Long-term (<i>n</i> = 116)	<i>p</i> value
VSD (days)	13 (9–18)	12 (9–15)	26 (24–30)	< 0.001
Age (years)	36 (25–51)	36 (24–50)	39 (28–56)	0.012
≤ 50	593 (74.5%)	516 (77.5%)	65 (56.0%)	< 0.001
> 50	203 (25.5%)	152 (22.4%)	51 (44.0%)	< 0.001
Male, <i>n</i> (%)	456 (57.3%)	391 (57.5%)	65 (56.0%)	0.768
Obesity, <i>n</i> (%)	205 (25.8%)	168 (24.7%)	37 (31.9%)	0.102
Smoking, <i>n</i> (%)	138 (17.3%)	118 (17.4%)	20 (17.2%)	0.977
Comorbidity, <i>n</i> (%)	134 (16.8%)	102 (15%)	32 (27.6%)	0.001
Fever, <i>n</i> (%)	411 (51.6%)	338 (49.7%)	73 (62.9%)	0.008
SOTA (days)	4 (3–8)	4 (3–7)	4 (3–8)	0.684
WBC (10 ⁹ /L)	5.5 (4.4–6.8)	5.6 (4.4–6.9)	5.0 (4.1–5.8)	< 0.001
≤ 4.0	136 (17.1%)	109 (16.0%)	27 (23.3%)	0.055
> 4.0	660 (82.9%)	571 (84.0%)	89 (76.7%)	0.055
Lymphocyte (10 ⁹ /L)	1.5 (1.1–1.9)	1.5 (1.1–2.0)	1.4 (1.0–1.7)	0.004
≤ 1.0	183 (23.0%)	146 (21.5%)	37 (31.9%)	0.014
> 1.0	613 (77.0%)	534 (78.5%)	79 (68.1%)	0.014
CD4 + T cell (cells/ul)	579 (414–775)	585 (421–795)	517 (353–700)	0.008
≤ 410	193 (24.2%)	151 (22.2%)	42 (36.2%)	0.001
> 410	603 (75.8%)	529 (77.8%)	74 (63.8%)	0.001
CRP (mg/L)	0.5 (0.5–6.0)	0.5 (0.5–6.0)	1.5 (0.5–7.8)	0.003
≤ 10	650 (81.7%)	580 (85.3%)	70 (60.3%)	< 0.001
> 10	146 (18.3%)	100 (14.7%)	46 (39.7%)	< 0.001
LDH (U/L)	200 (177–232)	201 (178–232)	198 (173–232)	0.454
D-dimer (ng/mL)	0.3 (0.2–0.5)	0.3 (0.2–0.5)	0.3 (0.2–0.5)	0.316
ESR (mm/h)	29 (10–73)	27 (10–74)	35 (19–70)	0.060
Corticosteroids use	48 (6.0%)	29 (4.3%)	19 (16.4%)	< 0.001

VSD viral shedding duration; SOTA symptom onset to admission; WBC white blood count; CRP C-reactive protein; LDH lactate dehydrogenase; ESR erythrocyte sedimentation rate

CCCCA score was 0.87 (95% CI 0.85–0.89). Using a cut-off value of 8 points, the sensitivity was 75.9%, the specificity was 87.7%, the PPV was 51.2%, and the NPV was 95.5% (Table 4).

Clinical Characteristics of Patients in the Validation Set

Clinical characteristics of patients in the validation set are shown in Table 5. Of 415 patients

Table 2 Cox regression analysis of prolonged viral RNA shedding duration

	Univariate Cox analysis		Multivariate Cox analysis	
	HR (95% CI)	<i>p</i> value	HR (95% CI)	<i>p</i> value
Age (years)				
≤ 50	1	–	1	–
> 50	2.83 (1.81–4.10)	< 0.001	2.67 (1.54–3.69)	0.001
Comorbidity				
Without	1	–	1	–
With	2.16 (1.37–3.41)	0.001	1.83 (1.03–3.25)	0.039
Fever				
Without	1	–	1	–
With	1.72 (1.15–2.58)	0.009	1.28 (0.83–1.98)	0.271
WBC ($\times 10^9/L$)				
≤ 4.0	1	–		
> 4.0	1.55 (0.97–2.50)	0.069		
Lymphocyte ($\times 10^9/L$)				
> 1.0	1	–	1	–
≤ 1.0	1.71 (1.11–2.64)	0.014	1.07 (0.57–2.01)	0.825
CD4 + T cell (cells/ul)				
> 410	1	–	1	–
≤ 410	1.99 (1.31–3.03)	0.001	1.77 (1.05–2.89)	0.042
CRP (mg/L)				
≤ 10	1	–	1	–
> 10	3.81 (2.48–5.85)	< 0.001	3.39 (2.06–5.61)	< 0.001
Corticosteroids use				
Without	1	–	1	–
With	4.40 (2.37–8.15)	< 0.001	3.91 (2.02–7.56)	< 0.001

