Discriminant Models for the Prediction of Viral Shedding Time and Disease Progression in COVID-19

Wen Yang Li
   First hospital of Chinese Medical University

Yuhao Guo
   Xi'an Jiaotong University

Xiaowei Zheng
   Karolinska Institutet

Hongwen Zhao
   The first hospital of China Medical University

Jian Kang
   the first hospital of China Medical University

Daqing Wang
   Liaoning province peaples's Hospital

Wei Wang (✉ wwbycmu@126.com)

Research article

Discriminant models for the prediction of viral shedding time and disease progression in COVID-19

Wen-Yang Li[1], Yuhao Guo[4], Xiaowei Zheng[3], Hongwen Zhao[1], JianKang[1], Daqing Wang[2#], Wei Wang[1#]

[1] Department of Respiratory and critical care, the First Hospital of China Medical University, Shenyang, China. [2] Department of Respiratory disease, Liaoning province peoples' hospital, Shenyang, China. [3] Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden [4]. Department of Mathematics and Statistics, Xian Jiaotong University, Xi`an, China 710049

Summary: this study fully characterize the clinical features associated with postponed viral shedding and disease progression, then develop and validate two prognostic models with satisfactory discriminant performance.

Abstract

Background: COVID-19 infection can cause life-threatening respiratory disease. This study aimed to fully characterize the clinical features associated with postponed viral shedding and disease progression, then develop and validate two prognostic discriminant models.

Methods: This study included 125 hospitalized patients with COVID-19. 44 parameters were recorded, including age, gender, underlying comorbidities, epidemic features, laboratory indexes, imaging characteristics and therapeutic regimen, et al. F-test and χ^2 test were used for feature selection. All models were developed with 4-fold cross-validation, and the final performances of each model were compared by the Area Under Receiving Operating Curve (AUROC). After optimizing the parameters via L_2 regularization, prognostic discriminant models were built to predict postponed viral shedding and disease progression of COVID-19 infection. The test set was then used to detect the predictive values via assessing models sensitivity and specificity.

Results: 69 patients had a postponed viral shedding time (>14 days), and 28 of 125 patients progressed into severe cases. Eleven and six demographic, clinical features and therapeutic regimen were significantly associated with postponed viral shedding and disease progressing, respectively ($p<0.05$). The optimal discriminant models are: y_1 (postponed viral shedding) = -0.244 + 0.2829x_1 (the interval from the onset of symptoms to antiviral treatment) + 0.2306x_4 (age) + 0.234x_{28} (Urea) - 0.2847x_{34} (Dual-antiviral therapy) + 0.3084x_{38} (Treatment with antibiotics) + 0.3025x_{21} (Treatment with Methylprednisolone); y_2 (disease progression) = -0.348 - 0.099x_2 (interval from Jan 1st,

2020 to individualized onset of symptoms) + $0.0945x_4$ (age) + $0.1176x_5$ (imaging

characteristics) + $0.0398x_8$ (short- term exposure to Wuhan) - $0.1646x_{19}$ (lymphocyte

counts) + $0.0914x_{20}$ (neutrophil counts) + $0.1254x_{21}$ (neutrphil/lymphocyte ratio) +

$0.1397x_{22}$ (C-Reactive Protein)+ $0.0814x_{23}$ (Procalcitonin) + $0.1294x_{24}$ (Lactic

dehydrogenase) + $0.1099x_{29}$ (Creatine kinase). The output ≥ 0 predicted postponed viral

shedding or disease progressing to severe/critical state. These two models yielded the

maximum AUROC, and faired best in terms of prognostic performance (sensitivity of

73.3%, 75%, and specificity of 78.6%, 75% for prediction of postponed viral shedding

and disease severity, respectively).

Conclusion: The two discriminant models could effectively predict the postponed viral

shedding and disease severity, and be used as early-warning tools for COVID-19.

Key words: prognostic discriminant model, postponed viral shedding time, disease

progression, COVID-19

Introduction

The prevalence of coronavirus disease 2019 (COVID-19) has put a huge burden to medical resources [1]. Although patients with COVID-19 infection mostly manifested as non-severe cases, it can also cause life-threatening conditions before or during hospitalization, such as severe pneumonia, adult respiratory distress syndrome or multiple organ failure, which are all related to worse outcomes [2]. Compare to the other epidemic disease, such as the previous outbreaks of SARS-CoV and MERS-CoV, COVID-19 progresses and spreads more rapidly, with peculiar epidemiological traits. High viral loads of SARS-CoV-2 were observed in the upper respiratory specimens of patients with little or no symptoms, this indicated that inapparent-transmission plays a major but underestimated role in sustaining the outbreak of COVID-19 [3].

