



A clinical-radiomics nomogram based on automated segmentation of chest CT to discriminate PRISm and COPD patients

TaoHu Zhou^{a,b,1}, Yu Guan^{a,1}, XiaoQing Lin^{a,c,1}, XiuXiu Zhou^a, Liang Mao^d, YanQing Ma^e, Bing Fan^f, Jie Li^{a,c}, WenTing Tu^a, ShiYuan Liu^a, Li Fan^{a,*}

^a Department of Radiology, Second Affiliated Hospital of Naval Medical University, No. 415 Fengyang Road, Shanghai 200003, China

^b School of Medical Imaging, Shandong Second Medical University, Weifang, Shandong 261053, China

^c College of Health Sciences and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China

^d Department of Medical Imaging, Affiliated Hospital of Ji Ning Medical University, Ji Ning 272000, China

^e Department of Radiology, Zhejiang Provincial People's Hospital, Affiliated People's Hospital of Hangzhou Medical College, Hangzhou, ZJ, China

^f Department of Radiology, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang, China

ARTICLE INFO

Keywords:
Radiomics
COPD
PRISm

ABSTRACT

Purpose: It is vital to develop noninvasive approaches with high accuracy to discriminate the preserved ratio impaired spirometry (PRISm) group from the chronic obstructive pulmonary disease (COPD) groups. Radiomics has emerged as an image analysis technique. This study aims to develop and confirm the new radiomics-based noninvasive approach to discriminate these two groups.

Methods: Totally 1066 subjects from 4 centers were included in this retrospective research, and classified into training, internal validation or external validation sets. The chest computed tomography (CT) images were segmented by the fully automated deep learning segmentation algorithm (Unet231) for radiomics feature extraction. We established the radiomics signature (Rad-score) using the least absolute shrinkage and selection operator algorithm, then conducted ten-fold cross-validation using the training set. Last, we constructed a radiomics signature by incorporating independent risk factors using the multivariate logistic regression model. Model performance was evaluated by receiver operating characteristic (ROC) curve, calibration curve, and decision curve analyses (DCA).

Results: The Rad-score, including 15 radiomic features in whole-lung region, which was suitable for diffuse lung diseases, was demonstrated to be effective for discriminating between PRISm and COPD. Its diagnostic accuracy was improved through integrating Rad-score with a clinical model, and the area under the ROC (AUC) were 0.82 (95 %CI 0.79–0.86), 0.77(95 %CI 0.72–0.83) and 0.841(95 %CI 0.78–0.91) for training, internal validation and external validation sets, respectively. As revealed by analysis, radiomics nomogram showed good fit and superior clinical utility.

Conclusions: The present work constructed the new radiomics-based nomogram and verified its reliability for discriminating between PRISm and COPD.

1. Introduction

Preserved ratio impaired spirometry (PRISm), called the unclassified spirometry or restrictive pattern as well, is defined as the FEV1/predicted < 80 %, even though the ratio of forced expiratory volume in 1 s (FEV1)/forced vital capacity (FVC) ratio (≥ 0.70) is preserved [1]. The global prevalence of PRISm is remarkably high, with estimates ranging from

5 % to 20 % in various populations [1,2]. PRISm is a precursor state to Chronic Obstructive Pulmonary Disease (COPD), a leading cause of morbidity and mortality worldwide [3]. Despite the high prevalence of PRISm, there is a significant lack of recognition among patients due to the non-specific or mild symptoms that are often present in the early stages of the condition. This lack of awareness, coupled with the absence of routine pulmonary function test, leads to a substantial number of

* Corresponding author.

E-mail address: fanli0930@163.com (L. Fan).

¹ These authors contributed equally to this work.

individuals with PRISm remaining undiagnosed. Identifying PRISm is crucial for early intervention and prevention strategies, which can mitigate the individual and societal impact of this disease. PRISm has been assumed as the precursor for COPD, which may be the candidate interventional target for preventing to develop COPD [4]. PRISm, that is identified, offers an opportunity for interventions that may prevent or delay the onset of COPD, thus reducing the significant morbidity and mortality associated with this disease [5]. As previously reported, over 50 % of PRISm patients may progress into COPD in 5 years; but 15 % of people return to the normal spirometry [6,7]. As a result, PRISm may be seen as the early form of COPD occurrence. The distinction between PRISm and COPD is very important. Due to the mild clinical presentations and a lack of routine PFT test, the clinical diagnosis of PRISm is often underdiagnosed. In contrast to PFT, the chest CT is more popular in the routine physical examinations. Moreover, PRISm is related to the airway disease and emphysema based on the imaging research, which may influence the progression into COPD [8,9]. According to one COPDGene research [10], PRISm patients are more likely to develop disease deterioration than mild COPD patients and experience more severe disease. Therefore, precise identification can provide valuable information for clinical classification and treatment selection, so as to help make decision of pretreatment.

In clinical practice, PRISm and COPD are diagnosed through pulmonary function test (PFT). However, the conventional approach is limited by insensitivity and lack of reproducibility [11,12]. To be specific, PFT is measured in a quite complicated process, and patients can difficultly understand or abide by the doctors' requirements [13]. Additionally, PFT is unable to offer the intuitive anatomical details or morphological alterations, like bronchial wall thickening and emphysema subtype [14,15]. At present, an accurate, and repeatable tool is needed to make a differential diagnosis of PRISm and COPD patients more comprehensively based on chest computed tomography (CT) examination. Relative to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria or additional imaging tools, CT is deemed to exhibit the highest effectiveness on COPD characterization and quantification [16]. Experienced radiologists may not always capture the subtle changes in lung function that characterize PRISm. Furthermore, these methods can be subjective and vary between different radiologists, leading to potential misclassification. Radiomics features, which are unavailable to human naked-eyes, on the imaging of lung disease are considered to be the most updated for clinicians [17]. Nonetheless, a critical requirement for radiomics is the precise delineation of the region of interest (ROI); however, in COPD, the widespread distribution of pathological changes throughout the lungs makes it challenging to accurately outline the ROIs. This difficulty in defining ROIs limits the application of radiomics in COPD. To address this task, we used an automatic whole-lung algorithm for chest CT image segmentation. Previous research has demonstrated the predictive validity of fully automatic algorithm in COPD [18–20].

The present work focused on developing and validating a CT-based radiomics model based on the multicenter database to discriminate PRISm from COPD to overcome the underdiagnosis of PRISm due to the mild symptoms and lack of routine PFT test.

