
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Medical Informatics 154 (2021) 104545

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

# An AI-based radiomics nomogram for disease prognosis in patients with COVID-19 pneumonia using initial CT images and clinical indicators

Mudan Zhang [a,b], Xianchun Zeng [b], Chencui Huang [c], Jun Liu [d,e], Xinfeng Liu [b], Xingzhi Xie [d], Rongpin Wang [a,b,*]

[a] Medical College of Guizhou University, Guiyang, Guizhou Province 550000, China
[b] Department of Medical Imaging, International Exemplary Cooperation Base of Precision Imaging for Diagnosis and Treatment, NHC Key Laboratory of Pulmonary Immune-related Diseases, Guizhou Provincial People's Hospital, Guiyang 550002, China
[c] AI Lab, Deepwise & League of PhD Technology Co.LTD, Beijing, China
[d] Department of Radiology, the Second Xiangya Hospital, Central South University, No.139 Middle Renmin Road, Changsha, Hunan 410011, China
[e] Department of Radiology Quality Control Center, Changsha, Hunan Province 410011, China

ARTICLE INFO

*Keywords:*
COVID-19 pneumonia[1]
Radiomics[2]
Nomogram[3]
AI[4]
CT[5]

ABSTRACT

*Background:* This study utilized a comprehensive nomogram to evaluate the prognosis of patients with COVID-19 pneumonia. **Methods:** COVID-19 pneumonia data was divided into training set (256 of 321, 80%), internal validation set (65 of 321, 20%) and independent external validation set (n = 188). After image processing, lesion segmentation, feature extraction and feature selection, radiomics signatures and clinical indicators were used to develop a radiomics model and a clinical model respectively. Combining radiomics signatures and clinical indicators, a radiomics nomogram was built. The performance of proposed models was evaluated by the receiver operating characteristic curve (AUC). Calibration curves and decision curve analysis were used to assess the performance of the radiomics nomogram. **Results:** Two clinical indicators that were age and chronic lung disease or asthma and 21 radiomics features were selected to build the radiomics nomogram. The radiomics nomogram yielded an Area Under The Curve[1] (AUC) of 0.88 and accuracy of 0.80 in the training set, an AUC of 0.85 and accuracy of 0.77 in internal testing validation set and an AUC of 0.84 and accuracy of 0.75 in independent external validation set. The performance of radiomics nomogram was better than clinical model (AUC = 0.77, $p < 0.001$) and radiomics model (AUC = 0.72, $p = 0.025$) in independent external validation set. **Conclusions:** The radiomics nomogram may be used to assess the deterioration of COVID-19 pneumonia.

## 1. Background

Although the severity of COVID-19 pandemic has been mitigated [1], the world still faces a shortage of healthcare workers and medical supplies. Thus, having access to the prediction of disease prognosis can enable severe patients to get special care timely, which may prevent their disease from worsening and ultimately to reduce the mortality rate [2].

The American CDC has recommended using reverse transcription polymerase chain reaction (RT-PCR) approaches for detection of SARS-

*Abbreviations:* WHO, World Health Organization; CDC, Centers for Disease Control and Prevention; ARDS, Acute Respiratory Distress Syndrome; COVID-19, coronavirus disease 2019; CT, computed tomography; AI, artificial intelligence; ROC, receiver operating characteristic; LASSO, least absolute shrinkage and selection operator; LR, logistic regression; PCR, multiplexed polymerase chain reaction; RT-PCR, reverse transcription-PCR; ROI, region of interest; AUC, area under the curve; DC, dice coefficient; GLCM, gray-level co-occurrence matrix; GLRLM, gray-level run length matrix; GLSZM, gray-level size zone matrix; GLDM, gray-level dependence matrix; ACC, accuracy; SEN, sensitivity; SPE, specificity; DCA, decision curve analysis; CLD, chronic lung disease or asthma; CRP, C-reactive protein; WBC, white blood cells count; N, neutrophil count; L, lymphocyte count; Cr, creatinine; CK, creatine kinase; LDH, lactate dehydrogenase; VIF, variance inflation factor; EHR, selectronic health records.

* Corresponding author at: Department of Medical Imaging, International Exemplary Cooperation Base of Precision Imaging for Diagnosis and Treatment, NHC Key Laboratory of Pulmonary Immune-related Diseases, Guizhou Provincial People's Hospital. No. 83 Zhongshan East Road, Nan Ming District, Guiyang, Guizhou Province 550002, China.
*E-mail address:* wangrongpin@126.com (R. Wang).

[1] Area Under The Curve (AUC)

https://doi.org/10.1016/j.ijmedinf.2021.104545
Received 8 March 2021; Received in revised form 7 July 2021; Accepted 29 July 2021
Available online 10 August 2021

CoV-2 [3–4]. Studies have shown that the signs of COVID-19 pneumonia including fever, dry cough, fatigue, and other related symptoms. Abnormal laboratory findings including lymphopenia, prolonged prothrombin time and elevated lactate dehydrogenase were also observed in some patients [5–6]. However, there were no clear evidence that showed how RT-PCR, clinical symptoms, and laboratory tests were correlated with the severity of the COVID-19 disease.

In China, CT scans were used as criteria for clinical diagnosis of COVID-19 because of its higher sensitivity of detection of COVID-19 pneumonia [7–9]. A few studies reported that chest CT scans could accurately locate the lesions and the severity or changes in the lesion area during the course of the disease [10–11]. Yuan et al. published a simple CT scoring method to predict mortality of patients with COVID-19 [12]. However, there are limitations of using CT scans to understand disease prognosis. Firstly, evaluation of the disease severity based on routine CT images relies on radiologists' expertise. Additionally, the number and appearance of different types of lesions in chest CT images are often varied and irregular [13]. Furthermore, CT images may appear to be normal during early infection or abnormal even in the absence of symptoms [10]. Thus, more researches are needed to understand the correlation of CT findings with the severity and progression of the disease [14]. Artificial intelligence (AI) technology has been used to improve the efficiency of clinicians in the radiology field. A recent study showed that AI surpassed human-level performance in automatic detection of lung diseases during the COVID-19 outbreak [15]. Zheng et al. developed a deep learning based model for automatic detection of COVID-19 lesions on chest CT [16]. Another study also reported that AI systems performed well the prognosis of COVID-19 pneumonia [17].

