



Research Article

<https://doi.org/10.1631/jzus.B2000608>

Development and Validation of An Individualized Nomogram for Early Prediction of the Duration of SARS-CoV-2 Shedding in COVID-19 Patients with Non-severe Disease

Shijin YUAN^{1*}, Yong PAN^{2*}, Yan XIA¹, Yan ZHANG³, Jiangnan CHEN⁴, Wei ZHENG⁵, Xiaoping XU⁶, Xinyou XIE^{1✉}, Jun ZHANG^{1✉}

¹Department of Clinical Laboratory, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310000, China

²Department of Clinical Laboratory, Wenzhou Central Hospital, Wenzhou 325000, China

³Department of Clinical Laboratory, Xixi Hospital of Hangzhou, Hangzhou 310000, China

⁴Department of Clinical Laboratory, Affiliated Hospital of Shaoxing University, Shaoxing 312000, China

⁵Department of Clinical Laboratory, The third people's Hospital of Yueqing, Wenzhou 325000, China

⁶Department of Clinical Laboratory, Jinhua municipal central Hospital, Jinhua 321000, China

Abstract: With the number of cases of coronavirus disease-2019 (COVID-19) increasing rapidly, the World Health Organization has recommended that patients with mild or moderate symptoms could be released from quarantine without nucleic acid retesting, and self-isolate in the community. This may pose a potential virus transmission risk. We aimed to develop a nomogram to predict the duration of viral shedding for individual COVID-19 patients. This retrospective multicentric study enrolled 135 patients as a training cohort and 102 patients as a validation cohort. Significant factors associated with the duration of viral shedding were identified by multivariate Cox modeling in the training cohort and combined to develop a nomogram to predict the probability of viral shedding at 9, 13, 17 and 21 days after admission. The nomogram was validated in the validation cohort and evaluated by a concordance index (C-index), area under the curve (AUC), and calibration curves. A higher absolute lymphocyte count ($P=0.001$) and lymphocyte-to-monocyte ratio ($P=0.013$) were correlated with a shorter duration of viral shedding, while a longer activated partial thromboplastin time ($P=0.007$) prolonged the viral shedding duration. The C-indices of the nomogram were 0.732 (95% CI: 0.685-0.777) in the training cohort and 0.703 (95% CI: 0.642-0.764) in the validation cohort. The AUC showed a good discriminative ability (training cohort: 9 days: 0.879; 13 days: 0.762; 17 days: 0.738; 21 days: 0.715; validation cohort: 9 days: 0.855; 13 days: 0.758; 17 days: 0.728; 21 days: 0.706), and calibration curves were consistent between outcomes and predictions in both cohorts. A predictive nomogram for viral shedding duration based on three easily accessible factors was developed to help estimate appropriate self-isolation times for patients with mild or moderate symptoms, and to control virus transmission.

Key words: COVID-19; SARS-CoV-2; Duration of viral shedding; Nomogram

1 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is an enveloped ribonucleic acid (RNA) virus and the pathogen of coronavirus disease 2019 (COVID-19), which broke out in Wuhan, China in December 2019 and then rapidly spread nationwide and around the world (Li et al., 2020c; Zhu et al., 2020). On February 11, 2020, the Chinese Center for Disease Control and Prevention (CDC) reported that in the largest case series of COVID-19 in mainland China, 81% of patients were classified as non-severe cases (including mild and moderate disease) (Epidemiology Working Group for Ncip Epidemic Response and Prevention, 2020). As of August 16, 2020, the world health organization (WHO) reported over 21,000,000 accumulated confirmed COVID-19 cases globally, with over 760,000 death cases (WHO, 2020a).

✉ Jun ZHANG, jameszhang2000@zju.edu.cn

Xin-you XIE, scottxie@zju.edu.cn

* The two authors contributed equally to this work

Jun ZHANG, <https://orcid.org/0000-0002-2613-3165>

During the pandemic, a large number of studies have summarized the epidemiological history, clinical characteristics and laboratory abnormalities of COVID-19 patients, and identified several risk factors associated with disease deterioration or fatal outcomes. This could help clinicians to recognize risk factors early and optimize the management of inpatients to prevent disease progression (Chen et al., 2020; Guan et al., 2020; Liang et al., 2020; Wu et al., 2020; Zhou et al., 2020a). Based on the initial guidance of the WHO and the National Health Commission of China (NHC), to confirm total clearance of the virus, patients with symptoms resolved needed two consecutive negative results from SARS-CoV-2 nucleic acid tests (24 hours apart) via real-time reverse-transcriptase polymerase-chain-reaction (RT-PCR) assay before discharge (WHO, 2020b; NHC, 2020).

However, in light of the shortage of medical resources in areas with intense transmission, the initial discharge criteria were extremely difficult to meet. Besides, previous studies indicated that persistent detection of SARS-CoV-2 RNA generally did not reflect active viral replication and asymptomatic patients with positive RT-PCR results may not have the capacity to transmit virus (Bullard et al., 2020; Wolfel et al., 2020). Hence, the WHO recommended patients with non-severe disease should undertake self-isolation in a community health institution and updated the discharge criteria based on the time from symptom onset to resolution, while removing the requirement for SARS-CoV-2 nucleic acid retesting (WHO, 2020c; WHO, 2020d). Symptomatic patients could be released from quarantine 10 days after symptom onset plus at least 3 days without symptoms, while asymptomatic patients could be released 10 days after receiving a positive SARS-CoV-2 test result.

Although the potential risk of viral transmission after symptom resolution seemed minimal, it could not be completely ruled out. Hence, identifying the duration of SARS-CoV-2 viral shedding could help to determine the appropriate length of self-isolation for COVID-19 patients with non-severe disease and have significant implications for controlling viral transmission. In this study, we aimed to explore the risk factors that might prolong the duration of viral shedding and develop an individualized nomogram to predict the probability of viral shedding at a specific time. This could help to estimate the appropriate length of isolation for each non-severely affected patient.

2 Materials and Methods

2.1 Study Design and Participants

This retrospective multicentric study aimed to evaluate the factors associated with the duration of viral shedding and develop a model to predict the probability of viral shedding for COVID patients with non-severe disease. Patients diagnosed as COVID-19 cases from five designated hospitals in Zhejiang province were available to enroll in this study. The exclusion criteria included: (1) patients with influenza A or B infection. (2) patients classified as having severe or critical disease on admission or having progressive disease during the treatment (at least one of the following criteria should be met: respiratory rate ≥ 30 per min; arterial oxygen saturation $\leq 93\%$ at rest; oxygenation index [arterial oxygen partial pressure/fraction of inspired oxygen] ≤ 300 mmHg). (3) patients with potential coagulation disorders such as hemophilia A or B, hepatic disease or who had received any anticoagulant drug in the preceding three months.

Between January 17, 2020 and February 15, 2020, COVID-19 patients from four hospitals (Xixi Hospital of Hangzhou, Jinhua Municipal Central Hospital, Affiliated Hospital of Shaoxing University and the Third People's Hospital of Yueqing) were recruited to form the training cohort to identify significant factors related to the duration of SARS-CoV-2 viral shedding and develop the predictive model. Between January 23, 2020 and February 25, 2020, COVID-19 patients from another hospital (Wenzhou Central Hospital) were recruited to form the external validation cohort to examine the generalizability of the model. The study was approved by the Ethics Committee of Sir Run Run Shaw Hospital (Certificate no: Scientific Research 20200331-45), while written informed consent waived due to the use of retrospective data.