WBC white blood count; CRP C-reactive protein; HR hazard ratio; CI confidence interval

in the validation set, 81 (19.5%) were over 50 years, 64 (15.4%) had comorbidity, 53 (12.8%) had CD4 + T cell \leq 410 cells/ul, 17 (4.1%) had CRP > 10 mg/L, 6 (1.4%) received corticosteroid therapy, and 29 (7.0%) met criteria for long-term SARS-CoV-2 RNA shedding. The median CCCCA score was higher in long-

term group compared with short-term group (9 vs. 5, $p < 0.001$).

Predictive Power of the CCCCA Score in the Validation Set

In the validation set, the predictive power of the CCCCA score was also good, with an AUROC of

Table 3 Calculator of the CCCCA score

	Points
Age (years)	
≤ 50	1
> 50	3
Comorbidity	
Without	1
With	2
CD4+ T cell (cells/ul)	
> 410	1
≤ 410	2
CRP (mg/L)	
≤ 10	1
> 10	3
Corticosteroids use	
Without	1
With	4

0.83 (95% CI 0.79–0.87) (Table 4; Fig. 2b). Using a cut-off value of 8 points, we found sensitivity, specificity, PPV, and NPV of 51.7%, 92.2%, 33.3%, and 96.2%, respectively. Long-term SARS-CoV-2 RNA shedding duration increased from 14/370 (3.8%) in patients with CCCCA < 8 points to 15/45 (33.3%) in patients with CCCCA ≥ 8 points.

DISCUSSION

Defining viral shedding of SARS-CoV-2 is of utmost importance for determining an appropriate period of isolation. Negative SARS-CoV-2 RNA tests may be useful in providing certainty about non-infectiousness before ending isolation in COVID-19 patient groups with known risk factors for prolonged viral shedding. In this study, we derived and validated a score model, called the CCCCA score, based on five variables (comorbidity, CRP, CD4 + T cell, corticosteroid use, age), aimed at predicting prolonged SARS-CoV-2 RNA shedding. For the prediction of long-term SARS-CoV-2 RNA shedding, the CCCCA score was excellent, with an AUROC of