Our study developed an AI-based radiomics nomogram to assess the disease prognosis of COVID-19 pneumonia by integrating radiomics signatures from initial CT images with clinical indicators. We hope that the radiomics nomogram can be used by hospital teams for the management of the COVID-19 epidemic, especially in hospitals with a shortage of medical resources.

## 2. Materials and Methods

### 2.1. Ethical approval and patient resources

The data of patients with COVID-19 pneumonia in training set and internal validation set was collected from the Radiology Quality Control Center database of Hunan province, Optics Valley Hospital of Hubei Province [18] and four hospitals in Guizhou province. The data of patients in independent external validation set was from Huoshenshan Hospital, China. This multicenter study was approved by the ethics committees of all hospitals (2020, NO.01 listed in supplements). Because of its retrospective nature, the need to obtain informed consent from the patients was waived. The study was performed according to the principles of the declaration of Helsinki. Fig. 1 shows the workflow of our study.

### 2.2. Diagnostic criteria of COVID-19 pneumonia

The patient data for the study is from confirmed COVID-19 pneumonia patients hospitalized between January 12 and April 30, 2020. In training set and internal validation set, 185 patients have moderate pneumonia while 136 patients have severe pneumonia. In independent external validation set, 101 patients have moderate pneumonia and 87 have severe pneumonia. We confirmed the diagnosis and clinical classification of patients by two associate chief physicians according to the data from electronic health records (EHRs), laboratory information system and Diagnosis and Treatment Protocol for Novel Coronavirus Pneumonia (Trial Version 7, listed in the supplements). If the two doctors are divided, there will be a chief physician to make a final diagnosis that used as the gold standard. Data from patients with a normal initial chest CT and patients with mild symptoms [19], were not included.

### 2.3. Data selection and processing

All images are non-enhanced chest CT images and collected from Picture Archiving and Communication system then be reconstructed at a slice thickness of 1.00 mm. Details of the CT characteristics are listed in Supplementary Table 1. We chose the chest CT images scanned within four days of initial diagnosis [20] as well as the clinical features. Clinical examination and CT scan were finished within 24 hours after admission. If the CT scan or examinations were done more than once, we chose the one closer to initial diagnosis.

Before any data pre - processing steps and model construction, data of 321 patients is randomly split into two individual sets as 80% ($n = 256$) and 20% ($n = 65$) in training and internal validation sets respectively. Data of 188 patients from Huoshenshan hospital is used as independent external validation set. Figure s1 demonstrates the inclusion and exclusion criteria. Secondly, features that had higher than 10% missing rate were excluded. For features with <10% of missing rate, we imputed the average value based on the train set to replace the missing values. The same steps were applied to internal and independent external validation sets.

### 2.4. Image processing

First, regions of interest (ROI) volumes were segmented by an automated segmentation architecture based on three deep learning algorithms. The evaluation of the auto-segmentation accuracy was completed before image segmentation. A B-spline interpolation resampling was used to normalize the voxel size, and the anisotropic voxels were resampled to form isotropic voxels of 1.0 mm × 1.0 mm × 1.0 mm.

### 2.5. Radiomic feature extraction, selection, and radiomics model building

Radiomic feature extraction was carried out using Pyradiomics[2]. Based on the original images, six common feature groups were extracted. They were first-order features ,shape features, gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and gray-level dependence matrix (GLDM) [21]. Next, the training set was standardized with the standard scaler package[3], and the standardized model in the training set was applied to the two validation sets.

We performed a feature dimension reduction process, as high-dimensional features were extracted to select the most relevant features. Additionally, the intra-class correlation coefficient and inter-class correlation coefficient were used to evaluate consistency of measurements made by different observers measuring the same quantity and the same observer measuring different quantity, respectively. (Supplementary Fig. 2). Features with an intra-class correlation coefficient $> 0.75$ and an inter-class correlation coefficient $> 0.75$ were considered to have a satisfactory agreement and were selected for further analysis.

Next, a univariable analysis named K-best was employed [22]. This test selected features according to the K highest scores as computed through the ANOVA F-value between the label and the feature. Features with a significant difference ($p < 0.05$) were selected. The least absolute shrinkage and selection operator (LASSO) feature-selection algorithm was used to screen the most informative image features to avoid the "curse of dimensionality".

After feature extraction and selection, logistic regression (LR) algorithms were trained to construct a radiomics model for the disease prognosis by using a five-fold cross-validation strategy. This process including feature extraction, selection and model construction was all finished in train set then we applied them into internal validation set and independent external validation set.

[2] https://pyradiomics.readthedocs.io/en/latest/. Accessed 6 July 2019
[3] https://scikit-learn.org/stable/modules/preprocessing.html