Thus, in order to optimize therapeutic strategies and to effectively control the transmission resources on the imported cities or regions, to promptly identify the viral shedding time and disease severity is required but challenging. Traditional evaluation scoring tools, such as CURB-65, qSOFA, and NEWS, could be adopted to assess disease severity, but not for the early assessment of COVID-19 severity [4]. To date, a certain number of models predict the risk of severe COVID-19 has been developed [5-7]. However, there is geographic discrepancy of severity and mortality rate in patients with COVID-19 infection, and most of the previous prediction models were established to predict survival risk or progression to severe or critical state in the south part of China. Few of them was designed to predict the postponed viral shedding time. Besides, the models established previously were mainly consist of variables that extracted from the

clinical and laboratory parameters, with few of them incorporated epidemic features or therapeutic regimen. Since the first case emerged in Liaoning province in Jan 22th, 2020, especially the rebounding outbreak of imported cases from abroad in the northeast of China (Jilin province) in May, 2020, there is an urgent need to construct a simple, efficient and accurate "early-warning prediction model" for disease progression at early stage once the patients was admitted to the hospital.

This study aimed to fully characterize the demographic, epidemic, clinical features and therapeutic regimens and to detect their association with postponed viral shedding time and disease progression among patients with COVID-19 in Liaoning province, China. Then, to specifically design and validate two prognostic discriminant models incorporating the associated features. These models can serve as early-warning prediction tools to estimate the postponed viral shedding and to identify the risk of progressing to severe stage in advance among patients with COVID-19 infection.

Methods

Study design

This retrospective multi-center cohort study included consecutive patients that were laboratory-confirmed of COVID-19 infection and enrolled from Jan 22th to Mar 22th, 2020 in eight designated hospitals throughout Liaoning province. All patients enrolled were diagnosed with COVID-19 according to the WHO interim guidance [8]. Laboratory confirmation of COVID-19 was achieved by the nucleic acid test using real-time reverse transcriptase-polymerase chain reaction (RT-PCR) assay at Liaoning municipal Center for

Disease Prevention and Control (CDC). Samples were collected using a nose swab and/or throat swab from each suspected patients. This study was approved by the institutional review board of each participating site, and the writing informed consent was waived.

Data collection

The date of disease onset (defined as the day when any symptom was noticed by the patients) and hospital admission date, the first day for nucleic acid detected to be positive or negative were all recorded. Virus detection was repeated twice every 24 hours. All the patients were hospitalized and the clinical outcomes were monitored for at least 8 weeks. All the clinical data on epidemiology (recent exposure history), symptoms, signs, underlying comorbidities, laboratory results (on admission), imaging findings (on admission) and clinical progression were recorded and retrospectively double-extracted from electronic medical records, with two independent reviewers extracted the data and evaluated the eligibility of the original data.

Regarding antiviral treatment, 14 patients with confirmed COVID-19 were treated with lopinavir–ritonavir (400 mg/100 mg), 11 with arbidol (200mg t.i.d.). Dual-antiviral therapy of nebulized Interferon-α (IFN-α) (5mU b.i.d.) with lopinavir–ritonavir or arbidol were used in 42 and 8 patients. Triple-antiviral therapy of IFN-α, lopinavir–ritonavir and arbidol were used in 20 patients. Other therapies such as antibiotics and corticosteroids was used in 38 and 22 patients.

Definition of disease progression and postponed virus shedding time

Disease progression were recorded for at least 8 weeks after admission. Severity of COVID-19 was defined according to the American Thoracic Society (ATS) guidelines for

community-acquired pneumonia (CAP) [9]. Severe/critical cases of COVID-19 should

meet one major criterion (septic shock with need for vasopressors or respiratory failure

requiring mechanical ventilation) or at least three minor criteria: (a) respiratory distress

with respiratory frequency ≥30/min; (b) oxygenation index (artery partial pressure of

oxygen/inspired oxygen fraction, PaO_2/FiO_2) ≤ 250 mmHg; (c) multilobar infiltrates,

confusion/disorientation; (d) uremia (blood urea nitrogen≥20 mg/dL); (e) leukopenia

(white blood cell count <400 cells/μL); (f) thrombocytopenia (platelet count

<100,000/μL); (g) hypothermia (core temperature <36˚C); (h) hypotension requiring

aggressive fluid resuscitation. Non-severe patients were defined as a confirmed case with

fever, respiratory symptoms, with or without radiographic evidence of pneumonia.

The median viral shedding time was 14 days (IQR, 11–19), those ≥14 days were

deemed as postponed virus shedding.

Variables

Disease severity and postponed viral shedding discriminant variables were determined

according to the assessment of the existing medical records (Table 1) listed as follows: the

interval from onset of symptoms to antiviral treatment (x_1), interval from Jan 1st ,2020

(the day for the first case emerged in Liaoning) to individualized onset of symptoms (x_2),

gender (x_3), age (x_4), imaging characteristics (x_5); long-term exposure to Wuhan (x_6),

local transmitted history (x_7), short- term exposure to Wuhan (x_8); respiratory symptoms

(x_9), digestive symptoms (x_{10}), general malaise (x_{11}), fever (x_{12}); comorbidities with

chronic respiratory disease (x_{13}), hypertension(x_{14}), diabetes mellitus (x_{15}), surgery

history (x_{16}), other comorbidities (x_{17}), White blood cell (WBC) (x_{18}), lymphocyte counts

