Age-associated changes in aging lungs: A study with auto-segmentation and radiomics based on CT images

Sijia Cui  
Zhejiang Provincial People's Hospital

Zhenyu Shu  
Zhejiang Provincial People's Hospital

Yuan Shao  
Zhejiang Provincial People's Hospital

Peipei Pang  
GE Healthcare

XiangYang Gong (✉️ cjrgxy@hotmail.com)  
Zhejiang Provincial People's Hospital

Research

Keywords: aging lung, texture analysis, computed tomographic, pulmonary function

Posted Date: February 11th, 2020

DOI: https://doi.org/10.21203/rs.2.23216/v1

License: ☑️ Ⓡ This work is licensed under a Creative Commons Attribution 4.0 International License. 
Read Full License
Abstract

**Background:** The senile lung undergoes physiologic changes that are well known but have not been investigated with computed tomographic (CT) texture analysis. The thin-section pulmonary computed tomographic (CT) microstructure features change in asymptomatic elderly individuals were not explored.

**Methods:** We retrospectively selected 259 subjects under-going chest computed tomography (CT) between April 2018 and June 2019, as group A(consist of group1 and group2), without a history of smoking within the past 5 years, respiratory symptoms or any known chronic pulmonary disease. There were 118 patients in group 1(age ≥ 60 years-elderly) and 141 patients in group 2 (age≤50 years-young). Furthermore, 273 patients with pulmonary function tests(PFTs) were included as group B, which divided into two cohorts, chronic obstructive pulmonary disease (COPD) (n=83) and non-COPD (healthy smoker [HS], n=90; healthy non-smoker [HNS], n=100) cohort. The radiomic features were extracted and selected from group A, trialed in group B, using the LASSO algorithm.

**Results:** A total of 233 features were significant in group A. Among these features, 17 features exhibited distinct differences between COPD and non-COPD patients, 18 features exhibit distinct differences between HSs and HNSs. Meanwhile, five features were shared in group B. A negative correlation was determined between carbon monoxide diffusing capacity(DL\textsubscript{CO}) and the two features: ShortRunLowGreyLevelEmphasis\_AllDirection\_offset9\_SD (\(\rho = -0.70\)) and LowGreyLevelRunEmphasis\_AllDirection\_offset2\_SD (\(\rho = -0.63\)). Similarly, a positive correlation was found between FEV1/FVC and HighGreyLevelRunEmphasis\_AllDirection\_offset8\_SD (\(\rho = 0.74\)).

**Conclusion:** Radiomic features, which associated with the ages and significant in COPD patients and smokers, maybe reveal the microstructure changes of the aging lungs.

Introduction

As it is known, aging has become a public health problem, which influences all aspects of human biology, such as aging lung. The aging lung undergoes anatomic changes, with a deterioration in terms of function. Furthermore, the aging of lung considered as a risk factor for respiratory diseases, and strongly correlates with the onset and progression of chronic respiratory diseases [1-3]. The structure of aging lung is characterized as alveolar space enlargement, alveolar wall thinning and bronchial basement membrane thickening. This usually presents with decreasing thoracic volume, reduced lung volumes and alterations in respiratory muscles [4]. However, the degree of lung aging is inconsistent and greatly varies among individuals [5, 6]. Several factors, such as age, race, environmental exposure and living conditions, have been reported to have a potential effect on lung aging.

The lung aging eventually cause a decrease of pulmonary function[7]. However, prior to the function change, structural change perhaps occurs in the lungs [8]. It was previously reported that computed tomography (CT) could detect the smoking-related inflammation of small and/or distal bronchi (bronchiolitis), and even the changes in lung structure before airflow limitation [7, 9, 10]. In addition, CT
can also reveal the structural change of the lung before PF starts to decrease. To the date, the role of imaging in lung aging remains unclear, and there is a lack of imaging biomarkers for lung aging.

In previous studies, Copley et al found in the older group, there is a high prevalence of a basal subpleural reticular pattern (60% of subjects), small cysts (25%), bronchial dilatation (60%), and wall-thickening (55%) and confirmed interlobular septal thickening (18%), as described in previous studies[11]. A.M. Chiesa et al revealed the absence of A-lines in healthy elderly subjects[12]. Recently, radiomic signature has proven to be a significant biomarker for biological, clinical condition, or disease progression [13-14]. The aim of the present study was to explore the age-associated changes based on the radiomics texture analysis for lung aging and validate or trial this in another two cohorts: a cohort of COPD patients and a cohort of smoking patients.

