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

Xiaoli Lei
Henan Provincial People's Hospital, People's Hospital of Zhengzhou University
Junli Wang
Fuwai Central China Cardiovascular Hospital
Lijie Kou
Henan Provincial People's Hospital, People's Hospital of Zhengzhou University
Zhigang Yang (18203679389@163.com)
Henan Provincial People's Hospital, People's Hospital of Zhengzhou University

Research Article

Keywords: Prediction model, Coronavirus disease-19, Severe acute respiratory syndrome coronavirus 2, Viral shedding, Retrospective cohort study

Posted Date: October 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-940304/v1

Abstract

Background: Because of the lack of compelling evidence for predicting the duration of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA shedding, the purpose of this retrospective study was to establish a predictive model for long-term SARS-CoV-2 RNA shedding in non-death hospitalized patients with coronavirus disease-19 (COVID-19).

Methods: 97 non-death hospitalized patients with COVID-19 admitted to two hospitals in Henan province of China from February 3, 2020 to March 31, 2020 were retrospectively enrolled. Multivariate logistic regression was performed to identify the high risk factors associated with long-term SARS-CoV-2 RNA shedding and a predictive model was established and represented by a nomogram. Its performance was assessed with discrimination and calibration.

Results: 97 patients were divided into the long-term (>21 days) group (n = 27, 27.8%) and the short-term (≤ 21 days) group (n = 70, 72.2%) based on their viral shedding duration. Multivariate logistic regression analysis showed that time from illness onset to diagnosis (OR 1.224, 95% CI 1.070-1.400, P = 0.003) and interstitial opacity in chest computerized tomography(CT) scan (OR 6.516, 95% CI 2.041-20.798, P = 0.002) were independent risk factors for long-term SARS-CoV-2 RNA shedding. A prediction model, which is presented with a nomogram, was established by incorporating the two risk factors. The goodness-of-fit statistics for the nomogram was not statistically significant (χ^2 = 8.292; P = 0.406), and its area under the receiver operator characteristic curve was 0.834 (95% CI 0.731- 0.936; P < 0.001).

Conclusion: The established model has a good predictive performance on the long-term viral RNA shedding in non-death hospitalized patients with COVID-19, but it still needs further validation by independent data set of large samples in the future.

Background

Since December 2019, a novel coronavirus, named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has spread rapidly across the globe, resulting in a global outbreak of coronavirus disease-19 (COVID-19) [1]. As of January 12, 2021, more than 88 million cases and 1.9 million deaths have been reported globally [2]. Due to the rapid spread of COVID-19, it has caused a huge shortage of medical resources and social and economic burden. On 30th January 2020, the World Health Organization (WHO) declared the outbreak of COVID-19 as a public health emergency of international concern and strongly recommended that all countries should take action to detect infections and prevent transmission [3].

As is known to all, controlling the source of infection, cutting off transmission routes and protecting vulnerable population are important principles for controlling infectious diseases. Early isolation and management of patients with COVID-19 is an important strategy to reduce community transmission and spread of the disease.

SARS-CoV-2 RNA in respiratory specimens is an important biomarker. A positive test for SARS-CoV-2 RNA is not only the gold standard for identifying a COVID-19 patient, but also often a means that the patient is infectious and needs to be isolated. Determining the length of isolation of the patient depends on the time of transmission of the patient's virus. In clinical practice, it is often determined by the duration of SARS-CoV-2 RNA positive test in respiratory specimen such as throat swab. Negative SARS-CoV-2 RNA test of respiratory tract specimens indicates that the patient is no longer infectious and can be released from quarantine and discharged.

Many studies have suggested that the shedding duration of SARS-CoV-2 varies among different populations. Li *et al.* reported that the median duration of viral shedding was 11.5 days, 28 days and 31 days for presymptomatic, asymptomatic, and mildly symptomatic patients with COVID-19, respectively [4]. Similarly, Liu *et al.* found that the median duration of viral shedding was 25.0 days (IQR 20.0-30.0) in 140 health care workers with COVID-19 in Wuhan, China [5]. In view of this, it is necessary to study the risk factors associated with SARS-CoV-2 RNA shedding in order to help clinicians identify those patients who may have prolonged viral shedding early and predict the isolation time. Buetti *et al.* found that diabetes was associated with long-term SARS-CoV-2 viral shedding (HR: 0.31, 95% CI 0.11 0.89, P = 0.029) [6]. Another study reported that time from onset to admission, administration of corticosteroid, arbidol, and oseltamivir were independent risk factors associated with duration of SARS-CoV-2 RNA shedding [7].

Although the above studies have initially explored independent risk factors associated with SARS-CoV-2 RNA shedding, there have been few studies on the establishment of models to predict the long-term SARS-CoV-2 viral shedding in non-death inpatients with COVID-19. Therefore, we aimed to identify the high risk factors associated with long-term SARS-CoV-2 RNA shedding, and to establish a model for the prediction of prolonged shedding of SARS-CoV-2 RNA, which was shown as a nomogram, in the specific populations mentioned above.

Methods

Study Population

This retrospective cohort study included two cohorts of consecutive inpatients with COVID-19 from Henan Provincial People's Hospital (Zhengzhou, Henan Province, China) and Anyang Infectious Disease Hospital (Anyang, Henan Province, China). The two hospitals were designated hospitals for the treatment of COVID-19 in Henan Province. A total of 123 patients with confirmed COVID-19 including 58 cases from Henan Provincial People's Hospital and 65 cases from Anyang Infectious Disease Hospital who were transferred to the above two hospitals from February 3, 2020 to March 31, 2020 were included.

Diagnostic Criteria

Patients with COVID-19 can be confirmed when throat swab specimens are tested positive for SARS-CoV-2 RNA by real-time reverse transcription polymerase chain reaction (RT-PCR), which is the gold standard for testing for SARS-CoV-2 infection as recommended by WHO interim guidelines [8]. Cessation of SARS-CoV-2 RNA shedding was defined as negative RT-PCR on two consecutive throat swab specimens at least 24 hours apart. The duration of SARS-CoV-2 RNA shedding was defined as the time from illness onset to cessation of SARS-CoV-2 RNA shedding. At present, there is no unified time standard for the cut-off value of prolonged SARS-CoV-2 RNA shedding in academia. Based on previous research results, the median duration of virus shedding was 12-21 days [4, 9]. Therefore, the long-term SARS-CoV-2 RNA shedding was defined as SARS-CoV-2 RNA shedding time greater than 21 days in this study.