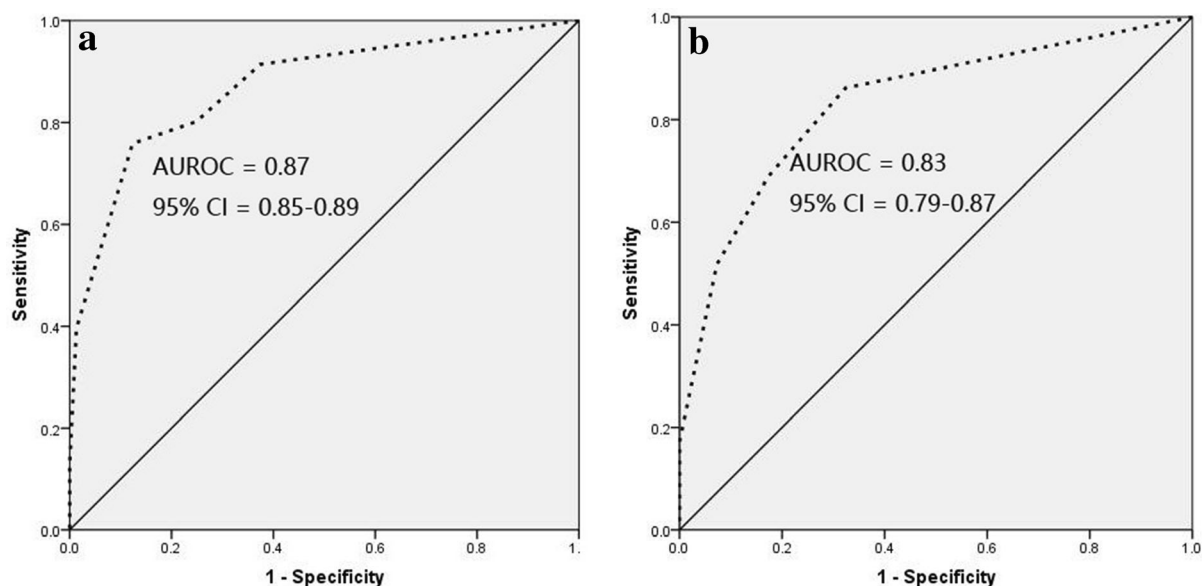


Fig. 2 ROC curves of the CCCCA score in the training set (**a**) and the validation set (**b**). For the prediction of prolonged viral RNA shedding duration, the AUROC of

the CCCCA score was 0.87 (95% CI 0.85 to 0.89) in the training set (**a**) and 0.83 (95% CI 0.79 to 0.87) in the validation set (**b**), respectively

Table 4 Accuracy of the CCCCA score for the prediction of prolonged viral RNA shedding

	Training set (<i>n</i> = 796)		Validation set (<i>n</i> = 415)	
AUROC	0.87 (0.85–0.89)		0.83 (0.79–0.87)	
Cut-off value	8		8	
	Sensitivity	75.9	Sensitivity	51.7
	Specificity	87.7	Specificity	92.2
	PPV	51.2	PPV	33.3
	NPV	95.5	NPV	96.2

PPV, Positive predictive value; NPV, Negative predictive value;

0.87 (95% CI 0.85–0.89) in the training set and an AUROC of 0.83 (95% CI 0.79–0.87) in the validation set.

In this study, age > 50 years, comorbidity, CD4 + T cell ≤ 410 cells/ul, CRP > 10 mg/L, and corticosteroid use were identified as independent high-risk factors associated with long-term SARS-CoV-2 RNA shedding. Previous studies have found several factors to be associated with prolonged shedding. These include older age, male sex, concomitant hypertension or diabetes, delayed admission to hospital after illness onset, severe illness at admission, invasive mechanical ventilation, administration of corticosteroid, and so on [12, 14–17]. In the study by Zhou et al., multivariable regression showed prolonged viral shedding independently associated with older age (odds ratio 1.02, $p = 0.003$) [12]. Another study reported time from onset to admission (HR = 0.829, $p < 0.001$), and administration of corticosteroid (HR = 0.496, $p = 0.002$), arbidol (HR = 2.605, $p = 0.008$) and oseltamivir (HR = 0.416, $p < 0.001$) were independently associated with duration of viral shedding [14]. In a retrospective study, risk factors associated with SARS-CoV-2 RNA shedding were evaluated in a cohort of 113 symptomatic patients from two hospitals outside Wuhan, and the results showed that prolonged SARS-CoV-2 RNA shedding was associated with male sex, old age, concomitant hypertension,

delayed admission to hospital after illness onset, severe illness at admission, invasive mechanical ventilation, and corticosteroid treatment [15]. In a study by Chen et al., older age (HR 0.99, 95% CI 0.98–1.00; $p = 0.04$), time lag from illness onset to hospital admission (HR 0.91, 95% CI 0.88–0.94; $p < 0.001$), and corticosteroid treatment (HR 0.60, 95% CI 0.39–0.94; $p = 0.024$) were significantly and independently associated with prolonged viral RNA shedding [16]. In order to identify risk factors for prolonged viral shedding, Buetti et al. used univariate and multivariate Cox proportional hazards models, and found that type two diabetes was associated with a prolonged viral RNA shedding (HR: 0.31, 95% CI 0.11–0.89, $p = 0.029$) [17].

Although previous studies have investigated the risk factors associated with the duration of SARS-CoV-2 RNA shedding, they did not establish a prediction model of long-term SARS-CoV-2 RNA shedding. There are currently considerable gaps regarding a scoring model for the prediction of prolonged SARS-CoV-2 RNA shedding. In this study, we derived a risk factors scoring system (the CCCCA score), which could identify a group of patients with a low risk of long-term SARS-CoV-2 RNA shedding. Over 95% of patients with CCCCA score of 5–8 points will not experience long-term SARS-CoV-2 RNA shedding. Long-term SARS-CoV-2 RNA shedding increased from 3.8% in patients with CCCCA score < 8 points to 33.3% in patients with a CALL score ≥ 8 points.