2.2 Detection of SARS-CoV-2, Disease Classification and Treatment

Nasopharyngeal swab specimens were collected and transported to the local CDC for laboratory testing via real-time RT-PCR assay (Shanghai ZJ Bio-Tech Co., Ltd). A cycle threshold value (Ct-value) of ≤ 43 was defined as a positive result, and a Ct-value of >43 as a negative result.

The classification of disease severity, including mild, moderate, severe or critical disease, was based on the

Clinical Diagnosis and Treatment Guidance 2019 for Pneumonia Caused by Novel Coronavirus (5th edition) (NHC, 2020). In our study, patients with mild or moderate disease were combined as “non-severe patients”, while those with severe or critical disease were combined as “severe patients”.

All patients received general symptomatic treatment. Antiviral therapy included lopinavir/ritonavir (LPV/r, 400 mg/100 mg, orally, twice daily) combined with interferon alpha inhalation. Antibiotic therapy was available following identification of specific bacterial infections.

2.3 Data Collection

Data related to epidemiological history, clinical characteristics and laboratory investigations were reviewed and extracted from electronic medical records by a well-trained team of doctors. Epidemiology history was examined by physicians in detail, including whether the patients were local residents of Wuhan, had recently been to Wuhan, or had had contact with people from Wuhan. Clinical characteristics included comorbidities (hypertension and diabetes), symptoms and signs, incubation period (the date of exposure based on epidemiology history to the date of symptoms or signs of onset), abnormal findings from chest CT, and therapeutic regimens. Laboratory investigations were undertaken following hospital admission, including red blood cell count (RBC), white blood cell count (WBC), absolute neutrophil count (ANC), absolute monocyte count (AMC), absolute lymphocyte count (ALC), absolute eosinophil count (AEC), platelet count, aspartate aminotransferase (AST), alanine aminotransferase (ALT), serum albumin concentration, C-reactive protein (CRP), lactate dehydrogenase (LDH), activated partial thromboplastin time (APTT) and prothrombin time (PT). A combined index reflecting systemic inflammation and immune status, called the lymphocyte-to-monocyte ratio (LMR, calculated as ALC divided by AMC), was also investigated.

2.3 Statistical Analysis

Categorical variables are expressed as frequencies and percentages, and significance was analyzed by a chi square or Fisher's exact test. The Shapiro-Wilk test was used to test the normality of the continuous variables. Continuous variables with normal distribution are presented as mean \pm standard deviation (SD), and the significance of differences was analyzed by Student's t-test. Continuous variables with abnormal distribution are reported as medians and interquartile ranges (IQR), and the significance of differences was tested by the Mann-Whitney U test.

The duration of viral shedding was defined as the interval from the date of a positive viral nucleic acid result to the first date of a negative result, without converting to positive thereafter. The Cox proportional hazards regression model was used to select the factors associated with the duration of viral shedding, and calculate the hazard ratios (HRs) with 95% confidence intervals (95% CI). Variables (including clinical characteristics and laboratory investigations) with a *P* value less than 0.05 in the univariate Cox model were incorporated in a multivariate Cox model to identify significant factors via a backward stepdown selection process. Variables of laboratory investigations are presented as continuous variables in the Cox model.

Based on the results of the multivariate Cox model in the training cohort, a nomogram to predict the probability of viral shedding at 9, 13, 17 and 21 days after admission for each COVID-19 patient was developed. The performance of the nomogram was measured based on the concordance index (C-index), the area under the curve (AUC), and the calibration curves by using a bootstrap method with 1000 resamples. The C-index was used to evaluate the predictive accuracy. The AUC and its 95% CI from receiver operating characteristic (ROC) curves were used to assess the discriminative ability, and the calibration curves were used to estimate the coincidence between the actual outcomes and the predicted probabilities. The AUCs of the nomogram and separate significant factors were compared based on the method described by DeLong et al. (DeLong et al., 1988). All statistical analyses were conducted using SPSS version 26.0 and R software (version 3.6.0). A two-tailed *P* value of less than 0.05 was considered statistically significant.

3 Results

3.1 Baseline Clinical Characteristics and Laboratory Findings

A total of 237 patients were enrolled in this study, including 135 from four hospitals as the training cohort and 102 from another hospital as the validation cohort. All enrolled patients had non-severe symptoms, but 186 (78.5%) were classified as having moderate disease. The mean age was 43.3 (\pm 15.2) years, and 141 patients (58.2%) were male. Forty-six patients had comorbidities, including 38 (16%) with hypertension and 14 (5.8%) with diabetes. Fever (176 patients; 74.3%) and dry cough (143 patients; 60.3%) were the most common

symptoms noted on admission to hospital. The epidemiological data showed that 86 patients had a history of exposure in Wuhan, including 16 (6.6%) who were local residents and 70 (29.6%) who had been to Wuhan recently. Another 151 (63.7%) patients had not been to Wuhan, but had had contact with people from Wuhan. Laboratory investigations on admission revealed the following abnormalities: 94 patients (39.7%) with decreasing AEC; 86 (36.3%) with decreasing ALC; 71 (30.0%) with hypoalbuminemia; and 59 (24.9%) with prolonged APTT. Only 11 patients (4.6%) presented a normal chest CT image. For the treatment, 229 patients (96.6%) received LPV/r. There were no significant differences in most baseline clinical characteristics or laboratory findings between the training and validation cohorts (Table 1).

3.2 Factors Associated with Duration of Viral Shedding

Among the 135 patients in the training cohort, the median duration of viral shedding was 13 days (IQR 8-18 days). The univariate Cox model showed that a higher ALC ($P<0.001$) and a higher LMR ($P<0.001$) were related to a shorter duration of viral shedding, while a higher APTT ($P=0.001$) was related to a longer duration of viral shedding (Table 2). These three variables were then incorporated into the multivariate Cox model, which showed that a higher ALC (HR=1.713, 95% CI: 1.262-2.325, $P=0.001$) and a higher LMR (HR=1.169, 95% CI: 1.034-1.322, $P=0.013$) independently remained significantly associated with a shorter duration of viral shedding. However, a higher APTT (HR=0.938, 95% CI: 0.896-0.983, $P=0.007$) was significantly associated with a longer duration of viral shedding (Table 2).

3.3 Nomogram Development and Evaluation

Through combining the three significant factors, ALC, LMR and APTT, from the multivariate Cox model in the training cohort, a nomogram was developed to predict the probability of viral shedding at 9, 13, 17 and 21 days after admission for each non-severe COVID-19 patient (Fig. 1). Each significant factor is shown in the variable axis and assigned a corresponding score on the points scale at the top of the nomogram. The length of the segment line of each factor represents the distribution of the assigned score. Hence, the nomogram demonstrates that the ALC value contributed the most to the duration of viral shedding. The three obtained scores were then added up to give a summed score at the bottom total points scale, which estimates the specific probability of viral shedding at 9, 13, 17 and 21 days for an individual patient. Here is an example to illustrate clinical usage of the nomogram (indicated in Fig. 1 by a vertical red dotted line and a solid line with an arrow). If the ALC, LMR and APTT results of a COVID-19 patient on admission were $0.8 \times 10^9/L$, 2.29 and 27 s, respectively, the corresponding scores of ALC, LMR and APTT for this patient were about 25.8, 31.2 and 51.0, respectively. Hence, the total score was 108, which indicated probabilities of 0.294, 0.501, 0.678 and 0.842 for viral shedding at 9, 13, 17 and 21 days, respectively, after admission.