Inclusion criteria: (1) patients with confirmed COVID-19 diagnosis; (2) patients with definite outcome of viral shedding duration; (3) patients with complete clinical data.

Exclusion criteria: (1) patients presenting with death while in hospital; (2) patients with incomplete clinical data.

Of the 123 confirmed COVID-19 patients, 19 patients with incomplete clinical data and 7 dead cases were excluded. Finally, 97 patients were enrolled as study subjects. The flow chart of the study population is shown in Fig. 1.

Data collection

We used standardized data collection form to extract demographic, epidemiological, clinical, laboratory, radiological data and SARS-CoV-2 RNA shedding duration from electronic medical records.

Laboratory procedures

Methods for laboratory confirmation of SARS-CoV-2 infection have been described elsewhere [8]. After the clinical symptoms were relieved, throat swab samples were taken every other day for SARS-CoV-2 RT-PCR review, but only qualitative information was available. Routine blood tests included complete blood count, coagulation spectrum, serum biochemical tests [including liver function, kidney function, glucose, creatine kinase (CK), lactate dehydrogenase (LDH)] and inflammatory markers such as C-reactive protein (CRP), calcitonin (PCT) were performed. All hospitalized patients underwent CT scans. The frequency of examinations was determined by the doctor treating the patient.

We referred to the Chinese management guideline for COVID-19 (version 8.0) to define the severity of COVID-19 [10]. In this study, severe and critical cases were defined as severe cases. Mild and normal cases were defined as non-severe cases.

Statistical analysis

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 n (%). The Mann-Whitney U test, χ^2 test, or Fisher's exact test were used to compare differences between long-term viral shedding group and short-term viral shedding group where appropriate. The significance of each variable was assessed by univariate and multivariate logistic analyses to identify the independent risk factors associated with long-term SARS-CoV-2 RNA shedding.

SPSS software v.15.0 (IBM Inc., Chicago, IL, USA) and R software v.3.6.1 (Foundation for Statistical Computing, Vienna, Austria) were used for statistical analysis. All significance tests were two-tailed, and $P < 0.05$ was considered as statistically significant.

Feature selection and nomogram establishment

The statistically significant variables from the multivariable logistic regression analysis were selected to be the most useful predictive features from our cohort data set. To provide the clinician with a quantitative tool to predict individual probability of long-term viral shedding, we built a nomogram, which was established with the R rms package, on the basis of multivariable logistic analysis.

Predicting Performance of the Nomogram

We used logistic regression to analyze the predicted long-term virus shedding probability, and then constructed a receiver operating characteristic curve (ROC). In order to quantify the discrimination performance of the nomogram, the area under the ROC (AUROC) curve is calculated. Calibration curves were plotted to evaluate the calibration of the nomogram, accompanied with the Hosmer-Lemeshow test.

Results

Clinical Characteristics of Patients

Clinical characteristics of patients in our cohort data set were shown in Table 1. Overall, 97 patients with COVID-19 were enrolled. The mean age was 49.3±16.1 years, 44 patients (45.4%) were male, 17 patients (17.5%) were older than 65 years, and the median SARS-CoV-2 RNA shedding duration was 17 days (IQR 14-22). 16 patients (16.5%) had diabetes, 24 patients (24.7%) had hypertension, 6 patients (6.2%) had coronary heart disease, 6 patients (6.2%) had cerebrovascular disease, and 8 patients (8.2%) had pulmonary disease.

97 patients were divided into long-term (> 21 days) group (n = 27, 27.8%) and short-term (≤ 21 days) group (n = 70, 72.2%) based on their viral shedding duration. The age of long-term group was older than that of short-term group (55.7±14.1 vs. 46.7±16.2) (P = 0.013). The viral shedding duration of long-term

group was longer than that of short-term group [26 (23-31) vs. 15 (12-18)] (P < 0.001). The time from illness onset to diagnosis of long-term group was longer than that of short-term group [10 (6-12) vs. 4 (3-6)] (P < 0.001). The number of cases and composition ratios of long-term group and short-term group with hypertension, dyspnea, illness severity, procalcitonin greater than 0.05 ng/ml, and interstitial opacity in chest CT scan were 12 (44.4%) vs. 12 (17.1%), 13 (48.1%) vs. 15 (21.4%), 16 (59.3%) vs. 19 (27.1%), 15 (55.6%) vs. 18 (25.7%), and 14 (51.9%) vs. 8 (11.4%), respectively (P < 0.05 for all) (Table 1).

Table 1

Demographic, clinical, laboratory, and radiographic findings of non-death hospitalized patients with COVID-19 on admission

	Total (n = 97)	Short-term (n = 70)	Long-term (n = 27)	*P* value
Demographics and clinical characteristics				
Age, years	49.3±16.1	46.7±16.2	55.7±14.1	0.013
Age, ≥65years	17(17.5%)	10(14.3%)	7(25.9%)	0.292
Sex				
Female	53(54.6%)	42(60.0%)	11(40.7%)	0.088
Male	44(45.4%)	28(40%)	16(59.3)	
Time from illness onset to diagnosis, days	5(3-9)	4(3-6)	10(6-12)	<0.001
Exposure history	88(90.7%)	62(88.5%)	26(96.3%)	0.433
Coexisting underlying diseases at admission				
Diabetes	16(16.5%)	9(12.9%)	7(25.9%)	0.212
Hypertension	24(24.7%)	12(17.1%)	12(44.4%)	0.005
Coronary heart disease	6(6.2%)	5(7.1%)	1(3.7%)	0.873
Cerebrovascular disease	6(6.2%)	3(4.3%)	3(11.1%)	0.435
Pulmonary disease	8(8.2%)	6(8.6%)	2(7.4%)	1.000
Hepatopathy	4(4.1%)	3(4.3%)	1(3.7%)	1.000
Tumor	2(2.1%)	2(2.9%)	0(0.0%)	0.928
Fever (temperature ≥37.3°C)	81(83.5%)	59(84.3%)	22(81.5%)	0.977
Headache	11(11.3%)	9(12.9%)	2(7.4%)	0.688
Muscular soreness	15(15.5%)	11(15.7%)	4(14.8%)	1.000
Fatigue	43(44.3%)	33(47.1%)	10(37.0%)	0.503
Rhinobyon	7(7.2%)	6(8.6%)	1(3.7%)	0.695
Rhinorrhoea	5(5.2%)	4(5.7%)	1(3.7%)	1.000
Sore throat	11(11.3%)	11(15.7%)	0(0.0%)	0.067
Cough	58(59.8%)	40(57.1%)	18(66.7%)	0.391
Sputum	29(29.9%)	18(25.7%)	11(40.7%)	0.147

