



Development and multicenter validation of a CT-based radiomics signature for predicting severe COVID-19 pneumonia

Liang Li¹ · Li Wang² · Feifei Zeng¹ · Gongling Peng³ · Zan Ke¹ · Huan Liu⁴ · Yunfei Zha¹ 

Received: 13 August 2020 / Revised: 23 December 2020 / Accepted: 28 January 2021 / Published online: 30 March 2021
© European Society of Radiology 2021

Abstract

Objectives To develop and validate a radiomics nomogram for timely predicting severe COVID-19 pneumonia.

Materials and methods Three hundred and sixteen COVID-19 patients (246 non-severe and 70 severe) were retrospectively collected from two institutions and allocated to training, validation, and testing cohorts. Radiomics features were extracted from chest CT images. Radiomics signature was constructed based on reproducible features using the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm with 5-fold cross-validation. Logistic regression modeling was employed to build different models based on quantitative CT features, radiomics signature, clinical factors, and/or the former combined features. Nomogram performance for severe COVID-19 prediction was assessed with respect to calibration, discrimination, and clinical usefulness.

Results Sixteen selected features were used to build the radiomics signature. The CT-based radiomics model showed good calibration and discrimination in the training cohort (AUC, 0.9; 95% CI, 0.843–0.942), the validation cohort (AUC, 0.878; 95% CI, 0.796–0.958), and the testing cohort (AUC, 0.842; 95% CI, 0.761–0.922). The CT-based radiomics model showed better discrimination capability (all $p < 0.05$) compared with the clinical factors joint quantitative CT model (AUC, 0.781; 95% CI, 0.708–0.843) in the training cohort, the validation cohort (AUC, 0.814; 95% CI, 0.703–0.897), and the testing cohort (AUC, 0.696; 95% CI, 0.581–0.796). Decision curve analysis demonstrated that in terms of clinical usefulness, the radiomics model outperformed the clinical factors model and quantitative CT model alone.

Conclusions The CT-based radiomics signature shows favorable predictive efficacy for severe COVID-19, which might assist clinicians in tailoring precise therapy.

Key Points

- Radiomics can be applied in CT images of COVID-19 and radiomics signature was an independent predictor of severe COVID-19.
- CT-based radiomics model can predict severe COVID-19 with satisfactory accuracy compared with subjective CT findings and clinical factors.
- Radiomics nomogram integrated with the radiomics signature, subjective CT findings, and clinical factors can achieve better severity prediction with improved diagnostic performance.

Keywords COVID-19 · Nomograms · Pneumonia · Tomography, X-ray computed

Liang Li and Li Wang contributed equally to this work.

✉ Yunfei Zha
zhayunfei999@126.com

¹ Department of Radiology, Renmin Hospital of Wuhan University, Wuhan 430060, China

² Department of Infection Prevention and Control, Renmin Hospital of Wuhan University, Wuhan 430060, China

³ Department of Thyroid and Breast Surgery, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430014, China

⁴ GE Healthcare, Shanghai 201203, China

Abbreviations

COVID-19	Corona Virus Disease 2019
CTSS	CT severity score
CTLP	CT lesion percentage
DCA	Decision curve analysis
GGO	Ground-glass opacity
GLM	Generalized linear model
GLCM	Gray-level co-occurrence matrix
GLRLM	Gray-level run length matrix
ICC	Intraclass correlation coefficients
LASSO	Least absolute shrinkage selection operator

NGTDM	Neighboring gray tone difference matrix
NSD	Non-severe disease
ROC	Receiver operating characteristic
SD	Severe disease

Introduction

Corona Virus Disease 2019 (COVID-19) has spread out in the world, posing a critical threat to public health [1]. Initial report of the disease indicated that from November 16, 2019, to November 5, 2020, there were 47,596,852 confirmed cases of COVID-19, including 1,216,357 deaths, with a global fatality of approximately 2.56% [2].

According to the “Diagnosis and Treatment Program of Pneumonia of New Coronavirus Infection (Trial 7th Edition)” recommended by China’s National Health Commission, severe COVID-19 patients were more likely to develop poor clinical outcomes including acute respiratory distress syndrome (ARDS), acute cardiac/kidney injury, and multiple organ failure [3]. Current epidemiological data also suggests that the mortality rate of severe COVID-19 is about 20 times higher than that of non-severe COVID-19 [4]. Therefore, it is very crucial for early identification of severe cases, which prevent disease progression and reduce the mortality.

Computed tomography (CT) is widely used to diagnose, evaluate, and monitor COVID-19 pneumonia in high-risk areas [5]. However, the evaluation of these conventional CT features depends heavily on the radiologist’s experience and is non-objective. Furthermore, some scholars have tried using quantitative CT method to assess the severity of lung injury in COVID-19 pneumonia, such as visual CT severity score (CTSS) [6] and the CT lesion percentage (CTLP) [7] in each lung lobe, which were alternative ways to stratify COVID-2019 cases. The radiomics approach has drawn increased attention in recent years, which may help the detection, diagnosis, monitoring, and prognostic assessment of lung disease [8]. To our knowledge, there has been rarely radiomics study for the accurate prediction of the clinical severity of COVID-19 pneumonia to date.

Therefore, the aim of our study is to (1) investigate the role of radiomics features for predicting severe patients with COVID-19 pneumonia, and (2) examine whether the addition of quantitative CT characteristics and/or clinical factors can improve the performance of the predictive model.

Materials and methods

Patient population and groups

Our institutional review board approved this retrospective study, and the requirement for informed consent was waived.

We searched the electronic database at two institutions ((I) Renmin Hospital of Wuhan University, (II) Central Hospital of Wuhan optical valley) and retrospectively reviewed records for patients between January 30, 2020, and April 30, 2020, and identified 449 patients infected with COVID-19. All patients with COVID-19 were proven using real-time reverse transcriptase–polymerase chain reaction. The whole pipeline of our study was shown in Fig. 1.

According to the “Diagnosis and Treatment Program of Pneumonia of New Coronavirus Infection (Trial 7th Edition)” recommended by China’s National Health Commission, all patients are classified as having minimal, ordinary, severe, and critical type [9]. In our study, ordinary cases were included in the non-severe disease (NSD) group, while severe and critical cases were merged into the severe disease (SD) group. All patients in the SD group should meet any of the following criteria: (1) respiratory rate ≥ 30 breaths per minute; (2) finger of oxygen saturation $\leq 93\%$ in a resting state; (3) arterial oxygen tension (PaO_2)/inspiratory oxygen fraction (FiO_2) ≤ 300 mmHg (1 mmHg = 0.133 kPa); (4) respiratory failure occurred and mechanical ventilation required; (5) shock; (6) other organ failure needing intensive care unit (ICU) monitoring treatment.

Image acquisition and lesion segmentation

Non-enhanced chest CT scans of 316 patients were carried out from the lung apex to the lung base using multi-detector CT (MDCT) scanners (Brightspeed CT or Optima 680 CT, GE Healthcare) at the end of inspiration. Breath-hold training was carried out before each examination. Parameters for chest CT scanning were listed as follows: field of view (FOV), 36 cm; tube voltage, 120 kV; tube current adjusted automatically; noise index, 13; section thickness, 5 mm; slice interval, 5 mm; pitch, 1.375; collimation 64×0.625 mm; gantry rotation speed, 0.7 s; matrix, 512×512 ; the mediastinal window: window width of 200 HU with a window level of 35 HU, and the lung window: window width of 1500 HU with a window level of - 700 HU.