The performance of the nomogram, as estimated by the C-index, was 0.732 (95% CI: 0.685-0.777). The AUC showed good discriminative ability (9 days: 0.879, 95% CI: 0.847-0.949; 13 days: 0.762, 95% CI: 0.707-0.853; 17 days: 0.738, 95% CI: 0.665-0.834; 21 days: 0.715, 95% CI: 0.612-0.836, respectively, Table 3). The calibration curves showed that the probabilities predicted by the nomogram were consistent with the actual observed results (Fig. 2A-D).

3.4 External Validation

In the external validation cohort, the median duration of viral shedding was 16 days (IQR 8-20 days). In this validation cohort, the C-index of the nomogram was 0.703 (95% CI: 0.642-0.764). The AUC also showed the good discriminative ability of the nomogram (9 days: 0.855, 95% CI: 0.781-0.937; 13 days: 0.758, 95% CI: 0.741-0.896; 17 days: 0.728, 95% CI: 0.666-0.838; 21 days: 0.706, 95% CI: 0.571-0.842, respectively, Table 3). Moreover, the curves of predicted viral shedding probability at 9, 13, 17 and 21 days closely matched those of the observed results, indicating that the nomogram also showed good calibration in the external validation cohort (Fig. 2E-H).

3.5 Predictive performance comparison of the ALC, LMR, APTT separately and in the nomogram

ROC curve analysis was used to compare the performance of the three significant factors (ALC, LMR and APTT), separately and in the nomogram, in predicting viral shedding duration (Fig. 3). In the training cohort, the AUC values of the nomogram at 9, 13, 17 and 21 days were all higher than those of the ALC, LMR and APTT alone (all P values <0.05 ; Fig. 3A-D). Similarly, in the external validation cohort, the AUC values of the

nomogram at 9, 13, 17 and 21 days were also higher than those of each independent factor (all P values <0.05 ; Fig. 3E-H). These results indicate that the nomogram had a better predictive performance than the ALC, LMR or APTT alone in predicting the duration of viral shedding.

4 Discussion

SARS-CoV-2 RNA viral load and the duration of viral shedding are two major factors affecting viral spread and transmission. Previous studies had investigated dynamic changes of viral load in different specimens between patients with severe or non-severe symptoms, and predictive factors associated with the duration of viral shedding (Lin et al., 2020; Xiao et al., 2020; Xu et al., 2020; Zheng et al., 2020b; Zou et al., 2020). However, an accurate model to predict viral shedding was lacking. In this study, we identified three factors, ALC, LMR and APTT, readily amenable to laboratory investigation, associated with the duration of SARS-CoV-2 viral shedding, and developed a nomogram to predict the probability of viral shedding that could help to estimate early the appropriate length of isolation needed for each non-severe COVID-19 patient.

In this study, we found higher ALC and LMR on admission were associated with a shorter duration of viral shedding, while higher APTT was associated with a longer duration. SARS-CoV-2 infection could activate the antiviral response of the host immune system, including T lymphocytes, B lymphocytes, natural killer cells and complementary systems (Li et al., 2020a). However, immune cells, especially T lymphocytes, would be damaged by direct cytopathic effects induced by SARS-CoV-2. In addition, an uncontrolled inflammatory response and an impaired immune system may lead to systemic damage during the period of COVID-19 infection in patients with severe symptoms. A previous study examined the cellular level of T cells and found that several functional molecules, including interferon- γ and tumor necrosis factor- α , showed a significant reduction in expression on CD4+ T cells, but an increased expression on CD8+ T cells. This indicates that SARS-CoV-2 could destroy the function of CD4+ T cells and exhaust the quantity of CD8+ T cells (Zheng et al., 2020a).

A nationwide retrospective analysis of 1099 patients showed that 83.2% of patients presented decreased ALC on admission, and patients with severe disease had more prominent lymphocytopenia compared with those with non-severe disease (Guan, et al., 2020). Two meta-analyses also reported that lymphocyte count was significantly lower in patients with severe or non-survivable disease, respectively (Henry et al., 2020; Zhao et al., 2020). In our present study, we found a higher ALC on admission was associated with a shorter duration of viral shedding. This is consistent with the results of another retrospective study that found that higher counts of lymphocytes and CD8+ T cells were related to a significantly shorter duration of viral shedding (Lin, et al., 2020).

Tissue macrophages, a population of innate immune cells derived from monocytes in peripheral blood, can control infection by eliminating pathogens and promoting tissue repair. Therefore, they could play an important role in alleviating an excessive SARS-CoV-2-induced inflammation status. A previous study compared the bronchoalveolar fluid of patients with either severe or non-severe COVID-19 disease and found that two chemokines, CC-chemokine ligand 2 and CC-chemokine ligand 7, which can recruit CC-chemokine receptor 2-positive monocytes, were enriched in the bronchoalveolar lavage fluid (BALF) of patients with severe disease (Zhou et al., 2020b). Another study investigated the levels of 14 cytokines in peripheral blood and found that the expression level of monocyte chemotactic protein-3 (MCP-3) was highest in patients in critical condition, followed by those with severe disease, and then those with moderate disease, indicating that MCP-3 was highly associated with disease severity (Yang et al., 2020b). A significantly increasing population of CD14+CD16+ monocytes in peripheral blood was also observed in patients with severe disease compared with those with non-severe symptoms (Zhou Y et al., 2020). Hence, a high concentration of monocytes may be considered a risk factor for disease progression and death in COVID-19. LMR, calculated as the ALC divided by the AMC, is a novel combined indicator of the systematic inflammatory response and immune status. An elevated level of LMR means increasing lymphocytes or decreasing monocytes, which may represent an active immune response and mild inflammation. Previous studies showed that LMR was lower in COVID-19 patients with severe disease than in those with non-severe symptoms, indicating that a higher LMR may predict a better clinical outcome (Sun et al., 2020; Yang et al., 2020a). Similar to those results, we found a higher LMR was associated with a shorter duration of viral shedding, which suggests that an active immune response could help to accelerate the process of viral clearance.

Coagulation dysfunction and subsequent disseminated intravascular coagulation (DIC) are prominent in patients with acute respiratory distress syndrome caused by coronavirus infections including SARS-CoV,

SARS-CoV-2 and middle east respiratory syndrome coronavirus (MERS) (Giannis et al., 2020). Coagulation dysfunction may be explained as an imbalance between procoagulant and anticoagulant homeostatic mechanisms triggered by viral infection, involving multiple pathophysiological mechanisms including endothelial damage, elevated plasma level of von Willebrand factor, and activation of the Toll-like receptor and tissue-factor pathways (Subramaniam and Scharrer, 2018). A previous study showed non-survivors had a markedly prolonged APTT than survivors on hospital admission (Tang et al., 2020). In this study, we also found prolonged APTT was associated with a longer duration of viral shedding. This suggests that SARS-CoV-2 could influence the activity of the endogenous coagulation system, and coagulation dysfunction may reflect a weak ability for viral clearance.