(x_{19}), neutrophil counts (x_{20}), neutrophil/lymphocyte ratio (N/L) (x_{21}), C-Reactive Protein (CRP) (x_{22}), Procalcitonin (PCT) (x_{23}), Lactic dehydrogenase (LDH) (x_{24}), Aspartate amino transferase (AST) (x_{25}), Alanine aminotransferase (ALT) (x_{26}), Creatinine (Cr) (x_{27}), Urea nitrogen (Urea) (x_{28}), Creatine kinase (CK) (x_{29}), oxygenation index(x_{30}), disease severity (x_{31}), treatment with Lobinavi/ritonavir alone (x_{32}), treatment with Arbidol alone (x_{33}), combined treatment of nebulized IFN-α with lopinavir–ritonavir (x_{34}), combined treatment of nebulized IFN-α, with Arbidol (x_{35}), combined treatment of nebulized IFN-α, lopinavir–ritonavir and Arbidol (x_{36}), treatment with Oseltamivir phosphate alone (x_{37}), Treatment with antibiotics (x_{38}), treatment with ribavirin (x_{39}), treatment with Chinese traditional medicine (Xuebijing) (x_{40}), treatment with Methylprednisolone (x_{41}), treatment with γ-globulin (x_{42}), antiviral treatment course (x_{43}). The outcome were viral shedding time (y_1) and disease progression (y_2).

Discriminate factors were also quantitatively assigned, some variables such as the imaging characteristics were assigned from 0, and the order was based on their influence on the progression of disease.

Statistical analysis

Categorical variables were summarized as frequencies and percentages. Continuous variables were described using median and interquartile ranges (IQR) values. Comparisons between groups were tested by F test for continuous variable, or by χ^2 test for categorical data. Features significantly different ($p < 0.05$) in both algorithms were selected into the models. Setting disease progression and postponed viral shedding time as the goal for discriminant models, logistic regression, linear discriminant analysis, K-

nearest neighbor, support vector machine (SVM) and decision tree were constructed through Python software. After comparing effectiveness among them, the most optimal discriminant models incorporating multiple related factors were established to reflect the probability of disease progressing to severe stage or postponed viral shedding. The models were constructed using the output as an outcome, while the output ≥ 0 indicated disease progressing to severe/critical stage or postponed viral shedding time. The precision of the prediction models were further evaluated and validated using the area under the receiver-operating characteristic curve (AUROC).

Results

Demographic and clinical characteristics of patients with COVID-19 infection

A total of 125 hospitalized patients diagnosed as COVID-19 infection were included in this study. Disease progression were recorded during the 8-weeks follow-up after admission. Ten patients were categorized into mild cases, 97 developed into moderate cases and 28 patients developed into severe cases (including 3 critically ill) during hospitalization. Symptoms were shown in Table 2. Among 125 patients enrolled, 75 (62%) were imported cases (with an exposure history to Wuhan), which further divided into 37 (29.6%) short-term exposure (meeting, transfer or travel et al.) and 38 (30.4%) long-term exposure (residence or study, work in Wuhan) history to Wuhan. The rest 50 (40%) cases had no history of Wuhan exposure, thus were categorized into local transmitted cases who had contact with symptomatic cases. The median age for all patients were 44 years, for non-severe patients was 48 years old, and for severe patients

was 50 years old. About half (55.2%) of patients were male. Among all patients, up to 61 patients had at least one underlying comorbidities, the most common of which were chronic diseases, such as hypertension and diabetes, et al. The median interval from the onset of symptoms to hospital admission was 4 days (IQR, 2-7). The median viral shedding time was 14 days (IQR, 11–19).

Clinical and laboratory features associated with disease progression or postponed viral shedding time

A total of 44 laboratory and clinical records on admission and during hospitalization were obtained and analyzed, including but not limited to the demographics, symptoms, signs, images, blood routine, immunocytochemistry, enzymatic and liver/renal function. These data was acquired within 24 hours on admission. In order to eliminate the overfitting effect and regularize the models, z-score standardization [$x^* = (x-mean)$ /standard deviation] was conducted on all continuous variables in the data set. Thus, each corresponding feature was converted into a normal distribution with mean value of 0 and variance of 1, so as to eliminate the dimensional influence. The pre-conditioned data were then compared between groups divided by disease severity (non-severe vs. severe group) or viral shedding time (with cutoff value 14 days) using F test or χ^2 test. Features were considered to be significantly associated with disease progression or postponed viral shedding time when $p < 0.05$ (Table 2).

For the postponed viral shedding discriminant model, six clinical, laboratory and treatment features were selected to be the significant indicators of the postponed virus clearance time, including the interval from the onset of symptoms to antiviral treatment

(day) (x_2), age (x_4), CK (x_{29}), Combined treatment of nebulized IFN-α with

lopinavir/ritonavir (x_{34}), Treatment with antibiotics (x_{38}), Treatment with

Methylprednisolone (x_{41}) (Table 3).

