Performances of Clinical Characteristics and Radiological Findings in Identifying COVID- 19 From Suspected Cases

Xuanxuan Li
  Huashan Hospital, Fudan University
Yajing Zhao
  Huashan Hospital, Fudan University
Yiping Lu
  Huashan Hospital, Fudan University
Yingyan Zheng
  Huashan Hospital, Fudan University
Nan Mei
  Huashan Hospital, Fudan University
Qiuyue Han
  Huashan Hospital, Fudan University
Zhuoying Ruan
  Huashan Hospital, Fudan University
Anling Xiao
  Fu Yang No.2 People`s Hospital
Xiaohui Qiu
  Bozhou People`s Hospital
Dongdong Wang
  Huashan Hospital, Fudan University
Bo Yin (✉ yinbo@fudan.edu.cn)
  Huashan Hospital, Fudan University

Abstract

Background: To identify effective factors and establish a model to distinguish COVID-19 patients from suspected cases.

Methods: The clinical characteristics, laboratory results and initial chest CT findings of suspected COVID-19 patients in 3 institutions were retrospectively reviewed. Univariate and multivariate logistic regression were performed to identify significant features. A nomogram was constructed, with calibration validated internally and externally.

Results: 239 patients from 2 institutions were enrolled in the primary cohort including 157 COVID-19 and 82 non-COVID-19 patients. 11 features were included for multivariate logistic regression analysis after LASSO selection. We found that the COVID-19 group are more likely to have fever (OR, 4.22), contact history (OR, 284.73), lower WBC count (OR, 0.63), left lower lobe involvement (OR, 9.42), multifocal lesions (OR, 8.98), pleual thickening (OR, 5.59), peripheral distribution (OR, 0.09), and less mediastinal lymphadenopathy (OR, 0.037). The nomogram developed accordingly for clinical practice showed satisfactory internal and external validation.

Conclusions: In conclusion, fever, contact history, decreased WBC count, left lower lobe involvement, pleural thickening, multifocal lesions, peripheral distribution and absence of mediastinal lymphadenopathy are able to distinguish COVID-19 patients from other suspected patients. The corresponding nomogram is a useful tool in clinical practice.

1. Introduction:

In December 2019, a few pneumonia cases of unknown etiology were reported in Wuhan, Hubei Province, China.(1) The disease, now named coronavirus disease 2019 (COVID-19) then spread at a striking speed worldwide. The causative organism was identified as a novel coronavirus named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) due to the phylogenetic similarity to SARS-CoV.(2) As of Oct 15th, there were a total of 238,940,176 cumulative cases and 4,882,066 cumulative deaths worldwide. COVID-19 was declared as a public health emergency of international concern (PHEIC) by the World Health Organization (WHO) as early as January 30, 2020.(3)(4)

The confirmation of COVID-19 relies on the positive result of the nucleic acid amplification test (NAAT) of the respiratory tract or blood specimens using the real-time reverse transcriptase–polymerase chain reaction (RT-PCR) tests.(5) However, the limitations of RT-PCR tests include: 1) The severity and progression of the disease cannot be quantitatively judged. 2) They have long turnaround times, especially in less developed regions; 3) They require certified laboratories, expensive equipments and trained technicians. (6, 7)

Chest CT scan is relatively easy to perform with fast diagnosis high in sensitivity, thus is considered an ideal primary tool for COVID-19 detection.(8–10) The Diagnosis and Treatment Program of COVID-19 (trail version 7)(11, 12) fomulated by the National Health Commission of China has summarized the typical CT manifestations of COVID-19 as follows and incorporated it in the diagnosis criteria: multiple small patchy shadows and interstitial changes, notably at the peripheral zone, at the early stage. As the disease progresses, multiple ground-glass opacities (GGO) and infiltration occur bilaterally, and consolidation is found in severe cases. Pleural effusion is rarely seen.

Patients with above-said CT manifestations are suspected as COVID-19 infectors therefore need further examinations. Before the RT-PCR result is available, the patient needs isolation, but the quarantine of the patients may lead to a waste of medical resources and a possible delay of essential treatment. Hence, effective and convenient methods to better distinguish COVID-19 patients are needed.

The aim of our study is to identify the useful clinical, laboratory and radiographic features that are able to distinguish COVID-19 patients from other suspected cases and generate a nomogram as a useful tool for clinical practice.

2. Materials And Methods:

The schematic workflow is depicted in Figure 1.

2.1. Patient cohort:

Data were de-identified to guarantee the patients' confidentiality. From Jan 21 to Mar 5, 2020, patients admitted to a hospital in Anhui province, China and our institution in Shanghai, China who met the following requirements were enrolled as the primary cohort in our study: 1) Patients with chest CT manifestations suggested by the Diagnosis and Treatment Program of COVID-19 (trail version 7) (12) that had a suspicion of COVID-19. 2) Patients that took laboratory examination at admission. 3) Patients with the diagnosis of COVID-19 ruled out or confirmed by the RT-PCR. Exclusion criteria included: 1) Patients who were hospitalized before (n=4). 2) Significant motion artefacts in CT images (n=12). 3) Patients lacking essential data (n=21). The epidemiological history, the symptoms, the laboratory test results and the imaging features of their first CT scan after onset were recorded.

From Feb 6 to Mar 13, 2020, an independent cohort of CT-suspected patients from another institution in Anhui Province was prospectively studied, using the same inclusion and exclusion criteria. These patients formed the validation cohort.

The laboratory tests were carried out in the outpatient department or in the wards on admission, mostly on the same day when CT scan was done. Collected laboratory indices included the white blood cell (WBC) count, lymphocyte count, lactate dehydrogenase (LDH), C reactive protein (CRP), procalcitonin (PCT), alanine aminotransferase (ALT), aspartate aminotransferase (AST).

2.2. CT protocol

105 patients from Huashan Hospital Affiliated to Fudan University were imaged with 1.5mm-thickness with a 256-slice spiral CT scanner (Philips). 134 patients from Fuyang No.2 People

$sHosπtalwereima \geq dwith1mm\text{-}thick \neq sswitha64\text{-}sectionCTsca \cap er(Aquilion64, ToshibaMedicalSystems). 59patientsomBozhouPeop \leq$ s Hospital in the validation cohort were imaged with 5mm-thickness with a 64-section CT scanner (Siemens Somantom Sensation).