All images were segmented on the commercial segmentation software (Lung Intelligence Kit 2.1, LK 2.1, GE Healthcare) [10]. First, pre-processing was executed and included the following steps: resampling adjust the *x*-spacing, *y*-spacing, *z*-spacing size (spatial resolution = 1 mm \times 1 mm \times 1 mm). Gaussian filter with a standard deviation of 0.5 was applied for signal smoothing. Then, the lung was automatically segmented into five lobes; the volumes of interest (VOIs) were automatically contoured for each lobe. The segmentation results are manually corrected by a radiologist (Z.K., 2 years of experience in radiology) and then confirmed by another radiologist (F.Z., 5 years of experience in radiology). CTLP was defined as the volume of the lesions (including ground-glass opacity (GGO), consolidation, and reticulation) divided by the volume of the entire lung. CTSS was used to estimate pulmonary involvement of all

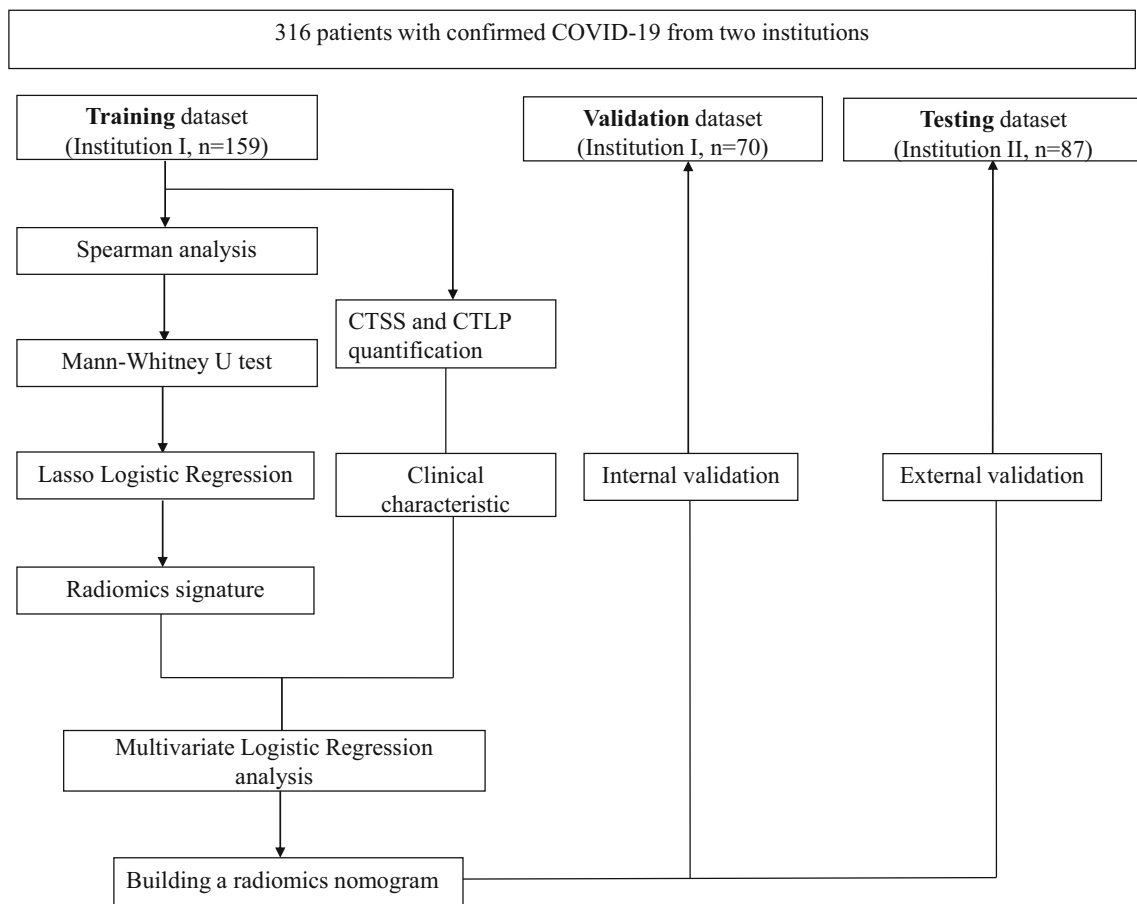


Fig. 1 Flowchart of the study

abnormalities on the basis of the area involved [11]. Each of the five lung lobes was visually scored from 0 to 5 as follows: 0, no involvement; 1, less than 5% involvement; 2, less than 25% involvement; 3, 26–49% involvement; 4, 50–75% involvement; or 5, more than 75% involvement [12].

Collection of clinical data and evaluation of CT radiological features

Clinical data were recorded, including the following 10 characteristics: age, sex, duration of onset, comorbidity, clinical type, treatments, respiratory support strategies (RSS), ICU admission, length of ICU stay, and length of hospital stay. We also recorded other clinical parameters, such as oxygen saturation and temperature, but which were not used in the clinic’s model, because the former was one of the known diagnostic criteria for severe COVID-19, and the latter was a variable parameter due to uncertain medical history. Comorbidities included hypertension, diabetes, chronic obstructive pulmonary disease (COPD), chronic kidney disease (CKD), malignant tumors, and surgery history (on any part of the body in the last 10 years). The number of comorbidities was from

0 to 4; 0 means no complications, and 4 means there are 4 kinds of diseases. Treatments included antiviral therapy, antibiotic therapy, glucocorticoid therapy, immunoglobulin therapy, and Chinese medicine therapy. RSS included nasal catheter, high-flow nasal cannula oxygen therapy, non-invasive mechanical ventilation, invasive mechanical ventilation, and extracorporeal membrane oxygenation (ECMO).

All CT images were evaluated by 2 radiologists (Z.K. and L.L.) who were blinded to each subject’s clinical data. For disagreement between the two primary radiologist interpretations, a third experienced thoracic radiologist with 25 years of experience (Y.Z.) adjudicated a final decision. Ten CT radiological features were assessed, namely GGO, consolidation, reticular pattern, interlobular septal thickening, air bronchogram sign, lesion location, distribution, involved lobe, thickening of pleura, and pleural effusion.

Radiomics feature extraction

A total of 851 radiomics features were extracted from the VOIs segmented based on the L. K software, including first-order statistics parameters (n = 18), morphological parameters (n = 14),

gray-level co-occurrence matrix (GLCM) parameters ($n = 24$), gray-level run length matrix (GLRLM) parameters ($n = 16$), gray-level size zone matrix (GLSZM) parameters ($n = 16$), gray-level dependence matrix (GLDM) parameters ($n = 14$), neighboring gray tone difference matrix (NGTDM) parameters ($n = 5$), and wavelet parameters ($n = 744$). All the features defined were in compliance with feature definitions as described by the Imaging Biomarker Standardization Initiative (IBSI) [13]. The detailed workflow of radiomics analysis can be found in Fig. 2. Intra- and interclass correlation coefficients (ICC) were used to assess the intra- and inter-observer reproducibility of radiomics feature extraction.

Radiomics features selection and radiomics signature construction

The outlier values were replaced by the median value of the particular variance vector once the values were beyond the range of the mean and standard deviation. And standardization was performed to scale the data in a specific interval. Spearman correlation, generalized linear model (GLM), and least absolute shrinkage and selection operator (LASSO) were used to reduce the redundancy or selection bias of the features, thereby removing a high correlation. A radiomics score (Rad-score) was calculated for each patient via a linear combination of selected features that were weighted by their respective coefficients.

Development of predictive models

The most significant features were investigated to construct radiomics model based on logistic regression. The likelihood ratio test with backward step-down selection was applied to the multivariate logistic regression model. We grouped the selected features into seven models—the radiomics model

(radiomics features), the CTSS model (semi-quantitative CTSS), CTLP model (quantitative CTLP), the clinical model (clinical features), the integrated A model (CTSS + CTLP + clinical features), the integrated B model (clinical features + radiomics features), and the integrated C model (radiomics features + CTSS + CTLP + clinical features). The calibration curves were used to investigate the performance characteristics of the nomograms.

Statistical analysis

Statistical analyses were performed with the Institute of Precision Medicine Statistics (IPMs, version 1.1, GE Healthcare). The differences in all variables between NSD and SD groups were assessed using the Mann-Whitney U test or independent samples t test for continuous variables, and the chi-square test or Fisher's exact test for categorical variables. Univariate analysis was used to estimate the relationship between clinical factors and the identification of the two subtypes. The performances of the seven models were assessed by area under the receiver operating characteristic curve (AUC), specificity, and sensitivity. The optimal cut-off points to predict the severity of COVID-19 were determined by Youden's index. The DeLong test was used for pairwise comparisons among the seven models using the R software (v. 3.6.0; <http://www.Rproject.org>). A two-sided $p < 0.05$ was considered statistically significant throughout the study.