Previous studies reported that older patients with comorbidities have a higher risk of disease progression and death (Chen, et al., 2020; Li et al., 2020b). However, we did not find any association between old age, comorbidity and prolonged duration of viral shedding in this study. The main reason may be that patients enrolled in our study were all classified as having non-severe disease, and the mean age of all 237 patients was merely 43 years, with only 25 patients older than 70 years, which is far younger than the mean age of other studies. We also did not observe any association between an elevated level of CRP and prolonged duration of viral shedding. We found that most patients had mild elevation of CRP, which may be explained by SARS-CoV-2 inducing only a mild inflammatory reaction in patients with non-severe symptoms.

Here, we identified three easily accessible factors, ALC, LMR and APTT, associated with the duration of viral shedding, and developed a nomogram to predict the probability of viral shedding at 9, 13, 17 and 21 days. Based on these results, clinicians could recognize patients with risk factors associated with prolonged viral shedding duration at an early stage of disease, and estimate an appropriate length of isolation for those patients with non-severe disease, which is important in containing virus transmission in the community.

This study had several limitations. First, due to the retrospective nature of the study, this model needs further verification through a larger prospective study. Second, viral load data were lacking, yet viral load is another important feature of virus spread and transmission. Third, SARS-CoV-2 was tested based on a nasopharyngeal swab specimen in this study. However, viral shedding time differs when based on nasopharyngeal swabs, oropharyngeal swabs, blood specimens, fecal specimens or urine specimens. Hence, the results of our study may not be applicable to those patients who had been tested for the virus with samples taken from other sites. Lastly, the AUC values in this study were not very high, indicating the value of our model is relatively limited. More indicators, such as flow cytometry data, and multi-indicator combinations need to be considered and analyzed in further studies.

In conclusion, we identified three factors associated with the duration of viral shedding, and developed a nomogram to predict the probability of viral shedding at 9, 13, 17 and 21 days for COVID-19 patients with non-severe disease. The results could be used to estimate an appropriate length of self-isolation for each patient with non-severe disease, and help to control virus transmission in the community.

Author contributions

Jun ZHANG and Xinyou XIE designed the study and reviewed the manuscript prior to submission. Shijin YUAN coordinated the work and took the lead in drafting the manuscript and interpreting. Yong PAN and Yan XIA developed the statistical methods. Yong PAN, Jiangnan Chen, Yan ZHANG, Wei ZHENG and Xiaoping Xu participated in the collection of experimental data. The corresponding author attests that all listed authors met authorship criteria and that no others meeting the criteria have been omitted. All authors have read and approved the final manuscript and, therefore, have full access to all the data in the study and take responsibility for the integrity and security of the data.

Acknowledgements

This research was supported by the Medical and Health Science and Technology Project of Zhejiang Province, China (2018KY116). We would like to acknowledge Dr. Yu-zhen Gao (Department of Clinical Laboratory, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China) and Dr. Yan-zhong Wang (Department of Clinical Laboratory, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China) for his writing assistance and proof reading the article.

Compliance with ethics guidelines

Shijin YUAN, Yong PAN, Yan XIA, Yan ZHANG, Jiangnan CHEN, Wei ZHENG, Xiaoping XU, Xinyou XIE, and Jun ZHANG declare that they have no conflict of interest.

This study was approved by the Institutional Ethics Committee of Sir Run Run Shaw Hospital (No. Scientific Research 20200331-45). All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008(5). Informed consent was waived due to retrospective nature of the study.

Table 1. Demographics, clinical characteristics and laboratory investigations of patients

Parameters	All patients (N=237)	Training cohort (N=135)	Validation cohort (N=102)	P-value
Characteristics				
Age; Years, Mean (\pm SD)	43.3 (15.2)	42.7 (16.3)	44.2 (13.8)	0.455
Gender, N (%)				0.026
Male	138 (58.23)	87 (64.44)	51 (50)	
Female	99 (41.77)	48 (35.56)	51 (50)	
Incubation period, Days (25 th , 75 th)	4 (2, 7)	3 (2, 6)	5 (2, 7)	0.132
Clinical type, N (%)				0.107
Mild	51 (21.52)	24 (17.78)	27 (26.47)	
Moderate	186 (78.48)	111 (82.22)	75 (73.53)	
Comorbidity, N (%)				
Hypertension	38 (16.03)	19 (14.07)	19 (18.63)	0.344
Diabetes	14 (5.79)	5 (3.70)	9 (8.41)	0.098
Epidemiological history, N (%)				
Local residents of Wuhan	16 (6.61)	14 (10.37)	2 (1.87)	0.011
Non-local residents: Recently been to Wuhan	70 (29.54)	43 (31.85)	27 (26.47)	0.369
Non-local residents: Had contact with people from Wuhan	151 (63.71)	78 (57.78)	73 (71.57)	0.029
Symptoms, N (%)				
Fever	176 (74.26)	104 (77.04)	72 (70.59)	0.261
Cough	143 (60.34)	86 (63.70)	57 (55.88)	0.223
Fatigue	80 (33.76)	42 (31.11)	38 (37.25)	0.322
Diarrhea	27 (11.39)	15 (11.11)	12 (11.21)	0.875
Abnormalities on chest CT, N (%)				
Local patchy shadowing	64 (27.00)	53 (39.26)	11 (10.78)	<0.001

Bilateral patchy shadowing	134 (56.54)	61 (45.19)	73 (71.57)	<0.001
Ground-glass opacity	28 (11.81)	12 (8.89)	16 (14.95)	0.108
No obvious abnormalities	11 (4.64)	9 (6.67)	2 (1.87)	0.088
Treatment, N (%)				
Antiviral therapy	229 (96.62)	131 (97.04)	98 (96.08)	0.686
Use of corticosteroids	14 (5.91)	9 (6.67)	5 (4.90)	0.568
Interferon inhalation	211 (89.03)	112 (82.96)	99 (97.06)	0.001
Antibiotic therapy	56 (23.63)	23 (17.04)	33 (32.35)	0.006
Laboratory parameters, (25th, 75th)				
WBC, ×10 ⁹ /L	4.80 (3.84, 6.23)	5.00 (4.00, 6.40)	4.65 (3.70, 5.83)	0.063
ANC, ×10 ⁹ /L	2.97 (2.23, 3.93)	3.05 (2.27, 4.08)	2.93 (2.21, 3.79)	0.298
ALC, ×10 ⁹ /L	1.31 (1.01, 1.96)	1.28 (0.90, 1.91)	1.36 (1.09, 2.00)	0.073
AMC, ×10 ⁹ /L	0.43 (0.31, 0.55)	0.42 (0.30, 0.55)	0.44 (0.32, 0.56)	0.370
AEC, ×10 ⁹ /L	0.02 (0.01, 0.05)	0.01 (0.00, 0.05)	0.02 (0.01, 0.06)	0.054
RBC, ×10 ¹² /L	4.59 (4.29, 5.00)	4.56 (4.25, 4.94)	4.67 (4.31, 5.03)	0.178
Platelets, ×10 ⁹ /L	189.0 (153.0, 231.0)	199.0 (156.0, 233.0)	182.5 (150.8, 228.8)	0.190
CRP, mg/L	7.90 (3.00, 17.00)	8.00 (3.90, 14.60)	7.25 (2.20, 19.45)	0.776
APTT, seconds	30.20(28.76, 32.65)	30.30 (29.30, 32.70)	30.00 (27.99, 32.45)	0.060
PT, seconds	12.30 (11.70, 12.70)	12.20 (11.50, 12.80)	12.30 (12.00, 12.70)	0.214
Albumin, g/L	41.10 (39.25, 43.59)	41.00 (38.50, 43.70)	41.23 (40.11, 43.58)	0.653
LDH, U/L	189 (153, 238)	188 (145, 252)	193 (157, 225)	0.442
ALT, UL	19 (12, 32)	18 (12, 31)	21 (13, 35)	0.335
AST, UL	23 (18, 30)	23 (18, 29)	24 (19, 32)	0.098
LMR	3.21 (2.37, 4.42)	3.16 (2.29, 4.45)	3.24 (2.47, 4.40)	0.577