For the disease progression discriminant model, eleven demographic and

imaging/laboratory features were significantly associated with the progression of severe

disease. These features include the interval from Jan 1st ,2020 (the day for the first case

emerged in Liaoning) to individualized onset of symptoms (x_2), age(x_4), imaging

characteristics (x_5), epidemic history of short-term exposure to Wuhan (x_8), and immune

features (Lymphocyte counts (x_{19}), N/L ratio (x_{21}), CRP (x_{22}), PCT (x_{23})), Lactic

dehydrogenase (x_{24}), creatine kinase (x_{29}).

No multi-collinearity was found in the screened variables, since VIF (Variance Inflation

Factor) values of the screened dependent variables were all less than 10. Thus, the model

passed the multi-collinearity test (Table 4, 5).

Establishment and verification of the optimal discriminant models for disease

progression and postponed viral shedding time

All samples were stratified and randomly divided into training and testing datase

ts, and all models were developed with 4-fold cross-validation [10]. During constru

ction of the discriminant model of postponed viral shedding time, the training data

sets contained 55 postponed virus clearance cases and 45 non-postponed cases, whi

le the testing dataset was consist of 14 postponed virus clearance cases and 12 po

stponed cases. During construction of the discriminant model of disease progressio

n, the training datasets contained 31 severe/critical cases and 69 non-severe cases,

while the testing dataset was consist of 8 severe/critical cases and 18 non-severe cases. Then, the discriminant models were constructed in training set via the selected indicators using several methods, including logistic regression, linear discriminant analysis, decision tree, K-nearest neighbor, and support vector machine method. The outcome was predicted in testing data set.

By analyzing total accuracy in both training and testing data set, the optimal discriminant models for the prediction of disease progression and postponed viral shedding time of COVID-19 infection were established via the logistic regression with the selected eleven and six features used as independent variables. In order to eliminate the influence of model overfitting on the prediction results and to improve the generalization ability of the model, L_2 regularization [$\min_{\beta,C} \frac{1}{2}\beta^T\beta + C\sum_{i=1}^{n}\log(\exp(-y_i(X_i\beta+c)+1))$] was used to constrain the objective function in the logistic regression optimization process.

"Viral shedding postponed or not" was used as a dependent variable (y_1) for logistic regression, which yielded the equation:

$$Ln\left(\frac{p}{1-p}\right) = -0.244 + 0.2829x_1 + 0.2306x_4 + 0.234x_{28} - 0.2847x_{34} + 0.3084x_{38} + 0.3025x_{41}$$

"Severe or not" was used as a dependent variable (y_2) for logistic regression. This yielded the equation:

$$Ln\left(\frac{p}{1-p}\right) = -0.348 - 0.099x_2 + 0.0945x_4 + 0.1176x_5 + 0.0398x_8 - 0.1646x_{19} + 0.0914x_{20} + 0.1254x_{21} + 0.1397x_{22} + 0.0814x_{23} + 0.1294x_{24} + 0.1099x_{29}$$

Notably, all data needs to be equally standardized [$x^{*} = \dfrac{x - \mu}{\sigma}$] before being plugged into the equation, and the corresponding mean and standard deviation were determined and shown in Table 6,7.

Performance of the discriminant models

When applying the discriminant model of postponed viral shedding time onto the validation of the training set (Table 8), sensitivity was 0.727, specificity was 0.733, positive predictive value was 0.769, and negative predictive value was 0.688. In the test set, sensitivity was 0.786, specificity was 0.667, with positive predictive value 0.733 and negative predictive value 0.727. Recall rate was 78.6%. Accuracy was 0.732. The AUROC of the combinations of 7 clinical features was 0.73 in the training dataset (Figure 1a) and 0.73 in the testing dataset (Figure 1b).

According to the confusion matrix of discriminant model for disease progression (Table 9), during the validation of the training set, sensitivity was 0.774, specificity was 0.884, positive predictive value was 0.75, and negative predictive value was 0.897. In the test set, sensitivity was 0.75, specificity was 0.889, with positive predictive value 0.75 and negative predictive value 0.889. The recall rate was 75%, and accuracy was 0.846. The AUROC was also constructed to evaluate the effectiveness of the discriminant models The AUROC of the combinations of 11 demographic, clinical and imaging/laboratory features was 0.829 in the training dataset (Figure 2a) and 0.819 in the testing dataset (Figure 2b).

Discussion

This retrospective study tentatively developed two discriminant models consisting of

several clinical features that could be quickly obtained on admission. For the first time, the impact of epidemic features and therapeutic regimen on the disease progression and postponed viral shedding were described and integrated into the predicted polynomial equations among patients confirmed as COVID-19 infection. These two discriminant models can be applied as "early warning tools" to robustly and effectively predict the postponed viral shedding and the severe/critical progression in patients with COVID-19 infection.