Results

Clinical and CT radiological features

The detailed distribution of clinical characteristics in the NSD and SD group was summarized in Table 1. Clinical factors (gender and duration of onset) were found not significantly different

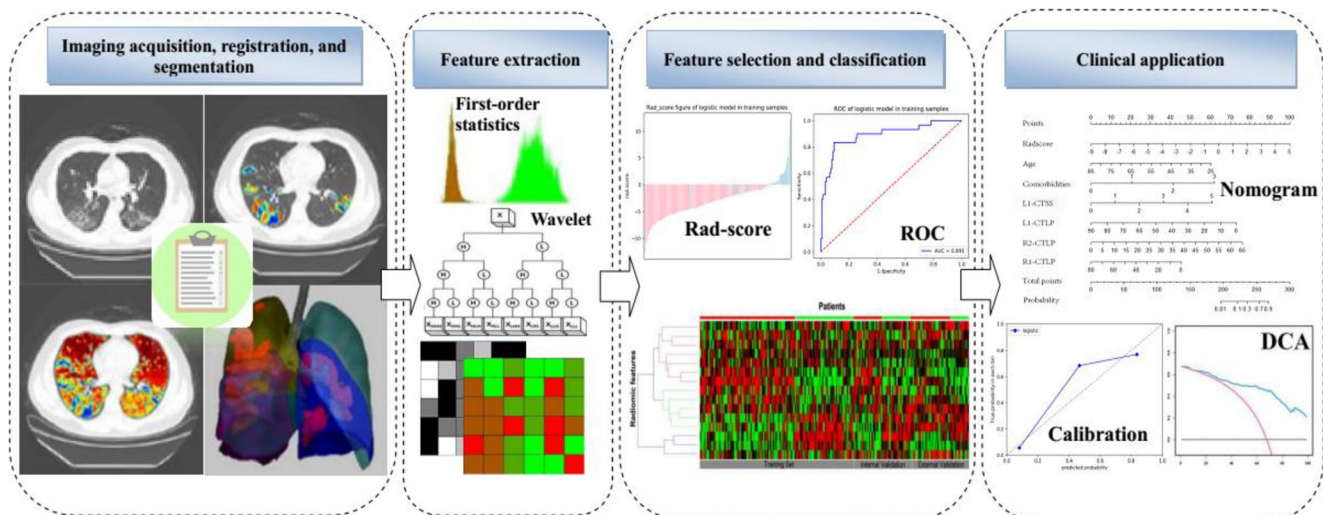


Fig. 2 Radiomics framework of predicting the severe patients with COVID-19

Table 1 Clinical and CT features of training, validation, and testing datasets in two groups

Characteristic	Training cohort (institution I, <i>n</i> = 159)			Validation cohort (institution I, <i>n</i> = 70)			Testing cohort (institution II, <i>n</i> = 87)		
	NSD (<i>n</i> = 129)	SD (<i>n</i> = 30)	<i>p</i> value	NSD (<i>n</i> = 56)	SD (<i>n</i> = 14)	<i>p</i> value	NSD (<i>n</i> = 61)	SD (<i>n</i> = 26)	<i>p</i> value
Age, year	54.00 (32.00,60.00)	57.00 (39.00,72.00)	0.003*	50.50 (36.90,60.65)	54.50 (39.55,62.10)	0.602	56.58 (43.73,59.11)	63.96 (49.34,71.00)	0.029*
Male (%)	67 (51.94%)	15 (50.00%)	0.848	38 (67.86%)	8 (57.14%)	0.659	37 (60.66%)	11 (42.31%)	0.115
Comorbidity no.									
0	80 (62.02%)	10 (33.33%)	0.006*	32 (57.14%)	4 (28.57%)	0.069	31 (50.82%)	8 (30.77%)	0.023*
1	33 (25.58%)	9 (30.00%)		20 (35.71%)	6 (42.86%)		20 (32.79%)	13 (50.00%)	
2	10 (7.75%)	6 (20.00%)		3 (5.36%)	3 (21.43%)		8 (13.11%)	1 (3.85%)	
3	5 (3.88%)	3 (10.00%)		1 (1.79%)	1 (7.14%)		1 (1.64%)	4 (15.38%)	
4	1 (0.78%)	2 (6.67%)		0 (0.00%)	0 (0.00%)		1 (1.64%)	0 (0.00%)	
Days onset (%)									
1-3 days	16 (12.40%)	6 (20.00%)	0.424	4 (7.14%)	3 (21.43%)	0.051	3 (4.92%)	2 (7.69%)	0.484
4-7 days	47 (36.43%)	12 (40.00%)		21 (37.50%)	8 (57.14%)		20 (32.79%)	5 (19.23%)	
> 7 days	66 (51.16%)	12 (40.00%)		31 (55.36%)	3 (21.43%)		38 (62.30%)	19 (73.08%)	
L1-CTSS (%)									
0	31 (24.03%)	1 (3.33%)	0.001*	13 (23.64%)	2 (14.29%)	0.229	9 (14.75%)	3 (11.54%)	0.591
1	61 (47.29%)	10 (33.33%)		24 (43.64%)	3 (21.43%)		34 (55.74%)	12 (46.15%)	
2	29 (22.48%)	11 (36.67%)		11 (20.00%)	5 (35.71%)		15(24.59%)	8 (30.77%)	
3	5 (3.88%)	3 (10.00%)		3 (5.45%)	1 (7.14%)		3 (4.92%)	3 (11.54%)	
4	2 (1.55%)	4 (13.33%)		3 (5.45%)	2 (14.29%)		0 (0.00%)	0 (0.00%)	
5	1 (0.78%)	1 (3.33%)		1 (1.82%)	1 (7.14%)		0 (0.00%)	0 (0.00%)	
L2-CTSS (%)									
0	13 (10.08%)	0 (0.00%)	0.057	4 (7.27%)	0 (0.00%)	0.119	2 (3.28%)	0 (0.00%)	0.516
1	30 (23.26%)	3 (10.00%)		17 (30.91%)	1 (7.14%)		14 (22.95%)	7 (26.92%)	
2	42 (32.56%)	13 (43.33%)		14 (25.45%)	4 (28.57%)		29 (47.54%)	9 (34.62%)	
3	29 (22.48%)	6 (20.00%)		11 (20.00%)	3 (21.43%)		12 (19.67%)	9 (34.62%)	
4	12 (9.30%)	7 (23.33%)		8 (14.55%)	4 (28.57%)		4 (6.56%)	1 (3.85%)	
5	3 (2.33%)	1 (3.33%)		1 (1.82%)	2 (14.29%)		0 (0.00%)	0 (0.00%)	
R1-CTSS (%)									
0	39 (30.23%)	3 (10.00%)	0.001*	9 (16.36%)	0 (0.00%)	0.129	6 (9.84%)	0 (0.00%)	0.016*
1	44 (34.11%)	11 (36.67%)		21 (38.18%)	4 (28.57%)		33 (54.10%)	9 (34.62%)	
2	29 (22.48%)	8 (26.67%)		15 (27.27%)	4 (28.57%)		19 (31.15%)	11 (42.31%)	
3	15 (11.63%)	2 (6.67%)		5 (9.09%)	3 (21.43%)		3 (4.92%)	6 (23.08%)	
4	1 (0.78%)	5 (16.67%)		4 (7.27%)	1 (7.14%)		0 (0.00%)	0 (0.00%)	
5	1 (0.78%)	1 (3.33%)		1 (1.82%)	2 (14.29%)		0 (0.00%)	0 (0.00%)	
R2-CTSS (%)									
0	59 (45.74%)	5 (16.67%)	0.01*	21 (38.18%)	1 (7.14%)	0.04*	20 (32.79%)	8 (30.77%)	0.068
1	40 (31.01%)	9 (30.00%)		17 (30.91%)	4 (28.57%)		31 (50.82%)	8 (30.77%)	
2	23 (17.83%)	11 (36.67%)		10 (18.18%)	5 (35.71%)		9 (14.75%)	7 (26.92%)	
3	6 (4.65%)	4 (13.33%)		6 (10.91%)	2 (14.29%)		1 (1.64%)	3 (11.54%)	
4	1 (0.78%)	1 (3.33%)		1 (1.82%)	2 (14.29%)		0 (0.00%)	0 (0.00%)	
5	0 (0.00%)	0 (0.00%)		0 (0.00%)	0 (0.00%)		0 (0.00%)	0 (0.00%)	
R3-CTSS (%)									
0	16 (12.40%)	1 (3.33%)	0.05*	4 (7.27%)	0 (0.00%)	0.008*	3 (4.92%)	0 (0.00%)	0.435
1	32 (24.81%)	4 (13.33%)		12 (21.82%)	0 (0.00%)		13 (21.31%)	5 (19.23%)	
2	41 (31.78%)	7 (23.33%)		19 (34.55%)	3 (21.43%)		24 (39.34%)	7 (26.92%)	
3	24 (18.60%)	10 (33.33%)		12 (21.82%)	5 (35.71%)		18 (29.51%)	11 (42.31%)	
4	13 (10.08%)	5 (16.67%)		7 (12.73%)	2 (14.29%)		3 (4.92%)	3 (11.54%)	