Dyspnea	28(28.9%)	15(21.4%)	13(48.1%)	0.009
Diarrhoea	8(8.2%)	5(7.1%)	3(11.1%)	0.822
Disturbance of consciousness	5(5.2%)	2(2.9%)	3(11.1%)	0.256
Systolic pressure, <90mmHg	0(0.0%)	0(0.0%)	0(0.0%)	-
Diastolic pressure, <60mmHg	5(5.2%)	3(4.3%)	2(7.4%)	0.912
Respiratory rate, ≥30 breaths per min	0(0.0%)	0(0.0%)	0(0.0%)	-
Pulse , ≥100 beats per min	12(12.4%)	10(14.3%)	2(7.4%)	0.563
Disease severity status				
Severe	35(36.1%)	19(27.1%)	16(59.3%)	0.003
Non-severe	62(63.9%)	51(72.9%)	11(40.7%)	
Laboratory findings				
White blood cell count □109/L	4.6(3.6-6.6)	4.5(3.4-6.5)	5.5(3.8-7.4)	0.265
>4.0	68(70.1%)	49(70.0%)	19(70.4%)	0.792
<4.0	29(29.9%)	21(30.0%)	8(29.6%)	
Lymphocyte count□10^9/L	1.2(0.9-1.5)	1.1(0.9-1.6)	1.3(0.9-1.5)	0.643
>0.8	84(86.6%)	61(87.1%)	23(85.2%)	1.000
<0.8	13(13.4%)	9(12.9%)	4(14.8%)	
CURB-65				
0	72(74.2%)	54(77.1%)	18(66.7%)	0.362
1	18(18.6%)	12(17.1%)	6(22.2%)	
2	6(6.2%)	4(5.7%)	2(7.4%)	
3	1(1.0%)	0(0.0%)	1(3.7%)	
Procalcitonin,≥ 0.05ng/mL	33(34.0%)	18(25.7%)	15(55.6%)	0.005
CRP, mg/L	16.3(6.8-57.9)	15.1(3.9-46.2)	24.0(10.1-76.7)	0.170
CRP, >10mg/L	66(68.0%)	45(64.3%)	21(77.8%)	0.202

ALT, U/L	22.0(17.7-33.3)	23.0(18.4-33.1)	22.0(16.0-46.0)	0.917
ALT, >40U/L	18(18.6%)	11(15.7%)	7(25.9%)	0.246
AST, U/L	28.0(20.0-39.0)	28.0(20.1-37.4)	28.3(17.9-43.0)	0.949
AST, >40U/L	20(20.6%)	13(18.6%)	7(25.9%)	0.422
TBIL, mmol/L	10.8(8.6-15.1)	10.7(8.6-14.4)	10.9(8.3-20.8)	0.729
TBIL, >21mmol/L	13(13.4%)	8(11.4%)	5(18.5%)	0.558
LDH, U/L	201.0(174.0-252.0)	200.0(175.8-247.0)	212.0(155.0-331.0)	0.834
LDH, >245U/L	26(26.8%)	17(24.3%)	9(33.3%)	0.367
CK, mmol/L	54.0(34.0-95.5)	54.0(38.5-91.0)	43.0(24.0-101.0)	0.533
CK, >185 mmol/L	9(9.3%)	6(8.6%)	3(11.1%)	1.000
BUN, mmol/L	3.8(2.9-4.8)	3.6(2.8-4.8)	4.0(3.3-5.2)	0.161
BUN, >7.0mmol/L	8(8.2%)	6(8.6%)	2(7.4%)	1.000
CREA, mmol/L	55.0(47.0-65.0)	55.0(48.0-66.3)	53.0(44.0-63.0)	0.361
CREA, >133mmol/L	1(1.0%)	1(1.4%)	0(0.0%)	1.000
d-dimer, mg/L	0.3(0.2-0.6)	0.2(0.2-0.4)	0.4(0.2-0.7)	0.203
d-dimer, >0.5mg/L	25(25.8%)	15(21.4%)	10(37.0%)	0.115
Blood glucose, mmol/L	5.8(4.8-7.1)	5.7(4.8-7.1)	5.9(4.6-7.2)	0.732
Blood glucose, >7.0mmol/L	25(25.8%)	18(25.7%)	7(25.9%)	0.983
Imageological change				
Sites of imaging abnormalities				
No lesion on either side	8 (8.2%)	5 (7.1%)	3 (11.1%)	0.818
Unilateral lesion	8 (8.2%)	6 (8.6%)	2 (7.4%)	
Bilateral lesions	81 (83.5%)	59 (84.3%)	22 (81.5%)	
Morphology of the imaging changes				
Consolidation	32 (33.0%)	21 (30.0%)	11 (40.7%)	0.313

Interstitial opacity	22 (22.7%)	8 (11.4%)	14 (51.9%)	<0.001
Viral shedding duration, days	17 (14-22)	15 (12-18)	26 (23-31)	<0.001

Abbreviations: COVID-19, coronavirus disease 2019; CRP, C reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; LDH, lactate dehydrogenase; CK, creatinine kinase; BUN, urea nitrogen; CREA, creatinine.

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

Logistic regression analysis of long-term SARS-CoV-2 RNA shedding duration is shown in Table 2. Multivariate logistic regression analysis revealed that time from illness onset to diagnosis (OR 2.67, 95% CI 1.070-1.400, P = 0.003) and interstitial opacity in chest CT (OR 6.516, 95% CI 2.041-20.798, P = 0.002) were independent risk factors associated with long-term SARS-CoV-2 RNA shedding.