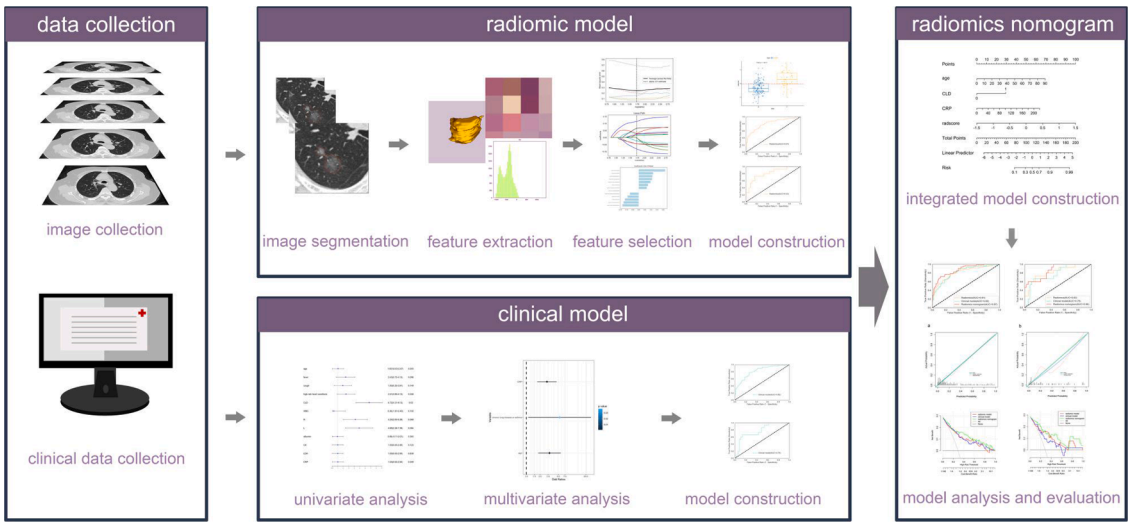


**Fig. 1.** The workflow of radiomics nomogram.

**Table 1**
Clinical characteristics of patients in the training and internal validation sets (n = 321).

| Variable | Training set (n = 256) | | | Internal Validation set (n = 65) | | | |
|---|---|---|---|---|---|---|---|
| | Moderate pneumonia | Severe pneumonia | *p* | Moderate pneumonia | Severe pneumonia | *p* | *p* |
| Age (yr,mean ± SD) | 43.676 ± 16.078 | 64.482 ± 17.303 | 0.000** | 40.000 ± 14.121 | 65.393 ± 17.508 | 0.000** | 0.890* |
| Sex n(%) | | | 0.377# | | | 0.313# | 0.214# |
| Men | 72(48.6%) | 59(54.6%) | | 20(54.1%) | 19(67.9%) | | |
| women | 76(51.4%) | 49(45.4%) | | 17(45.9%) | 9(32.1%) | | |
| Fever(n) | 91(54.5%) | 76(45.5%) | 0.147# | 24(52.2%) | 22(47.8%) | 0.279# | 0.463# |
| Cough(n) | 86(53.8%) | 74(46.2%) | 0.117# | 21(52.5%) | 19(47.5%) | 0.444# | 0.887# |
| Other symptoms(n) | 74(52.1%) | 68(47.9%) | 0.042# | 22(53.7%) | 19(46.3%) | 0.606# | 0.326# |
| high-risk heart conditions(n) | 19(29.2%) | 46(70.8%) | 0.000# | 5(22.7%) | 17(77.3%) | 0.000# | 0.211# |
| CLD(n) | 3(20.0%) | 12(80.0%) | 0.003# | 0(0.00%) | 0(0.00%) | NA | 0.048# |
| WBC (10^9/L) | 5.227 ± 2.955 | 13.922 ± 44.406 | 0.004** | 5.482 ± 2.561 | 15.413 ± 26.147 | 0.030** | 0.886* |
| N (10^9/L) | 3.441 ± 2.746 | 81.182 ± 204.596 | 0.000** | 3.526 ± 2.227 | 134.309 ± 266.247 | 0.000** | 0.037* |
| L (10^9/L) | 1.310 ± 0.590 | 2.580 ± 11.324 | 0.003* | 1.474 ± 0.875 | 0.778 ± 0.434 | 0.000** | 0.314* |
| D dimer (μg/mL) | 89.217 ± 169.431 | 48.046 ± 194.655 | 0.056* | 74.807 ± 153.752 | 46.288 ± 194.845 | 0.487* | 0.244* |
| albumin (g/L) | 40.465 ± 3.909 | 45.293 ± 50.910 | 0.000* | 40.841 ± 4.687 | 44.894 ± 60.191 | 0.051* | 0.741* |
| Cr(μmol/L) | 62.746 ± 33.771 | 95.242 ± 133.566 | 0.010** | 58.138 ± 19.012 | 65.034 ± 19.493 | 0.529* | 0.068* |
| CK (U/L) | 107.980 ± 114.984 | 327.280 ± 1336.189 | 0.007* | 133.097 ± 205.248 | 94.606 ± 102.351 | 0.422* | 0.248* |
| LDH (U/L) | 219.476 ± 86.180 | 618.741 ± 3181.724 | 0.001** | 211.938 ± 105.688 | 253.951 ± 108.985 | 0.630* | 0.041* |
| CRP(mg/L) | 20.301 ± 26.260 | 341.612 ± 1783.586 | 0.000** | 15.297 ± 19.802 | 232.289 ± 767.839 | 0.005* | 0.548* |
| co-infection | 7(18.4%) | 31(81.6%) | 0.000# | 1(8.3%) | 11(91.7%) | 0.000# | 0.450# |

Clinical characteristics and serum biomarkers of patients in the training and internal validation set. Other symptoms including muscle aches, fatigue, headache, nausea, diarrhea, abdominal pain, shortness of breath and vomiting. High-risk heart conditions means companying with any of hypertension, hyperglycemia and dyslipidemia. CLD = chronic lung disease or asthma, WBC = white blood cells count, N = neutrophil count, L = lymphocyte count, Cr = creatinine, CK = creatine kinase, LDH = lactate dehydrogenase, CRP = C reactive protein and co-infection means co-infected with other pathogen infection.
Note: yr = year; std = standard deviation; P-value < 0.05 was considered as a significant difference.
* student's *t* test
** Mann–Whitney *U* test
\# chi-square test

### 2.6. Clinical model building

We used univariate analysis to assess the relationship between clinical factors, serum biomarkers, and disease outcome. The features with $p < 0.05$ were introduced into a multivariable logistic regression analysis to select a combination of clinical factors and serum biomarkers. Next, we built a clinical model with the selected clinical indicators to predict the disease prognosis.

### 2.7. Radiomics nomogram construction and evaluation

A radiomics nomogram was constructed, based on the radiomic features along with the clinical indicators, using a multivariate logistic regression model in the training set. To detect the multi-collinearity among variables in the radiomics nomogram, the collinearity diagnosis was conducted by calculating the variance inflation factor (VIF) for variables in the radiomics nomogram. In the end, the radiomics nomogram was verified in the validation sets. The calibration curves and Hosmer–Lemeshow test were used to assess the relationship between the predicted risks and the actual results. Finally, decision curve analysis (DCA) was used to evaluate the performance of radiomics nomogram.