2. Materials and methods

2.1. Patients

The present work gained approval from ethics committees of four study centers. No informed consent was needed since this was a retrospective study (ClinicalTrials.gov registration number ChiCTR2300069929, CSD-COPD cohort). During the period between February 2013 and December 2022, 1066 patients from 1 center received complete PFT and were diagnosed as PRISm or COPD. Medical records of all patients were reviewed, which included clinical features, and serial chest CT scans. Patients with all the following criteria were

included: 1) chest CT and PFT in one hospital; 2) chest CT and PFT with an interval less than 2 weeks; and 3) those with available thin-slice (< 2 mm) chest CT images. Patients with any of the following criteria were excluded: 1) those with additional underlying thoracic disorder (pulmonary atelectasis, pneumonia, lung nodules > 6 mm or masses, pleural effusion or asthma); 2) cancers; and 3) those undergoing spine implants or with extensive image artifacts. Finally, 1066 patients (789 males, 277 women; mean age, 65.01±11.64 years) were enrolled and randomized as training (n = 651) and internal validation (n = 278) in a 7:3 ratio. Patients recruited from the other 3 centers were used assigned to the validation (n =137) cohort.(Fig. 1)

2.2. Acquisition of CT images and PFTs

The CT acquisition parameters are presented in Supplement. The PFT apparatus (Ganshorn Medizin Electronic GmbH; CHEST Multifunction Spirometer HI-801, Japan; Masterscreen PFT Pro, Carefusion, Netherlands; Carefusion GmbH, Hoechberg, Germany) was used to measure pulmonary function parameters (FEV1, FVC). In COPD, the PFT diagnostic criteria were shown below: FEV1/FVC < 0.7 and FEV1 increase by < 200 ml after bronchodilator use. On the contrary, patients with FEV1/FVC ≥ 0.70 and FEV1 < 80 % were deemed as PRISm. Based on the above criteria, patients in training, internal validation, and independent external validation sets were further classified as COPD or PRISm group.

2.3. Image segmentation and preprocessing of whole-lung CT images

The deep-learning model of open access U-net (R231) (<https://github.com/JoHof/lungmask>) was used to automatically segment CT images. The model is trained with diverse large datasets covering various visual variability, and its reliability has been demonstrated [21].

After automatic segmentation of both left and right lungs, both lungs were integrated into the pooled ROI (Fig. S1).

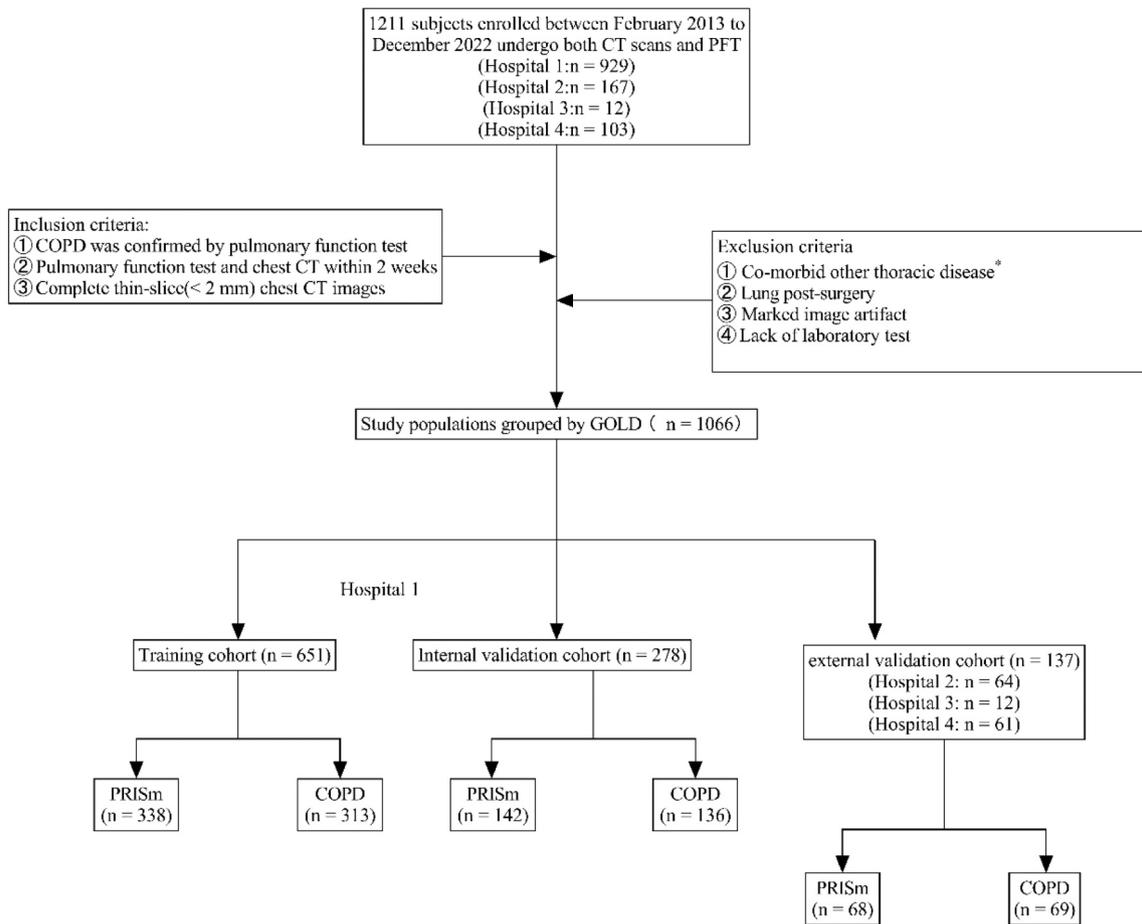
Manual segmentation has been frequently recognized as the ground truth. Therefore, this study evaluated whether automatic segmentation results were consistent with manual segmentation results based on 20 patients randomly selected from our datasets. CT images from totally 20 patients were manually segmented with ITK-SNAP software (version 3.8.0, www.itksnap.org). Next, the Dice index, which helps objectively quantify spatial overlap of both contours, was evaluated to explore whether fully automatic segmentation results conformed to manual counterparts. Afterwards, automatic segmentation was performed in the rest patients.

Before the extraction of radiomic features, image preprocessing was completed in three steps. At first, linear interpolation was used for resampling images to 1 mm×1 mm×1 mm. Secondly, a bin width of 25 was used for grayscale discretization to reduce the effect of imaging noise. Thirdly, the mixed noise during image digitization was removed through log and wavelet image filters, while high- or low-frequency features were acquired.

2.4. Collection and screening of radiomic features

Using open-source package PyRadiomics (version 3.0.1, <https://pyradiomics.readthedocs.io/en/latest/>), totally 1218 lung radiomic features were obtained in every ROI, including the first-order, gray level co-occurrence matrix, gray level size zone matrix, gray level run length matrix, gray level dependence matrix, and shape features. Detailed radiomics features are supplemented in the [supplementary materials](#). All radiomic features obtained using the software were in consistence with the image biomarker standardization initiative. Then, the features were normalized with the Z-score method. Differences at the numerical scale were removed.

Then, the optimal radiomic features were selected using in steps below. Firstly, we eliminated unnecessary features whose correlation



* Indicates other thoracic disease: pneumonia, pulmonary atelectasis, lung nodules larger than 6 mm or masses, asthma, and pleural effusion.

Fig. 1. Patient recruitment process at four centers.

coefficient with other features was >0.90 . Secondly, unnecessary and unrelated features were removed through the minimal redundancy maximal relevance algorithm, which has been demonstrated with high effectiveness and reliability in selecting radiomic features and contributes to considering feature importance and correlation for selecting the optimal feature subset [22,23]. At last, ten-fold cross-validation was conducted through penalty parameter adjustment and the least absolute shrinkage and selection operator (LASSO) regression algorithm. Notably, we chose the optimal feature dataset that had the lowest cross-validation binomial deviation, and defined non-zero coefficients as weights of chosen features, which represented the association of feature with COPD. LASSO has been extensively utilized as an embedding approach to select radiomic features from high-dimensional data[21]. Eventually, the chosen features were linearly combined with coefficient vectors to determine the Rad-score values of all patients, so as to construct the radiomic model.