Table 2

Logistic regression analysis of long-term SARS-COV-2 RNA shedding of non-death hospitalized patients with COVID-19

	Univariable Logistic Analysis		Multivariable Logistic Analysis	
	OR(95% CI)	P value	OR(95% CI)	P value
Demographics and clinical characteristics				
Age, years*	1.039(1.007-1.073)	0.016		
Age, ≥65years	2.100(0.706-6.249)	0.182		
Sex				
Female	1（ref）			
Male	2.182(0.883-5.390)	0.091		
Time from illness onset to diagnosis, days*	1.276(1.119-1.455)	<0.001	1.224(1.070-1.400)	0.003
Exposure history	3.355(0.399-28.194)	0.265		
Coexisting underlying diseases at admission				
Diabetes	2.372(0.782-7.193)	0.127		
Hypertension	3.867(1.450-10.314)	0.007	3.210(0.985-10.465)	0.053
Coronary heart disease	0.500(0.056-4.489)	0.536		
Cerebrovascular disease	2.792(0.527-17.783)	0.227		
Pulmonary disease	0.853(0.161-4.514)	0.852		
Hepatopathy	0.859(0.085-8.637)	0.897		
Tumor	0.000(0.000-)	0.999		
Fever (T ≥ 37.3°C)	0.820(0.256-2.630)	0.739		
Headache	0.542(0.109-2.689)	0.454		

Muscular soreness	0.933(0.269-3.299)	0.913
Fatigue	0.660(0.265-1.641)	0.371
Rhinobyon	0.410(0.047-3.577)	0.420
Rhinorrhoea	0.635(0.068-5.948)	0.690
Sore throat	0.000(0.000-)	0.999
Cough	1.500(0.592-3.801)	0.393
Sputum	1.986(0.779-5.065)	0.151
Dyspnea	3.405(1.321-8.773)	0.011
Diarrhoea	1.625(0.360-7.326)	0.527
Disturbance of consciousness	4.260(0.669-26.995)	0.125
Systolic pressure, <90 mmHg	--	--
Diastolic pressure, <60 mmHg	1.787(0.282-11.331)	0.538
Respiratory rate, ≥30 breaths per min	--	--
Pulse, ≥ 100 beats per min	0.480(0.098-2.350)	0.365
Disease severity status		
Severe	1(ref)	
Non-severe	3.904(1.539-9.906)	0.004
Laboratory findings		
White blood cell count($\times 10^9$/L) *	1.161(0.988-1.363)	0.070
<4.0	0.982(0.372-2.595)	0.972
Lymphocyte count$\times 10^9$/L *	0.887(0.506-1.556)	0.676

<0.8	1.179(0.330-4.204)	0.800
CURB-65		0.890
0	1(ref)	
1	1.500(0.491-4.578)	0.476
2	1.500(0.253-8.888)	0.655
3	4846424594 (0.000-)	1.000
Procalcitonin, ≥ 0.05 ng/mL	3.611(1.426-9.146)	0.007
CRP, mg/L*	1.003(0.995-1.011)	0.504
CRP, >10 mg/L	1.944(0.694-5.450)	0.206
ALT, U/L*	1.004(0.988-1.021)	0.618
ALT, >40U/L	1.877(0.641-5.499)	0.251
AST, U/L*	1.003(0.980-1.027)	0.798
AST, >40U/L	1.535(0.537-4.388)	0.424
TBIL, mmol/L*	1.004(0.988-1.019)	0.626
TBIL, >21 mmol/L	1.761(0.521-5.958)	0.363
LDH, U/L*	1.001(0.997-1.004)	0.760
LDH, >245U/L	1.559(0.592-4.107)	0.369
CK, mmol/L*	1.003(0.997-1.008)	0.324
CK, >185 mmol/L	1.333(0.309-5.759)	0.700

BUN, mmol/L*	0.974(0.866-1.095)	0.656		
BUN, >7.0 mmol/L	0.853(0.161-4.514)	0.852		
CREA, mmol/L*	0.992(0.968-1.016)	0.493		
CREA, >133 mmol/L	0.000(0.000-)	1.000		
d-dimer, mg/L*	1.109(0.894-1.376)	0.347		
d-dimer, >0.5 mg/L	2.157(0.820-5.675)	0.119		
Blood glucose, mmol/L*	1.005(0.864-1.170)	0.944		
Blood glucose, >7.0 mmol/L	1.011(0.367-2.787)	0.983		
Imageological change		0.812		
Sites of imaging abnormalities				
No lesion on either side	1(ref)			
Unilateral lesion	0.556(0.065-4.755)	0.592		
Bilateral lesions	0.621(0.137-2.821)	0.538		
Morphology of the imaging changes				
Consolidation	1.604(0.638-4.035)	0.315		
Interstitial opacity	8.346(2.907-23.959)	<0.001	6.516(2.041-20.798)	0.002
Abbreviations: OR, odds ratio; CRP, C-reactive protein; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; LDH, lactate dehydrogenase; CK, creatinine kinase; BUN, urea nitrogen; CREA, creatinine; *Per 1 unit increase.				

Development of a Prediction Model

Time from illness onset to diagnosis (TIOD) and interstitial opacity in chest CT were identified as independent predictors by multivariate logistic regression analysis. Incorporating these two risk factors, a prediction model was constructed and displayed as a nomogram (Fig.2).

Predicting Performance of the Nomogram

In our cohort, the calibration curve of nomogram of the probability of long-term SARS-CoV-2 RNA shedding showed good consistency between prediction and observation (Fig.3). The Hosmer-Lemeshow test calculated a nonsignificant statistics (χ^2 = 8.292; P = 0.406), which showed that there was no deviation from perfect fitting. The area under receiver operating characteristic (AUROC) for nomogram predicting the probability of long-term SARS-CoV-2 viral shedding was 0.834 (95% CI 0.731-0.936, P < 0.001) (Fig.4).

Discussion

Determining viral shedding for SARS-CoV-2 is critical to determine the appropriate quarantine time. Negative SARS-CoV-2 RNA testing may be helpful in determining that COVID-19 patients are no longer non-infectious until isolation is ended for those patients with known risk factors for long-term virus shedding. Our findings revealed that time from illness onset to diagnosis and interstitial opacity in chest CT scan as independent risk factors associated with long-term SARS-CoV-2 RNA shedding. Based on these two variables, a prediction model presented as a nomogram was to be established for predicting long-term SARS-CoV-2 viral shedding.