In a previous retrospective study, the median duration of SARS-CoV-2 RNA detection was 17 (13–22) days as measured from illness onset [15]. Liu et al. reported the median duration of viral shedding was 25.0 days (IQR 20.0–30.0) in 140 medical workers with COVID-19 in Wuhan, China [18]. In a retrospective, multicenter cohort study, Zhou et al. included all adult inpatients with COVID-19 from Jinyintan Hospital and Wuhan Pulmonary Hospital, and found that the median duration of viral shedding was 20.0 days (IQR 17.0–24.0) in survivors, but SARS-CoV-2 was detectable until death in non-survivors [12]. In this study, the median duration of SARS-CoV-2 RNA shedding was 13 (9–18) days. In contrast, the median duration of

Table 5 Clinical characteristics of patients in the validation set

	All (<i>n</i> = 415)	Short-term (<i>n</i> = 386)	Long-term (<i>n</i> = 29)	<i>p</i> value
VSD (days)	9 (9–13)	9 (9–12)	29 (24–36)	< 0.001
Age (years)	35 (28–48)	35 (28–47)	42 (31–55)	0.012
≤ 50	334 (80.5%)	316 (81.9%)	18 (62.1%)	0.010
> 50	81 (19.5%)	70 (18.1%)	11 (37.9%)	0.010
Male, <i>n</i> (%)	287 (69.2%)	268 (69.4%)	19 (65.5%)	0.670
Comorbidity, <i>n</i> (%)	64 (15.4%)	52 (13.5%)	12 (41.4%)	< 0.001
Fever, <i>n</i> (%)	85 (20.5%)	72 (18.7%)	13 (44.8%)	< 0.001
SOTA (days)	3 (2–5)	3 (2–5)	3 (2–6)	0.477
WBC (10 ⁹ /L)	6.6 (5.4–7.8)	6.6 (5.5–7.8)	5.3 (4.2–6.8)	< 0.001
Lymphocyte (10 ⁹ /L)	1.7 (1.3–2.2)	1.8 (1.4–2.2)	1.4 (0.9–1.7)	< 0.001
CD4 + T cell (cells/ul)	692 (525–887)	695 (533–900)	572 (374–769)	0.011
≤ 410	53 (12.8%)	44 (11.4%)	9 (31.0%)	0.002
> 410	362 (87.2%)	342 (88.6%)	20 (69.0%)	0.002
CRP (mg/L)	0.5 (0.5–0.5)	0.5 (0.5–0.5)	0.5 (0.5–1.5)	0.003
≤ 10	398 (95.9%)	373 (96.3%)	25 (86.2%)	0.006
> 10	17 (4.1%)	13 (3.7%)	4 (13.8%)	0.006
LDH (U/L)	174 (156–203)	174 (156–204)	170 (157–204)	0.730
D-dimer (ng/mL)	0.2 (0.2–0.3)	0.2 (0.2–0.3)	0.2 (0.2–0.4)	0.179
ESR (mm/h)	12 (9–32)	12 (9–32)	14 (11–46)	0.064
Corticosteroids use	6 (1.4%)	4 (1.0%)	2 (6.9%)	0.011
CCCCA Score	5 (5–6)	5 (5–6)	9 (7–10)	< 0.001

VSD viral shedding duration; SOTA symptom onset to admission; WBC white blood count; CRP C-reactive protein; LDH lactate dehydrogenase; ESR erythrocyte sedimentation rate

viral shedding in this study was shorter than that reported by previous studies. There are several possible explanations for the differences. First, the statistical difference may have been caused by the size of the sample. We provided a large sample size (1211 patients) compared with previous studies (113 patients, 140 patients, and 384 patients, respectively). Second, there are some differences in the study populations regarding age, comorbidity, administration of corticosteroid, and the degree of severity, which are the confirmed risk factors related to SARS-CoV-2 RNA shedding.