Table 1 (continued)

Characteristic	Training cohort (institution I, <i>n</i> = 159)			Validation cohort (institution I, <i>n</i> = 70)			Testing cohort (institution II, <i>n</i> = 87)		
	NSD (<i>n</i> = 129)	SD (<i>n</i> = 30)	<i>p</i> value	NSD (<i>n</i> = 56)	SD (<i>n</i> = 14)	<i>p</i> value	NSD (<i>n</i> = 61)	SD (<i>n</i> = 26)	<i>p</i> value
5	3 (2.33%)	3 (10.00%)		1 (1.82%)	4 (28.57%)		0 (0.00%)	0 (0.00%)	
Total CTSS	7.00 (3.00,10.30)	10.00 (7.00,15.05)	< 0.001*	8.00 (4.00,10.00)	13.00 (8.90,18.00)	0.005*	7.00 (5.00,9.30)	10.00 (6.00,12.00)	0.068
Quantitative CTLP									
L1-CTLP	1.00 (0.00, 6.03)	7.75 (2.67, 25.70)	< 0.001*	0.25 (0.00, 3.55)	6.70 (1.06, 21.02)	0.013*	2.90 (0.17, 5.72)	3.35 (0.69,15.13)	0.243
L2-CTLP	11.20 (1.54, 31.49)	21.55 (15.89,50.4-7)	0.006*	6.20 (1.30,18.21)	29.90 (15.87,54.2-6)	0.003*	10.70 (4.62,27.64)	22.50 (4.38,32.82)	0.241
R1-CTLP	1.10 (0.00, 11.52)	8.05 (1.09, 34.70)	0.003*	1.15 (0.04, 5.13)	13.35 (1.72, 25.51)	0.016*	1.40 (0.30, 9.24)	10.00 (1.49,18.60)	0.006*
R2-CTLP	0.10 (0.00,4.70)	6.65 (1.29, 16.04)	< 0.001*	0.00 (0.00, 3.64)	7.00 (0.57, 37.11)	0.004*	0.30 (0.00, 2.45)	2.10 (0.00, 6.93)	0.143
R3-CTLP	11.40 (1.37,32.32)	32.10 (14.53,50.6-9)	0.004*	7.85 (1.94,19.45)	40.20 (25.16,50.2-9)	< 0.001*	13.90 (4.43,30.42)	28.30 (7.96,44.04)	0.067
Total CTLP	5.20 (1.54,18.43)	15.70 (7.77, 35.99)	0.001*	3.35 (1.24,10.68)	21.65 (8.76, 33.91)	0.001*	5.70 (3.07,14.56)	14.05 (4.35,24.47)	0.069

Note—Age, total CTSS, and CTLP are interquartile range; other data are the number of patients with the percentage in parentheses. *CTSS*, CT severity score; *CTLP*, CT lesion percentage; *L1*, left upper lobe; *L2*, left lower lobe; *R1*, right upper lobe; *R2*, right middle lobe; *R3*, right lower lobe; *NSD*, non-severe disease group; *SD*, severe disease group; *IQR*, interquartile range. 0, no involvement; 1, less than 5% involvement; 2, less than 25% involvement; 3, 26–49% involvement; 4, 50–75% involvement; or 5, more than 75% involvement. Comorbidities included hypertension, diabetes, COPD, CKD, malignant tumors, and surgery history. The number of comorbidities from 0 to 4, as follows: 0, no comorbidity; 4, 4 comorbidities. *Data with statistical significance

between the two groups in all cohorts, while the comorbidities and age were the only similarity between the two groups in the validation cohort.

Three hundred sixteen patients (NSD, 246 vs. SD, 70) from two hospitals were included in the study. The patients in both groups showed GGO with mainly peripheral distribution; there was no statistical difference (78% vs. 84%,

$p > 0.05$). In the SD group, 45% of patients had consolidation, and 60% had multiple lung lobes involved, but only 35% and 50% in the NSD group ($p < 0.05$). More patients had pleural effusions in the SD group compared to that in the NSD group (54.28% vs. 7.72%, $p < 0.05$). More imaging details between the two groups in different hospitals are shown in Table 2.

Table 2 CT radiological features of training, validation, and testing datasets in two groups

CT features	Training cohort (<i>n</i> = 159)			Validation cohort (<i>n</i> = 70)			Testing cohort (<i>n</i> = 87)		
	NSD (<i>n</i> = 129)	SD (<i>n</i> = 30)	<i>p</i> value	NSD (<i>n</i> = 56)	SD (<i>n</i> = 14)	<i>p</i> value	NSD (<i>n</i> = 61)	SD (<i>n</i> = 26)	<i>p</i> value
GGO	111 (86.05%)	25 (83.33%)	0.077	49 (87.5%)	12 (85.71%)	0.122	54 (88.52%)	22 (84.62%)	0.163
Consolidation	66 (51.16%)	26 (86.67%)	0.004*	25 (44.64%)	11 (78.57%)	< 0.01*	25 (40.98%)	21 (80.77%)	< 0.01*
Reticular pattern	34 (26.36%)	17 (56.67%)	0.031*	14 (25%)	6 (42.56%)	0.017*	18 (29.51%)	11 (42.31%)	0.025*
Interlobular septal thickening	55 (42.64%)	13 (43.33%)	0.069	23 (41.07%)	8 (57.14%)	0.042*	23 (37.7%)	13 (50%)	0.059
Air bronchogram sign	43 (33.33%)	19 (63.34%)	< 0.01*	16 (28.57%)	7 (50%)	< 0.01*	19 (31.15%)	15 (57.7%)	< 0.01*
Bilateral involvement	108 (83.72%)	26 (86.67%)	0.141	50 (89.29%)	13 (92.86%)	0.192	54 (88.52%)	23 (88.46%)	0.299
Peripheral distribution	92 (71.32%)	24 (80%)	0.222	42 (75%)	12 (85.71%)	0.257	47 (77.05%)	19 (73.08%)	0.184
Multilobar involvement	88 (68.22%)	20 (66.67%)	0.178	40 (71.43%)	11 (78.57%)	0.199	41 (67.21%)	20 (76.92%)	0.208
Adjacent pleura thickening	69 (53.49%)	16 (53.34%)	0.454	28 (50%)	8 (57.14%)	0.577	26 (42.62%)	18 (69.23%)	0.049
Pleural effusion	7 (5.43%)	17 (56.67%)	< 0.01*	4 (7.14%)	7 (50%)	< 0.01*	8 (13.11%)	14 (53.85%)	< 0.01*

Note—Except where indicated, data are numbers of patients, with percentages in parentheses. *GGO*, ground-glass opacity; *NSD*, non-severe disease group; *SD*, severe disease group. *Data with statistical significance

Radiomics features selection and radiomics signature building

In total, 99 features with statistical significance ($p < 0.05$) between the NSD and SD groups were selected in the training dataset. A radiomics signature was further constructed based on sixteen features with respective non-zero coefficients selected from these 99 features. One first-order feature and fifteen wavelet-based features were used to represent the radiomics signature. Details of the procedure for the construction of the radiomics signature are showed in Fig. 3a, b. The lists of the selected features and their associated coefficients in the logistic regression model are illustrated in Fig. 3c. The intra-observer ICCs ranged from 0.822 to 0.957 and the inter-observer ICCs ranged from 0.769 to 0.936, indicating favorable intra- and inter-observer feature extraction reproducibility. Figure 4 shows representative images and lesion segmentation result of severe COVID-19

The utility of severity prediction using developed radiomics signature

The developed initial CT-based radiomics signature model showed a favorable result in predicting the severity (NSD vs. SD) that produced an AUC of 0.9 in the training set (95% CI, 0.843 to 0.942), 0.878 in the internal validation set (95% CI, 0.796 to 0.958), and 0.842 in the testing set (95% CI, 0.761 to 0.922).