Although risk factors associated with the duration of SARS-CoV-2 RNA shedding have been explored, there is currently a lack of predictive model that can predict long-term SARS-CoV-2 RNA shedding in non-death hospitalized COVID-19 patients. In non-severe hospitalized patients with COVID-19, Huang *et al.* recently reported a prediction model for long-term viral shedding, which was named the CCCCA score, that included five variables (comorbidity, CRP, $CD4^+$ T cell, corticosteroid use and age) [11]. The CCCCA score could identify patients with low risk of prolonged SARS-CoV-2 viral shedding. More than 95% of patients with a CCCCA score of 5-8 are less likely to have long-term SARS-CoV-2 RNA shedding [11]. Although this study has important implications for early identification of patients with long-term transmission of the virus, there were still some limitations. First, the study population was mild and did not include patients with severe COVID-19, so the results of this study might be biased by population selection. Second, there was no calibration test for this model, and it was impossible to verify whether the observed values and predicted values have good goodness of fit. In our study, we developed a clinical prediction model, which was presented in the form of a nomogram, has a very good predictive performance in predicting patients with long-term viral shedding. Area under the receiver operative characteristics (AUROC) of the model was 0.834 [(95% CI 0.731-0.936), P < 0.001] in our cohort. The calibration curve of the nomogram produced

insignificant statistics (χ^2 = 8.292; P = 0.406), indicating no deviation from the perfect fit. Our developed model took non-death patients with confirmed COVID-19 as the research population, and the research object covers a wider range, so our conclusion may avoid population selection bias and have better extensibility. At the same time, our model contains only two risk factors, time from illness onset to diagnosis and interstitial opacity in chest CT scan, which further make the model to be more practical and convenient for clinicians to use.

In this study, time from illness onset to diagnosis (OR 2.67, 95% CI 1.070-1.400, P = 0.003) was identified as an independent high risk factor associated with long-term SARS-CoV-2 viral shedding. The result is similar to those reported by previous researchers. Hu *et al*. reported that time from onset to admission, corticosteroid use, arbidol use and oseltamivir use were independently associated with the time of virus shedding [7]. 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 use (HR 0.60, 95% CI 0.39–0.94; P = 0.024) were independently associated with long-term SRAS-CoV-2 viral shedding [12]. Both the prolonged time from onset to diagnosis and the delay in hospitalization after illness onset indicates that patients are not diagnosed and treated in time due to various reasons, which leads to prolonged viral shedding. Therefore, in order to reduce the time for viral shedding, symptomatic patients or suspected asymptomatic infected persons should receive nucleic acid testing to be diagnosed, be admitted to the isolation ward as early as possible, and receive antiviral treatment and organ monitoring and supportive treatment.

In addition, we found that interstitial opacity in chest CT scan (OR 6.516, 95% CI 2.041-20.798, P = 0.002) was identified as another independent high risk factor associated with prolonged SARS-CoV-2 RNA shedding. This finding has rarely been reported in the literatures. Several studies have shown that imaging examination plays an irreplaceable role in screening, diagnosis and monitoring the therapeutic effect for patients with COVID-19. Among various imaging methods, chest CT has attracted much attention because of its high sensitivity and specificity [13]. The typical CT findings are bilateral ground-glass opacity (GGO), consolidation with prominent posterior and peripheral distribution of the lung [13]. Another study found that GGO and vascular enhancement were the most common ones followed by interlobular septal thickening and air bronchus sign as well as consolidation, fibrosis and gas trapping [14]. Other study has also found that GGO was the most common manifestation of COVID-19 pneumonia, followed by consolidation and fibrosis [15]. GGO, vascular enhancement, septal thickening, and pulmonary fibrosis are all manifestations of interstitial opacity. Our study suggests that interstitial opacity in chest CT may indicate prolonged time of virus shedding. For patients with interstitial opacity, SARS-CoV-2 RNA testing and isolation should be strengthened to reduce the spread of the virus in the community.

Our study has some limitations. First, shedding of SARS-CoV-2 RNA in throat swab samples does not necessarily represent true viral shedding, as recent studies have shown that SARS-CoV-2 RNA persists longer in stool samples [16]. Second, we used RT-PCR to detect SARS-CoV-2 RNA rather than virus isolation. According to a recent study, SARS-CoV-2 RNA can be detected long after the infectious virus

has disappeared. The immune system can neutralize viruses such as SARS-CoV-2, but it cannot eliminate RNA and degrades slowly [17]. Third, this study is a retrospective cohort study, and there may be type II errors. Therefore, the results of this study should be interpreted with caution when applied. Finally, the sample size of our study was not large enough, and there was a lack of independent sample set for internal or external verification. Therefore, prospective, independent, multicenter and large sample size studies should be still needed to evaluate the performance of our established model in predicting long-term SARS-CoV-2 viral shedding.

Conclusion

In conclusion, time from illness onset to diagnosis and interstitial opacity in chest CT were identified to be independent risk factors associated with long-term SARS-CoV-2 RNA shedding. Based on these two variables, the model established seems to be good predictive performance for predicting prolonged SARS-CoV-2 RNA shedding. Prospective, independent, multicenter and large sample size studies are needed to further assess the performance of the model for the prediction of prolonged viral shedding.

Abbreviations

SARS-CoV-2:severe acute respiratory syndrome coronavirus 2; RNA : ribonucleic acid; COVID-19:coronavirus disease-19; CT :computerized tomography ; OR: odds ratio; WHO: World Health Organization; IQR: interquartile range; RT-PCR: real-time reverse transcription polymerase chain reaction; CK :creatine kinase; LDH :lactate dehydrogenase; CRP:C-reactive protein; PCT : calcitonin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; TBIL, total bilirubin; BUN, urea nitrogen; CREA, creatinine; SPSS: Statistical Product and Service Solutions; ROC: receiver operating characteristic curve; AUROC: area under the ROC; TIOD: time from illness onset to diagnosis; GGO :ground-glass opacity; LTRS: long-term RNA shedding;

Declarations

Acknowledgements

We thank all doctors and nurses of Henan Provincial People's Hospital and Anyang Infectious Disease Hospital for their efforts in the diagnosis and treatment of patients with COVID-19. We are also very much thankful to Anqi Yang (School of Clinical Medicine, Tianjin Medical University) who helped us to check the data and proofread the manuscript.