2.3. CT manifestation analysis

All imaging data were analyzed with consensus by two experienced radiologists (12 and 7 years of experience). 23 features were collected as listed below: a) The involved pulmonary lobes including five features: right upper, right middle, right lower, left upper, left lower lobes; b) Main distribution of lesions including anterior and posterior part of lungs; c) The location of lesions that is set as dummy variables: peripheral, central or both; d) The extent of the lesions that is set as dummy variables: unifocal, multifocal and diffuse. e) An extent score was semi-quantitatively calculated. Both lungs were divided into upper (above tracheal carina), lower (below inferior pulmonary vein) and middle (in between) zones, and involved percentage in each zone was scored: 0, 0%; 1, < 25%;2, 25% - 49%; 3, 50% - 74%; 4, > 75%, and they added up to the extent score(range 0-24). f) The existence of opacification set as dummy variables included GGO, mixed (mainly GGO), mixed (mainly consolidation) and consolidation; g) The shape of the lesions, including nodular, linear, patchy and large patchy; h) The halo sign; i) The reversed halo sign ; j) Reticulated changes; k) The existence of vascular enlargement; l) The existence of air bronchogram; m) Bronchiectasis; n) Pleural thickening; o) Pleural traction; p) Pleural effusion; q) Mediastinal lymphadenopathy (the maximal axial diameter >1 cm). The description of the radiological features followed the definition compiled by the Fleischner Society.(13) Five 1cm^2 regions of interests (ROI) were drawn in the liver and spleen parenchymal to obtain the mean CT values of liver and spleen. Liver spleen ratio (LS ratio) was calculated as CT_{liver}/CT_{spleen} to indicate the relative density.

2.4. Feature selection

The clinical (8), laboratory (7) and CT features (23) were analyzed altogether, but with the limited sample size, a total of 38 features would lead to overfitting in multivariate analysis. Thus, the least absolute shrinkage and selection operator (LASSO) method was adopted to select the most relevant features. This method is able to shrink the coefficients and diminish some to zero, thus can be used for feature reduction and selection. The R software and the "glmnet" package (version 3.6.0; R foundation for Statistical computing) were used.

2.5. Statistical analysis

All statistical analyses were executed with R software. The Shapiro-Wilk test was used to evaluate the distribution type and Bartlett`s test was used to evaluate the homogeneity of variance. Normally distributed data were displayed as mean ± standard deviation. Non-normally distributed data and ordinal data were displayed as median (inter-quartile range). Categorical variables were summarized as counts and percentages. Both univariate and multivariate logistic regression were analyzed to demonstrate the correlation of the features with COVID-19 diagnosis. The regression coefficient (β) was calculated using the odds ratio (OR). The model was estimated as follows:

$$\beta = \log(OR)$$
$$\mathrm{logitP} = \beta_1\chi_1 + \beta_2\chi_2 + \cdots + \beta_i\chi_i$$

A nomogram was established. The calibration ability was internally assessed with the bootstraping method and the Hosmer-Lemeshow test (HL test) was performed to test the goodness of fit.

For the external validation of the nomogram, the prediction value of each case were calculated according to the nomogram and compared with the observed diagnosis. The accuracy was validated by correctly predicted case proportion, and the HL goodness-of-fit test. A *P*-value of <0.05 was defined as statistical significance.

2.6. IRB approval

This multi-center retrospective study was approved by the institutional review board (IRB) and the requirement of written informed consent was waived.

3. Results:

3.1 Clinical information:

In the primary cohort, 239 patients with COVID-19 (134 males and 105 females) were included in this study with an average age of 46.31±15.90 years old. 28.87% of the patients had a direct contact with confirmed COVID-19 patients before the onset or had travelled/lived in the Hubei Province. 17.57% of the patients had indirect contact. Most common symptoms the patients presented were fever (70.29%), cough (44.35%), and chest distress (11.30%). Some patients had digestive symptoms such as diarrhea 2.09%) and anorexia (2.09%). The median interval between the onset and the date of CT scan was 8 (range 1-22) days (Table 1). 157 patients were confirmed as COVID-19 by RT-PCR and were allocated to the COVID-19 group. They were put in quarantine and treated with the antiviral therapy based on the evolving recommendations.(12) The other 82 patients had negative RT-PCR results. They were eventually diagnosed as other viral pneumonia, bacterial infection, or other respiratory conditions. Clinical information of two groups were compared using univariate analysis (Table 3). COVID-19 patients were found to be younger (*P* = 0.037), more likely to have fever (*P* = 0.001) or cough (*P* <0.001), and more likely to have contact history (*P* < 0.001).

Clinical Characteristics	Primary cohort (n = 239)	Validation cohort (n = 59)	*P* value
Age, mean ± SD	46.30±15.90	45.64±16.57	0.614
Gender			
Male	134 (56.07%)	31 (52.54%)	0.733
Female	105 (43.93%)	28 (47.46%)	
Epidemiological history			
Direct contact	69 (19.12%)	19 (32.20%)	0.546
Indirect contact	42 (48.53%)	13 (22.03%)	
None contact	128 (32.35%)	27 (45.76%)	
Symptom			
Fever	168 (70.29%)	47 (79.66%)	0.202
Cough	106 (44.35%)	31 (52.54%)	0.097
Chest distress	27 (11.30%)	6 (10.17%)	0.988
Diarrhea	5 (2.09%)	5 (8.47%)	0.042*
Anorexia	5 (2.09%)	1 (1.69%)	1.000
Laboratory Test, median (inter-quartile range)			
WBC, median (range), ×10^9/L	5.28 (4.30-10.44)	5.96 (3.91-6.00)	0.101
Lymphocyte count, median (range), ×10^9/L	1.19 (0.90-1.63)	1.21 (0.85-1.44)	0.746
LDH, median (range), U/L	233.00 (193.00-271.40)	234 (199-290)	0.158
CRP, median (range), mg/L	14.80 (4.8-42.93)	25.90 (3.7-30.30)	0.038*
PCT, median (range), ng/mL	0.05 (0-0.19)	0.04 (0.02-0.06)	0.743
ALT, median (range), U/L	30.00 (20.00-51.50)	29.90 (17.30-37.70)	0.558
AST, median (range), U/L	28.00 (21.00-46.75)	28.00 (20.40-34.70)	0.450

Abbreviations: WBC: White blood cell count; LDH: Lactate dehydrogenase; CRP: C-reactive protein; PCT: Procalcitonin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase

3.2 Laboratory Tests

Compared with the non-COVID-19 group, COVID-19 group showed lower WBC (*P* < 0.001) and lymphocyte count(*P* = 0.002), lower levels of PCT(*P* = 0.002). (Table 3)

3.3 Chest CT imaging findings

Imaging characteristics were assessed and compared between two groups (Table 2, 3). Regarding the location and the distribution of the lesions, COVID-19 patients were found to be more located in posterior part of the lungs (P < 0.001) compared with non-COVID-19 patients. They had more involvement in every lobe of the lung (P < 0.05) due to more multifocal distribution (P < 0.001). Besides, they are more likely to have specific signs including reticular changes (*P* = 0.04), vascular enlargement (*P* < 0.001), air bronchogram (*P* = 0.043), and pleural thickening (*P* < 0.001). They are less likely to show pleural effusion (OR = 0.16, *P* = 0.007) or mediastinal lymphadenopathy (*P* < 0.001). Other parameters were not significantly different.