As shown in Fig. 5, predictive nomogram and correlation coefficients were built by combining Rad-score, age, comorbidity, L1-CTSS, L1-CTLP, R1-CTLP, and R2-CTLP. The predictive nomogram had the best differentiation ability of the severe cases with an AUC of 0.918 (95% CI, 0.864–0.956) in the training set, an AUC of 0.934 (95% CI, 0.848–0.979) in the validation set, and an AUC of 0.854 (95% CI, 0.762–0.92) in the testing set, as shown in Fig. 6a–c. The calibration curves of the nomogram also showed that the predictions agreed well with the observations in Fig. 6d–f.

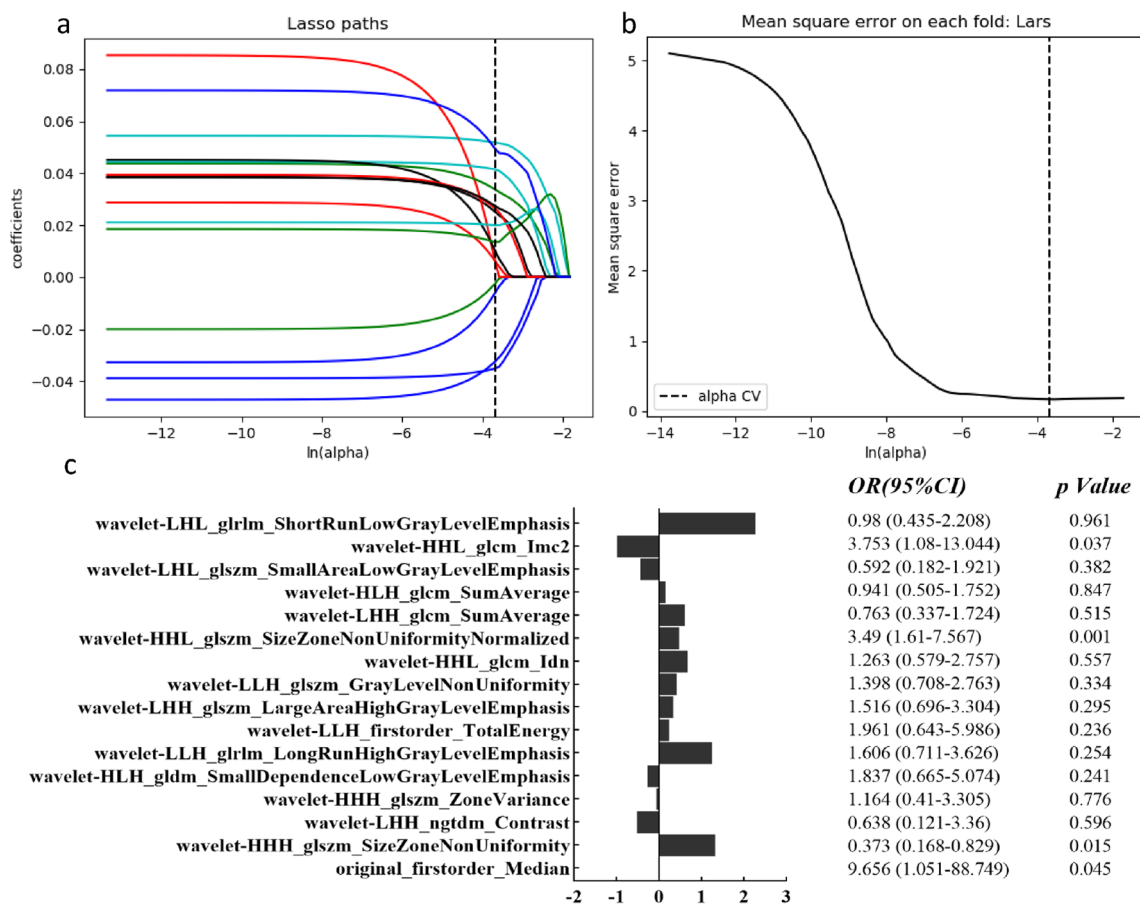
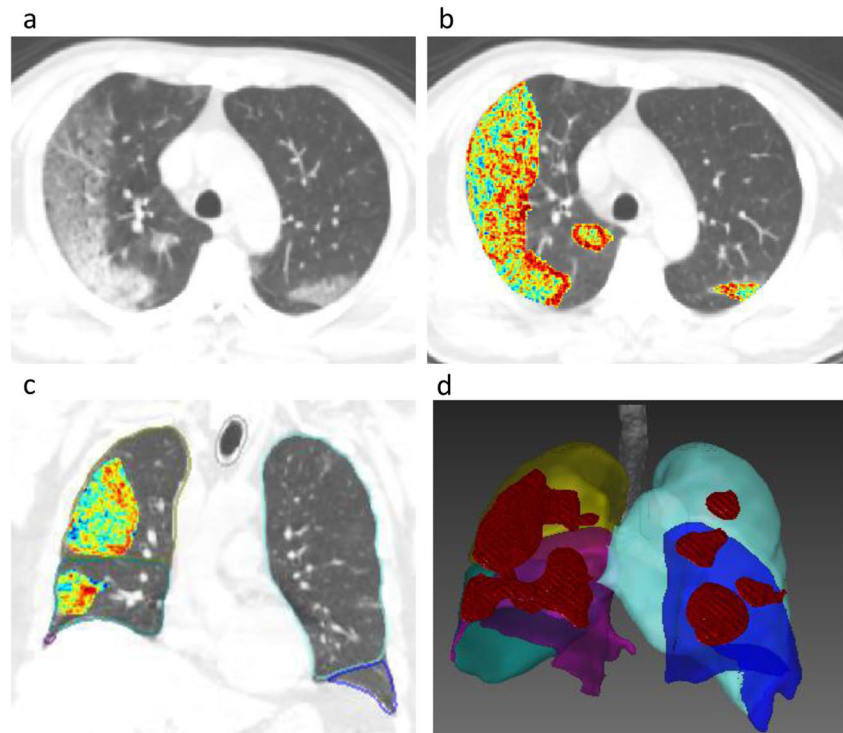


Fig. 3 Texture feature selection using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. **a** LASSO coefficient profiles of the radiomics features. Vertical line was drawn at the value selected using 5-fold cross-validation in the $\ln(\alpha)$ sequence, and 16 non-zero coefficients are indicated. **b** The tuning parameter λ

selection in the LASSO model used 5-fold cross-validation via the minimum criteria. Mean square error was plotted vs. $\log(\lambda)$. The dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1-SE criteria. **c** Multivariate logistic of the predictive radiomics features. OR, odds ratio