Disclosures

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors’contributions

XL and ZY: wrote the manuscript. XL and ZY: statistical analysis. XL, JW, LK, and ZY: data collection. XL and ZY: data preparation and analysis. ZY: revised the manuscript. ZY: design and conceptualization. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Key projects of Science and Technology of Henan Province (142102310418). The funding organizations are public institutions and had no role in the design and execution of the study including data collection, management, and analysis or preparation, review, and approval of the manuscript.

Availability of data and materials

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation

Ethical approval and consent to participate

The study was approved by the Clinical Research Ethics Committee of Henan provincial people's hospital, and the requirement for informed consent was waived by the Ethics Committee of Henan Provincial People's Hospital.All methods in our study were carried out in accordance with relevant guidelines and regulations.

Consent for publication

There are no human identifiable images provided in the manuscript, therefore not applicable.

Competing interests

The authors declare that there is no conflict of interest.

Author details

[1]Department of Pulmonary and Critical Care Medicine, Henan Provincial People's Hospital, People's Hospital of Zhengzhou University,7,Weiwu Road, Zhengzhou 450051, Henan, People’s Republic of China.

[2]Department of Cardiopulmonary Function, Fuwai Central China Cardiovascular Hospital, 1, Fuwai Avenue, Zhengdong New District, Zhengzhou 450051, Henan, People's Republic of China.

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. World Health Organization. Weekly operational update on COVID-19. Web link: https://www.who.int/publications/m/item/weeklyepidemiological-update. Accessed January 12, 2021.
3. World Health Organization Director-General's opening remarks at the media briefing on COVID-19 - 24 February 2020 https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—24-february-202. Accessed on February 26, 2020.
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. 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.
6. Buetti N, Trimboli P, Mazzuchelli T, et al. Diabetes mellitus is a risk factor for prolonged SARS-CoV-2viral shedding in lower respiratory tract samples of critically ill patients. Endocrine. 2020; 70:454–60.
7. 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.
8. World Health Organization. Laboratory diagnostics for novel coronavirus. Web link: https://www.who.int/health-topics/coronavirus/laboratorydiagnostics-for-novel-coronavirus. Accessed February 6, 2020.
9. 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.
10. National Health Commission of China. New coronavirus pneumonia prevention and control program (the eighth edition). Web link: http://www.nhc.gov.cn/xcs/zhengcwj/202008/0a7bdf12bd4b46e5bd28ca7f9a7f5e5a.shtml. Accessed on May 23, 2021.
11. Huang CL, Fei L, Li W, et al. 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. Infect Dis Ther. 2021; 10:897-909.
12. 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.

13. Li B, Li X, Wang Y, et al. Diagnostic value and key features of computed tomography in Coronavirus Disease 2019. Emerg Microbes Infect. 2020; 9:787-93.
14. Dai H, Zhang X, Xia J, et al. High-resolution Chest CT Features and Clinical Characteristics of Patients Infected with COVID-19 in Jiangsu, China. Int J Infect Dis. 2020; 95:106-12.
15. Shang Y, Xu C, Jiang F, et al. Clinical characteristics and changes of chest CT features in 307 patients with common COVID-19 pneumonia infected SARS-CoV-2: A multicenter study in Jiangsu, China. Int J Infect Dis. 2020; 96:157-62.
16. 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
17. Atkinson B, Petersen E. SARS-CoV-2 shedding and infectivity. Lancet. 2020; 395:1339-40.