Table 2
Imaging manifestations on chest CT of the primary and validation cohort

Imaging manifestation	Primary cohort (n = 239)	Validation cohort (n = 59)	P value
Involved lobes:			
Right Upper Lobe	144 (60.25%)	39 (66.1%)	0.498
Right Middle Lobe	129 (53.97%)	32 (54.24%)	1.000
Right Lower Lobe	179 (74.9%)	39 (66.1%)	0.230
Left Upper Lobe	143 (59.83%)	40 (67.8%)	0.329
Left Lower Lobe	176 (73.64%)	44 (74.58%)	1.000
Main distribution:			
Anterior Part of Lungs	44 (18.41%)	18 (30.51%)	0.061
Posterior Part of Lungs	168 (70.29%)	40 (67.8%)	0.847
Location of lesions:			
Peripheral	158 (66.11%)	33 (55.93%)	0.191
Central	16 (6.69%)	2 (3.39%)	0.516
Both	65 (27.2%)	24(40.68%)	0.482
Extent of lesions:			
Unifocal	58 (24.27%)	16 (27.12%)	0.775
Multi-focal	141 (59%)	26 (44.07%)	0.055
Diffuse	40 (16.74%)	17 (28.82%)	0.971
Extent score:	4 (2-5)	5 (3-7)	0.057
Density of lesions:			
GGO	77 (32.22%)	11(18.64%)	
Mixed (Mainly GGO)	98 (41.00%)	27 (45.76%)	0.606
Mixed (Mainly Consolidation)	57 (23.85%)	20 (33.9%)	0.158
Consolidation	7 (2.93%)	1(1.69%)	0.940
Shape of lesions:			
Nodular	1 (0.42%)	1 (1.69%)	0.853
Linear	5 (2.09%)	3 (5.08%)	0.410
Patchy	161 (67.6%)	41 (69.49%)	0.875
Large patchy	72 (30.13%)	14 (23.73%)	
Halo sign	67 (28.03%)	22 (37.29%)	0.218
Reverse halo sign	11 (4.60%)	2 (3.39%)	0.958
Reticulation	61 (25.52%)	11 (18.64%)	0.349
Air bronchogram	85 (35.56%)	26 (44.07%)	0.289
Bronchiectasis	25 (10.46%)	2 (3.39%)	0.150
Vascular enlargement	82 (34.31%)	21 (35.59%)	0.974
Pleural thickening	101 (42.26%)	27 (45.76%)	0.734
Pleural traction	60 (25.10%)	15 (25.42%)	1.000
Pleural effusion	12 (5.02%)	6 (10.17%)	0.237
Mediastinal Lymphadenopathy	23 (9.62%)	7 (11.86%)	0.787
Liver-spleen CT value ratio	1.17 (1.05-1.27)	1.19 (1.07-1.37)	0.278

Abbreviations: GGO: Ground-glass opacities

Table 3
Univariate logistic regression analysis of features for differentiating COVID-19 patients and non-COVID patients in Primary cohort

Features	Non-COVID-19 (n=82)	COVID-19 (n=157)	Coefficient	OR	P value
Clinical characteristics					
Age, mean ± SD	49.29±17.49	44.75±14.82	-0.02	0.98	0.037*
Gender, male/female	50/32	84/73	-0.31	0.74	0.270
Epidemiological history#					
Direct contact	1 (1.22%)	68 (43.31%)	4.13	61.89	<0.001*
Indirect contact	3 (3.66%)	39 (24.84%)	2.16	8.70	<0.001*
None contact	78 (95.12%)	50 (31.85%)	-3.73	0.02	<0.001*
Symptom					
Fever	42 (51.22%)	126 (80.25%)	1.35	3.87	<0.001*
Cough	24 (29.27%)	82 (52.23%)	0.47	1.60	0.084
Chest distress	9 (10.98%)	18 (11.46%)	0.05	1.05	0.910
Diarrhea	1 (1.22%)	4 (2.55%)	0.75	2.12	0.505
Anorexia	1 (1.22%)	5 (2.55%)	0.75	2.12	0.505
Laboratory Test, median (range)					
WBC, ×10^9/L	8.72±4.15	5.068±1.80	-0.54	0.58	<0.001*
Lymphocyte count, ×10^9/L	1.42±0.68	1.18±0.47	-0.77	0.46	0.002*
LDH, U/L	231.78±109.50	250.66±72.02	0.003	1.00	0.114
CRP, mg/L	31.08±40.56	23.06±29.40	-0.01	0.99	0.089
PCT, ng/mL	0.91±4.28	0.07±0.13	-3.56	0.03	0.002*
ALT, U/L	47.80±32.60	38.51±61.19	-0.003	1.00	0.226
AST, U/L	44.95±40.05	34.38±43.01	-0.01	0.99	0.091
Imaging manifestation					
Involved lobes:					
Right Upper Lobe	35 (42.68%)	109 (69.43%)	1.12	3.05	<0.001*
Right Middle Lobe	36 (43.90%)	93 (59.24%)	0.62	1.86	0.025*
Right Lower Lobe	48 (58.54%)	131 (83.44%)	1.27	3.57	<0.001*
Left Upper Lobe	36 (43.90%)	107 (68.15%)	1.01	2.73	0.001*
Left Lower Lobe	42 (52.44%)	123 (84.71%)	1.62	5.03	<0.001*
Main distribution:					
Anterior Part of Lungs	19 (23.17%)	25 (15.92%)	-0.47	0.63	0.172
Posterior Part of Lungs	45 (54.88%)	123 (78.34%)	1.06	2.88	<0.001*
Location of lesions: #					
Peripheral	49 (59.76%)	109 (69.43%)	0.43	1.53	0.135
Central	12 (14.63%)	4 (2.55%)	-1.88	0.15	0.002*
Both	21 (25.61%)	44 (28.02%)	0.12	1.13	0.690
Extent of lesions: #					