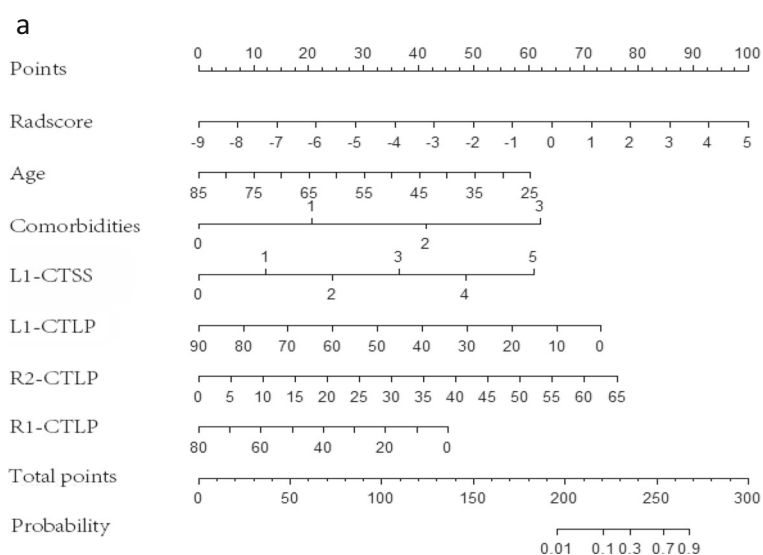
Fig. 4 A case of confirmed severe COVID-19. A 62-year-old female presented with a 7-day history of fever and cough. First CT imaging revealed diffuse pure GGO with mainly peripheral distribution in the bilateral lobes (a). The area of the lesions was delineated on the axial, coronal, and reconstructed three-dimensional images (b, c, d). CTSS = total 10 scores, predicted probability for severe COVID-19 = 90.1%



Evaluation of models and comparison of predictive model performance

The diagnostic performance of each model is shown in Table 3 and the results of ROC curve analysis are shown in Fig. 6. In all cohorts, the integrated C model achieved the best performance and radiomics model outperformed clinic's model for predicting severe

COVID-19. Moreover, the AUC of the integrated B model was better than CTSS model and CTLP model (all $p < 0.05$). In the training set, although the integrated C model showed the highest AUC (0.918) among the seven predictive models, no statistical difference was found between any of the two models using the DeLong test ($p = 0.832$ for CTSS vs. CTLP; $p = 0.165$ for CTSS vs. clinics; $p = 0.323$ for CTLP vs.



Variable	Coefficient	<i>P</i> Value
Age	-0.094	0.047
Comorbidities	1.940	0.010
L1-CTSS	1.141	0.253
L1-CTLP	-0.076	0.315
R2-CTLP	0.110	0.077
R1-CTLP	-0.053	0.344
Radscore	0.669	0.003
Constant	2.225	0.262

Fig. 5 Radiomics-based nomogram (a) and their correlation coefficients (b) were developed in the training set, including the Rad-score, age, comorbidities, CTSS, and CTLP

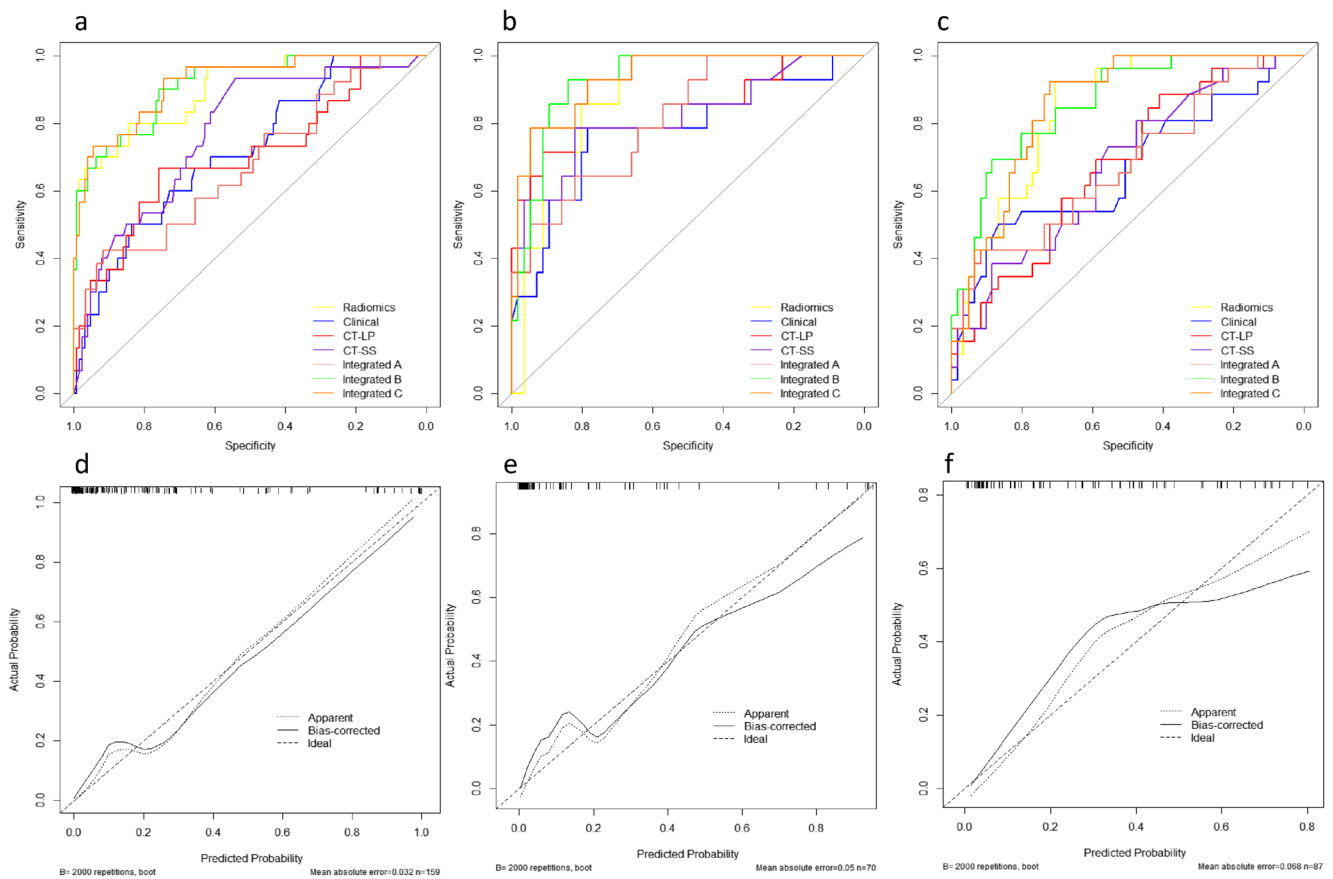


Fig. 6 The ROC curves of the seven prediction models that indicate severe COVID-19 cases in the training cohort (a), validation cohort (b), and testing cohort (c). Calibration curves of the combined nomogram in the training cohort (d), internal validation cohort (e), and testing cohort

(f). Calibration curves depict the calibration of the nomogram in terms of agreement between the predicted risk and actual probability for severe COVID-19

clinics). No statistical difference in AUC was also found between any of the two models ($p = 0.09$ for CTSS vs. integrated A; $p = 0.137$ for CTLP vs. integrated A; $p = 0.751$ for clinics vs. integrated A), but a statistical

difference in AUC was found between the radiomics model and the integrated A model ($p = 0.017$), and between radiomics model and clinics model ($p = 0.00075$)

Table 3 Comparison of predictive model performance for identifying severe COVID-19 pneumonia

Model	Training cohort ($n = 159$)			Validation cohort ($n = 70$)			Testing cohort ($n = 87$)		
	AUC (95% CI)	Sensitivity	Specificity	AUC (95% CI)	Sensitivity	Specificity	AUC (95% CI)	Sensitivity	Specificity
Radiomics model	0.90 (0.84–0.94)	0.80	0.85	0.88 (0.80–0.96)	1.00	0.70	0.84 (0.76–0.92)	0.92	0.71
CTSS model	0.77 (0.70–0.83)	0.93	0.54	0.82 (0.71–0.90)	0.79	0.82	0.67 (0.56–0.77)	0.73	0.56
CTLP model	0.71 (0.63–0.78)	0.68	0.77	0.84 (0.74–0.92)	0.71	0.91	0.68 (0.57–0.77)	0.88	0.41
Clinical model	0.72 (0.64–0.78)	0.50	0.84	0.78 (0.66–0.87)	0.79	0.79	0.67 (0.54–0.76)	0.52	0.87
Integrated A model	0.78 (0.71–0.84)	0.73	0.77	0.82 (0.70–0.90)	0.79	0.82	0.70 (0.58–0.80)	0.70	0.67
Integrated B model	0.91 (0.85–0.95)	0.97	0.69	0.93 (0.84–0.98)	0.93	0.82	0.84 (0.75–0.91)	0.92	0.71
Integrated C model	0.92 (0.86–0.96)	0.73	0.95	0.93 (0.85–0.98)	0.79	0.95	0.84 (0.76–0.92)	0.92	0.72