Data are expressed as: N (%), mean (\pm SD) and median (IQR).

Abbreviations: WBC, white blood cell count; ANC, absolute neutrophil count; ALC, absolute lymphocyte count; AMC, absolute monocyte count; AEC, absolute eosinophil count; RBC, red blood cell count; CRP, C-reactive protein; CT, computerized tomography; APTT, activated partial thromboplastin time; PT, prothrombin time; LDH, lactate dehydrogenase; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; LMR, lymphocyte-to-monocyte ratio; SD, standard deviation; IQR, interquartile range.

Table 2. Univariate and multivariate analysis of duration of viral shedding in training cohort

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95 % CI)	P value
Sex (male vs female)	1.184 (0.828, 1.694)	0.354		
Age	0.999 (0.989, 1.009)	0.838		
Hypertension (yes vs no)	1.066 (0.654, 1.737)	0.799		
Diabetes (yes vs no)	1.153 (0.471, 2.825)	0.755		
Incubation period	0.993 (0.952, 1.037)	0.761		
WBC	1.027 (0.964, 1.095)	0.409		
ANC	1.037 (0.964, 1.117)	0.330		
ALC	2.213 (1.737, 2.82)	<0.001	1.713 (1.262, 2.325)	0.001
AMC	0.802 (0.338, 1.902)	0.617		
AEC	0.855 (0.607, 1.204)	0.370		
RBC	1.038 (0.810, 1.329)	0.770		
Platelets	1.000 (0.998, 1.002)	0.939		
CRP	0.994 (0.986, 1.003)	0.184		
APTT	0.921 (0.879, 0.966)	0.001	0.938 (0.896, 0.983)	0.007
PT	0.926 (0.734, 1.169)	0.520		
Albumin	0.987 (0.945, 1.031)	0.559		
LDH	1.001 (0.999, 1.002)	0.389		
ALT	1.002 (0.996, 1.008)	0.475		
AST	0.996 (0.988, 1.003)	0.282		
LMR	1.347 (1.220, 1.488)	<0.001	1.169 (1.034, 1.322)	0.013

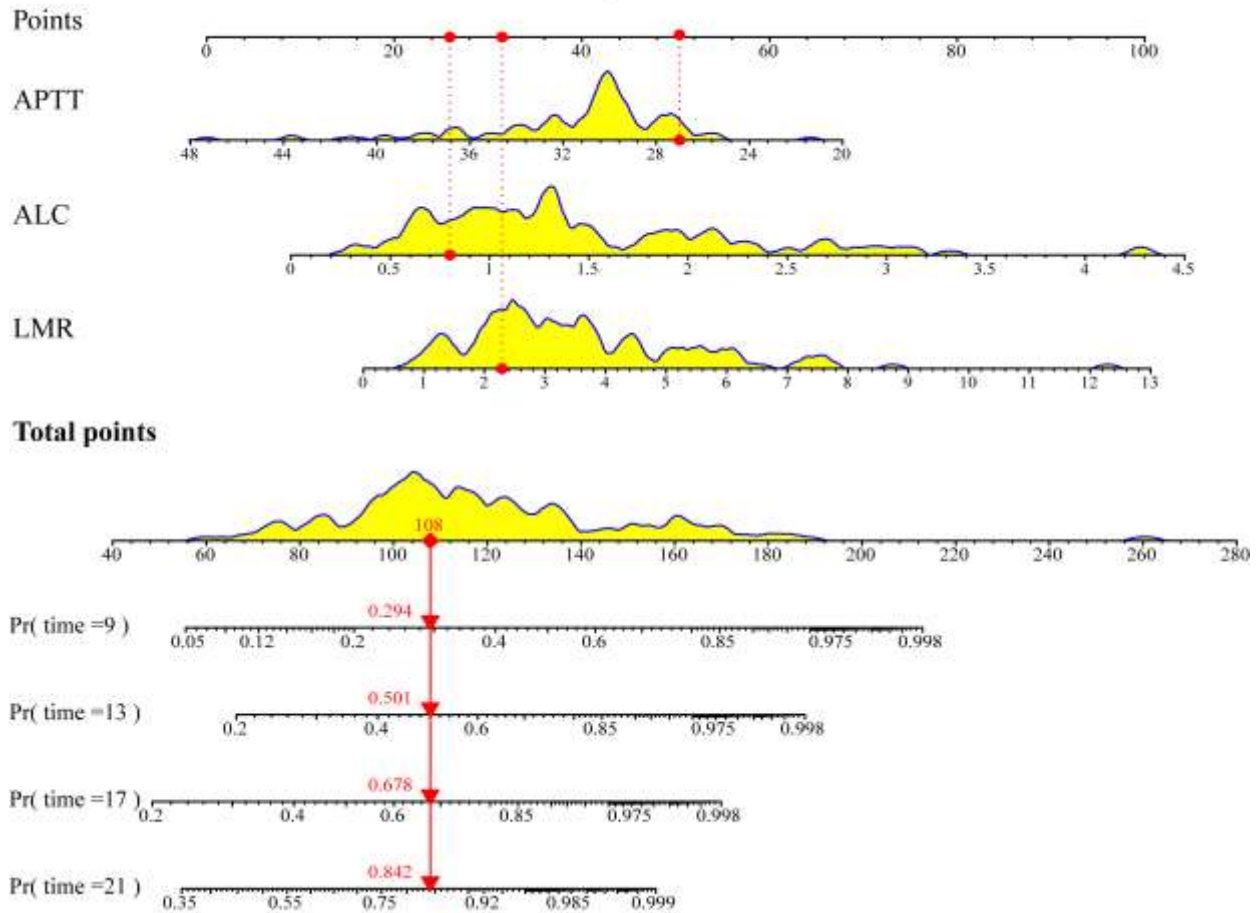
Abbreviations: WBC, white blood cell count; ANC, absolute neutrophil count; ALC, absolute lymphocyte count; AMC, absolute monocyte count; AEC, absolute eosinophil count; RBC, red blood cell count; CRP, C-reactive protein; CT, computerized tomography; APTT, activated partial thromboplastin time; PT, prothrombin time; LDH, lactate dehydrogenase; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; LMR, lymphocyte-to-monocyte ratio.