In this study, the median viral shedding time was 14 days (IQR, 11–19). For hospitalized patients, one of the isolation release and discharge criteria is a sputum/oral swab testing negative twice with 24h interval [5], however, SARS-CoV-2 in the respiratory tract, especially sputum, has been observed to be associated with a prolonged viral shedding and high viral load, when compared with the stool specimens [11]. In order to optimize therapeutic strategies and to effectively control the transmission resources on the imported cities or regions, it is essential to identify factors that associated with the COVID-19 PCR negative conversion time and to establish a prediction model that could individually estimate the risk of postponed viral shedding time among patients with COVID-19 upon admission to hospitals or shelters. To the best of our knowledge, no previous studies have evaluated discriminant models for predicting the virus shedding time in patients with different severities of COVID-19 infection outside Wuhan. In this study, we initially selected 43 variables probably associated with disease progression and postponed viral shedding time respectively according to the published literatures [5], then detected the risk factors by F-test and χ^2 test analysis. As a result, 6 variables were

identified as discriminatory factors and were devised to discriminant models for prediction of postponed viral shedding of COVID-19 infection.

Both older age and delayed antiviral treatment could give rise to the postponed virus clearance time. Consistently, the association between delayed initiation of antiviral treatment and the prolonged virus shedding for influenza A (H7N9) and SARS-CoV-2 was also observed in previous studies, indicating that timely initiation of antiviral treatment is necessary for viral clearance [12-14]. In addition, dual-antiviral therapy of nebulized IFN-α with lopinavir/ritonavir (x_{33}) was negatively associated with the viral shedding time, whereas treatment with antibiotics (x_{37}) and methylprednisolone (x_{40}) were related to postponed viral shedding. So far, no antiviral drug targeted the virus or the host cell has been proved to be effective for the treatment of COVID-19 [15], but few existing antiviral drugs have brought hopes when the reduction of viral load is concerned. Evidence has emerged that SARS-CoV-2 is more susceptible to IFNs when compared to SARS-CoV, as the inhalation of Interferon-α (IFN-α) 2b could reduce the infection rate significantly and it can be used for prophylaxis of SARS-CoV-2 infection [16, 17]. In addition, Lopinavir/ritonavir (Kaletra) was found to have anti-SARS-CoV efficacy in vitro [18], but presented controversial therapeutic effects as compared to the standard care in vivo [19-21]. Lopinavir/ritonavir is recommended by the National Health Commission of China for the treatment of COVID-19 at present [22]. For the first time, this study observed that combined treatment of IFN-α inhalation and lopinavir/ritonavir was related to the shortened viral shedding time. Nevertheless, there is no additional benefit on virus clearance from an extra use of Arbidol, when combined with IFN-α and

lopinavir/ritonavir. Anyway, statistical analysis can only stress association but cannot explain causality. A recent study also observed that early initiation of dual-antiviral treatment with lopinavir/ritonavir + IFN-α combination therapy could help shorten the duration of SARS-CoV-2 shedding when compared with triple antiviral treatment (opinavir/ritonavir + IFN-α + arbidol) [14]. This conclusion may provide a rationale for clinicians to optimize and early initiate the antiviral treatments. Conversely, treatment with antibiotics (x_{37}) and methylprednisolone (x_{40}) could give rise to the prolonged viral shedding time. This finding is consistent with previous studies, which observed that high-dose corticosteroids were associated with increased mortality and longer viral shedding in patients with influenza A (H7N9) viral pneumonia and MERS [23-25]. Systemic corticosteroids could increase the risk of opportunistic infections (such as bacterial or fungal) that occur secondary to immunosuppression, and eventually hinders the virus clearance ability [26]. Besides, potential bacterial infections secondary to influenza viral infection that has been commonly seen in this study (32.8%) could also prolong the viral clearance time, as indicated by the evidence that the use of antibiotics was associated with postponed viral shedding time.

Noteworthy, this study for the first time observed that intervals from the first case emerged in Liaoning province (Jan 22st, 2020) to the individualized onset of symptoms (x_2) could serve as an important prognostic feature in our early-warning model of disease progression. In this study, nearly half of the confirmed cases at the early stage (in January) of COVID-19 outbreak were severe cases, whereas in the latter period (after February), the percentage of non-severe cases became dominant (76.9%) in Liaoning province.

Another explanation is that pathogens tend to reduce their virulence overtime in order to maximize their between-host transmission, which could result in the gradually lowered severity of COVID-19 infection on the imported regions [27, 28]. Besides, the human intervention efforts in China, such as to promptly admit suspected patients to the designated shelter hospitals, have effectively contributed to the decreased number of severe cases of COVID-19. Interestingly, short-term exposure to the epidemic area (Wuhan) during traveling or transfer could also result in a higher likelihood of progressing into severe stage of disease. One explanation is that transiently migrant individuals may have been primed by one or more prior coronavirus exposures during traveling or transfer around the epidemic area, thus have experienced the effects of antibody dependent enhancement (ADE) antigenic epitope heterogeneity due to antigenic epitope heterogeneity [29]. ADE hinders the ability to manage inflammation, and result in disease progression. Another explanation would be attributed to the SARS-CoV-2 strains of L type, which are evolutionarily more aggressive and contagious. This virus strains of L type with altered virulence could probability be the underlying causal pathogen for patients who acquired infections via short exposure to epidemic area [30]. In our discriminant models, both immune features (lymphocytes, neutrophils, N/L ratio, CRP) and enzymatic index (LDH, CK) obtained on admission were observed as the most significant prognostic factors for disease severity, which is consistent with the previous literatures. This illustrate an earlier exhibition of abnormal laboratory features prior to the disease progression [31, 32].