### 2.8. Statistical analysis

Before modelling, the differences in clinical factors and serum biomarkers between the moderate and severe pneumonia sets were assessed using the Mann–Whitney *U* test or independent *t*-test for continuous variables and the chi-square test for categorical variables (SPSS for

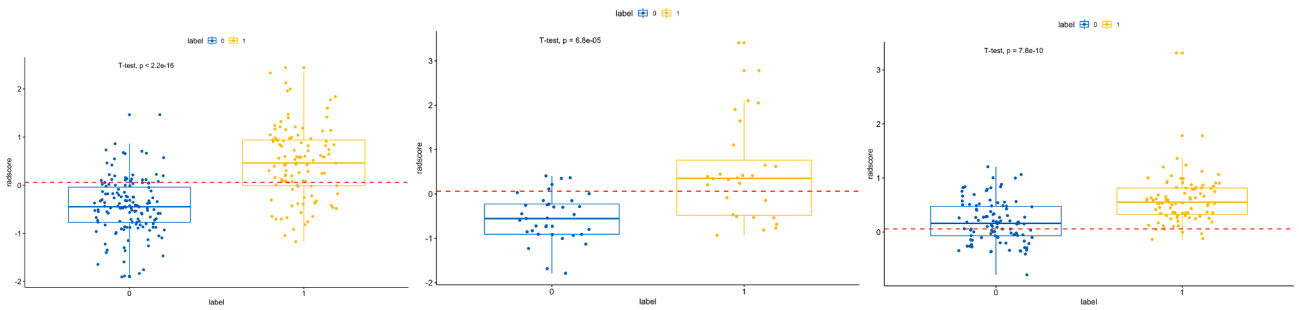


**Fig. 2.** **A,B,C** the box plots of the radomics scores in the training set, internal validation set and independent external validation set. The optimal cut-off value was 0.059 according to the maximized Youden index in the training set. The difference radiomics scores between moderate pneumonia set and severe pneumonia set was computed with *t*-test.

Windows, v.20.0; Chicago, IL). A p-value $< 0.05$ was considered a statistically significant difference.

The area under the curve (AUC) of receiver-operating characteristics (ROC) with 95% confidence interval (95% CI) was used to evaluate the performance of the models. Accuracy was calculated to assess the prediction performance. Differences in the AUC values between different models were estimated by the DeLong test. Youden index was used to classify the patients into the high-risk or the low-risk group.

## 3. Results

### 3.1. Patient characteristics

A total of 509 patient data were included in this study. 321 patient data were included in training and internal validation sets, 188 were included in independent external validation set. Patient characteristics in the training and internal validation sets are listed in Table 1. No significant differences were observed between the training set and the internal validation set in age ($p = 0.890$) and sex ($p = 0.214$). All patients were Asian. High-risk heart conditions (including hypertension, hyperglycemia, and dyslipidemia), chronic lung disease or asthma, white blood cells count (WBC), Neutrophil count (N), Lymphocyte count (L), albumin, creatinine (Cr), Creatine kinase (CK), lactate dehydrogenase (LDH), C-reactive protein (CRP) and co-infected with other pathogen infection differed significantly between the moderate pneumonia and severe pneumonia sets in the training set (p-value $< 0.05$).

### 3.2. Development of clinical model

Seventeen clinical factors and serum biomarkers were included in our study (Table 1). A total of 12 factors were selected from univariate logistic regression analysis, and two predictive indicators were selected from multivariate logistic regression analysis (Supplementary Fig. 3). The clinical model to predict and assess COVID-19 pneumonia was developed based on the following two independent predictive factors: age and chronic lung disease or asthma. Higher total points based on the sum of the assigned number of points for each factor in this model were associated with the risk of severe COVID-19 pneumonia. Clinical model obtained an AUC of 0.83 and accuracy of 0.77(sensitivity, 0.77; specificity, 0.78; F1,0.74) in the training set, an AUC of 0.83 and accuracy of 0.68 (sensitivity, 0.68; specificity, 0.68; F1,0.64) in the internal validation set and an AUC of 0.72 and accuracy of 0.66 (sensitivity, 0.70; specificity, 0.62; F1,0.66) in the independent external validation set. (Fig. 3, Supplementary Table 4)

### 3.3. Feature selection and radiomics model building

Analysis of pneumonia segmentation using CT images on 30 patients randomly selected from the entire data set was done using the Dice coefficient (DC) as the evaluation metric. In 30 patients, the average DC value was $0.825 \pm 0.047$, suggesting a good segmentation result.

In the intra-reader class, 1157 out of 1218 (95%) radiomic features had a good agreement with the ICCs ranging from 0.754 to 0.999. In the inter-reader class, 955 out of 1218 (78%) radiomic features had a good agreement with the ICCs ranging from 0.750 to 0.999. A total of 925 features were selected for further analysis (Supplementary Fig. 2).

In the training set, the univariate analysis identified 347 features with statistically significant association with severe pneumonia. Next, the LASSO logistic regression model was used to minimize the number of features. (Supplementary Fig. 4). The number of predictive features was reduced to 21 (Fig. 2). These 21 predictive features were then evaluated to construct the radiomics model, which included 3 first-order features, one sharp feature, and 17 texture features (GLCM = 6, GLSZM = 4, GLDM = 4, and GLRLM = 3); 9 of the 21 features were transformed by wavelet filters. The box plot of the radomics scores in the training set and internal validation set are shown in Fig. 2. The calculation formula of radiomics scores is listed in supplementary. The difference of radiomics scores between moderate pneumonia set and severe pneumonia set is significant ($p < 0.001$). The details of these features are shown in Supplementary Table 2.