**Materials And Methods**

**Patients**

The present retrospective study was approved by the Ethics Committee of Zhejiang Provincial People’s Hospital. The informed consent was waived due to the retrospective nature of the study. Between April 2018 and June 2019, we retrospectively selected 259 patients undergoing low-dose chest CT scans (Somatom Definition AS, Siemens, Germany) from Zhejiang Provincial People’s Hospital as group A. The exclusion criteria were: current smoking, history of smoking within the past 5 years, respiratory symptoms, any known chronic pulmonary disease or known pulmonary pathologies and clinically relevant cardiac diseases (i.e. left heart failure), the CT scan showed parenchymal consolidations, ground-glass opacities, pleural effusion, fibro-/pneumothorax, pleural plaques, bronchiectasis, emphysema, pulmonary fibrosis, nodular pattern, outcomes of thoracic surgery or the poor image quality were excluded. Patients were divided into the following 2 groups by age: group 1 (age ≥ 60 years-elderly) and group 2 (age ≤ 50 years-young). We did not include patients between the ages of 50 and 60 because we wanted to consider 2 groups of subjects with different ages, as previously done by other authors[11, 12].

In the trial cohort, patients with lung CT examinations and PF tests were enrolled to group B for further analysis, which included two sub-groups: COPD (FEV1/FVC <0.70) and non-COPD (FEV1/FVC >0.70) patients [15]. Non-COPD group had no clinical evidence of obstructive airways disease, and had spirometry results of FEV1/FVC ratio >0.70 and FEV1 >80% of predicted, which included healthy smoker (HS, smoking behavior for over 10 years) and healthy non-smoker (HNS) [16]. The demographic, radiology and clinical evaluation were obtained by reviewing the medical records. Patients were excluded when they had any of the following conditions: missing clinical data and poor imaging quality.

**Lung CT**

The conditions for the lung CT examination were as follows: 120 kV, 10-40 mA, gantry rotation speed (0.5 seconds), and helical scan mode (pitch 1.2). All lung CT images were reconstructed at a slice thickness of
2 mm, with a spacing interval of 2 mm. All images were obtained at a lung window width of 1,500 HU and a window level of -400 HU.

**PF**

PF tests were evaluated using a flow spirometer (Vmax22; SensorMedics, Yorba Linda, CA), based on the American Thoracic Society guidelines [16] and European Community Lung Health Survey values [17]. Post-bronchodilator measurements were recorded at 15 minutes after bronchodilator use (such as albuterol). Pre- and post-bronchodilator spirometry, vital capacity (VC), forced vital capacity (FVC), forced expiratory volume in one second (FEV1), maximal mid-expiratory flow (MMEF), FEV1/FVC and single-breath DLco were obtained according to standard ATS/ERS (American Thoracic Society/European Respiratory Society) recommendations. The degree of airflow limitation was evaluated using the GOLD definition and spirometric classification [15].

**CT image feature extraction**

For processing, all images were standardized according to the following: (1) images resample: the resolution was resampled to 1.5×1.5×1.5 mm$^3$ voxel size; (2) intensity standardization: the grey value was re-ranged to 0-255. Subsequently, the software Lung Kit (LK Version V1.0.0.R, GE Healthcare) was used to segment the whole lung, automatically relying on the CT images, and the other lung tissues, bronchus and vessels were removed. Then, the automated segmentation of the lung was manually corrected by two experienced radiologists, who were blinded to the clinical data. The texture features were extracted from the segmented CT images using the Artificial Intelligence Kit Version 3.0.1.A (GE Healthcare) [18].

**Radiomic feature selection and model establishment**

The significance of the feature related to aging lung was evaluated by paired $t$-test. Then, the significant features ($P<0.05$) were selected to distinguish between COPD and non-COPD patients, and HSs and HNSs. In order to compare the features between groups, feature normalization was performed. The values of each feature were normalized using z-score normalization [19]. Similarly, the normalization was performed in the validation groups (COPD or HS) using the corresponding data calculated in the training cohort. A feature selection algorithm based on the least absolute shrinkage and selection operator (LASSO) method was adopted. This algorithm adopts the 10-fold cross validation method, and the features were selected by weighting the LASSO coefficients (Supplementary material). The radiomics signature used in the validation cohorts were calculated using the formula obtained from the training set.