Overall, these two discriminant models in the present study was demonstrated to have

satisfied sensitivity (>72.00%) and specificity (>73.00%), and they can be used as early warning tools of disease progression among patients with COVID-19 upon admission. A medical staff can easily predict in advance using these two discriminant models and conduct a timely and optimal medical intervention at an early stage. Anyway, some limitations should be noted in this study. First, the study design was retrospective and the sample size may be insufficient for characterization of an entire population. However, by including all patients from eight designated hospitals throughout Liaoning province, we considered patients recruited in this study are representative of cases diagnosed with COVID-19 in Liaoning, China. Secondly, not enough severe/critical cases were recruited for the present study, possibly because the morbidity of fatality rate of patients infected by COVID-19 in Liaoning province was lower (1.6%) than the whole national average level (3.2%) [33], and not resembling that in previous studies from Wuhan [2, 34]. Thirdly, we only included the initial antiviral treatments as factors for prolonged shedding duration, so as to minimize bias as much as possible. This discrepancy may have had an unknown influence on the efficacy of the models.

In conclusion, the discriminant models reported here is the first attempt of its kind to develop an early warning tool for both postponed viral shedding and disease progression in the northeast area of China. We believe that these models can help to judge the disease progression early enough in a great number of patients with COVID-19 infection, and this early judgement can facilitate a timely medical intervention, which will ultimately reduce the prevalence and mortality of COVID-19.

Acknowledgments

We thank all the efforts from staff of Liaoning provincial health commission and

Liaoning COVID-19 treatment team in the management of the confirmed patients.

Author contributions

Wen-Yang Li: Conceptualization, Formal analysis, Investigation, Methodology,

Resources, Supervision, Writing - original draft. Yuhao Guo: Formal analysis,

Methodology. Xiaowei Zheng: Formal analysis, Investigation, Supervision. Hongwen

Zhao: Formal analysis, Supervision. Wen-Yang Li: Data curation. Jian Kang: Validation,

Visualization. Wei Wang: Conceptualization, Funding acquisition, Investigation,

Validation, Visualization, Writing original draft, review & editing.

Competing interests

All authors declare no competing interests.

References

1. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in

   Wuhan, China, of Novel Coronavirus-Infected Pneumonia. N Engl J Med.

   2020;382(13):1199-1207.

2. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical

characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020; 395 (10223):507–13.

3. Wang Y, Wang Y, Chen Y, Quin Q. Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. J. Med. Virol. 92, 568–576 (2020).

4. Ito A, Ishida T, Tokumasu H, Washio Y, Yamazaki A, Ito Y, et al. Prognostic factors in hospitalized community-acquired pneumonia: a retrospective study of a prospective observational cohort. BMC Pulm Med. 2017; 17(1):78.

5. Zhou Y, He Y, Yang H, Yu H, Wang T, Chen Z, et al. Development and validation a nomogram for predicting the risk of severe COVID-19: A multi-center study in Sichuan, China. PLoS One. 2020;15(5):e0233328.

6. Sun L, Song F, Shi N, et al. Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19. J Clin Virol. 2020;128:104431.

7. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. BMJ. 2020 Jun 3;369:m2204

8. World Health Organization. (2020). Clinical management of COVID-19: interim guidance, 27 May 2020. World Health Organization. https://apps.who.int/iris/handle/10665/332196.

9. Metlay JP, Waterer GW, Long AC, Anzueto A, Brozek J, Crothers K, et al. Diagnosis and Treatment of Adults with Community-acquired Pneumonia An Official Clinical

Practice Guideline of the American Thoracic Society and Infectious Diseases Society of America. Am J Resp Crit Care. 2019; 200(7):E45–E67.

10. Wolff RF, Moons KG, Riley RD, et al, for the PROBAST Group†. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019;170:51–58.

11. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, et al. Virological assessment of hospitalized patients with COVID-19. Nature. Apr 1 2020.

12. Wang Y, Guo Q, Yan Z, et al. Factors associated with prolonged viral shedding in patients with avian influenza A(H7N9) virus infection. J Infect Dis. 2018; 217(11):1708-1717.

13. Zheng S, Wang Y, Yu F, et al. Benefit of early initiation of neuraminidase inhibitor treatment to hospitalized patients with avian influenza A (H7N9) virus. Clin Infect Dis 2018; 66(7):1054-1060.

14. Zuo Y, Liu Y, Zhong Q, Zhang K, Xu Y, Wang Z. Lopinavir/ritonavir and interferon combination therapy may help shorten the duration of viral shedding in patients with COVID-19: A retrospective study in two designated hospitals in Anhui, China. J Med Virol. 2020;10.1002/jmv.26127.