2.5. Model establishment, radiomic nomogram, and performance assessment

Three models, namely, clinical, radiomic, and combined models were established. Risk variables with statistical significance were obtained from univariate logistic regression. Later, both clinical and combined models were established by multivariate regression. The combined model was visualized by producing the radiomic nomogram. It was developed by combining the significant variables of the clinical factors and the Rad-score. Its prediction performance was determined. Area under the curve (AUC), accuracy, sensitivity, specificity, positive

predictive value (PPV), and negative predictive value (NPV) were used to evaluate model performance. AUC of these three models were compared by DeLong tests. Afterwards, the combined model was calibrated based on calibration curves (Hosmer–Lemeshow test) to evaluate the goodness-of-fit of the nomogram. Decision curve analysis (DCA) was used for assessing its clinical utility.

2.6. Statistical analysis

In study work, statistical analysis was carried out with R software (version 4.2.2; <http://www.Rproj.org>) and IBM SPSS Statistics (version 26.0; IBM Corp., New York, USA). Measurement data were indicated by mean \pm standard deviation. Continuous data consistent with normal distribution were compared by Student's unpaired t-test, while those with non-normal distribution were explored by Mann–Whitney U test. Chi-square test was used for comparing categorical variables between groups. To obtain independent predicting factors, multivariate regression was carried out. $P < 0.05$ represented significant difference. “glmnet” package was utilized for LASSO regression. Moreover, “rms” package was used to draw calibration plots and conduct multivariate logistic regression. The receiver operating characteristic (ROC) curves were drawn with the ROC package, while DCA was carried out with “rmda” package.

3. Results

3.1. Evaluation of the consistency between fully-automatic and manual segmentation

The Dice index, which objectively measures spatial overlap of two regions, was used to evaluate segmentation results. Therefore, the average Dice index was (0.97 ± 0.06) between automatic and manual segmentation (Fig. S2).

3.2. Clinical features comparison between COPD and PRISm

Table 1 and Table S2 display patient characteristics in training and validation sets. There were 651, 278 and 137 in training, internal validation and external validation sets, respectively. Age between PRISm and COPD groups was of significant difference (p>0.05) among three sets. However, no significant difference was observed in height among three sets. Weight and body mass index (BMI) differed significantly between two groups of training and internal validation sets. Gender was significantly different between two groups only in the training set. In addition, the smoking status also only exhibited statistically significant difference in internal validation set.

3.3. Feature screening and radiomics signature establishment

There were 1218 features obtained on segmented chest CT images from each of the 1,066 subjects. Following the aforementioned screening methods, an optimized feature subset was selected based on LASSO, and the model was established after appropriate processing. When selecting the best radiomics features, LASSO method with 10-fold cross validation is applied to process the results, as shown in Fig. 2A, B. When the feature number was determined, we chose the feature subset with greatest prediction ability and assessed the associated coefficients. At last, 15 radiomics features with high importance were chosen, and the radiomics signature was established (Fig. 2C). The chosen features were then weighted by the corresponding coefficients, and the weights were added up to obtain the Rad-score. The formula used to calculate Rad-score can be obtained from the Supplementary Results. According to our results, AUCs of our constructed radiomics signature were 0.82 (95 % CI: 0.79–0.85), 0.77 (95 % CI: 0.71–0.82) and 0.80 (95 % CI: 0.72–0.87) in training, internal validation and external validation sets separately (Fig. 3).

3.4. Establishment and validation of the clinical model

According to univariate regression of clinical features, age, weight, BMI, smoking status and gender were dramatically associated with COPD. Based on multivariate regression, age, weight, and gender

independently predicted COPD (Table 2). Finally, logistic regression was adopted for constructing the clinical model by incorporating factors including age, weight, and gender. Based on our findings, AUCs were 0.68 (95 % CI: 0.64–0.72), 0.62 (95 % CI: 0.56–0.69) and 0.73 (95 % CI: 0.65–0.82) in training, internal validation and external validation sets, respectively.

3.5. Radiomics nomogram development and validation

Multivariate logistic regression was performed to construct the nomogram model by incorporating age, weight, gender, and radiomics signature (Fig. 4A). Using patients from the three sets, our constructed nomogram exhibited high calibration performance, and Hosmer-Lemeshow test did not reveal any obvious difference in training versus two validation sets (P>0.05), suggesting no deviation from the fit (Fig. 4B-D). The prediction performance in the discrimination between PRISm and COPD of our nomogram for training and two validation sets were 75 %, 70 % and 73 %, separately, the sensitivities were 68 %, 73 % and 90 %, while specificities were 83 %, 68 % and 56 %, separately. Besides, nomogram model demonstrated AUCs of 0.82, 0.77, and 0.84 in three datasets for distinguishing between PRISm and COPD patients (Table 3, Fig. 3). The Delong test indicated that both nomogram and radiomics models outperformed the clinical model in the training set with significant differences (Z=7.42 and 6.53, p<0.001), while their diagnostic performances were comparable (Z=0.86, p=0.39). Similarly, in the internal validation set, nomogram and radiomics models surpassed the clinical model with significant differences (Z=5.09 and 4.34, p<0.001), yet showed no significant difference from each other (Z=0.89, p=0.37). In the external validation set, nomogram model was superior to both the clinical and radiomics models with statistical significance (Z=4.63 and 2.96, p<0.001 and p=0.003), though the radiomics model's performance was not significantly different from the clinical model (Z=1.30, p=0.19). In conclusion, nomogram model outperforms the clinical model. The Decision Curve Analysis (DCA) is a novel and clinically practical method used for evaluating and comparing diagnostic models. It provides a quantitative approach to assess the clinical utility of a prediction model by integrating the predicted probabilities with the decision thresholds that matter to patients or physicians. According to DCA results, our constructed nomogram model produced greater net benefits than clinical model for training set within the probability range of 0.1–1 (Fig. 4F). It provides actionable information that can be used to tailor medical interventions to individual patients, thus enhancing the clinical practicality of the nomogram model. An example of the nomogram in use is shown in Fig. 5. Similar to the points scoring system, we assigned points for each predictor and then equated these predictors with the risk of PRISm. We can read the top score scale upward from the predictors to determine the points score associated with patient age, gender, weight, and the Rad-score. Once a

Table 1

Clinical Factors of PRISm and COPD Patients in the Training, Internal validation and External validation Sets.