Table 3. Predictive performance of nomogram and ALC, LMR, APTT alone

	9 days		13 days		17 days		21 days	
	AUC 95%CI	P-value	AUC 95%CI	P-value	AUC 95%CI	P-value	AUC 95%CI	P-value
Training Cohort								
Nomogram	0.879(0.847,0.949)	ref	0.762(0.706,0.853)	ref	0.738(0.665,0.834)	ref	0.715(0.612,0.836)	ref
APTT	0.634(0.531,0.715)	<0.001	0.663(0.564,0.749)	0.005	0.656(0.547,0.748)	0.012	0.611(0.503,0.769)	0.015
ALC	0.829(0.764,0.908)	0.036	0.696(0.604,0.776)	0.029	0.693(0.602,0.797)	0.021	0.651(0.593,0.825)	0.033
LMR	0.807(0.744,0.882)	0.010	0.685(0.597,0.770)	0.011	0.657(0.553,0.751)	0.015	0.626(0.515,0.776)	0.025
Validation Cohort								
Nomogram	0.855(0.781,0.937)	ref	0.758(0.741,0.896)	ref	0.728(0.666,0.838)	ref	0.706(0.571,0.842)	ref
APTT	0.755(0.653,0.836)	0.019	0.654(0.536,0.752)	0.022	0.653(0.535,0.753)	0.019	0.602(0.495,0.732)	0.015
ALC	0.763(0.664,0.865)	0.024	0.710(0.602,0.797)	0.034	0.613(0.507,0.722)	0.014	0.628(0.518,0.751)	0.021
LMR	0.692(0.569,0.786)	0.001	0.639(0.507,0.722)	0.016	0.688(0.566,0.798)	0.022	0.668(0.521,0.790)	0.029

Abbreviations: ALC, absolute lymphocyte count; APTT, activated partial thromboplastin time; LMR, lymphocyte-to-monocyte ratio; AUC, areas under the curve; CI, confidence interval.

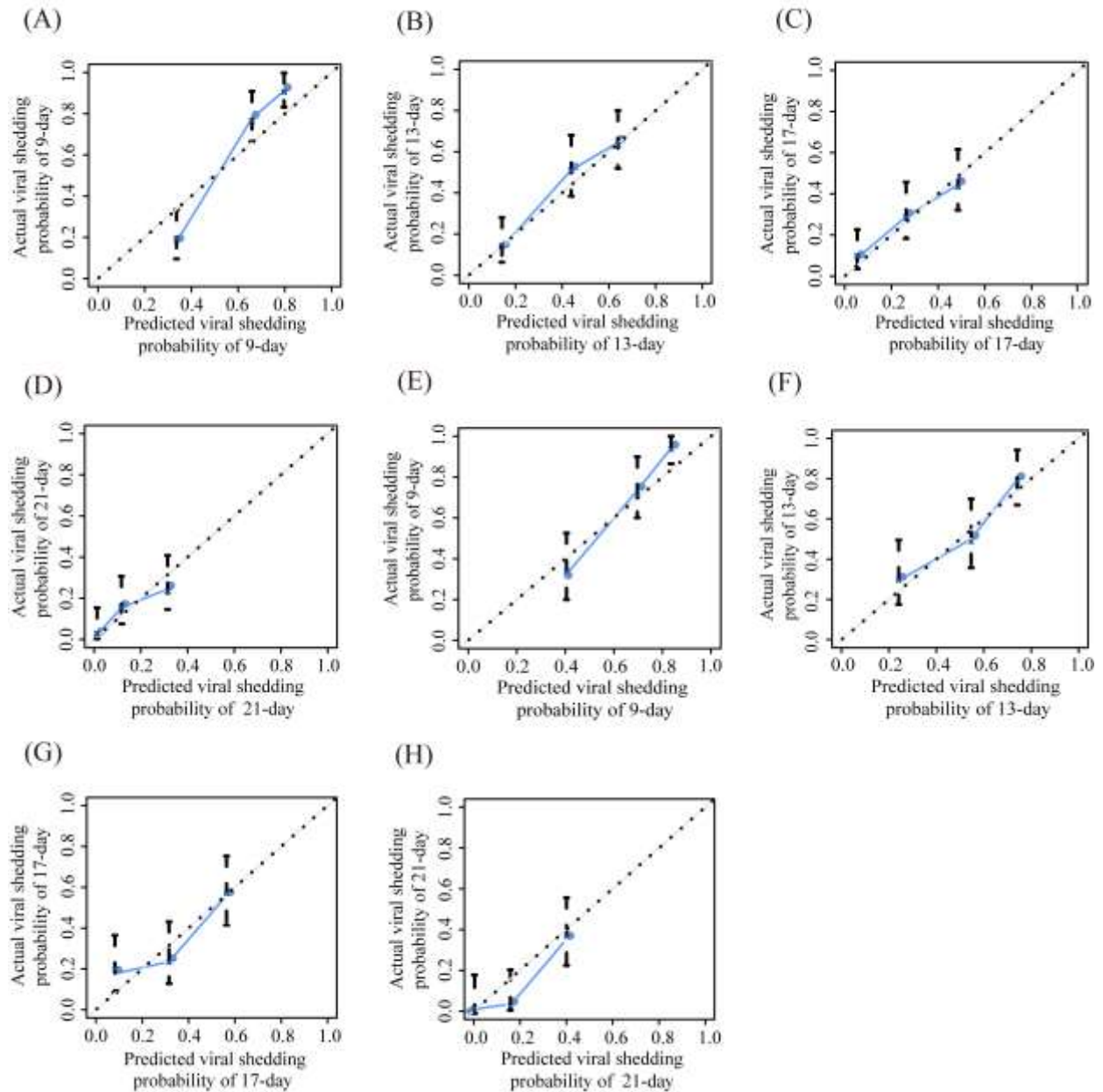
Figure 1. Nomogram to predict the probability of viral shedding at 9, 13, 17 and 21 days



The nomogram is constructed based on three factors: ALC, LMR and APTT. Each significant factor is shown in the variable axis and assigned a corresponding score via plotting an upward red dotted line to intersect with the points scale at the top of the nomogram. Then, the three obtained scores are added up to achieve a total score on the bottom total points scale. Yellow density plots represent the score distribution of significant factors and total points. Through the located summing score on the total points scale, a downward red solid line with an arrow intersecting with four probability line segment lines below can estimate a specific probability of viral shedding at 9, 13, 17 and 21 days for an individual patient. For example (indicated by a vertical red dotted line and solid line with arrow), if a COVID-19 patient on admission with an ALC of $0.8 \times 10^9/L$, LMR of 2.29 and APTT of 27 s, the corresponding scores of ALC, LMR and APTT for this patient were about 25.8, 31.2 and 51.0, respectively. Hence, the total score was 108, which indicated probabilities of 0.294, 0.501, 0.678 and 0.842 for viral shedding at 9, 13, 17 and 21 days, respectively, after admission.