15. Sheahan TP, Sims AC, Leist SR, et al. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. Nat Commun. 2020;11:222.

16. Shen K, Yang Y. Diagnosis and treatment of 2019 novel coronavirus infection in children: A pressing issue. World J Pediatr. 2020.

17. Lokugamage KG, Schindewolf C, Menachery VD. SARS-CoV-2 sensitive to type I interferon pretreatment. bioRxiv. 2020.

18. Chu CM, Cheng VCC, Hung IFN, Wong MML, Chan KH, Chan KS, Kao RYT, Poon LLM, Wong CLP, Guan Y, Peiris JSM, Yuen KY. Role of lopinavir/ritonavir in the treatment of SARS: Initial virological and clinical findings. Thorax. 2004; 59:252-256.

19. Cao B, Wang Y, Wen D, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. N Engl J Med. 2020.

20. Li Y, Xie Z, Lin W, Cai W, Wen C, Guan Y, et al. An exploratory randomized, controlled study on the efficacy and safety of lopinavir/ ritonavir or arbidol treating adult patients hospitalized with mild/moderate COVID-19 (ELACOI). medRxiv 2020.03.19.20038984.

21. Lim J, Jeon S, Shin HY, Kim MJ, Seong YM, Lee WJ, et al. Case of the Index Patient Who Caused Tertiary Transmission of COVID-19 Infection in Korea: the Application of Lopinavir/Ritonavir for the Treatment of COVID-19 Infected Pnuemonia Monitored by Quantitative RT-PCR. J Korean Med Sci. 2020; 35(6): e79.

22. National Health Commission of China. Guideline for COVID-19 pneumonia of China; 2020. [cited 2020 Feb 8].

23. Cao B, Gao H, Zhou B, Deng X, Hu C, Deng C, et al. Adjuvant Corticosteroid Treatment in Adults With Influenza A (H7N9) Viral Pneumonia. Crit Care Med. 2016;44(6):e318-e328.

24. Hu Y, Lu S, Song Z, Wang W, Hao P, Li J, et al. Association between adverse clinical

outcome in human disease caused by novel influenza A H7N9 virus and sustained

viral shedding and emergence of antiviral resistance. Lancet 2013; 381:2273–2279.

25. Arabi YM, Mandourah Y, Al-Hameed F, Sindi AA, Almekhlafi GA, Hussein MA, et al. Corticosteroid Therapy for Critically Ill Patients with Middle East Respiratory. Am J Respir Crit Care Med. 2018; 197(6):757-767.

26. Rice TW, Rubinson L, Uyeki TM, Vaughn FL, John BB, Miller RR 3rd, et al; NHLBI ARDS Network: Critical illness from 2009 pandemic influenza A virus and bacterial coinfection in the United States. Crit Care Med 2012; 40:1487–1498

27. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. N Engl J Med. 2020 Mar 5;382(10):929-936.

28. Lei H, Li Y, Xiao S, Lin CH, Norris SL, Wei D, et al. Routes of transmission of influenza A H1N1, SARS-CoV, and norovirus in air cabin: comparative analyses. Indoor Air 2018; 28: 394-403.

29. Jason A Tetro, Is COVID-19 Receiving ADE From Other Coronaviruses? Microbes Infect. 2020 Mar;22(2):72-73.

30. Guo YR, Cao QD, Hong ZS, Tan YY, Chen S-D, Jin H-J, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. Mil Med Res. 2020;7(1):11. Published 2020 Mar 13.

31. Sun L, Song F, Shi N, et al. Combination of four clinical indicators predicts the severe/critical symptom of patients infected COVID-19. J Clin Virol.

2020;128:104431.

32. Zhang H, Zhou P, Wei Y, Yue H, Wang Y, Hu M, et al. Histopathologic Changes and SARS-CoV-2 Immunostaining in the Lung of a Patient With COVID-19. Ann Intern Med. 2020;172(9):629-632.

33. National Health Commission of the People's Republic of China home page (http://www .nhc .gov .cn).

34. Wang Z, Yang B, Li Q, Wen L, Zhang R. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497-506.

Figure legend

Figure 1. ROC curve of the postponed viral shedding discriminant model for training set (a) and testing set (b)

Figure 2. ROC curve of the disease progression discriminant model for training set (a) and testing set (b)