Clinical Factors	Training Set (n =651)			Internal validation Set (n =278)			External validation Set (n =137)		
	PRISm (n = 338)	COPD (n = 313)	P value	PRISm (n = 142)	COPD (n = 136)	P value	PRISm (n = 69)	COPD (n = 68)	P value
Age	63.6±11.3	67.5±9.7	<0.001	63.6±13.6	67.4±10.1	0.008	55.9±1.9	68.2±1.0	<0.001
Gender	Male	222 (65.7)	258 (82.4)	108 (76.1)	109 (80.1)	0.497	44(63.8)	48(70.6)	0.722
	Female	116 (34.3)	55 (17.6)	34 (23.9)	27 (19.9)		25(36.2)	20(29.4)	
Height	162.9±8	163.2±7.5	0.606	174.4±126.2	163.6±7.7	0.320	164.5±0.9	162.9±0.9	0.293
Weight	67.1±13.4	64±11.3	0.002	67.9±11.9	62.8±11.1	0.0002	67.4±2.0	61.8±1.2	0.067
BMI	25.2±4.4	24±3.6	<0.001	25.1±4.5	23.4±3.4	0.0003	24.7±0.6	23.3±0.4	0.081
Smoking status	Current Smoker	77 (22.8)	85 (27.2)	34 (23.9)	39 (28.7)	0.666	42(60.9)	10(14.7)	<0.001
	Former Smoker	38 (11.2)	71 (22.7)	24 (16.9)	21 (15.4)		10(14.5)	18(26.5)	
	Non-smoker	223 (66.0)	157 (50.2)	84 (59.2)	76 (55.9)		17(24.6)	40(58.8)	

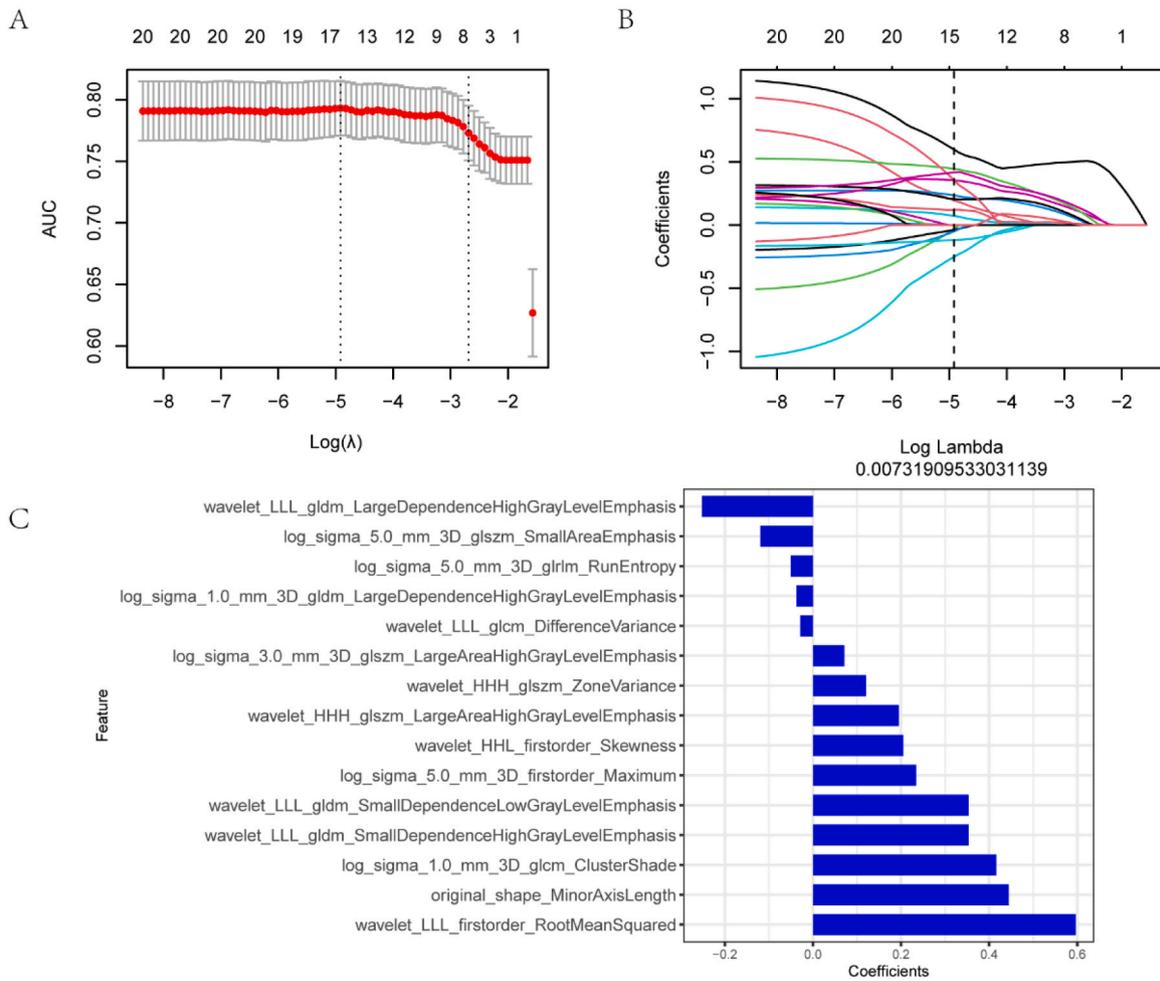


Fig. 2. Screening of radiomics features by the least absolute shrinkage regression and the histogram showing Rad-score of those selected features. A. 10-fold cross-validation conducted to select tuning parameter (λ) in LASSO model on the basis of minimum criteria. Binomial deviance obtained through cross-validation by LASSO regression was plotted with $\log(\lambda)$. The best λ value was chosen to be 0.007. B. 10-fold cross-validation was performed to select the value in λ for drawing the black vertical line. The 15 obtained features with nonzero coefficients were shown in the plot. C. The x- and y-axes stand for radiomics coefficients and those 15 selected radiomics features separately.

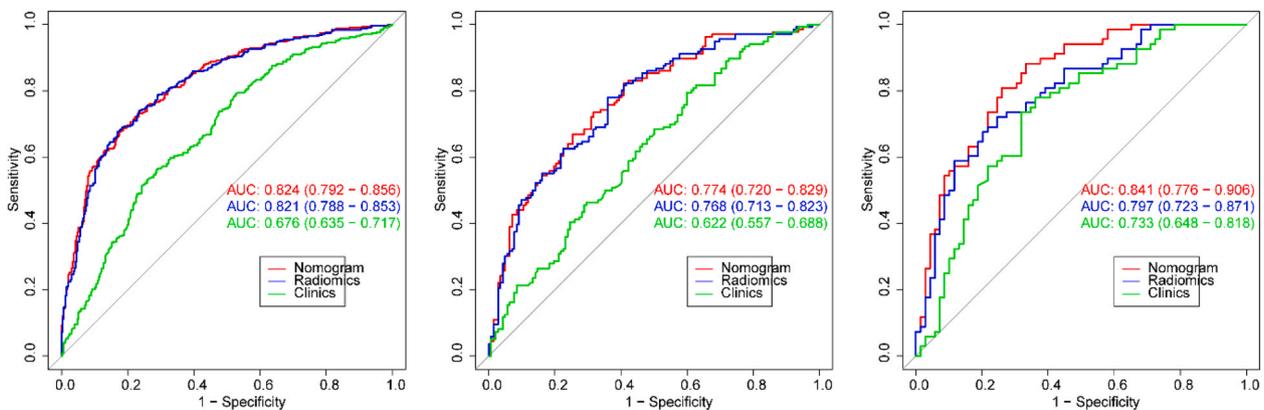


Fig. 3. AUC of Rad-score, clinical model, and combined model for training, internal validation, and external validation sets. The combined model outperformed clinical model and Rad-score in predictive performance for three sets.

score has been assigned to each predictor, an overall score is calculated. Then, the total score is converted to the probability of COPD by reading the associated probability of COPD from the total point scale.