Note—The integrated A model contained CTSS, CTLP, and clinical features. The integrated B model contained the selected radiomics features and clinical features. The integrated C model contained the selected radiomics features, CTSS, CTLP, and clinical features. CTSS, CT severity score; CTLP, CT lesion percentage; AUC, area under the receiver operating characteristic curve

Treatment and prognosis in the patients with COVID-19

All 316 patients with COVID-19 received hospitalization, and their treatment information was recorded. Compared with the non-severe group, severe patients have more comorbidities and longer ICU and hospital length of stay ($p < 0.05$). Most patients received antibiotics (166 [52.53%]), antiviral treatment (276 [87.34%]), and glucocorticoids (178 [56.33%]). One hundred twenty-nine (40.82%) patients received intravenous infusions of immunoglobulin. Two-thirds of the patients had oxygen therapy. Treatment and prognosis information of all three cohorts are provided in Table 4.

Discussion

In this study, statistical analysis was performed in comparing clinical data, quantitative CT, and radiomics features between non-severe and severe COVID-19 patients. Two clinical

features, four quantitative CT features, and sixteen radiomics features were identified to be significantly different between the two groups ($p < 0.05$). Furthermore, CT-based radiomics model had a significantly high area under the curve (AUC), sensitivity, and specificity to distinguish individuals with severe COVID-19 from non-severe COVID-19. The radiomics model alone was also validated with the independent dataset from the other institutions, suggesting the reliability and reproducibility of our new prediction model.

In general, the most common CT features for COVID-19 included GGO and consolidation involving the bilateral lungs in a peripheral distribution [14, 15]. In the early diagnostic stage, it is difficult for clinicians to judge whether it is COVID-19 pneumonia. Meanwhile, for non-severe and severe COVID-19, CT features are often highly overlapping and non-specific. Recently, some literatures reported that the proportion of lesions (GGO/consolidation) is different between NSD and SD patients [16–18]. Non-severe patients are more likely to have more GGO, while severe patients tend to have more consolidation in

Table 4 Treatments and outcomes of all patients with COVID-19 in two groups

Characteristic	Training cohort ($n = 159$)			Validation cohort ($n = 70$)			Testing cohort ($n = 87$)		
	NSD ($n = 129$)	SD ($n = 30$)	p value	NSD ($n = 56$)	SD ($n = 14$)	p value	NSD ($n = 61$)	SD ($n = 26$)	p value
Hospitalization information									
ICU admission (%)	13 (10.08%)	26 (86.87%)	< 0.001*	7 (12.5%)	11 (78.57%)	< 0.001*	9 (14.75%)	24 (92.31%)	< 0.001*
ICU length of stay, days	12 (9,20)	21 (18, 29)	< 0.001*	15 (9,19)	20 (16, 28)	< 0.001*	16 (9,20)	19 (13, 27)	< 0.001*
Hospital length of stay, days	25 (21,30)	39 (30,48)	0.001*	24 (21,30)	36 (32,47)	< 0.001*	26 (22,31)	37 (30,47)	< 0.001*
Treatments, n (%)									
Antibiotic treatment	65 (50.38%)	20 (66.67%)	0.031*	26 (46.43%)	11 (78.57%)	0.022*	30 (49.18%)	14 (53.85%)	0.011*
Antiviral treatment	110 (85.27%)	29 (96.67%)		50 (89.29%)	13 (92.86%)		50 (81.97%)	24 (92.31%)	
Glucocorticoid therapy	70 (54.26%)	21 (70%)		27 (48.21%)	10 (71.42%)		32 (52.46%)	18 (69.23%)	
Intravenous immunoglobulin	39 (30.23%)	20 (66.67%)		21 (37.5%)	9 (64.29%)		23 (37.7%)	17 (65.38%)	
Chinese medicine treatment	110 (85.27%)	19 (63.33%)		46 (82.14%)	9 (64.29%)		52 (85.25%)	16 (61.54%)	
Respiratory support strategies, n (%)									
Nasal catheter	88 (68.22%)	12 (40%)	0.007*	32 (57.14%)	7 (50%)	0.012*	34 (55.74%)	11 (42.31%)	0.044*
High-flow nasal cannula oxygen therapy	31 (24.03%)	9 (30%)		20 (35.71%)	4 (28.57%)		17 (27.87%)	7 (26.92%)	
Non-invasive mechanical ventilation	10 (7.75%)	4 (13.33%)		4 (7.14%)	1 (7.14%)		10 (16.39%)	3 (11.54%)	
Invasive mechanical ventilation	0 (0%)	2 (6.67%)		0 (0%)	1 (7.14%)		0 (0%)	2 (7.69%)	
ECMO	0 (0%)	3 (10%)		0 (0%)	1 (7.14%)		0 (0%)	3 (11.54%)	
Outcomes, n (%)									
Survival	129 (100%)	28 (93.33%)	0.033*	56 (100%)	13 (92.86%)	0.042*	61 (100%)	25 (96.15%)	0.496*
Death	0 (0%)	2 (6.67%)		0 (0%)	1 (7.14%)		0 (0%)	1 (3.85%)	

Note—ICU length of stay and hospital length of stay are interquartile range; other data are the number of patients with the percentage in parentheses. NSD, non-severe disease group; SD, severe disease group; ECMO, extracorporeal membrane oxygenation; ICU, intensive care unit. *Data with statistical significance

CT imaging. Although we have also observed this phenomenon, no statistical difference was found ($p > 0.05$). Therefore, it is also very difficult for the radiologist to judge whether it is severe pneumonia [19, 20].

Radiomics has great potential to capture useful medical information and to enhance the accuracy of clinical differential diagnosis. In this study, 851 candidate radiomics features were extracted from CT images and were reduced to only 16 potential predictors by using a LASSO regression model to develop the radiomics signatures. The selected radiomics features were divided into two types (first-order and wavelet features) and varied significantly between NSD and SD groups. These features reflect intrinsic information from the distribution of pixel intensity and the texture morphology that cannot be detected by radiologists [21]. For example, first-order features mainly reflect the internal texture of lesions. Wavelet features mainly reflect the change of time domain and frequency domain inside the lesion. Among the selected radiomics features in this study, *Wavelet-HHL_glm_Imc2* and *Wavelet-LHL_glrml_ShortRunLowGrayLevelEmphasis* were the most significant and robust features associated with severe cases, which reflect lesion's intensity and textural features within the high-intensity CT voxels.

Several studies have reported and analyzed the value of traditional CT characteristics and clinical features in the diagnosis of COVID-19. Chen et al [22] developed and validated a diagnostic model for COVID-19 based on CT imaging and clinical manifestations. Kang et al [23] introduced an AI-assisted model based on the radiological and clinical features to estimate the clinical progression to critical illness. In addition, in order to quantify the severity of COVID-19, CT quantification of pneumonia lesions is also used in clinical practice. The CTSS and CTLP are the two most common CT quantification methods, but the differences among observers are still inevitable [7, 16, 22]. In our model, L1-CTSS, L1-CTLP, R1-CTLP, and R2-CTLP were selected as independent risk factors for severe COVID-19 cases. The possible explanation is that the SARS-CoV-2 mainly invades the bilateral lower lobes in patients with COVID-19, so the difference in the distribution of lesions in the upper or middle lobe is often more statistically different among different patients with COVID-19.