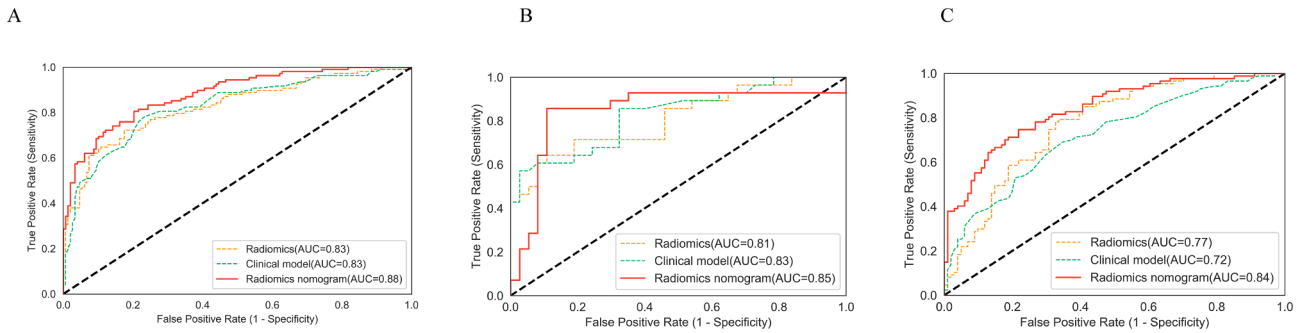


**Fig. 3.** ROC curves in the training set, internal validation set and independent external validation set.

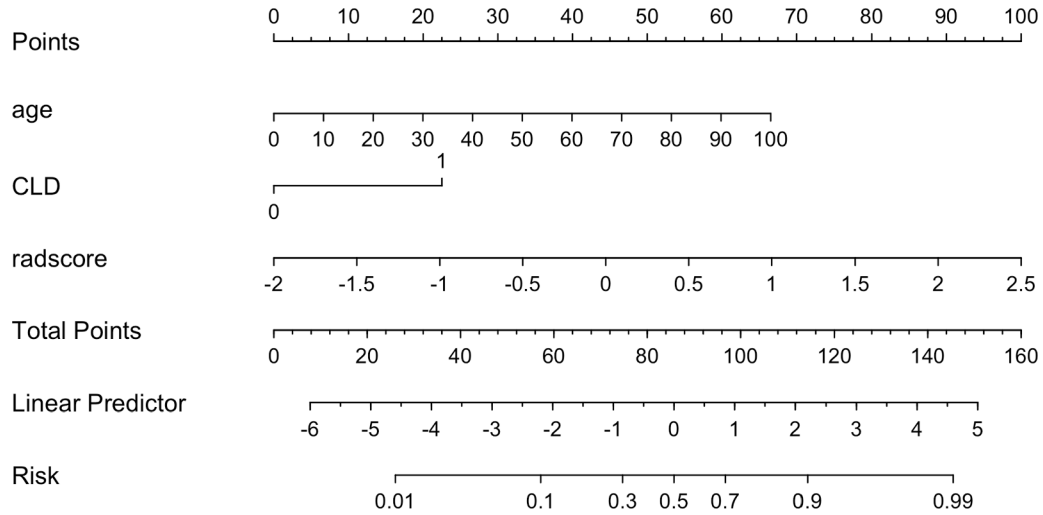


**Fig. 4.** Radiomics nomogram developed in the training set with radiomic features, age and chronic lung disease or asthma (CLD).. Points are assigned for each variable by drawing a line upward from the corresponding variable to the Points line. The sum of points plotted on the total Points line corresponds with the severity.

In general, the radiomics model achieved a satisfying performance with AUC = 0.83, SEN = 0.72, SPE = 0.82, F1 = 0.74, and ACC = 0.78 in the training set, AUC = 0.81, SEN = 0.71, SPE = 0.76, F1 = 0.70, and ACC = 0.74 in the internal validation set and AUC = 0.77, SEN = 0.75, SPE = 0.69, F1 = 0.71, and ACC = 0.77 in the independent external validation set. (Fig. 3, Supplementary Table 4)

### 3.4. Construction and validation of the radiomics nomogram

Multivariable analysis revealed that radiomics score and two clinical indicators were significant independent factors that accessed disease prognosis. A radiomics nomogram incorporating these two variables was built (Fig. 4). By using the collinearity diagnosis, the VIFs for the radiomics score and two clinical indicators were<10, indicating no severe collinearity existing in these factors (Supplementary Table 3). The radiomics nomogram showed good performance with AUC of 0.88 (95% CI, 0.840 to 0.922) in the training set, 0.85 (95% CI, 0.723 to 0.958) in internal validation set and 0.84 (95% CI, 0.781 to 0.894) in independent external validation set. It yielded an accuracy of 0.80 (sensitivity, 0.80; specificity, 0.80; F1, 0.77) in the training set, accuracy of 0.77 (sensitivity, 0.86; specificity, 60.70; F1, 0.76) in the internal validation set and accuracy of 0.74 (sensitivity, 0.83; specificity, 0.68; F1, 0.75) in the independent external validation set (Fig. 3, Supplementary Table 4).

DeLong's test was used to compare the performance of clinical model, radiomics model, and radiomics nomogram. The result showed that the radiomics nomogram was significantly better than clinical model ($p < 0.001$) and the radiomics model ($p = 0.025$) in the independent external validation set.

### 3.5. Clinical usefulness of the radiomics nomogram

The calibration curve showed the agreement between predicted and actual values. The Hosmer–Lemeshow tests were not significant in the training set ($p = 0.973$), internal validation set ($p = 0.932$) and independent external validation set ($p = 0.273$), which suggested there was no significant departure from actual values. The calibration curves of the radiomics nomogram in both sets were shown in Fig. 5. Decision curve analysis (DCA) was used to evaluate the performance of the radiomics nomogram (Fig. 6). If the threshold probability was>20%, the radiomic nomogram was more net benefits than other models and the treat-all or treat-none scheme, indicating its good performance with clinical application.