* *P* value < 0.05 indicates statistical significance,

Set as dummy variables in feature selection and Logistic model analysis

Abbreviations: WBC: White blood cell count; LDH: Lactate dehydrogenase; CRP: C-reactive protein; PCT: Procalcitonin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase

Features	Non-COVID-19 (n=82)	COVID-19 (n=157)	Coefficient	OR	*P* value
Unifocal	41 (50.00%)	17 (10.83%)	-2.11	0.12	<0.001*
Multi-focal	28 (34.15%)	113 (71.97%)	1.60	4.95	<0.001*
Diffuse	13 (15.85%)	27 (17.20%)	0.10	1.10	0.792
Extent score:	4.41±5.32	5.48±3.59	0.07	1.07	0.072
Density of lesions: #					
GGO	35 (42.68%)	42 (26.75%)	-0.71	0.49	0.013*
Mixed (Mainly GGO)	26 (31.70%)	72 (45.86(%)	0.60	1.82	0.036*
Mixed (Mainly Consolidation)	18 (21.95%)	39 (24.84%)	0.16	1.18	0.619
Consolidation	3 (3.66%)	4 (2.54%)	-0.37	0.69	0.631
Shape of lesions: #					
Nodular	0 (0%)	1 (0.63%)	13.92	1113402.31	0.987
Linear	0 (0%)	5 (3.18%)	14.95	3106188.55	0.982
Patchy	56 (68.29%)	106 (66.88%)	-0.07	0.94	0.825
Large patchy	26 (31.71%)	46 (29.30%)	-0.11	0.89	0.700
Halo sign	22 (26.83%)	45 (28.66%)	0.09	1.10	0.765
Reverse halo sign	2 (2.44%)	9 (5.73%)	0.89	2.43	0.263
Reticulation	11 (13.41%)	50 (31.85%)	1.10	3.02	0.003*
Air bronchogram	22 (26.83%)	63 (31.85%)	0.60	1.83	0.043*
Bronchiectasis	8 (9.76%)	17 (10.83%)	0.12	1.12	0.797
Vascular enlargement	14 (17.07%)	68 (43.31%)	1.31	3.71	<0.001*
Pleural thickening	17 (20.73%)	84 (53.50%)	1.48	4.40	<0.001*
Pleural traction	16 (19.51%)	44 (28.03%)	0.47	1.61	0.152
Pleural effusion	9 (10.98%)	3 (1.91%)	-1.85	0.16	0.007*
Mediastinal Lymphadenopathy	20 (24.39%)	3 (1.91%)	-2.81	0.06	<0.001*
Liver-spleen CT value ratio	1.18(1.02-1.29)	1.17(1.06-1.35)	0.11	1.12	0.826
* *P* value < 0.05 indicates statistical significance,					
# Set as dummy variables in feature selection and Logistic model analysis					
Abbreviations: WBC: White blood cell count; LDH: Lactate dehydrogenase; CRP: C-reactive protein; PCT: Procalcitonin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase					

3.4. Feature selection

In LASSO model, the λ value of 0.0376 with log (λ) of -3.280 chosen (1-SE criteria), and a total of 38 features were reduced to 11 potential features with nonzero coefficients on the basis of 239 patients (21.7:1 ratio; Figure 2). These features were further incorporated in the multivariate logistic analysis (Table 4). Eight features were found to be statistically significant. COVID-19 group tended to have more fever (OR, 4.22; 95% CI [confidence interval], 1.09-18.63; *P* = 0.043), less probability of no contact history (meaning higher probability of indirect or direct contact history [OR, 284.73; 95% CI, 38.17-4214.18; *P* < 0.001]), lower WBC count (OR, 0.63; 95% CI, 0.47-0.77; *P* < 0.001), more involving left lower lobe (OR, 9.42; 95% CI, 1.95-62.80; *P* = 0.010), more exhibiting multifocal lesions (OR, 8.98; 95%CI, 1.58-61.36; *P* = 0.017), more pleural thickening (OR, 5.59; 95%CI, 1.32-28.85; *P* = 0.026), less located in central part (OR, 0.09; 95%CI, 0.01-0.75; *P* = 0.043), and less mediastinal lymphadenopathy (OR, 0.037; 95% CI, 0.00-0.29; *P* = 0.004).

Table 4
Multivariate logistic regression analysis of features for differentiating COVID-19 patients
and non-COVID patients

Features	Coefficient	OR	95%CI	P value
Fever	1.44	4.22	(1.09,18.63)	0.043*
Epidemiological history: None contact	-5.65	0.00	(0.00,0.03)	<0.001*
WBC count	-0.47	0.63	(0.48,0.77)	<0.001*
Lesion involvement: Unifocal	0.11	1.12	(0.12,10.58)	0.919
Lesion involvement: Multi-focal	2.19	8.98	(1.59,61.36)	0.017*
Involved lobes: Right Upper lobe	1.12	3.05	(0.75,13.21)	0.121
Involved lobes: Left Upper Lobe	0.77	2.16	(0.51,9.52)	0.295
Involved lobes: Left Lower Lobe	2.24	9.42	(1.95,62.80)	0.010*
Pleural thickening	1.72	5.59	(1.32,28.85)	0.026*
Mediastinal lymphadenopathy	-3.30	0.04	(0.00,0.29)	0.004*
Distribution Central	-2.45	0.09	(0.01,0.75)	0.043*
* P value < 0.05 indicates statistical significance,				
Abbreviations: WBC: White blood cell count.				

3.5. Nomogram

A nomogram was constructed based on the multivariate Logistic analysis model. The adjusted C-index of the nomogram was 0.97 (Figure 3A). The calibration curve was determined with bootstrap analysis to get bias-corrected estimation. It indicated great agreement between the prediction and the actual diagnosis in the probability (Figure 3B). The HL goodness-of-fit test showed good calibration as well (P = 0.4797). The CT images of two cases illustrated the application of the nomogram. (Figure 4)

3.6. External validation

The validation cohort included 59 cases with 43 COVID-19 and 16 non-COVID. The baseline data were collected in Table 1 and Table 2. 56 out of 59 cases were correctly predicted using the nomogram, reaching an accuracy of 94.91%. Calibration was good (P = 0.9956 for the HL goodness-of-fit test).