The radiomics scores (rad-score) were calculated for each patient, and several methods, such as Logistic regression, support vector machine (SVM), Random Forest, Bayes, K-Nearest Neighbor (KNN), and Decision Tree [20], were used to establish the prediction models for grouping COPD and non-COPD subjects (Model 1) or HSs from HNSs (Model 2). Then, the discriminative accuracies were evaluated using the receiver
operating characteristic (ROC) curve analysis. The radiomic methods included the following steps (Figure 1), and the LASSO method and dimensionality reduction framework presented in Figure S1 and S2.

**Statistical analysis**

The statistical analysis was performed using the R-project software (Version 3.0.1), GraphPad (San Diego, CA) and SPSS 17.0 (IBM, Armonk, NY). The normality of the distribution was analyzed using the Kolmogorov–Smirnov test. The quantitative variables with normal distribution were compared using \( t \)-test, and nonparametric variables compared by Mann-Whitney \( U \)-test. Categorical variables were compared using chi-square test. The goodness-of-fit of the models in the training set were assessed using the Hosmer-Lemeshow test. The ROC analysis were used to evaluate the discriminative accuracy of the models. The relationship between the radiomics features and PF data were analyzed by Spearman's correlation coefficient. A \( P \)-value of <0.05 was considered statistically significant.

This work was supported by National Key Research and Development Project(2017YFC0114103) and Key Research and Development Project of Zhejiang Province(2020C01058)

**Results**

**Patient characteristics**

A total of 259 subjects were enrolled to group A, 118 subjects (66 females and 52 males), with a mean age of 64.78 ± 4.8 years in group1; 141 subjects (62 females and 79 males), with a mean age of 34.5 ± 9.5 years in group2. There was no significant difference in gender, BMI, smoking history between the two groups(Table1). A total of 273 patients were assigned to group B, which included 83 COPD patients (age range, 27-89 years old; median age, 68 years old) and 189 non-COPD patients (age range, 25-90 years old; median age, 65 years old). Among the non-COPD patients, 89 patients were HSs and 100 were HNSs. Subsequently, patients in group B were randomly divided into two groups: training set \( (n=190 / n=132) \) and validation set \( (n=82 / n=57) \). The proportion was 7:3.

No significant difference was observed in clinical characteristics (such as gender, age, hypertension, BMI, alcohol abuse, coronary heart disease and diabetes) between COPD and non-COPD patients, as well as between HSs and HNSs, irrespective of the cohorts (training or validation, all \( P>0.05 \)). The detailed characteristics are summarized in Tables S1-S3.

**Feature selection**

A total of 1,044 radiomic features were extracted from group A, which included histogram features \( (n=42) \), texture features \( (n=216) \), form factor features \( (n=20) \), gray-level co-occurrence matrix (GLCM) features \( (n=226) \) and run-length matrix (RLM) features \( (n=540) \) (Supplementary material). Among these, 233 features exhibited significant differences (using paired sample \( t \)-test, \( P<0.05 \)). Finally, 17 and 18 features were respectively used for classification purposes: COPD vs. non-COPD and HSs vs. HNSs. Interestingly, five features, which included one low-order texture feature \( (\text{Percentile70}) \) and four high-order
texture features \((\text{InverseDifferenceMoment\_AllDirection\_offset6\_SD}, \text{GreyLevelNonuniformity\_AllDirection\_offset1\_SD}, \text{RunLengthNonuniformity\_AllDirection\_offset1\_SD}, \text{ShortRunEmphasis\_AllDirection\_offset8\_SD})\) were shared by group B (Table 2).

**Diagnostic performance of the established models**

The rad-scores were calculated using the formula for the radiomics signature (Figure 2, Supplementary material). The alignment curve shows the calibration of the training and validation (COPD and HN) cohorts (Figure S3). Then, the diagnostic efficacy of several models, including logistic regression, SVM, Bayes, KNN, Decision Tree and Random forest, were evaluated in the training and validation cohorts (Tables 3A and 3B).

For the discrimination between COPD and non-COPD, the AUC areas for the Logistic regression, SVM, Bayes, KNN, Decision Tree and random forest models in the training cohort were 0.84, 0.67, 0.75, 0.86, 0.85 and 0.92, respectively. Furthermore, the AUC areas for these models in the validation cohort were 0.85, 0.64, 0.58, 0.76, 0.54 and 0.63, respectively. In terms of the logistic regression model, in the training cohort, the accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) in the discrimination between COPD and non-COPD was 0.77 (95% CI: 0.70-0.83), 75.94%, 79.66%, 89.38% and 59.49%, respectively. The Hosmer-Lemeshow test indicated that the model fits well \((P>0.05)\). In the validation cohort, the accuracy, sensitivity, specificity, PPV and NPV was 0.84 (95% CI: 0.74-0.91), 89.29%, 70.83%, 87.72% and 73.91%, respectively (Table 4, Figures 3A and 3B).