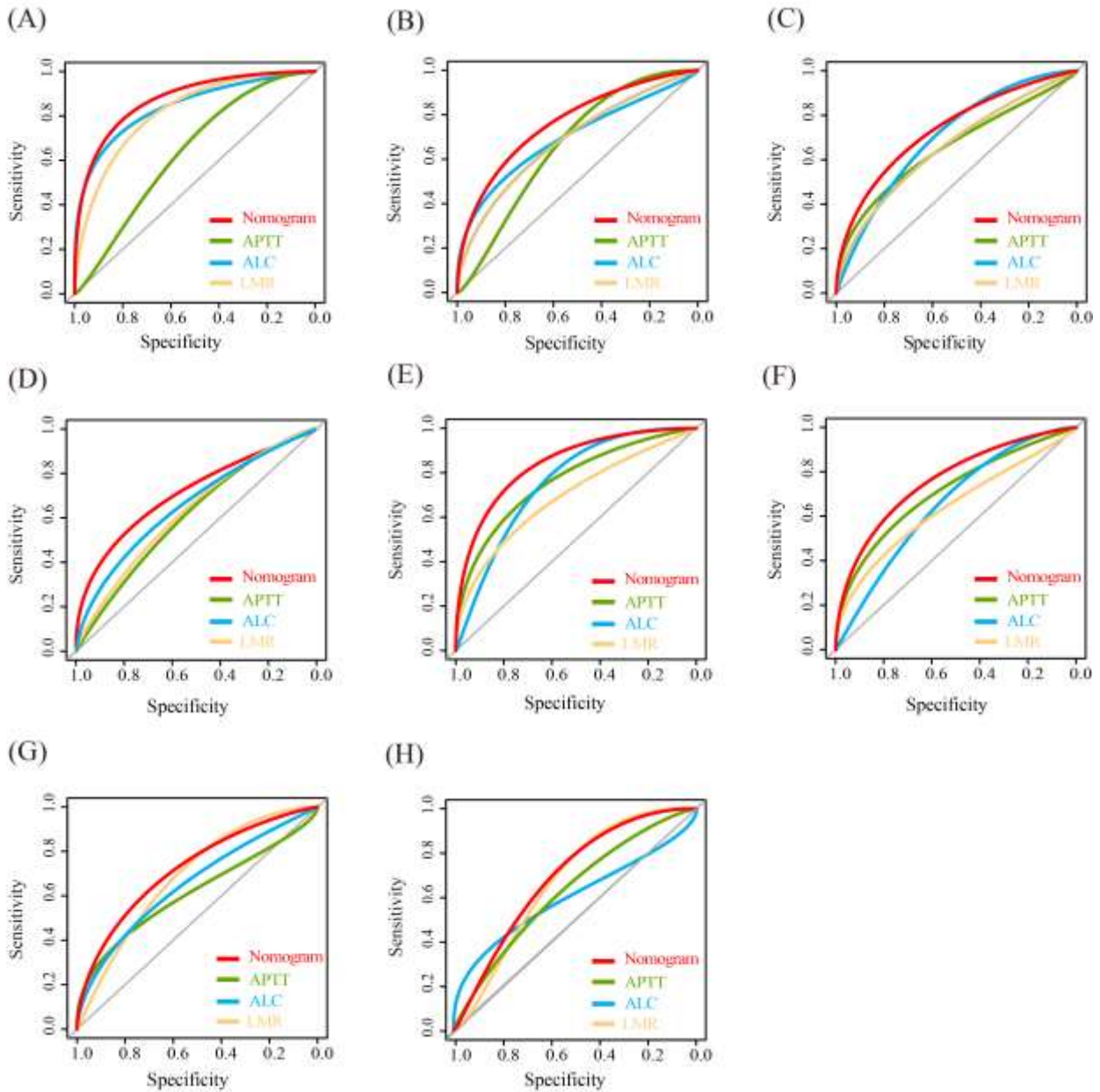
Abbreviations: ALC, absolute lymphocyte count; APTT, activated partial thromboplastin time; LMR, lymphocyte-to-monocyte ratio; Pr, Probability.

Figure 2. The calibration curves of the nomogram to predict viral shedding



Calibration curves of viral shedding probability at 9, 13, 17 and 21 days in the training cohort (A-D) and validation cohort (E-H). X-axis represents the nomogram-predicted viral shedding probability and the Y-axis represents the actual probability. The diagonal dotted line is an ideal line, which indicates perfect prediction of the nomogram. The solid blue line represents the actual predictive performance of the nomogram. The closer the solid blue line to the ideal line, the better the predictive accuracy of the nomogram.

Figure 3. ROC curve comparison of prediction between ALC, LMR, APTT and the nomogram



ROC curve analysis comparing the ALC, LMR, APTT and nomogram for predicting viral shedding probability in the training and validation cohorts at 9, 13, 17 and 21 days after admission. The ROC curve reflects the relationship between sensitivity and specificity, which shows the discriminative ability of the nomogram to identify patients with different outcomes. The X-axis is the 1-specificity (false positive rate) and the Y-axis is the sensitivity (true positive rate). The closer the value of the curve on the X-axis is to 0 and the greater the value of the curve on the Y-axis, the higher the value of the area under the curve (AUC), indicating a greater prediction accuracy. (A-D) Nomogram presenting AUCs of 0.879 (95%CI: 0.847-0.949), 0.762 (95%CI: 0.706-0.853), 0.738 (95%CI: 0.665-0.834) and 0.715 (95%CI: 0.612-0.836) at 9, 13, 17 and 21 days after admission, respectively, in the training cohort. These values were all higher than those of the AUCs of ALC, LMR and APTT alone (specific data shown in Table 3); (E-H). Nomogram presenting AUCs of 0.855 (95% CI: 0.781-0.937), 0.758 (95% CI: 0.741-0.896), 0.728 (95% CI: 0.666-0.838) and 0.706 (95% CI: 0.571-0.842) at 9, 13, 17 and 21 days after admission, respectively, in the validation cohort. These values were all higher than those of the AUCs of ALC, LMR and APTT alone (specific data shown in Table 3). Abbreviations: ROC, receiver operating characteristic curve; AUC, area under the ROC curve.