4. Discussion:

An ongoing outbreak of COVID-19 originated from Hubei province in China has been spreading worldwide. Experts in infectious and respiratory diseases, critical care, and radiology from all over the world have been making a joint effort to contain the epidemic situation.[14] Presently, RT-PCR is the standard confirmative method in spite of a few flaws including long turnaround time for the results in underdeveloped regions and low sensitivity especially in the early phase of the disease.[14][15] On the contrary, chest CT scan is able to recognize the lesions at earlier stages with high sensitivity, thus is considered an important tool for screening.[8] Typical radiographic features on chest CT in COVID-19 patients were reported to predominantly include bilateral and peripheral GGOs and consolidative pulmonary opacities.[8, 16–18] Less typical signs included linear opacities, "crazy-paving" pattern and the reverse halo sign, etc.[8, 19–23] Before the RT-PCR results are attainable, the quarantine needed, but the isolation site is insufficient, and it possibly delays essential treatment. Thus, we investigated the differential values of clinical characteristics, laboratory results and CT features to better distinguish COVID-19 patients.

The most common symptom in the patients we enrolled is fever, followed by cough and chest distress. As a differential feature, fever is significant in both univariate and multivariate analysis. This echoes previous studies, and fever is the leading symptom listed in the case definition for surveillance of COVID-19 by the Chinese Health Commission.[12, 24, 25] Therefore, it is necessary to monitor body temperature and at-home temperature measurement is a useful and easy way for the public to early notice. Additionally, we noticed a small portion of the patients with digestive disorders like diarrhea and anorexia, and it occurred more in the COVID-19 group. Increasing evidence shows the manifestation of COVID-19 is not always confined to respiratory symptoms, but may also involve other systems, e.g., the central nervous system.[26][27]

The contact history is another valuable factor for COVID-19, including direct contact with COVID-19 patients, direct exposure in Hubei Province or other districts with confirmed cases, and indirect contact with those who were exposed.[28] According to the National Health Commission of China, a patient with one exposure or contact history and two clinical conditions can be regarded as a suspected case. [12] However, with the swift spread of the disease, some contact history is unrevealed, making it harder to contain the epidemic.[29] More active precaution and isolation is needed.

Among the laboratory parameters, WBC count is significantly lower in COVID-19 group in both univariate and multivariate analysis, and lymphocyte count is lower in univariate analysis. This is consistent with previous findings and the criteria by the Chinese Health Commission.[1, 10, 12] We also found lower levels of CRP and PCT in the COVID-19 group. They are useful indicators of infection or inflammation, and CRP was previously reported to increase in COVID-

19 patients by some researchers.(8)(30) Our finding may result from higher extent of increased levels of these indices in non-COVID-19 patients since they had other inflammatory conditions including bacterial infection, while other studies used healthy controls.

The location of the lesions varied among studies, yet the peripheral site is most frequently reported. In this study, COVID-19 lesions were less seen in the central lung compared with non-COVID-19 group in both univariate and multivariate analyses, and the parameter of the left lower lobe involvement is a significant distinct finally included in nomogram construction. Besides, compared with non-COVID-19 cases, COVID-19 is more likely to exhibit multifocal distribution rather than unifocal changes, and more likely to have reticulated changes, vascular enlargement, and pleural thickening. COVID-19 patients are also less likely to have pleural effusion and mediastinal lymphadenopathy, which is consistent with prior researches.(19)

Fever, contact history, decreased WBC count, left lower lobe location, pleural thickening, multifocal lesions, peripheral distribution, and absence of mediastinal lymphadenopathy were found to be features independently associated to COVID-19 patients. On the basis of these parameters, a nomogram was built to better interpret our findings, which is popular in cancer research these years.(31) According to our nomogram, the point of each feature adds up to a total score with a corresponding probability of COVID-19.

A nomogram can be validated by both internal and external validation.(32) In this study, internal validation used the data of the same cohort for the generation of the nomogram, and external validation used the data from another institution. Both internal and external validation indicated good agreement between the prediction and the actual diagnosis in the probability.

In summary, this study is the first to investigate the features to distinguish confirmed COVID-19 patients from CT-suspected cases, which is a critical challenge in clinical practice before RT-PCR results are available. The nomogram can be used as an instant tool able to provide practical reference for individualized management for every suspected patient and is likely to offer effective and scientific basis for empirical treatment.

Our study had several limitations. Firstly, in this multi-center study, the normal range and results of the laboratory data might be different due to the differences in the kits, equipment, and environmental conditions. However, three institutions are all China's Grade-A Tertiary Hospitals, with laboratories of the highest qualifications, and similar protocols are adhered, thus the results are relatively stable. Secondly, the sample size is relatively small since no data was obtained from the epicenter of the outbreak. Besides, despite being the standard confirmative test, RT-PCR has false-negative probabilities, therefore our results might be biased since non-COVID-19 group might include infected patients. Further prospective investigation with larger sample size and evolved diagnostic techniques is expected.

5. Conclusion:

In conclusion, fever, contact history, decreased WBC count, left lower lobe involvement, pleural thickening, multifocal lesions, peripheral distribution, and absence of mediastinal lymphadenopathy are able to distinguish COVID-19 patients from other suspected patients. The nomogram based on these features is a useful tool in the clinical practice.

Abbreviations:

COVID-19: Coronavirus Disease 2019

SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2

WHO: World Health Organization

PHEIC: Public Health Emergency of International Concern

NAAT: Nucleic Acid Amplification Test

RT-PCR: Reverse Transcriptase-polymerase Chain Reaction

CT: Computed Tomography

IRB: Institutional Review Board

WBT: White Blood Count

CPR: C-reactive protein

LDH: Lactate Dehydrogenase

PCT: Procalcitonin

GGO: Ground-glass Opacity

RHS: Reversed Halo Sign

ROI: Region of Interest

LASSO: Least Absolute Shrinkage and Selection Operator

OR: Odds Ratio

LR ratio: Liver/Spleen ratio

95% CI: 95% confidence interval

Declarations:

Ethics approval and consent to participate

This multi-center retrospective study was approved by the institutional review board (IRB) and the requirement of written informed consent was waived.

Consent for publication

All co-authors have read and approved of its submission to this journal.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to ethical restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

XL: drafted the work

YZ (Yajing Zhao): drafted the work

YL: substantively revised the work

YZ (Yingyan Zheng): analysis

NM: interpretation of data

QH: interpretation of data

ZR: analysis

AX: acquisition

XQ: acquisition

DW: design of the work

BY: conception

All authors have read and approved the manuscript

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics Accordance

This study was approved by the institutional review board (IRB), and the methods were carried out in accordance with the Declaration of Helsinki.