## 4. Discussion

In this study, we found 21 radiomic features and 2 clinical indicators

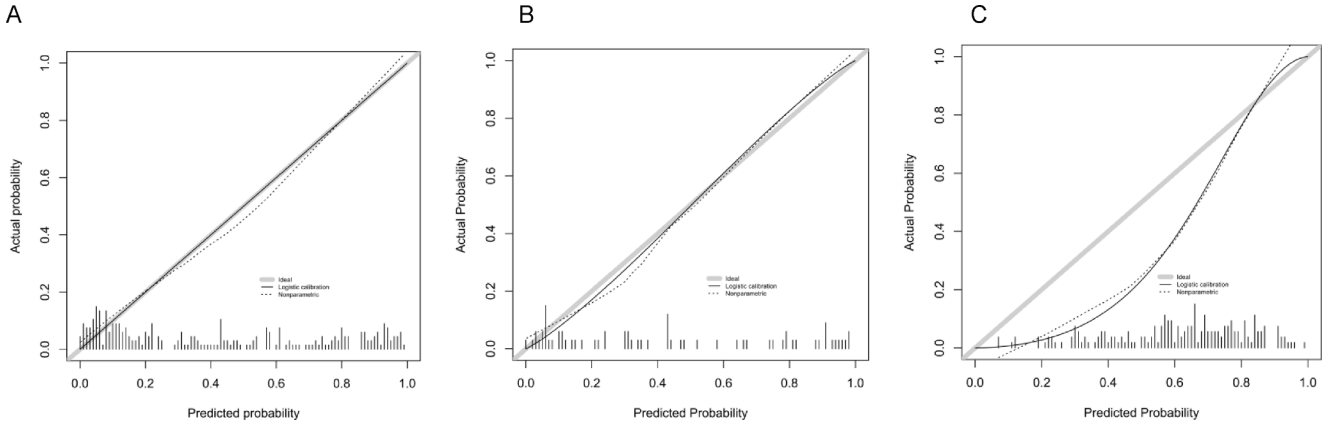


**Fig. 5.** Calibration curves of the radiomic nomogram for the disease prognosis of COVID-19 pneumonia in training set (A), internal validation set (B) and independent external validation set (C). The y-axis represents the actual probability of the COVID-19 pneumonia becoming severe, the x-axis represents the predicted risk. Dashed line was reference line where an ideal nomogram would lie. Dotted line was the performance of radiomics nomogram, while the solid line corrects for any bias in radiomics nomogram.

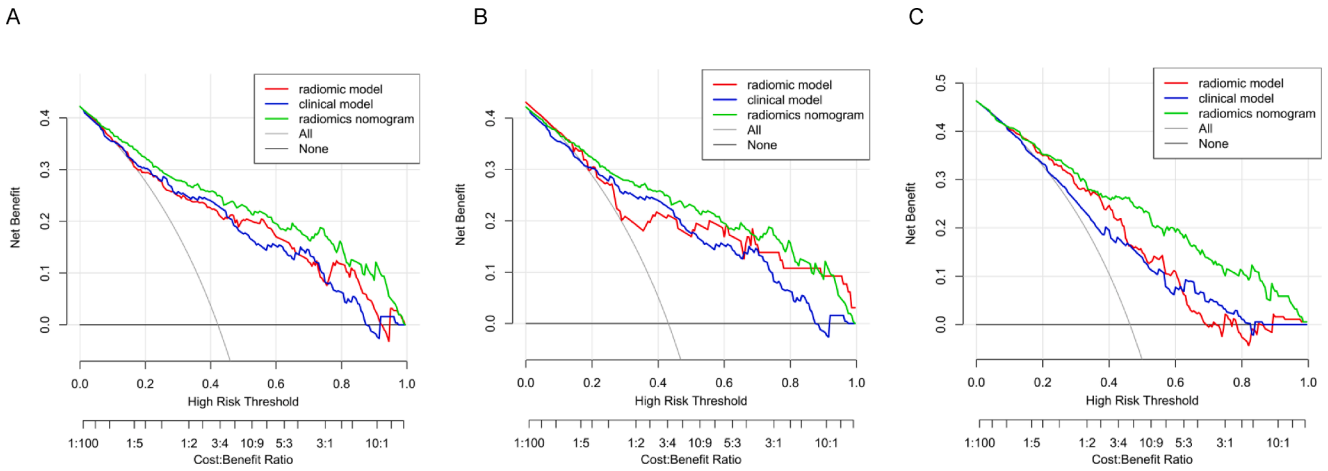


**Fig. 6.** Decision curve analysis(DCA) for the radiomic model, clinical model and radiomics nomogram in training set (A), internal validation set (B) and independent external validation set (C). The y-axis measures the net benefit. using the clinical model,radiomic model and radiomics nomogram in the study to predict COVID-19 pneumonia progress adds more benefit than the treat all patients as severity patients scheme or the treat none scheme. The net benefit of radiomics nomogram was better than clinical model and radiomic model in both two sets and with several overlaps in the training set.

that were significantly related to the disease prognosis of COVID-19 pneumonia. We then constructed and validated a radiomics nomogram for disease prognosis based on radiomics features extracted from initial CT images combined with clinical indicators. Results in our study indicated that the radiomics nomogram performed better than the radiomics model and clinical model. But there was tiny difference of performance between internal validation set and independent external validation set that perhaps caused by the presence of significant differences among the two populations, or maybe different instruments were used in the two sets.

Our results suggested that radiomics could also be a potential tool for evaluating the disease prognosis in COVID-19 pneumonia. FANG et al. investigated the value of radiomics in screening COVID-19, Chen et al. constructed a system based on deep learning for detecting COVID-19 pneumonia on high resolution CT[23]. Other researchers have done similar studies[24–26]. But in our study, we implemented an AI-based semiautomatic method that substantially reduced the time required for obtaining ROI as compared with a completely manual process. And we designed and implemented the models using MINIMAR (MINimum Information for Medical AI Reporting)[27] and checked using IJMEDI checklist[28] that to manage concerns in terms of accuracy and bias. Because with the huge increased number of AI researches of COVID-19, there were some critical doubt: Michael Roberts et al. [29] systematically reviewed 62 studies about COVID-19 modelling and find that none of the models are of potential clinical use due to methodological flaws and/or underlying biases. Similar questions have been raised by other researchers and they all suggested that in order to solve this problem, standard reporting list must be adhered to, for example TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) proposed in 2015 [30], MINIMAR or IJMEDI checklist [28] that we chose in our study. In addition, we used multi-center data to train and evaluate our model. But there are still some drawbacks in our study. For instance, we did not report the brand and model of the analyzer equips about the hematochemical parameters as well as the report about model interpretability and explainability. Moreover, the most meaningful evaluation of an algorithm's performance is to assess it in a clinical setting and take deep insight into the biological meaning of radiomic features [31], which are the future direction of our research.

Our results suggested that the inflammatory signs of patients with severe pneumonia as seen on the initial CT images were different from those of patients with moderate pneumonia, which may be related to the different pathological changes caused by the virus. Another study found that CT findings of viral pneumonia are diverse and may be affected by the immune status of the host and the underlying pathophysiology of the viral pathogen [32]. Mild and moderate cases of COVID-19 mimic common respiratory viral infections. However, histological examination from a patient who died of COVID-19 [33] showed that disease severity is related to ARDS. Thus, early pathological differences in lungs of patients with moderate versus severe symptoms are likely to be detected with chest CT images. Patients with COVID-19 pneumonia may get co-infected with other pathogen in the later stages of the disease, which could aggravate the disease. The pathological changes and CT signs of viral pneumonia are different from those of other pneumonia [13]. Fang et al. suggested that hospital-acquired pneumonia (HAP) is a possibility in the later stages of the disease [6], Bassetti et al. noted that in their study, bacterial infections (pneumonia or bloodstream infection) developed in 10% COVID patients [34].