The integrated C model composed of radiomics, quantitative CT, and clinical features had the highest AUC (0.92, 0.93, 0.84) in all three cohorts. The model with radiomics features alone reached an AUC of 0.9, 0.878, and 0.842, which is not inferior to the integrated B or C model (all $p > 0.05$). In other words, the addition of quantitative CT and clinical features to the integrated model did not increase the model's efficacy, suggesting the strong efficacy of radiomics as a tool to predict the severity of COVID-19. Furthermore, the predictive power of the radiomics model also outperformed that of integrated A model (clinics + quantitative CT) ($p < 0.05$). In fact, radiomics features and clinical or quantitative CT features were highly correlated.

This result was in accordance with that in other similar studies in gliomas [24], pancreatic neuroendocrine tumor (PNET) [23], lung adenocarcinoma [21], and breast cancer [25].

Several studies reported single or multiple risk factors of lung injury for severe COVID-19. Dong et al [26] found that underlying comorbidity, older age, higher LDH, and lower lymphocyte count were independent high-risk factors associated with COVID-19 progression in a multicenter study. Guan et al [27] reported that increasing age and comorbidity were associated with the disease. Zhou et al [28] showed that older age, high SOFA score, and D-dimer value are potential risk factors to identify patients with poor prognosis at an early stage. Chen et al [29] proposed the MuLBSTA score as an early warning indicator for predicting 90-day mortality in patients with COVID-19. Among the clinical features in our study, age and comorbidity were the most powerful factors to predict severe COVID-19. A possible reason is that poor immunosenescence of older patients with comorbidities could have higher levels of angiotensin-converting enzyme 2 in their alveoli [30], which was regarded as a receptor for novel coronaviruses [31].

Some limitations in this study should also be acknowledged. The biggest limitation is that no blood samples or some other clinical parameters are considered in the risk model. We believe that a reliable and robust multi-modality prediction model should be further developed to address this issue. In addition, the sample size was relatively small in order to build up a stable predicting model and the cases outside of Hubei province are not included. However, in this study, we tried to adopt a multicenter research method, including internal and external validation sets, subjective and objective CT evaluation, all of which ensure that the conclusions are reasonable. Furthermore, with the retrospective design, there may be potential biases in identifying and recruiting participants. Increasing a large number of different regional populations and imaging-pathologic control studies is needed in the future study.

In conclusion, initial CT-based radiomics features provide an excellent performance for the prediction of severe COVID-19. Sixteen features were significantly different between the two groups. This prediction model based on the radiomics features was developed and validated on the training, validation, and testing cohorts, which may have broad clinical applications.

Funding This study has received funding from the National Natural Science Foundation of China (Nos. 81871332, 81601461), Novel Coronavirus Pneumonia Emergency Key Project of Science and Technology of Hubei Province (No. 2020FCA015), and the Fundamental Research Funds for the Central Universities (No. 2042020kfxg10).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Yunfei Zha.

Statistics and biometry Statistician Huan Liu kindly provided all statistical work for this study.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Conflict of interest One of the authors of this manuscript (Huan Liu) is an employee of GE Healthcare. The remaining authors declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Methodology

- retrospective
- diagnostic or prognostic study
- multicenter study

References

1. Zhu N, Zhang D, Wang W et al (2020) A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 382:727–733
2. World Health Organization. Coronavirus disease (COVID-19) pandemic on 5 November 2020 [EB/OL]. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Accessed 5 Nov 2020
3. Huang C, Wang Y, Li X et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497–506
4. Henry BM, de Oliveira MHS, Benoit S, Plebani M, Lippi G (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*. <https://doi.org/10.1515/cclm-2020-0369>
5. Xiong Y, Sun D, Liu Y et al (2020) Clinical and high-resolution CT features of the COVID-19 infection: comparison of the initial and follow-up changes. *Invest Radiol* 55:332–339
6. Ding X, Xu J, Zhou J, Long Q (2020) Chest CT findings of COVID-19 pneumonia by duration of symptoms. *Eur J Radiol* 127:109009
7. Shen C, Yu N, Cai S et al (2020) Quantitative computed tomography analysis for stratifying the severity of Coronavirus Disease 2019. *J Pharm Anal*. <https://doi.org/10.1016/j.jpha.2020.03.004>
8. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48:441–446
9. The National Health Commission of the People's Republic of China. Diagnosis and treatment protocols of COVID-19 infection (trial version 7) [EB/OL]. <http://www.nhc.gov.cn/yzygj/s7653p/202002/8334a8326dd94d329df351d7da8aefc2s.html>. Accessed 3 Mar 2020
10. Wei W, Hu XW, Cheng Q, Zhao YM, Ge YQ (2020) Identification of common and severe COVID-19: the value of CT texture analysis and correlation with clinical characteristics. *Eur Radiol* 30:6788–6796
11. Pan F, Ye T, Sun P et al (2020) Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology*. <https://doi.org/10.1148/radiol.2020200370.200370>
12. Li K, Wu J, Wu F et al (2020) The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. *Invest Radiol* 55:327–331
13. Zwanenburg A, Vallieres M, Abdalah MA et al (2020) The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 295:328–338
14. Bao C, Liu X, Zhang H, Li Y, Liu J (2020) Coronavirus Disease 2019 (COVID-19) CT findings: a systematic review and meta-analysis. *J Am Coll Radiol*. <https://doi.org/10.1016/j.jacr.2020.03.006>
15. Zhou S, Wang Y, Zhu T, Xia L (2020) CT features of coronavirus disease 2019 (COVID-19) pneumonia in 62 patients in Wuhan, China. *AJR Am J Roentgenol*. <https://doi.org/10.2214/AJR.20.22975-1-8>
16. Li K, Fang Y, Li W et al (2020) CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *Eur Radiol* 30:4407–4416
17. Pan F, Ye T, Sun P et al (2020) Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19). *Radiology* 295:715–721
18. Wang Y, Dong C, Hu Y et al (2020) Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. *Radiology* 296:E55–E64
19. Lyu P, Liu X, Zhang R, Shi L, Gao J (2020) The performance of chest CT in evaluating the clinical severity of COVID-19 pneumonia: identifying critical cases based on CT characteristics. *Invest Radiol*. <https://doi.org/10.1097/RLI.0000000000000689>
20. Wu J, Wu X, Zeng W et al (2020) Chest CT findings in patients with coronavirus disease 2019 and its relationship with clinical features. *Invest Radiol* 55:257–261
21. Song L, Zhu Z, Mao L et al (2020) Clinical, conventional CT and radiomic feature-based machine learning models for predicting ALK rearrangement status in lung adenocarcinoma patients. *Front Oncol* 10:369
22. Chen X, Tang Y, Mo Y et al (2020) A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. *Eur Radiol*. <https://doi.org/10.1007/s00330-020-06829-2>
23. Zhang K, Liu XH, Shen J et al (2020) Clinically applicable AI system for accurate diagnosis, quantitative measurements and prognosis of COVID-19 pneumonia using computed tomography. *Cell*. <https://doi.org/10.1016/j.cell.2020.04.045>
24. Wang Y, Wei W, Liu Z et al (2020) Predicting the type of tumor-related epilepsy in patients with low-grade gliomas: a radiomics study. *Front Oncol* 10:235
25. Liu Z, Li Z, Qu J et al (2019) Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study. *Clin Cancer Res* 25:3538–3547
26. Ji D, Zhang D, Xu J et al (2020) Prediction for progression risk in patients with COVID-19 pneumonia: the CALL score. *Clin Infect Dis*. <https://doi.org/10.1093/cid/ciaa414>
27. Guan WJ, Ni ZY, Hu Y et al (2020) Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. <https://doi.org/10.1056/NEJMoa2002032>
28. Zhou F, Yu T, Du R et al (2020) Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395:1054–1062
29. Guo L, Wei D, Zhang X et al (2019) Clinical features predicting mortality risk in patients with viral pneumonia: the MuLBSTA score. *Front Microbiol* 10:2752
30. Pera A, Campos C, Lopez N et al (2015) Immunosenescence: implications for response to infection and vaccination in older people. *Maturitas* 82:50–55
31. Reynolds HR, Adhikari S, Pulgarin C et al (2020) Renin-angiotensin-aldosterone system inhibitors and risk of Covid-19. *N Engl J Med*. <https://doi.org/10.1056/NEJMoa2008975>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.