References:

1. Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 2020;395(10223):514–23. (DOI: 10.1016/S0140-6736(20)30154-9)

2. Munster VJ, Koopmans M, van Doremalen N, van Riel D, de Wit E. A Novel Coronavirus Emerging in China — Key Questions for Impact Assessment. N Engl J Med. 2020;2001017. (DOI: 10.1056/NEJMp2000929)

3. Guan W-J, Ni Z-Y, Hu Y, Liang W-H, Ou C-Q, He J-X, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;1–13. (DOI: 10.1056/NEJMoa2002032)

4. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). Int J Surg [Internet]. 2020;76(PG-):71–6. (DOI: 10.1016/j.ijsu.2020.02.034)

5. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance. 2020;25(3):1–8. (DOI: 10.2807/1560-7917.ES.2020.25.3.2000045)

6. Li Z, Yi Y, Luo X, Xiong N, Liu Y, Li S, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. J Med Virol [Internet]. 2020;92(9):1518–24. (DOI: 10.1002/jmv.25727)

7. Dai WC, Zhang HW, Yu J, Xu HJ, Chen H, Luo SP, et al. CT Imaging and Differential Diagnosis of COVID-19. Can Assoc Radiol J [Internet]. 2020;71(2):195–200. (DOI: 10.1177/0846537120913033)

8. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology. 2020 Feb;296(2):E32–40. (DOI: 10.1148/radiol.2020200642)

9. Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of chest CT for COVID-19: Comparison to RT-PCR. Radiology. 2020 Feb;296(2):E115–7. (DOI: 10.1148/radiol.2020200432)

10. Xiong Y, Sun D, Liu Y, Fan Y, Zhao L, Li X, et al. Clinical and High-Resolution CT Features of the COVID-19 Infection: Comparison of the Initial and Follow-up Changes. Invest Radiol. 2020;55(6):332–9. (DOI: 10.1097/RLI.0000000000000674)

11. WHO. Corona virus disease (COVID-19). World Health Organization. Advice for the Public. 2020. (http:// www .who .int)

12. Http://bgs.satcm.gov.cn/zhengcewenjian/2020-03-04/13594.html. General Office of National Health Committee. Office of State Administration of Traditional Chinese Medicine. Notice on the issuance of a program for the diagnosis and treatment of novel coronavirus (2019-nCoV) infected pneumonia (trial sixth edition)(2020 [Internet]. 2020. Available from: http://bgs.satcm.gov.cn/zhengcewenjian/2020-03-04/13594.html

13. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: Glossary of terms for thoracic imaging. Radiology. 2008;246(3):697–722. (DOI: 10.1148/radiol.2462070712)

14. Tan W, Zhao X, Ma X, Wang W, et al. A Novel Coronavirus Genome Identified in a Cluster of Pneumonia Cases — Wuhan, China 2019–2020. China CDC Wkly [Internet]. 2020;2(4):61–2. Available from: http://weekly.chinacdc.cn//article/id/a3907201-f64f-4154-a19e-4253b453d10c

15. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical 2019-nCoV Pneumonia: Relationship to Negative RT-PCR Testing. Radiology. 2020;200343. (DOI: 10.1148/radiol.2020200343)

16. Yang W, Cao Q, Qin L, Wang X, Cheng Z, Pan A, et al. Clinical characteristics and imaging manifestations of the 2019 novel coronavirus disease (COVID-19): A multi-center study in Wenzhou city, Zhejiang, China. J Infect. 2020 Apr;80(4):388-393. (DOI: 10.1016/j.jinf.2020.02.016)

17. Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, et al. CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). Radiology. 2020;200230. (DOI: : 10.1148/radiol.2020200230)

18. Pan Y, Guan H, Zhou S, Wang Y, Li Q, Zhu T, et al. Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan, China. Eur Radiol. 2020;30(6):3306–9. (DOI: 10.1007/s00330-020-06731-x)

19. Adam Bernheim. Chest CT findings in Coronavirus Disease-19: Relationship to Duration of Infection. Radiology. 2020;19. (DOI: 10.1148/radiol.2020200463)

20. Shi H, Han X, Jiang N, Cao Y, Alwalid O, Gu J, et al. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. Lancet Infect Dis. 2020;3099(20):1–10. (DOI: 10.1016/S1473-3099(20)30086-4)

21. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19) Pneumonia. Radiology. 2020;200370. (DOI: 10.1148/radiol.2020200370)

22. Xu X, Yu C, Qu J, Zhang L, Jiang S, Huang D, et al. Imaging and clinical features of patients with 2019 novel coronavirus SARS-CoV-2. Eur J Nucl Med Mol Imaging. 2020;(613):2–7. (DOI: 10.1007/s00259-020-04735-9)

23. Xu X, Yu C, Zhang L, Luo L, Liu J. Imaging features of 2019 novel coronavirus pneumonia. Eur J Nucl Med Mol Imaging. 2020;(613):1–2. (DOI: 10.1007/s00259-020-04735-9)

24. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. J Med Virol. 2020; (January):441–7. (DOI: 10.1002/jmv.25689)

25. Wu J, Liu J, Zhao X, Liu C, Wang W, Wang D, et al. Clinical Characteristics of Imported Cases of COVID-19 in Jiangsu Province: A Multicenter Descriptive Study. Clin Infect Dis. 2020; (DOI: 10.1093/cid/ciaa199)

26. Li YC, Bai WZ, Hashikawa T. The neuroinvasive potential of SARS-CoV2 may play a role in the respiratory failure of COVID-19 patients. J Med Virol. 2020 Jun;92(6):552-555. (DOI: 10.1002/jmv.25728)

27. Lu Y, Li X, Geng D, Mei N, Wu PY, Huang CC, et al. Cerebral Micro-Structural Changes in COVID-19 Patients – An MRI-based 3-month Follow-up Study: A brief title: Cerebral Changes in COVID-19. EClinicalMedicine. 2020;25(2). (DOI: 10.1016/j.eclinm.2020.100484)

28. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506. (DOI: 10.1016/j.virusres.2020.198043)

29. Zu ZY, Di Jiang M, Xu PP, Chen W, Ni QQ, Lu GM, et al. Coronavirus Disease 2019 (COVID-19): A Perspective from China. Radiology. 2020 Aug;296(2):E15-E25. (DOI: 10.1148/radiol.2020200490)

30. Li L quan, Huang T, Wang Y qing, Wang Z ping, Liang Y, Huang T bi, et al. COVID-19 patients' clinical characteristics, discharge rate, and fatality rate of meta-analysis. J Med Virol. 2020 Jun;92(6):577-583. (DOI: 10.1002/jmv.25757)

31. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol. 2016;34(18):2157−64. (DOI: 10.1200/JCO.2015.65.9128)

32. Lee SC, Chang MC. Development and validation of web-based nomograms to predict postoperative invasive component in ductal carcinoma in situ at needle breast biopsy. Healthc Inform Res. 2014;20(2):152−6. (DOI: 10.4258/hir.2014.20.2.152)
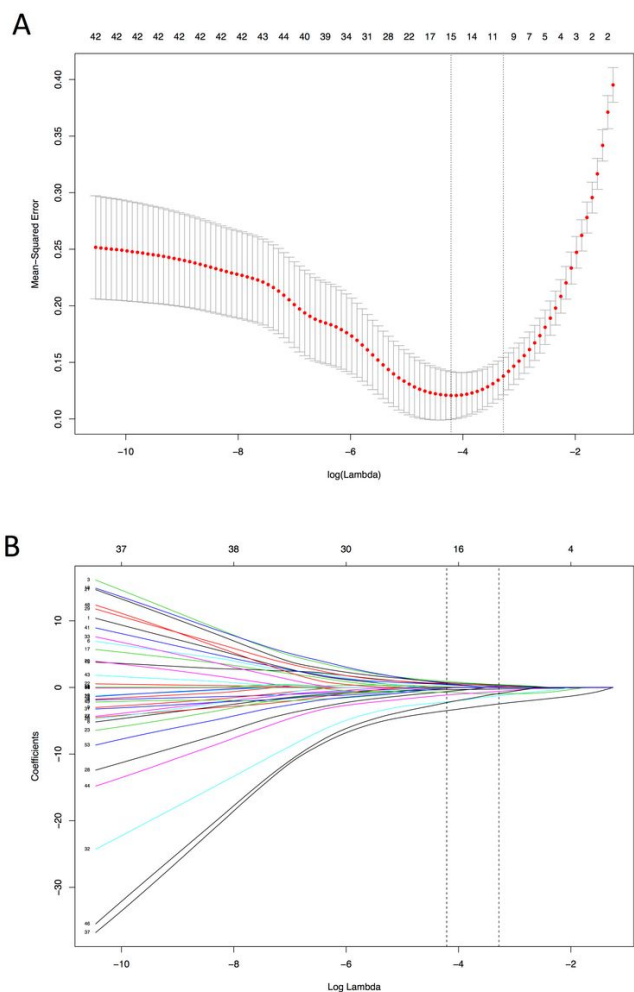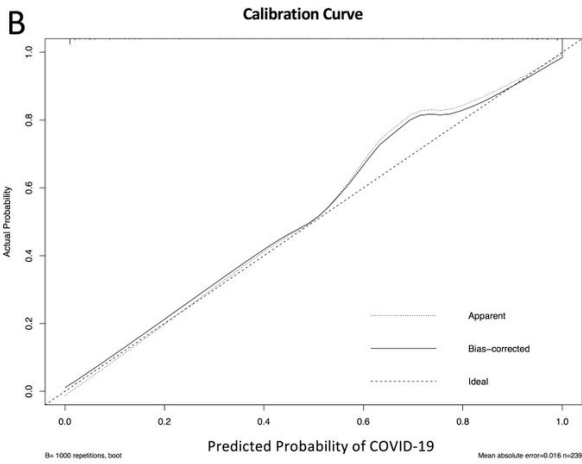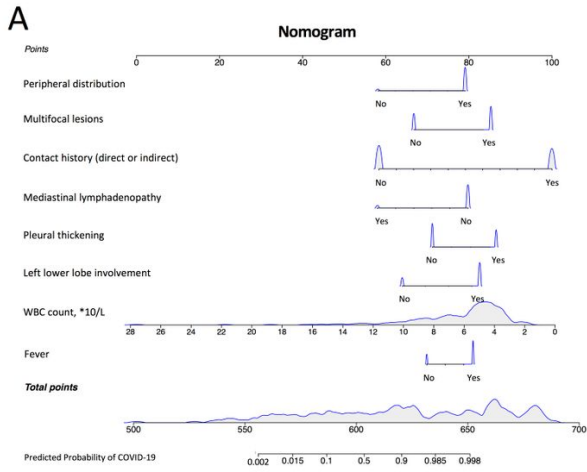
Figures
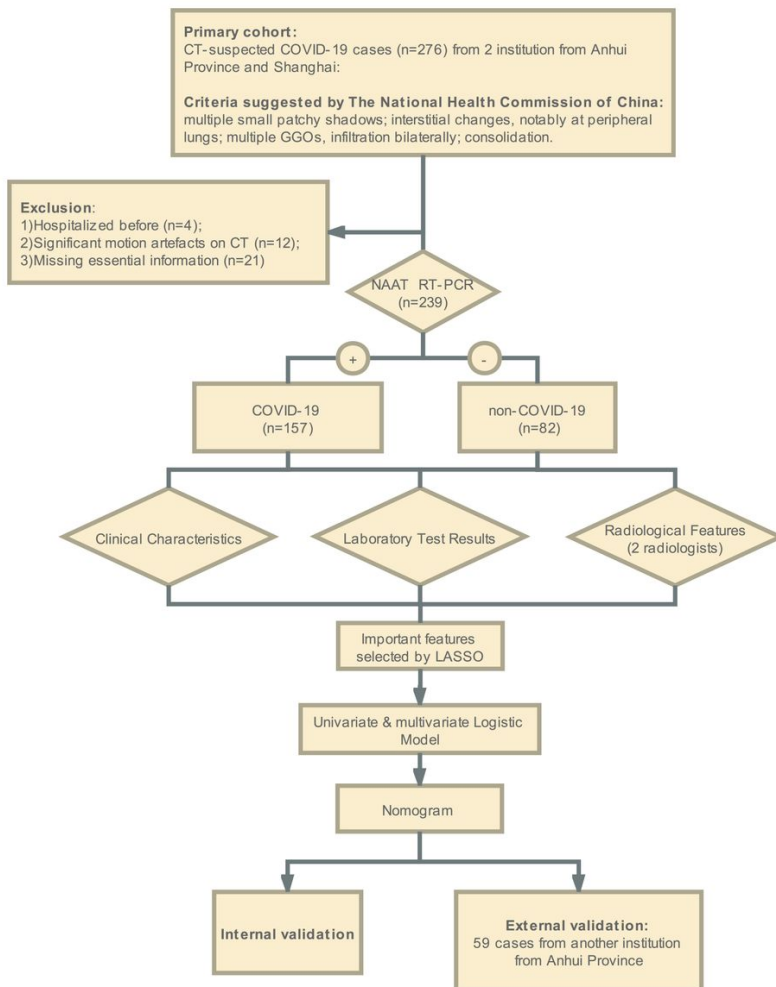
Figure 1

Workflow of the whole study.