4. Discussion

The present retrospective work applied the U-net-based deep learning (DL) model to segment the whole lung image and constructed the radiomics nomogram for differentiating PRISM from COPD subjects.

Table 2
Univariate and Multivariate Analysis in Relation to Risk Factors in the Training Set.

Clinical Risk Factors	Univariate Analysis		Multivariate Analysis (Clinical model parameters)		Multivariate Analysis (Combined model parameters)	
	OR (95 % CI)	p Value	OR (95 % CI)	p Value	OR (95 % CI)	p Value
Age	1.04(1.02–1.05)	<0.001	1.03(1.01–1.05)	<0.001	1.02(1.01–1.04)	0.01
Weight	0.98(0.97–0.99)	0.002	0.97(0.96–0.98)	<0.001	0.99(0.98–1.01)	0.45
Height	1.00(0.99–1.0)	0.508	-	-	-	-
BMI	0.92(0.89–0.96)	<0.001	-	-	-	-
Smoking	0.75(0.63–0.91)	0.003	-	-	-	-
Gender	0.41(0.28–0.59)	<0.001	0.34(0.23–0.51)	<0.001	0.93(0.58–1.48)	0.75
Radscore	3.14(2.59–3.80)	<0.001	-	-	3.01(2.46–3.70)	<0.001

The results showed that our constructed radiomics nomogram that incorporated clinical factors and radiomics signature exhibited the highest prediction performance in three sets, and the AUCs were 0.82, 0.77 and 0.84, separately. This radiomics nomogram can be the non-invasive, user-friendly, and individualized approach for differentiating PRISm from COPD sets with high performance.

Univariate and multivariate logistic regression identified age, gender and weight as independent predictive factors that discriminated PRISm from COPD. The findings suggest that older, less-weight male patients are more likely to be diagnosed with COPD, indicating a cause for concern in this demographic. In turn, young female patient with greater weight is associated with increased risk of PRISm. One study [24] found that PRISm patients are young, and have an increased risk of diabetes and BMI compared to COPD patients, conforming to our findings on age trajectories. The increases in body mass and adiposity have been recognized to affect lung volume and spirometry measurements[25–27]. Proportionate reduction of FEV1 and FVC and the resulting unchanged FEV1/FVC ratio can be detected among patients with obesity and overnight. Nonetheless, decreases in FEV1 and FVC levels can be quite small even though they reach the statistical significance level, and, their levels often remain in the normal range, even in extremely obese patients [26,27]. In an additional study of 2229 patients (including 141 PRISm patients and 1743 COPD patients), Sang[28] reported that the elderly and male patients accounted for greater percentages among COPD patients, conforming to prior results [29].

Firstly, CT scanning is widely available in clinical practice and can provide a wealth of high-resolution data that captures the complex structural changes within the lungs. Secondly, the prevalence of CT scans far exceeds that of PFT, particularly in the context of lung cancer screening. More importantly, CT image-based radiomics can objectively, reliably and quantitatively assess images, without being affected by inter-reader variability. The application of radiomics, especially whole lung radiomics, in COPD has been proven to be feasible [30], not only in the diagnosis of COPD, but also predicting comorbidities of COPD and severity evaluation of COPD [31,32]. Among 1,218 radiomics features collected on CT images, fourteen radiomics features extracted on Laplacian of Gaussian (LoG) and wavelet transformed images, and one shape radiomics feature were obtained to be significant elements for constructing a radiomics model, and the AUCs were 0.82, 0.77, and 0.80 for three sets, separately. Texture features contribute to quantifying data that can hardly be perceived intuitively, like tissue distribution and texture patterns [33]. To extract radiomic features, LoG was utilized to preprocess images. The LoG filter is capable of smoothing images by diverse parameter scale settings, thus contributing to decreasing the noise interference. Besides, it helps provide more textural details, thereby increasing the phenotypic feature capturing effectiveness for mapping to heterogeneity [34]. Additionally, when original images are turned to diverse frequency domains, wavelet transform helps obtain multiscale image and multifrequency domain information [35]. For disorders that can hardly be depicted using simple visual features, the wavelet transformed images-extracted high-dimensional abstract features usually offers diverse perspectives for capturing the concealed information not easy to be seen under visual evaluation.

In this work, our constructed clinical model achieved low diagnostic accuracy of 0.64, 0.58, and 0.67 for three sets, separately. While our developed nomogram outperformed this clinical model in all the three sets, and its AUCs were 0.82, 0.77, and 0.73, while its diagnostic accuracy was 0.75, 0.70, and 0.73, separately, in three sets.

Relative to prior reports, this work shows the following strengths. Firstly, based on the previous study [36], a discrimination model was established to diagnose and evaluate the PRISm according to CT quantitative parameters from inspiratory/expiratory CT images. The AUC of quantitative model for differentiating between mild-to-moderate COPD and PRISm was 0.852. We believe that the acquisition and analysis of paired inspiratory/expiratory CT images provide more valid information. However, expiratory scans are needed, which are usually not available in clinic and induce more radiation exposure to patients, has limited the generalizability of the technique. According to our results, our radiomics nomogram on the basis of an inspiration CT scan showed similar performance in differentiating between PRISm and COPD (AUC = 0.824). Secondly, nodule ROI segmentation accounts for a key process when extracting radiomics features. The U-net-based DL model was adopted to segment the whole lung CT images for improving segmentation accuracy and reliability. Typically, the whole lung segmentation results based on U-net-based DL model was consistent with manual segmentation. Thirdly, unlike conventional statistical models that present outcomes through numerical equations, our constructed radiomics nomogram, which can generate individual probabilities of COPD occurrence by integrating Rad-score and easily available deterministic clinical variables, was particularly valuable in situations that demand quick, reliable results and where the ability to visualize and interpret complex data was important. It outperformed the clinical model, and AUCs between the two models were significantly different in all the three sets (DeLong test: $p < 0.001$, $p < 0.001$, and $p < 0.001$, separately).

Nonetheless, certain limitations should be noted in this work. Firstly, there was a risk of selection bias in this retrospective research. Although selection bias is inevitable, we still adopt various methods (eg, data standardization or normalization and regularization) to avoid it as much as possible. However, retrospective studies can leverage the existing data to provide valuable insights for distinguishing COPD from PRISm, laying the foundation for future prospective studies. Secondly, we just assessed CT radiomics features, but not the commonly used CT qualitative and quantitative parameters useful for assessing COPD. At last, there might be certain side effects of CT images obtained using diverse scanners, while image resampling and data preprocessing might reduce rather than remove the biases. In our opinion, radiomic features facilitates to discriminate between these two diseases through the reproductive and quantitative use of several parameters, thus improving routine diagnosis and supporting clinical decision-making in a cost-effective manner. We suggest that future studies consider adapting our method to different patient populations, which may exhibit unique clinical characteristics that could influence the outcomes. This could include patients with comorbidities, or those from diverse geographical regions. More advanced AI algorithms, such as deep learning, regarding radiomics, especially those focusing on chest CT follow-up and radiomics feature alterations, are needed. Additionally, to increase the