Figures

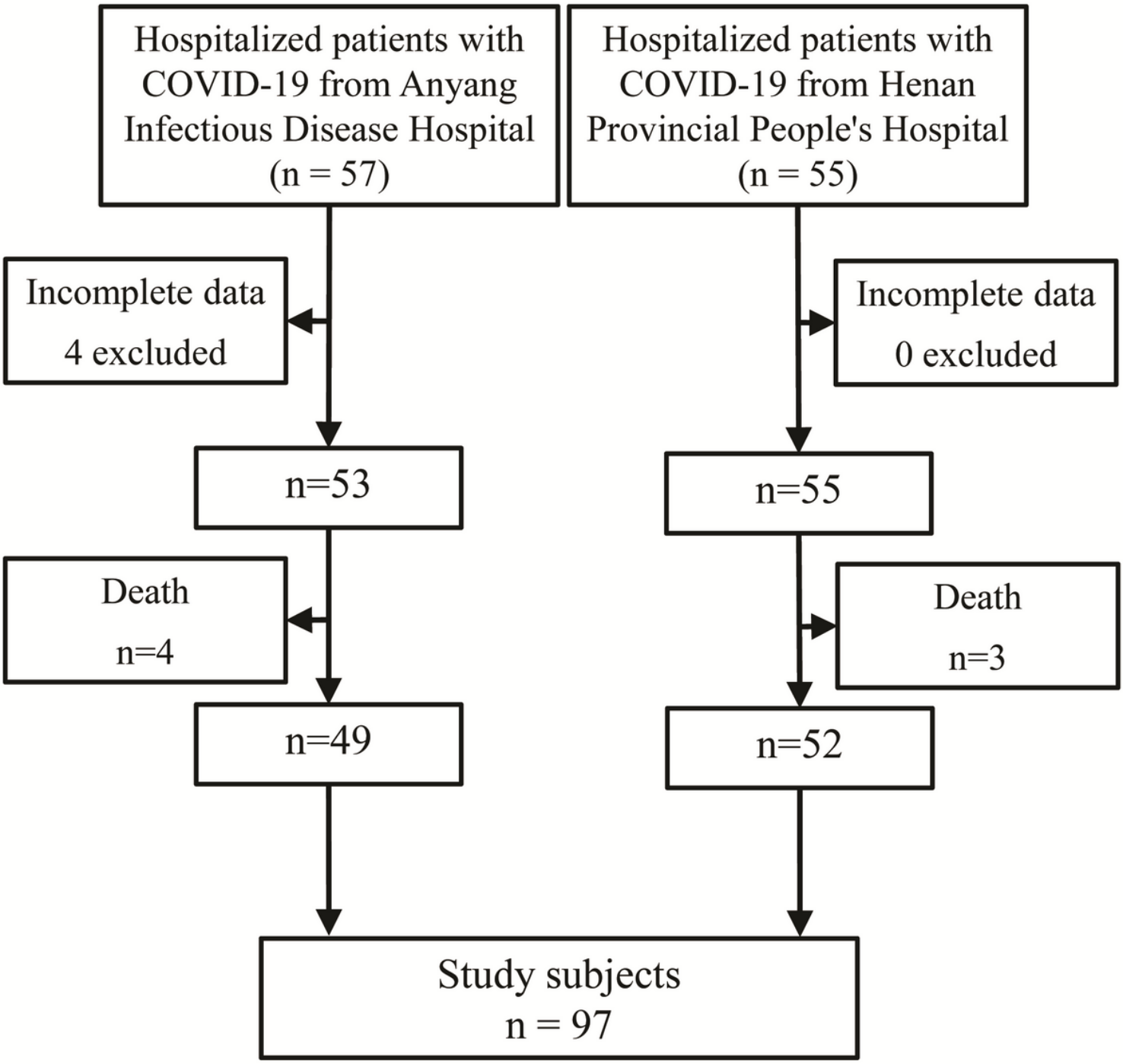

Figure 1

Flowchart of study population included

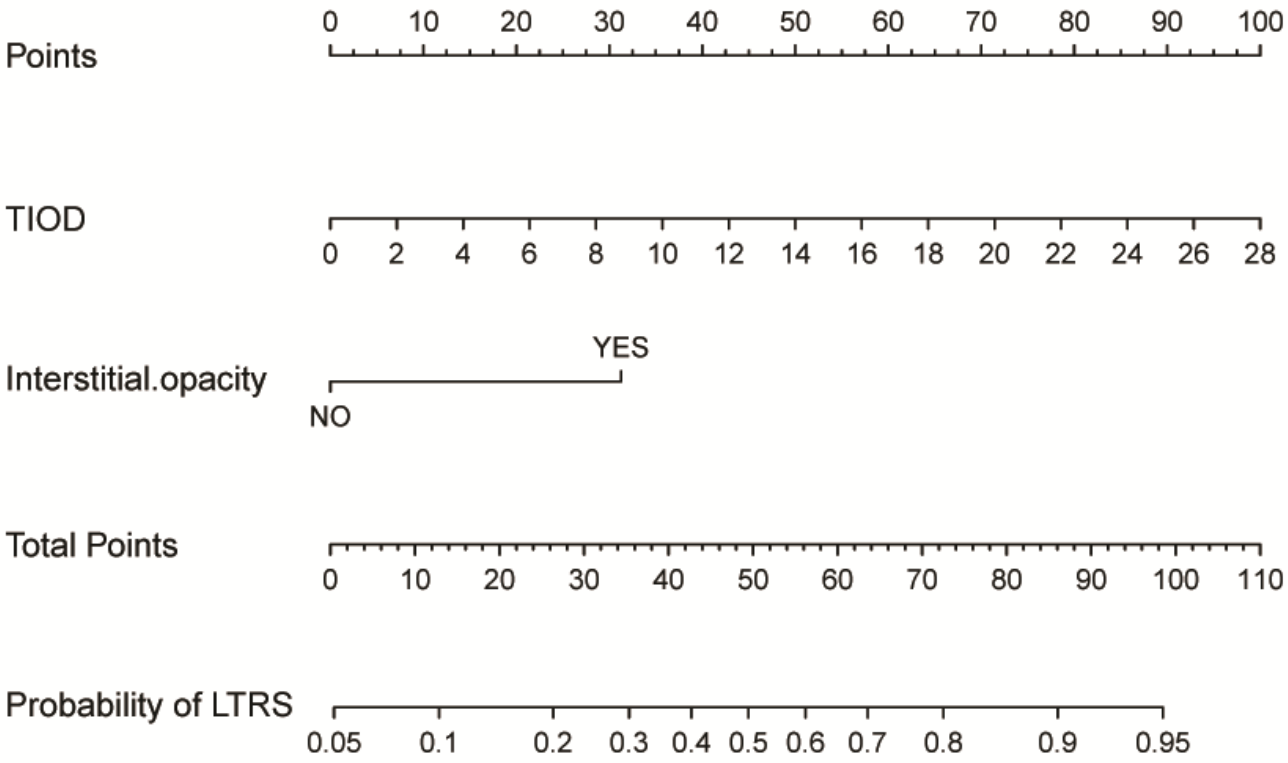

Figure 2

The nomogram for predicting the probability of long-term SARS-CoV-2 RNA shedding. Multivariate logistic regression analysis identified time from illness onset to diagnosis and interstitial opacity in CT scan as independent predictors. A model that incorporated these two independent predictors was developed and presented as a nomogram. To use this nomogram in clinical management, an individual patient's value is located on each variable axis, and a line is plotted upward to calculate the number of points received for each variable value. The sum of these scores is located on the total points axis and draw a line straight down to get the probability of long-term RNA shedding. TIOD, time from illness onset to diagnosis; LTRS, long-term RNA shedding