One previous study reported that fever (temperature > 38.5 °C) (OR 5.1, 95% CI 1.5–18.1) was associated with increased odds of prolonged duration of viral shedding [19]. Another study also reported that the highest temperature at admission (OR 5.200; 95% CI 1.190–22.726; *p* = 0.028) was one of the risk factors for prolonged duration of viral shedding [20]. In this study, multivariate Cox analysis showed that fever (HR 1.28, 95% CI 0.83–1.98, *p* = 0.271) was not an independent risk factor associated with long-term SARS-CoV-2 RNA shedding. There are several possible

explanations for the differences. First, the two previously mentioned studies included severe patients with COVID-19 (14–30.3%) [19, 20], while this study only included non-severe patients with COVID-19. Second, the prevalence of patients with fever is different (42.2%, 51.6%, and 78%, respectively). Third, the previous two studies had small sample sizes (66 and 147 patients, respectively). The statistical bias might be caused by the small size of the sample. In this study, we provided a much larger sample, including 796 patients in the training set and 415 patients in the validation set.

In this study, multivariate Cox analysis identified that $CD4 + T \text{ cell} \leq 410 \text{ cells/ul}$ (HR 1.77, $p = 0.042$), rather than lymphocyte $\leq 1.0 \times 10^9/L$ (HR 1.07, $p = 0.825$), was one of the independent risk factors associated with long-term SARS-CoV-2 RNA shedding. Although lymphocyte count is a more practical variable to use, $CD4 + T$ cell count is a more accurate marker of host's immunity, which is related with the SARS-CoV-2 RNA shedding.

Our study has several limitations. First, SARS-CoV-2 RNA shedding in nasopharyngeal or throat swab specimens does not necessarily represent true viral shedding, since recent research has indicated the prolonged presence of SARS-CoV-2 viral RNA in fecal samples [21]. In addition, the definition of starting point varies between studies (symptom onset and/or first positive RT-PCR). Second, we detected SARS-CoV-2 by RT-PCR but not by virus isolation. According to a recent report, RNA can be detected long after the disappearance of infectious virus. The immune system can neutralize viruses but not eliminate RNA, degrading slowly [22]. The presence of viral RNA does not always correlate with viability and transmissibility of virus, unless the virus is isolated and cultured from the specimens. However, the practical approach of discharging patients based on shedding cessation usually utilizes the widely available PCR tests rather than viral culture. Third, this study is a retrospective study, and might have the Type II error. Therefore, the results in this study should be interpreted with caution. Prospective, multicenter, large-scale studies are needed to further assess the performance of the CCCCA score for the prediction of

long-term SARS-CoV-2 RNA shedding. Four, in this study, pediatric patients were not seen. In Shanghai, China, under the regulations of the Shanghai Health Committee, all pediatric patients with COVID-19 were hospitalized at another designated hospital, the Children's Hospital of Fudan University.

CONCLUSIONS

In conclusion, the median time for SARS-CoV-2 RNA shedding was 13 days (IQR 9–18) in non-severe hospitalized adult patients with COVID-19. The CCCCA score seems to be a good prognosticator for long-term SARS-CoV-2 RNA shedding. Prospective, multicenter, large-scale studies are needed to further assess the performance of the CCCCA score for the prediction of long-term SARS-CoV-2 RNA shedding.

ACKNOWLEDGEMENTS

We thank all doctors who work in Shanghai Public Health Clinical Center for their efforts in the diagnosis and treatment of patients with COVID-19.

Funding. This study was supported by grant NO.19YF1441200 from Shanghai Association for Science and Technology. The journal's Rapid Service Fee funded by the authors. The funding organizations are public institutions and had no role in the design and conduct of the study; collection, management, and analysis of the data; or preparation, review, and approval of the manuscript.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole, and have given their approval for this version to be published.

Authors' Contributions. Study concept and design: Qiang Li. Data collection: ChenLu Huang, Ling Fei, WeiXia Li, Wei Xu, and XuDong Xie. Analysis and interpretation of

data: ChenLu Huang, Ling Fei, WeiXia Li, and Qiang Li. Drafting of the manuscript: Qiang Li. Critical revision of the manuscript: Liang Chen.

Disclosures. ChenLu Huang, Ling Fei, WeiXia Li, Wei Xu, XuDong Xie, Qiang Li, and Liang Chen have nothing to declare.