Figure 2

Feature selection using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. (A) The parameter (λ) in the LASSO model used 10-fold cross-validation based on minimum criteria. The mean squared error was plotted versus log(λ). Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria). (B) The plot of LASSO coefficient profiles was produced against the log (λ) sequence. The dotted vertical line was drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria), and the latter was chosen with the λ value of 0.0376 and log (λ) of -3.280 according to the 10-fold cross-validation that resulted in 11 nonzero coefficients.

Figure 3

The nomogram and calibration curves based on significant features in multivariate analysis. (A) A nomogram was built on the basis of eight significant features in multivariate Logistic model. If a patient is suspected to be COVID-19 by radiological diagnosis, the data needed includes whether he has fever, contact history, decreased WBC count, left lower lobe involvement, pleural thickening, multifocal lesions, peripheral distribution or absence of mediastinal lymphadenopathy. The point of each feature adds up to a total score with a corresponding probability of COVID-19. (B) The calibration curve was determined with bootstrap analysis to get bias-corrected estimation. It indicated great agreement between the prediction and the actual grouping in the probability.
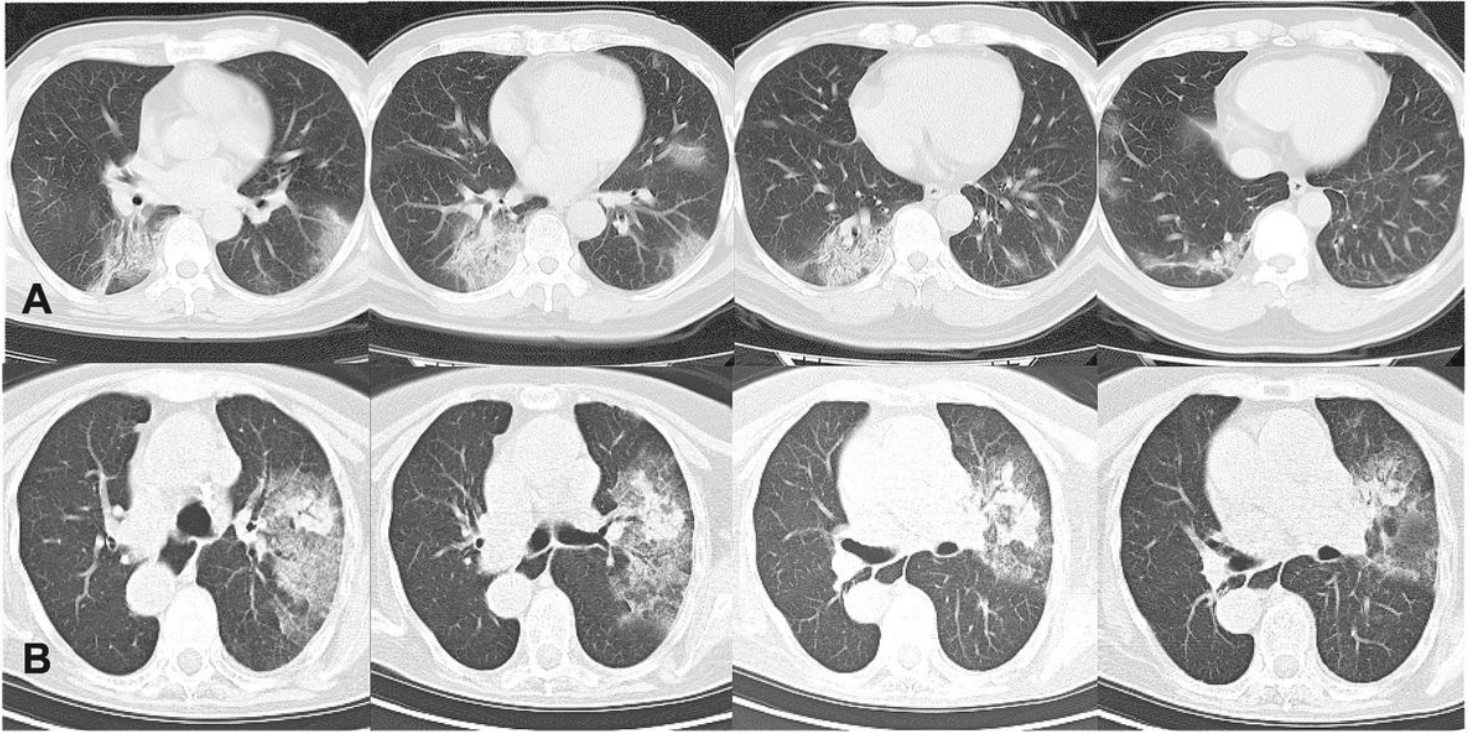
Figure 4

Two representative cases to illustrate the application of the nomogram. (A) A 40-year-old male patient complained of fever for 4 days (score ≈ 80). He had travelled to Huangshi, a city in Wuhan Province, China a week before the onset (score ≈ 100). His laboratory tests indicated leukocytopenia (1.99*10^9/L, score ≈ 92). His chest CT showed patchy ground glass opacities with vascular enlargement and reticular changes on bilateral lower lobes (left lower lobe involvement: score ≈ 83; multifocal: score ≈85). Lesions were located both central and peripheral (score ≈ 80). No mediastinal lymphadenopathy was observed (score ≈ 80). Slight pleural thickening was observed (score ≈ 85). Total estimated score reached around 687, indicating >99.8% probability to be a COVID-19 case. He was later confirmed by RT-PCR. (B) A 60-year-old female patient complained of fever for 3 days (score ≈ 80). She claimed no contact or exposure history (score ≈ 60). Her WBC count is slightly elevated (10.52*10^9/L, score ≈ 60). Her chest CT showed unifocal (score ≈ 68)large patchy ground glass opacities with consolidation only involving the left upper lobe (score ≈ 63), but with both central and peripheral distribution (score ≈ 80). Mediastinal lymphadenopathy was observed in mediastinal window (score ≈ 60). No pleural thickening(score ≈ 70) . Total estimated score reached around 541, indicating <0.2% probability to be a COVID-19 case. She was radiologically suspected as COVID-19, but the diagnosis of COVID-19 was ruled out by a negative RT-PCR. She was finally diagnosed with Respiratory syncytial virus infection.