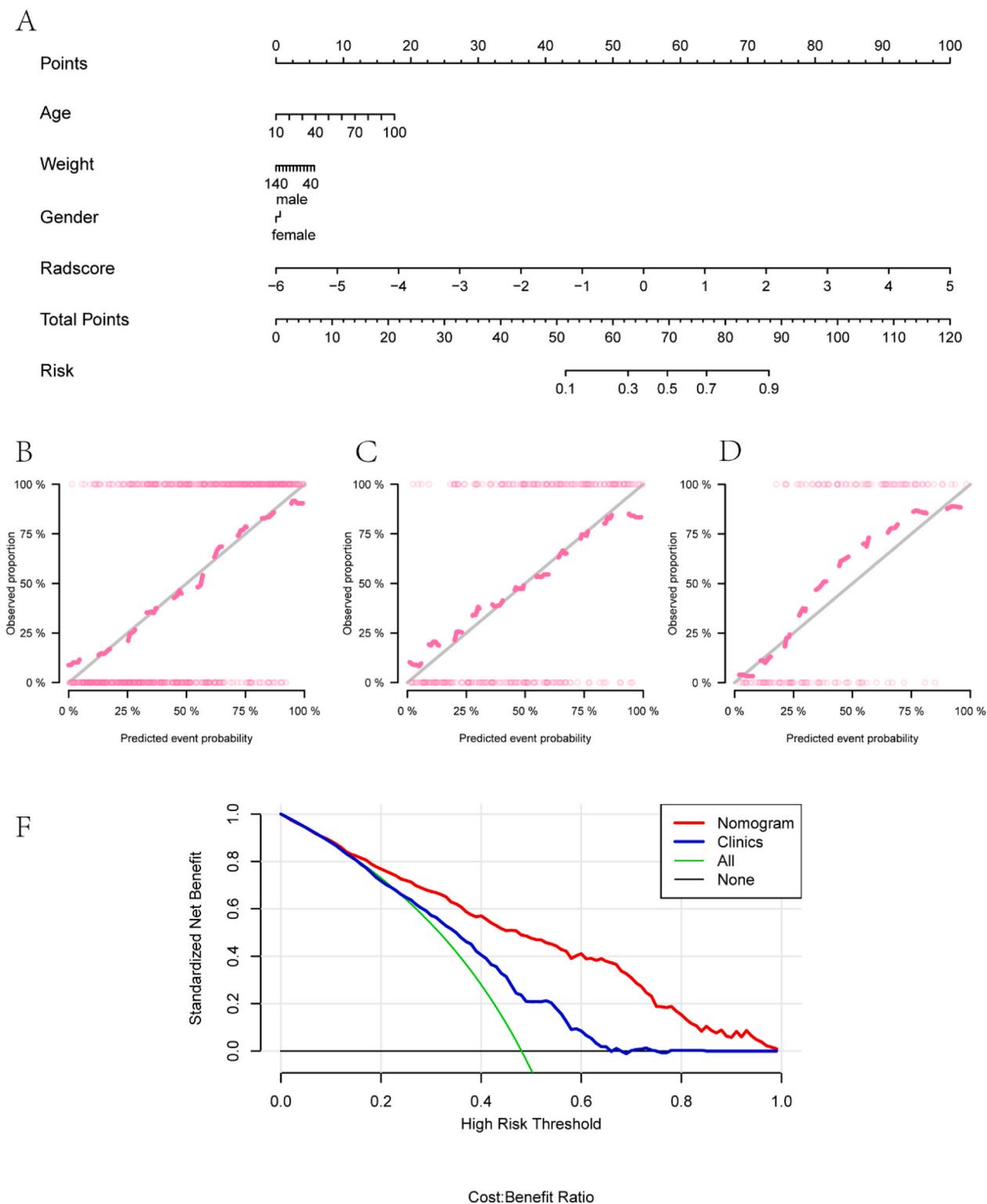


Fig. 4. The radiomics nomogram performance and DCA of diverse models. (A) Representation of radiomics nomogram constructed by incorporating clinical factors and radiomics signature. Calibration curves for radiomics nomogram of (B) training, (C) internal validation and (D) external validation sets, respectively. (E) DCA of diverse models.

generalizability of findings, we suggest that future studies use our method in more multicenter trials. This approach would help validate its applicability across different healthcare settings and potentially identify regional variations in treatment efficacy.

5. Conclusion

The radiomics nomogram is developed in this study for differentiating PRISm from COPD groups, which may be used as the virtual approach for clinical radiologists. Furthermore, based on the results of our study and future validation using larger samples, maybe we will

Table 3
Diagnostic performances of the three models in three sets.

Model		Accuracy	Accuracy Lower	Accuracy Upper	Sensitivity	Specificity	PPV	NPV	AUC (95 %CI)	p-value of DeLong-Test	
										vs Radiomics	vs Nomogram
Clinics	Training set	0.64	0.60	0.68	0.57	0.71	0.64	0.64	0.68(0.64–0.72)	<0.001	<0.001
	Internal validation set	0.58	0.52	0.63	0.55	0.60	0.57	0.58	0.62(0.56–0.69)	<0.001	<0.001
	External validation set	0.66	0.58	0.74	0.54	0.78	0.71	0.64	0.73(0.65–0.83)	0.19	<0.001
Radiomics	Training set	0.76	0.72	0.79	0.68	0.83	0.79	0.74	0.82(0.79–0.85)	-	0.39
	Internal validation set	0.70	0.64	0.75	0.63	0.77	0.72	0.68	0.77(0.71–0.82)	-	0.37
	External validation set	0.73	0.65	0.80	0.57	0.88	0.83	0.68	0.80(0.72–0.87)	-	-
Nomogram	Training set	0.75	0.72	0.79	0.68	0.83	0.78	0.73	0.82(0.79–0.86)	-	-
	Internal validation set	0.70	0.64	0.75	0.73	0.68	0.63	0.78	0.77(0.72–0.83)	-	-
	External validation set	0.73	0.65	0.80	0.90	0.56	0.67	0.84	0.84(0.78–0.91)	0.003	-