Compliance with Ethics Guidelines. The study was approved by the Clinical Research Ethics Committee of the Shanghai Public Health Clinical Center (No. 2020-S219-01). Although this is a retrospective study, at hospital admission, all patients provided verbal consent for their clinical data might be used for further medical study. The procedures followed were in accordance with the ethical standards of the Helsinki Declaration (1964, amended most recently in 2008) of the World Medical Association.

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

Open Access. This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China. *Nature*. 2020;579:265–9.
2. Zhong NS, Zheng BJ, Li YM, et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. *Lancet*. 2003;362:1353–8.
3. World Health Organization. Weekly operational update on COVID-19. Weblink: <https://www.who.int/publications/m/item/weekly-epidemiological-update>. Accessed 12 January 2021.
4. Li W, Su YY, Zhi SS, et al. Virus shedding dynamics in asymptomatic and mildly symptomatic patients infected with SARS-CoV-2. *Clin Microbiol Infect*. 2020;26:1556.e1-1556.e6.
5. Widders A, Broom A, Broom J. SARS-CoV-2: the viral shedding vs infectivity dilemma. *Infect Dis Health*. 2020;25:210–5.
6. Bennasrallah C, Bannour R, Jlassi O, et al. Three COVID-19 cases with a long-term viral shedding period in Tunisia. *Pan Afr Med J*. 2020;35:117.
7. Liu Y, Yan LM, Wan L, et al. Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect Dis*. 2020;20:656–7.
8. Zheng S, Fan J, Yu F, et al. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. *BMJ*. 2020;369:m1443.
9. World Health Organization. Laboratory diagnostics for novel coronavirus. Weblink: <https://www.who.int/health-topics/coronavirus/laboratory-diagnostics-for-novel-coronavirus>. Accessed 6 February 2020.
10. National Health Commission of China. New coronavirus pneumonia revention and control program (the fifth edition). Weblink: <http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440>. Accessed 4 February 2020.
11. Agarwal V, Venkatakrishnan AJ, Puranik A, et al. Long-term SARS-CoV-2 RNA shedding and its temporal association to IgG seropositivity. *Cell Death Discov*. 2020;6:138.
12. Zhou C, Zhang T, Ren H, et al. Impact of age on duration of viral RNA shedding in patients with COVID-19. *Aging (Albany NY)*. 2020;12:22399–404.

13. Lin A, He ZB, Zhang S, Zhang JG, Zhang X, Yan WH. Early risk factors for the duration of severe acute respiratory syndrome coronavirus 2 viral positivity in patients with coronavirus disease 2019. *Clin Infect Dis*. 2020;71:2061–5.
14. Hu F, Yin G, Chen Y, et al. Corticosteroid, oseltamivir and delayed admission are independent risk factors for prolonged viral shedding in patients with Coronavirus Disease 2019. *Clin Respir J*. 2020;14:1067–75.
15. Xu K, Chen Y, Yuan J, et al. Factors associated with prolonged viral RNA shedding in patients with coronavirus disease 2019 (COVID-19). *Clin Infect Dis*. 2020;71:799–806.
16. Chen X, Zhu B, Hong W, et al. Associations of clinical characteristics and treatment regimens with the duration of viral RNA shedding in patients with COVID-19. *Int J Infect Dis*. 2020;98:252–60.
17. Buetti N, Trimboli P, Mazzuchelli T, et al. Diabetes mellitus is a risk factor for prolonged SARS-CoV-2 viral shedding in lower respiratory tract samples of critically ill patients. *Endocrine*. 2020;70:454–60.
18. Liu W, Liu Y, Xu Z, et al. Clinical characteristics and predictors of the duration of SARS-CoV-2 viral shedding in 140 healthcare workers. *J Intern Med*. 2020;288:725–36.
19. Li TZ, Cao ZH, Chen Y, et al. Duration of SARS-CoV-2 RNA shedding and factors associated with prolonged viral shedding in patients with COVID-19. *J Med Virol*. 2021;93:506–12.
20. Qi L, Yang Y, Jiang D, et al. Factors associated with the duration of viral shedding in adults with COVID-19 outside of Wuhan, China: a retrospective cohort study. *Int J Infect Dis*. 2020;96:531–7.
21. Wu Y, Guo C, Tang L, et al. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *Lancet Gastroenterol Hepatol*. 2020;5:434–5.
22. Atkinson B, Petersen E. SARS-CoV-2 shedding and infectivity. *Lancet*. 2020;395:1339–40.