References

- WHO (world health organization), 2020c. Clinical management of covid-19 interim guidance. Available from <https://www.who.int/publications/i/item/clinical-management-of-covid-19>
- WHO (world health organization), 2020a. Coronavirus disease (covid-19) situation reports-209. Available from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200816-covid-19-sitrep-209.pdf?sfvrsn=5dde1ca2_2
- WHO (world health organization), 2020b. Laboratory testing of human suspected cases of novel coronavirus (ncov) infection (interim guidance). Available from <https://apps.who.int/iris/bitstream/handle/10665/330374/WHO-2019-nCoV-laboratory-2020.1-eng.pdf>
- NHC (national health commission of the people's republic of china), 2020. Clinical diagnosis and treatment guidance of 2019 novel coronavirus (2019-ncov) caused pneumonia (5th edition). Available from <http://www.nhc.gov.cn/yzygj/s7653p/202002/3b09b894ac9b4204a79db5b8912d4440/files/7260301a393845fc87fc6dd52965ecb.pdf>
- WHO (world health organization), 2020d. Criteria for releasing covid-19 patients from isolation. Available from <https://www.who.int/news-room/commentaries/detail/criteria-for-releasing-covid-19-patients-from-isolation>
- Bullard J, Dust K, Funk D, et al., 2020. Predicting infectious sars-cov-2 from diagnostic samples. *Clin Infect Dis*, <https://doi.org/10.1093/cid/ciaa638>
- Chen R, Liang W, Jiang M, et al., 2020. Risk factors of fatal outcome in hospitalized subjects with coronavirus disease 2019 from a nationwide analysis in china. *Chest*, 158(1):97-105. <https://doi.org/10.1016/j.chest.2020.04.010>
- Delong ER, Delong DM, Clarke-Pearson DL, 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3):837-845.
- Epidemiology Working Group for Ncip Epidemic Response CCFDC, Prevention, 2020. [the epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) in china]. *Zhonghua Liu Xing Bing Xue Za Zhi*, 41(2):145-151. <https://doi.org/10.3760/cma.j.issn.0254-6450.2020.02.003>
- Giannis D, Ziogas IA, Gianni P, 2020. Coagulation disorders in coronavirus infected patients: Covid-19, sars-cov-1, mers-cov and lessons from the past. *J Clin Virol*, 127:104362. <https://doi.org/10.1016/j.jcv.2020.104362>
- Guan WJ, Ni ZY, Hu Y, et al., 2020. Clinical characteristics of coronavirus disease 2019 in china. *N Engl J Med*, 382(18):1708-1720. <https://doi.org/10.1056/NEJMoa2002032>
- Henry BM, De Oliveira MHS, Benoit S, et al., 2020. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (covid-19): A meta-analysis. *Clin Chem Lab Med*, 58(7):1021-1028. <https://doi.org/10.1515/cclm-2020-0369>
- Li G, Fan Y, Lai Y, et al., 2020a. Coronavirus infections and immune responses. *J Med Virol*, 92(4):424-432. <https://doi.org/10.1002/jmv.25685>
- Li L, Zhang B, He B, et al., 2020b. Critical patients with coronavirus disease 2019: Risk factors and outcome nomogram. *J Infect*, 80(6):e37-e38. <https://doi.org/10.1016/j.jinf.2020.03.025>
- Li Q, Guan X, Wu P, et al., 2020c. Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia. *N Engl J Med*, 382(13):1199-1207. <https://doi.org/10.1056/NEJMoa2001316>
- Liang W, Liang H, Ou L, et al., 2020. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with covid-19. *JAMA Intern Med*, 180(8):1081-1089. <https://doi.org/10.1001/jamainternmed.2020.2033>
- Lin A, He Z-B, Zhang S, et al., 2020. Early risk factors for the duration of severe acute respiratory syndrome coronavirus 2 viral positivity in patients with coronavirus disease 2019. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 71(16):2061-2065. <https://doi.org/10.1093/cid/ciaa490>
- Subramaniam S, Scharer I, 2018. Procoagulant activity during viral infections. *Front Biosci (Landmark Ed)*, 23:1060-1081. <https://doi.org/10.2741/4633>
- Sun S, Cai X, Wang H, et al., 2020. Abnormalities of peripheral blood system in patients with covid-19 in wenzhou, china. *Clin Chim Acta*, 507:174-180. <https://doi.org/10.1016/j.cca.2020.04.024>
- Tang N, Li D, Wang X, et al., 2020. Abnormal coagulation parameters are associated with poor prognosis in patients with novel coronavirus pneumonia. *J Thromb Haemost*, 18(4):844-847. <https://doi.org/10.1111/jth.14768>
- Wolfel R, Corman VM, Guggemos W, et al., 2020. Virological assessment of hospitalized patients with covid-2019. *Nature*, 581(7809):465-469. <https://doi.org/10.1038/s41586-020-2196-x>
- Wu C, Chen X, Cai Y, et al., 2020. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in wuhan, china. *JAMA Intern Med*, 180(7):934-943. <https://doi.org/10.1001/jamainternmed.2020.0994>
- Xiao AT, Tong YX, Zhang S, 2020. Profile of rt-pcr for sars-cov-2: A preliminary study from 56 covid-19 patients. *Clin Infect Dis*, 71(16):2249-2251. <https://doi.org/10.1093/cid/ciaa460>
- Xu K, Chen Y, Yuan J, et al., 2020. Factors associated with prolonged viral rna shedding in patients with coronavirus disease 2019 (covid-19). *Clin Infect Dis*, 71(15):799-806. <https://doi.org/10.1093/cid/ciaa351>
- Yang AP, Liu JP, Tao WQ, et al., 2020a. The diagnostic and predictive role of nlr, d-nlr and plr in covid-19 patients. *Int Immunopharmacol*, 84:106504. <https://doi.org/10.1016/j.intimp.2020.106504>

- Yang Y, Shen C, Li J, et al., 2020b. Plasma ip-10 and mcp-3 levels are highly associated with disease severity and predict the progression of covid-19. *J Allergy Clin Immunol*, 146(1):119-127 e114. <https://doi.org/10.1016/j.jaci.2020.04.027>
- Zhao Q, Meng M, Kumar R, et al., 2020. Lymphopenia is associated with severe coronavirus disease 2019 (covid-19) infections: A systemic review and meta-analysis. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, 96:131-135. <https://doi.org/10.1016/j.ijid.2020.04.086>
- Zheng H-Y, Zhang M, Yang C-X, et al., 2020a. Elevated exhaustion levels and reduced functional diversity of t cells in peripheral blood may predict severe progression in covid-19 patients. *Cellular & molecular immunology*, 17(5):541-543. <https://doi.org/10.1038/s41423-020-0401-3>
- Zheng S, Fan J, Yu F, et al., 2020b. Viral load dynamics and disease severity in patients infected with sars-cov-2 in zhejiang province, china, january-march 2020: Retrospective cohort study. *BMJ*, 369:m1443. <https://doi.org/10.1136/bmj.m1443>
- Zhou F, Yu T, Du R, et al., 2020a. Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: A retrospective cohort study. *Lancet*, 395(10229):1054-1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)
- Zhou Y FB, Zheng X, Wang D, , Changcheng Z, Yingjie Q, et al., 2020. Pathogenic t-cells and inflammatory monocytes incite inflammatory storms in severe covid-19 patients. *National Science Review*, (6):6.
- Zhou Z, Ren L, Zhang L, et al., 2020b. Overly exuberant innate immune response to sars-cov-2 infection. *Social Science Electronic Publishing*,
- Zhu N, Zhang D, Wang W, et al., 2020. A novel coronavirus from patients with pneumonia in china, 2019. *N Engl J Med*, 382(8):727-733. <https://doi.org/10.1056/NEJMoa2001017>
- Zou L, Ruan F, Huang M, et al., 2020. Sars-cov-2 viral load in upper respiratory specimens of infected patients. *N Engl J Med*, 382(12):1177-1179. <https://doi.org/10.1056/NEJMc2001737>