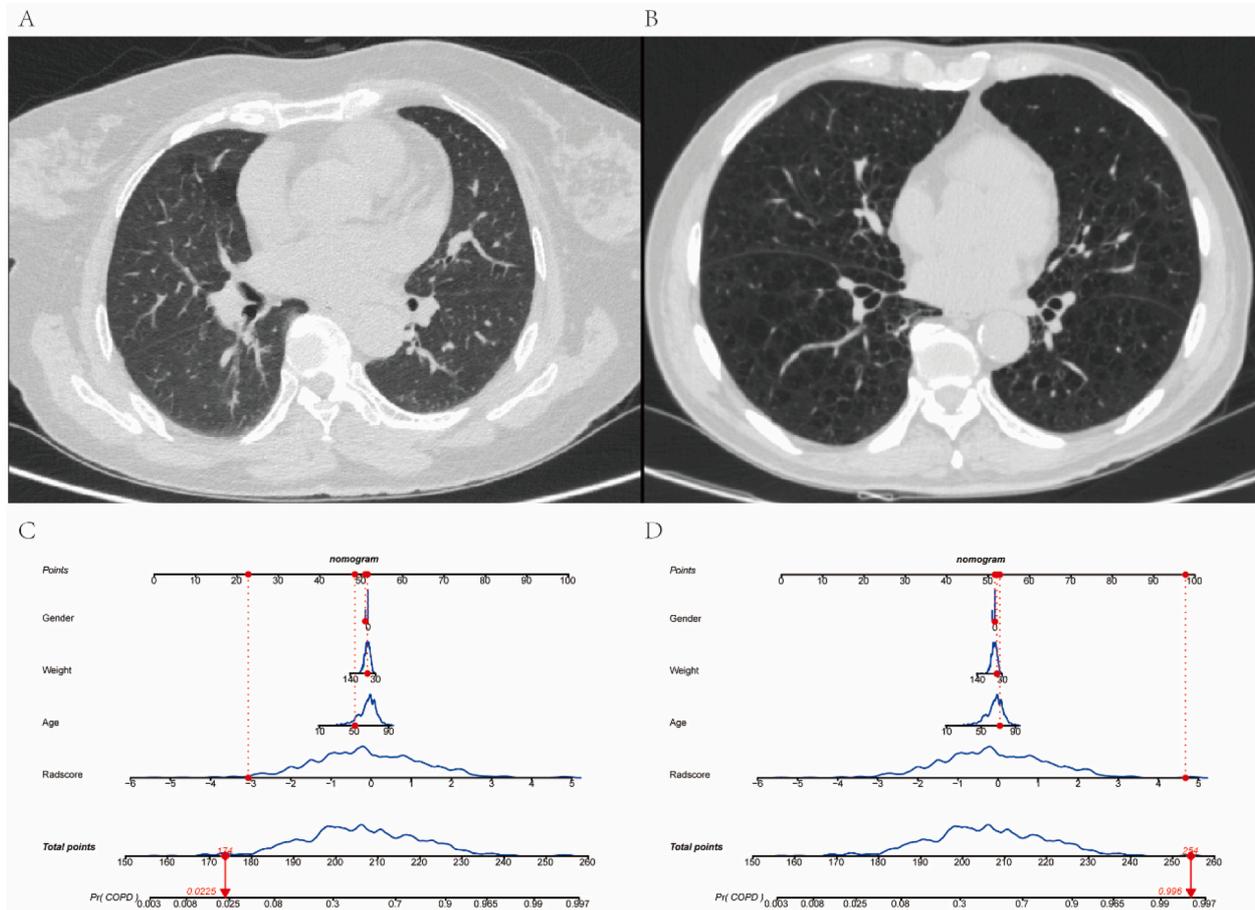


Fig. 5. The risk scores of COPD in two patients were calculated by using the nomogram. A. Thin-slice chest CT images of PRISM in a 51-year-old woman with height 163 cm, non-smoker, Radscore -3.06. C. The nomogram shows that the total score was 174 points, corresponding to the probability of developing COPD is approximately 2.25 %. B. Thin-slice chest CT image of COPD in a 72-year-old male subject. He is 162 cm tall, former smoker, and has a Radscore of 4.68. D. The total score of the nomogram was 254, corresponding to the probability of developing COPD of approximately 99.6 %. PRISM Preserved Ratio Impaired Spirometry, COPD chronic obstructive pulmonary disease, CT computed tomography, FEV1/FVC ratio of forced expiratory volume in 1 s to forced vital capacity.

translate the nomogram into a visual and operational application, which can be used for real-time discriminating of PRISM and COPD without PFT.

Ethics approval and consent to participate

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

Ethical approval: Institutional Review Board approval was obtained. ClinicalTrials.gov registration number ChiCTR2300069929, CSD-COPD cohort

Funding

This work was supported by the National Natural Science Foundation of China (81930049, 82171926 and 82202140), the National Key Research and Development Program of China (2022YFC2010000, 2022YFC2010002 and 2022YFC2010005), the Medical Imaging Database Construction Program of National Health Commission (YXFSC2022JJSJ002), the Clinical Innovative Project of Shanghai Changzheng Hospital (2020YLCYJ-Y24), the Program of Science and Technology Commission of Shanghai Municipality (21DZ2202600), and the Shanghai Sailing Program (20YF1449000).

CRediT authorship contribution statement

XiuXiu Zhou: Methodology, Investigation, Data curation. **XiaoQing Lin:** Methodology, Data curation. **Yu Guan:** Writing – original draft, Resources, Project administration, Data curation. **WenTing Tu:** Funding acquisition, Data curation. **Bing Fan:** Data curation. **YanQing Ma:** Data curation. **Liang Mao:** Data curation. **TaoHu Zhou:** Writing – review & editing, Writing – original draft, Methodology. **Li Fan:** Writing – review & editing, Supervision, Funding acquisition, Data curation, Conceptualization. **ShiYuan Liu:** Writing – review & editing, Supervision, Funding acquisition, Data curation, Conceptualization. **Jie Li:** Data curation.

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.

Data availability

Data will be made available on request.

Acknowledgements

Not applicable.

Authorship contribution statement

LF conceived, designed, and supervised the study. THZ, YG and XQL performed data analysis and drafted the manuscript. THZ conducted a major clinical experiment. XXZ, LM, YQM, BF, JL, and WTT contributed to the imaging and clinical data collection. LF and SYL supervised the literature review and data quality control. LF and THZ revised the manuscript. All the authors have read and approved the final manuscript.

Consent for publication

Not applicable.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.ejro.2024.100580](https://doi.org/10.1016/j.ejro.2024.100580).