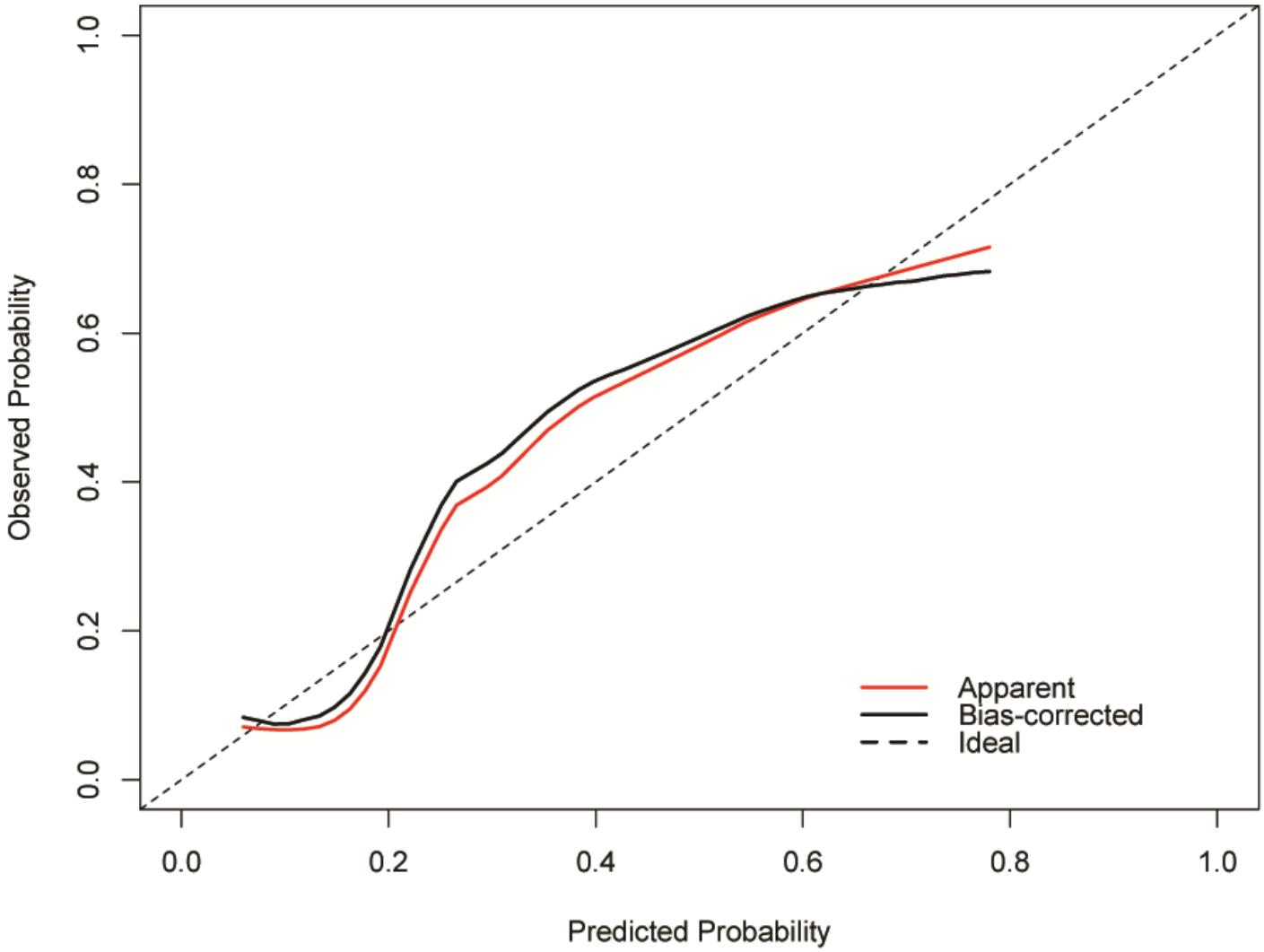

Figure 3

Calibration curves of the nomogram for predicting the probability of long-term SARS-CoV-2 RNA shedding. The red line shows the apparent calibration line. The black line shows the bias-corrected calibration line, and represents the performance of the nomogram, of which a closer fit to the diagonal line represents a better prediction. The goodness-of-fit statistics of the nomogram (x2 = 8.292; P = 0.406) was not statistically significant.

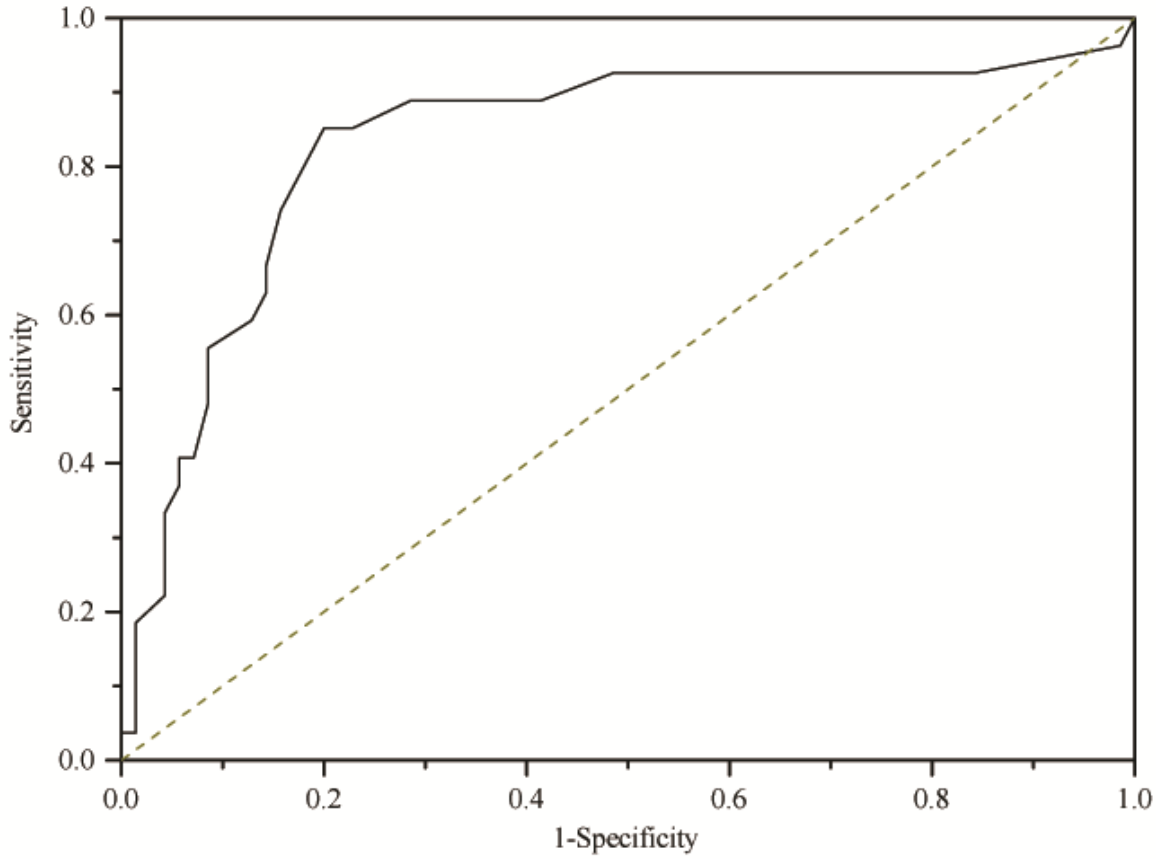

Figure 4

Receiver operating characteristic curve (ROC) for nomogram predicting the probability of long-term SARS-CoV-2 RNA shedding. Area under ROC was 0.834 (95% CI 0.731-0.936, P < 0.001).