In conclusion, by using an AI-based method, we established a radiomics nomogram for disease risk prediction based on the initial CT images and clinical indicators of patients with COVID-19 pneumonia. We believe that this radiomics nomogram can be used in the COVID-19 epidemic, especially in situations where there is a shortage of healthcare workers.

## Fundings

The study was supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2019PT320003); the Guizhou Science and Technology Project (QKHPTRC[2019]5803, QKHZC[2020]4Y002) and the Guiyang Science and Technology Project (ZKXM[2020] 4).

## Credit authorship contribution statement

**Mudan Zhang:** Writing – original draft, Conceptualization, Methodology. **Xianchun Zeng:** . **Chencui Huang:** . **Jun Liu:** . **Xinfeng Liu:** . **Xingzhi Xie:** . **Rongpin Wang:** Writing – review & editing, Project administration.

## Declaration of Competing Interest

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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijmedinf.2021.104545.

## References


[1] World Health Organization (WHO). WHO Director-General's statement on IHR Emergency Committee on Novel Coronavirus (2019-nCoV). Geneva: WHO; 2020. Available at: https://www.who.int/dg/speeches/detail/who-director-general-s-statement-on-ihr-emergency-committee-on-novel-coronavirus-(2019-ncov). Accessed Feb. 25, 2021.

[2] World Health Organization, Clinical management of COVID-19,Interim guidance. WHO,27 May 2020. Avalaible at:https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected. Accessed 27 May 2020.

[3] J.P. Broughton, X. Deng, G. Yu, C.L. Fasching, V. Servellita, J. Singh, X. Miao, J. A. Streithorst, A. Granados, A. Sotomayor-Gonzalez, K. Zorn, A. Gopez, E. Hsu, W. Gu, S. Miller, C.-Y. Pan, H. Guevara, D.A. Wadford, J.S. Chen, C.Y. Chiu, CRISPR–Cas12-based detection of SARS-CoV-2, Nat Biotechnol 38 (7) (2020) 870–874, https://doi.org/10.1038/s41587-020-0513-4.

[4] Buddhisha Udugama, Pranav Kadhiresan, Hannah N. Kozlowski, Ayden Malekjahani, Matthew Osborne, Vanessa Y. C. Li, Hongmin Chen, Samira Mubareka, Jonathan B. Gubbay, and Warren C. W. Chan. ACS Nano 2020 1(4), 3822-3835. doi: 10.1021/acsnano.0c02624.

[5] N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang, L.i. Zhang, novel coronavirus pneumonia in Wuhan, China: a descriptive study, The Lancet. 395 (10223) (2020) 507–513, https://doi.org/10.1016/S0140-6736(20)30211-7.

[6] L. Zhang, B. Huang, H. Xia, H. Fan, M. Zhu, L. Zhu, H. Zhang, X. Tao, S. Cheng, J. Chen, Retrospective analysis of clinical features in 134 coronavirus disease 2019 cases, Epidemiol Infect. 3 (148) (2020 Sep), e199, https://doi.org/10.1017/S0950268820002010.

[7] Yicheng Fang, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, Wenbin Ji.Sensitivity of Chest CT for COVID-19 Comparison to RT-PCR. Radiology. 19 Feb.2020. https://doi.org/10.1148/radiol.2020200432.

[8] Heshui Shi, Xiaoyu Han, Nanchuan Jiang, Yukun Cao, Osamah Alwalid, Jin Gu, Yanqing Fan, Chuansheng Zheng. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. The Lancet Infectious Diseases. 2020, 01 April .doi:10.1016/s1473-3099(20)30086-4.

[9] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, Wenzhi, Qian Tao, Ziyong Sun, Liming Xia, Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019, Radiology. 26 (Feb 2020), https://doi.org/10.1148/radiol.2020200642.

[10] Y.N. Duan, J. Qin, Pre- and Posttreatment Chest CT Findings: 2019 Novel Coronavirus (2019-nCoV) Pneumonia, Radiology. 295 (1) (2020 Apr) 21, https://doi.org/10.1148/radiol.2020200323.

[11] F. Pan, T. Ye, P. Sun, S. Gui, B.o. Liang, L. Li, D. Zheng, J. Wang, R.L. Hesketh, L. Yang, C. Zheng, Time Course of Lung Changes On Chest CT During Recovery From 2019 Novel Coronavirus (COVID-19) Pneumonia, Radiology. 13 (Feb 2020), https://doi.org/10.1148/radiol.2020200370.

[12] Mingli Yuan,Wen Yin,Zhaowu Tao,Weijun Tan,Yi Hu. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. PLoS ONE. March 19, 2020.15(3):e0230548. https://doi.org/10.1371/journal.pone.0230548.

[13] H.J. Koo, S. Lim, J. Choe, S.-H. Choi, H. Sung, K.-H. Do, Radiographic and CT Features of Viral Pneumonia, Radiographics. 38 (3) (2018) 719–739, https://doi.org/10.1148/rg.2018170048.

[14] E.Y.P. Lee, M.Y. Ng, P.L. Khong, COVID-19 pneumonia: what has CT taught us? Lancet Infect Dis. 20 (4) (2020 Apr) 384–385, https://doi.org/10.1016/S1473-3099(20)30134-1.

[15] Daniel Shu Wei Ting, Lawrence Carin, Victor Dzau & Tien Y. Wong. Digital technology and COVID-19. Nat 2020. Med 26, 459–461. https://doi.org/10.1038/s41591-020-0824-5.

[16] Xinggang Wang , Member, IEEE, Xianbo Deng, Qing Fu, Qiang Zhou, Jiapei Feng, Hui Ma, Wenyu Liu, "A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization From Chest CT," in IEEE Transactions on Medical Imaging, 2020, Aug, vol. 39, no. 8, pp. 2615-2625, doi: 10.1109/TMI.2020.2995965.