References

- [1] E.S. Wan, P.J. Castaldi, M.H. Cho, J.E. Hokanson, E.A. Regan, B.J. Make, et al., Epidemiology, genetics, and subtyping of preserved ratio impaired spirometry (PRISm) in COPDGene, *Respir. Res.* 15 (2014) 89.
- [2] D.M. Mannino, M.A. McBurnie, W. Tan, A. Kocabas, J. Anto, W.M. Vollmer, et al., Restricted spirometry in the burden of lung disease study, *Int. J. Tube Lung Dis.* 16 (2012) 1405–1411.
- [3] S. Chen, M. Kuhn, K. Prettnner, F. Yu, T. Yang, T. Bärnighausen, et al., The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020–50: a health-augmented macroeconomic modelling study, *Lancet Glob. Health* 11 (2023) e1183–e1193.
- [4] WHO, The Top 10 Causes of Death, World Health Organization, 2020.
- [5] B.R. Celli, A. Agustí, COPD: time to improve its taxonomy? *ERJ Open Res.* 4 (2018).
- [6] E.S. Wan, S. Fortis, E.A. Regan, J. Hokanson, M.K. Han, R. Casaburi, et al., Longitudinal phenotypes and mortality in preserved ratio impaired spirometry in the COPDGene study, *Am. J. Respir. Crit. Care Med.* 198 (2018) 1397–1405.
- [7] S.R.A. Wijnant, E. De Roos, M. Kavousi, B.H. Stricker, N. Terzikhan, L. Lahousse, et al., Trajectory and mortality of preserved ratio impaired spirometry: the rotterdam study, *Eur. Respir. J.* 55 (2020).
- [8] X. Wei, Q. Ding, N. Yu, J. Mi, J. Ren, J. Li, et al., Imaging features of chronic bronchitis with preserved ratio and impaired spirometry (PRISm), *Lung* 196 (2018) 649–658.
- [9] K.A. Young, M. Strand, M.F. Ragland, G.L. Kinney, E.E. Austin, E.A. Regan, et al., Pulmonary subtypes exhibit differential global initiative for chronic obstructive lung disease spirometry stage progression: the COPDGene® study, *Chronic Obstr. Pulm. Dis.* 6 (2019) 414–429.
- [10] M.T. Dransfield, K.M. Kunisaki, M.J. Strand, A. Anzueto, S.P. Bhatt, R.P. Bowler, et al., Acute exacerbations and lung function loss in smokers with and without chronic obstructive pulmonary disease, *Am. J. Respir. Crit. Care Med.* 195 (2017) 324–330.
- [11] A. Dirksen, N.H. Holstein-Rathlou, F. Madsen, L.T. Skovgaard, C.S. Ulrik, T. Heckscher, et al., Long-range correlations of serial FEV1 measurements in emphysematous patients and normal subjects, *J. Appl. Physiol.* 85 (1985) (1998) 259–265.
- [12] J.W. Gurney, K.K. Jones, R.A. Robbins, G.L. Gossman, K.J. Nelson, D. Daughton, et al., Regional distribution of emphysema: correlation of high-resolution CT with pulmonary function tests in unselected smokers, *Radiology* 183 (1992) 457–463.
- [13] J.D. Flesch, C.J. Dine, Lung volumes: measurement, clinical use, and coding, *Chest* 142 (2012) 506–510.
- [14] L. Fan, Y. Xia, Y. Guan, T.F. Zhang, S.Y. Liu, Characteristic features of pulmonary function test, CT volume analysis and MR perfusion imaging in COPD patients with different HRCT phenotypes, *Clin. Respir. J.* 8 (2014) 45–54.
- [15] D.A. Lynch, J.H. Austin, J.C. Hogg, P.A. Grenier, H.U. Kauczor, A.A. Bankier, et al., CT-Definable subtypes of chronic obstructive pulmonary disease: a statement of the fleischner society, *Radiology* 277 (2015) 192–205.
- [16] S. Bodduluri, J.M. Reinhardt, E.A. Hoffman, J.D. Newell Jr., S.P. Bhatt, Recent advances in computed tomography imaging in chronic obstructive pulmonary disease, *Ann. Am. Thorac. Soc.* 15 (2018) 281–289.
- [17] A.N. Frix, F. Cousin, T. Refaee, F. Bottari, A. Vaidyanathan, C. Desir, et al., Radiomics in lung diseases imaging: state-of-the-art for clinicians, *J. Pers. Med.* 11 (2021).
- [18] T.H. Zhou, X.X. Zhou, J. Ni, Y.Q. Ma, F.Y. Xu, B. Fan, et al., CT whole lung radiomic nomogram: a potential biomarker for lung function evaluation and identification of COPD, *Mil. Med. Res.* 11 (2024) 14.
- [19] Y. Yang, W. Li, Y. Guo, N. Zeng, S. Wang, Z. Chen, et al., Lung radiomics features for characterizing and classifying COPD stage based on feature combination strategy and multi-layer perceptron classifier, *Math. Biosci. Eng.* 19 (2022) 7826–7855.
- [20] X. Lin, T. Zhou, J. Ni, J. Li, Y. Guan, X. Jiang, et al., CT-based whole lung radiomics nomogram: a tool for identifying the risk of cardiovascular disease in patients with chronic obstructive pulmonary disease, *Eur. Radio.* (2024).
- [21] B. Remeseiro, V. Bolon-Canedo, A review of feature selection methods in medical applications, *Comput. Biol. Med.* 112 (2019) 103375.
- [22] Y.P. Zhang, X.Y. Zhang, Y.T. Cheng, B. Li, X.Z. Teng, J. Zhang, et al., Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling, *Mil. Med. Res.* 10 (2023) 22.
- [23] N. De Jay, S. Papillon-Cavanagh, C. Olsen, N. El-Hachem, G. Bontempi, B. Haibe-Kains, mRMRe: an R package for parallelized mRMRe ensemble feature selection, *Bioinformatics* 29 (2013) 2365–2368.
- [24] N. Zhao, F. Wu, J. Peng, Y. Zheng, H. Tian, H. Yang, et al., Preserved ratio impaired spirometry is associated with small airway dysfunction and reduced total lung capacity, *Respir. Res.* 23 (2022) 298.
- [25] R.L. Jones, M.M. Nzekwu, The effects of body mass index on lung volumes, *Chest* 130 (2006) 827–833.
- [26] D.D. Sin, R.L. Jones, S.F. Man, Obesity is a risk factor for dyspnea but not for airflow obstruction, *Arch. Intern. Med.* 162 (2002) 1477–1481.
- [27] L.M. Schachter, C.M. Salome, J.K. Peat, A.J. Woolcock, Obesity is a risk for asthma and wheeze but not airway hyperresponsiveness, *Thorax* 56 (2001) 4–8.
- [28] L. Sang, X. Gong, Y. Huang, J. Sun, Proportions and risk factors of chronic obstructive pulmonary disease and preserved ratio impaired spirometry, and association with small airway disease, in the positive screening older population from China: a cross-sectional study, *BMC Pulm. Med.* 24 (2024) 114.
- [29] X. Yin, Z. Zheng, Y. Dong, J. Li, S. Yang, Q. Xu, et al., Comparison of newly diagnosed COPD patients and the non-COPD residents in Shanghai minhang district, *Front Public Health* 11 (2023) 1102509.
- [30] T.H. Zhou, X.X. Zhou, J. Ni, Y.Q. Ma, F.Y. Xu, B. Fan, et al., CT whole lung radiomic nomogram: a potential biomarker for lung function evaluation and identification of COPD, *Mil. Med. Res.* 11 (2024) 14.
- [31] X. Lin, T. Zhou, J. Ni, J. Li, Y. Guan, X. Jiang, et al., CT-based whole lung radiomics nomogram: a tool for identifying the risk of cardiovascular disease in patients with chronic obstructive pulmonary disease, *Eur. Radio.* (2024).
- [32] Y. Yang, W. Li, Y. Guo, N. Zeng, S. Wang, Z. Chen, et al., Lung radiomics features for characterizing and classifying COPD stage based on feature combination strategy and multi-layer perceptron classifier, *Math. Biosci. Eng.* 19 (2022) 7826–7855.
- [33] S. Qin, B. Jiao, B. Kang, H. Li, H. Liu, C. Ji, et al., Non-contrast computed tomography-based radiomics for staging of connective tissue disease-associated interstitial lung disease, *Front Immunol.* 14 (2023) 1213008.
- [34] X. Zhou, Y. Yi, Z. Liu, W. Cao, B. Lai, K. Sun, et al., Radiomics-based pretherapeutic prediction of neo-adjuvant therapy in locally advanced rectal cancer, *Ann. Surg. Oncol.* 26 (2019) 1676–1684.
- [35] Z. Jiang, J. Yin, P. Han, N. Chen, Q. Kang, Y. Qiu, et al., Wavelet transformation can enhance computed tomography texture features: a multicenter radiomics study for grade assessment of COVID-19 pulmonary lesions, *Quant. Imaging Med. Surg.* 12 (2022) 4758–4770.
- [36] J. Lu, H. Ge, L. Qi, S. Zhang, Y. Yang, X. Huang, et al., Subtyping preserved ratio impaired spirometry (PRISm) by using quantitative HRCT imaging characteristics, *Respir. Res.* 23 (2022) 309.