Figures
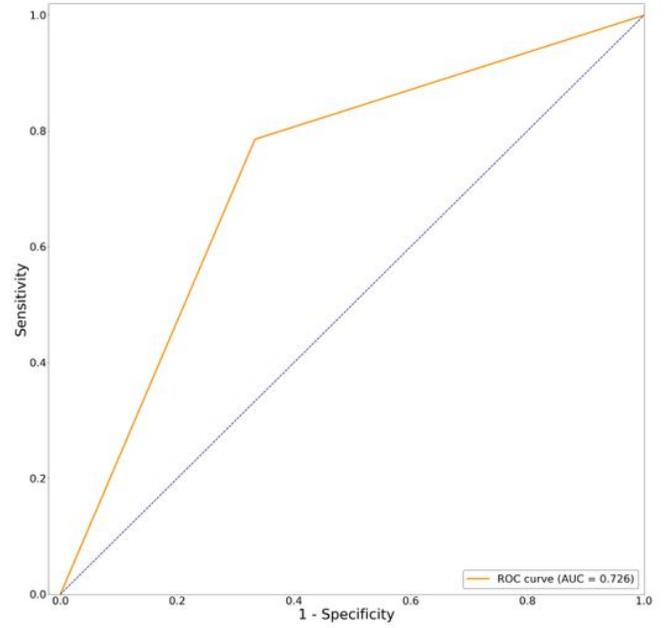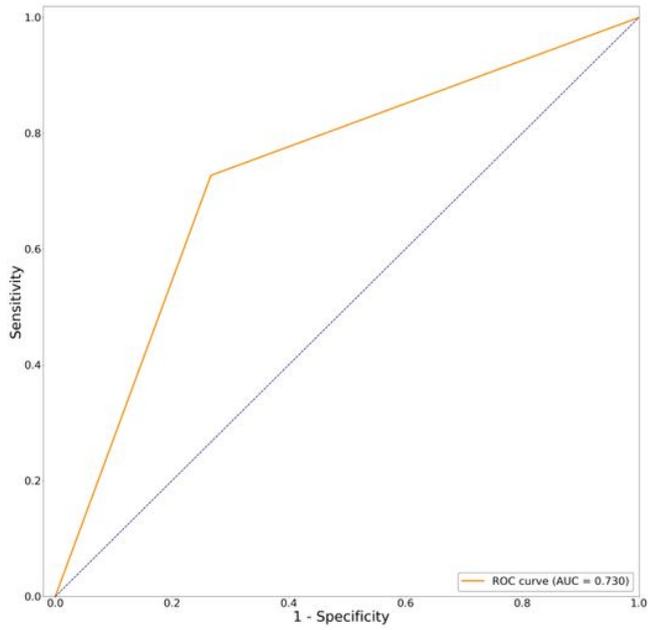
Figure 1

ROC curve of the postponed viral shedding discriminant model for training set (a) and testing set (b)
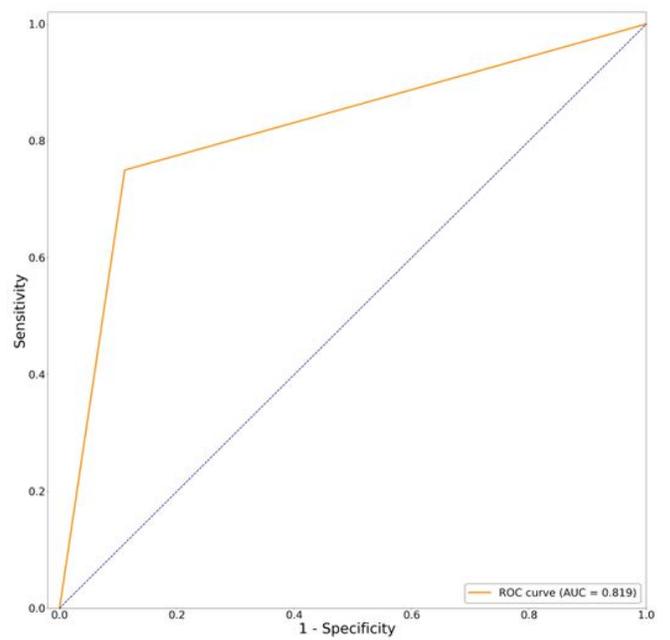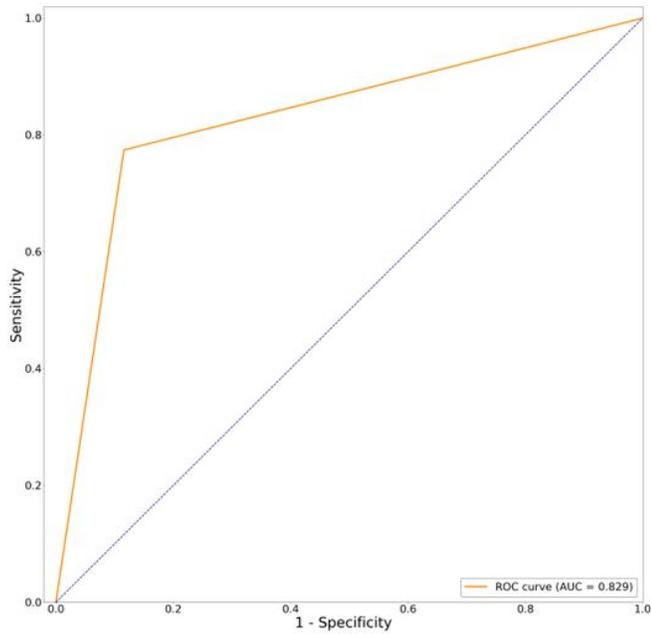

Figure 2

ROC curve of the disease progression discriminant model for training set (a) and testing set (b)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BMCTables.pdf