[17] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang, L. Ye, M. Gao, Z. Zhou, L. Li, J. Wang, Z. Yang, H. Cai, J. Xu, L. Yang, W. Cai, W. Xu, S. Wu, W. Zhang, S. Jiang, L. Zheng, X. Zhang, L. Wang, L. Lu, J. Li, H. Yin, W. Wang, O. Li, C. Zhang, L. Liang, T. Wu, R. Deng, K. Wei, Y. Zhou, T. Chen, J. Y. Lau, M. Fok, J. He, T. Lin, W. Li, G. Wang, Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography, Cell. 181 (6) (2020 Jun 11) 1423–1433. e11, https://doi.org/10.1016/j.cell.2020.04.045.

[18] J. Li, H. Luo, G. Deng, J. Chang, X. Qiu, C. Liu, B. Qin, Multidimensional Evaluation of All-Cause Mortality Risk and Survival Analysis for Hospitalized Patients with COVID-19, Int J Med Sci 18 (14) (2021) 3140–3149, https://doi.org/10.7150/ijms.58889.

[19] G.D. Rubin, C.J. Ryerson, L.B. Haramati, N. Sverzellati, J.P. Kanne, S. Raoof, N. W. Schluger, A. Volpi, J.J. Yim, I.B.K. Martin, D.J. Anderson, C. Kong, T. Altes, A. Bush, S.R. Desai, J. Goldin, J.M. Goo, M. Humbert, Y. Inoue, H.U. Kauczor, F. Luo, P.J. Mazzone, M. Prokop, M. Remy-Jardin, L. Richeldi, C.M. Schaefer-Prokop, N. Tomiyama, A.U. Wells, A.N. Leung, The Role of Chest Imaging in Patient Management During the COVID-19 Pandemic: A Multinational Consensus Statement From the Fleischner Society, Chest. 158 (1) (2020 Jul) 106–116, https://doi.org/10.1016/j.chest.2020.04.003.

[20] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R.L. Hesketh, L. Yang, C. Zheng, Time Course of Lung Changes at Chest CT during Recovery from Coronavirus Disease 2019 (COVID-19), Radiology. 295 (3) (2020 Jun) 715–721, https://doi.org/10.1148/radiol.2020200370.

[21] Alex Zwanenburg, Stefan Leger, Martin Vallières, Steffen Löck. Image biomarker standardisation initiative. Radiology. 2019.17 Dec. Available at:.arXiv: 1612.07003.

[22] Gaël Varoquaux AG. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12 (2011) 2825-2830.

[23] Jun Chen, Lianlian Wu, Jun Zhang, Liang Zhang, Dexin Gong, Yilin Zhao, Shan Hu, Yonggui Wang, Xiao Hu, Biqing Zheng, Kuo Zhang, Huiling Wu, Zehua Dong, Youming Xu, Yijie Zhu, Xi Chen, Lilei Yu, Honggang Yu. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study. March 01, 2020. MedRxiv 20021568 [Preprint] Available at:https://doi.org/10.1101/2020.02.25.20021568.

[24] Q. Wu, S. Wang, L. Li, Q. Wu, W. Qian, Y. Hu, L. Li, X. Zhou, H. Ma, H. Li, M. Wang, X. Qiu, Y. Zha, J. Tian, Radiomics Analysis of Computed Tomography helps predict poor prognostic outcome in COVID-19, Theranostics 10 (16) (2020) 7231–7244, https://doi.org/10.7150/thno.46428.

[25] X. Fang, X. Li, Y. Bian, X. Ji, J. Lu, Radiomics nomogram for the prediction of 2019 novel coronavirus pneumonia caused by SARS-CoV-2, Eur Radiol. 30 (12) (2020 Dec) 6888–6901, https://doi.org/10.1007/s00330-020-07032-z.

[26] Fatemeh Homayounieh, Shadi Ebrahimian, Rosa Babaei, Hadi Karimi Mobin, Eric Zhang, Bernardo Canedo Bizzo, Iman Mohseni, Subba R. Digumarthy, and Mannudeep K. Kalra. CT Radiomics, Radiologists and Clinical Information in Predicting Outcome of Patients with COVID-19 Pneumonia. Radiology: Cardiothoracic Imaging, 2020. 2(4), e200322.

[27] Tina Hernandez-Boussard, Selen Bozkurt, John P A Ioannidis, Nigam H Shah, MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care, Journal of the American Medical Informatics Association, 2020. Volume 27, Issue 12, December, Pages 2011–2015, https://doi.org/10.1093/jamia/ocaa088.

[28] Cabitza F, Campagner A. The need to separate the wheat from the chaff in medical informatics. Int J Med Inform. 2021 Jun 2:104510. doi: 10.1016/j.ijmedinf.2021.104510. Epub ahead of print.

[29] Roberts, M., Driggs, D., Thorpe, M. et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. Nat Mach Intell 2021. 3, 199–217 . https://doi.org/10.1038/s42256-021-00307-0.

[30] K.G. Moons, D.G. Altman, J.B. Reitsma, J.P. Ioannidis, P. Macaskill, E. W. Steyerberg, A.J. Vickers, D.F. Ransohoff, G.S. Collins, Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration, Ann Intern Med. 162 (1) (2015 Jan 6) W1–W73, https://doi.org/10.7326/M14-0698.

[31] Michal R. Tomaszewski, Robert J. Gillies, The Biological Meaning of Radiomic Features, Radiology 298 (3) (2021) 505–516.

[32] O. Ruuskanen, E. Lahti, L.C. Jennings, D.R. Murdoch, Viral pneumonia, Lancet. 377 (9773) (2011 Apr 9) 1264–1275, https://doi.org/10.1016/S0140-6736(10)61459-6.

[33] Carlos del Rio, Preeti N. Malani, COVID-19-New Insights on a Rapidly Changing Epidemic, JAMA. 323 (14) (2020) 1339, https://doi.org/10.1001/jama.2020.3072.

[34] M. Bassetti, A. Vena, D.R. Giacobbe, The novel Chinese coronavirus (2019-nCoV) infections: Challenges for fighting the storm, Eur J Clin Invest. 50 (3) (2020 Mar), e13209, https://doi.org/10.1111/eci.13209.