For the discrimination between HSs and HNSs, the AUC areas for Logistic regression, SVM, Bayes, KNN, Decision Tree and random forest models in the training cohort were 0.87, 0.79, 0.80, 0.81, 0.87 and 0.93, respectively. Furthermore, the AUC areas for these models in the validation cohort were 0.80, 0.84, 0.83, 0.72, 0.70 and 0.76, respectively. In terms of the logistic regression model, in the training cohort, the accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) in the discrimination between HSs and HNSs was 0.80 (95% CI: 0.71-0.86), 85.71%, 72.88%, 77.14% and 82.69%, respectively. The Hosmer-Lemeshow test indicated that the model fits well \((P>0.05)\). In the validation cohort, the accuracy, sensitivity, specificity, PPV and NPV was 0.76 (95% CI: 0.62-0.87), 76.92%, 75.00%, 76.92% and 75.00%, respectively (Table 4, Figures 4A and 4B).

**The associations between features and PF tests**

The associations between PF tests (six tests) and features (17 features) were analyzed. DLCO was found to be correlated with four coarse features: \(\text{ShortRunLowGreyLevelEmphasis\_AllDirection\_offset9\_SD}\) \((\rho=-0.70)\), \(\text{LowGreyLevelRunEmphasis\_AllDirection\_offset2\_SD}\) \((\rho=-0.63)\), \(\text{Inertia\_AllDirection\_offset2\_SD}\) \((\rho=0.41)\) and \(\text{HighGreyLevelRunEmphasis\_AllDirection\_offset8\_SD}\) \((\rho=0.41)\). FEV1/FVC was found to be correlated with three different radiomic features: \(\text{HighGreyLevelRunEmphasis\_AllDirection\_offset8\_SD}\) \((\rho=0.74)\), \(\text{Inertia\_AllDirection\_offset2\_SD}\) \((\rho=0.43)\) and \(\text{ClusterShade\_AllDirection\_offset1\_SD}\) \((\rho=0.33)\) (Figures 5 and 6).
**Discussion**

In the present study, the lung radiomics features associated with aging were investigated. Radiomics signatures were constructed, and its value was verified in two tested populations. To our knowledge, this is the first report to investigate the radiomics features of lung aging. Several advantages of the study have been identified. First, many molecules have been studied in the aging process. For example, in the lungs during aging, TLR, p53/p66/p21 and reactive oxygen species (ROS) generation are increased [21-22]. In addition, the telomere becomes short [23]. Unfortunately, an invasive procedure is required to obtain a tissue sample, and investigate these levels during aging [24]. However, the method for lung aging in the present study can be easily accessed, has little additional cost, and favors the discrimination of aging subgroups. These advantages make this method promising. Second, previously, clinicians have mainly focused on the analysis of imaging manifestations for lung aging. However, the present study revealed that the whole lung texture analysis could better reveal the micro-changes of lung structures, which is helpful in identifying lung aging and earlier initiating lifestyle interventions, and this may slow or delay the onset of lung aging [25]. Third, the present study also evaluated the relationship between texture features and PF tests, and found that a close relationship exists between these. Fourth, as it is known, lung aging can be affected by many factors. However, in the present study, efforts were made to alleviate the interference: (1) CT images from the same patient at different time points were analyzed; (2) automatic segmentation was performed for feature analysis. Compared with manual segmentation, the automatic method demonstrates good effectiveness and reliability, and has less dependence on humans.

A total of 30 radiomics features selected from 233 textural parameters may have the capability to reveal the characteristics of aging lungs, including histograms and morphological features, which are low- and higher-order texture parameters (such as GLCM, RLM and form factor features). Among these 30 textural parameters, five parameters were selected, which included one histogram, one GLCM, and three RLM parameters. It is noteworthy that these features could detect the invisible micro-structural changes of the lung. For example: histogram parameters could be used to assess the attenuation of the lungs [26]; GLCM could reflect the degree of heterogeneity of the whole lung, and the micro-structural changes may give uneven signals in the CT scan, which is irreversible of any anatomy change in the lung [27]; form factor features could reflect the changes in the region of interest, in terms of volume, area and shape, and these may be displayed more visually, allowing the changes in lung volume and shape to be easily observed and compared [28]; in addition to those mentioned above, the study also revealed that the RLM features, which reflect the roughness and directionality, were also associated with the changes in pulmonary structure [29].

Two populations, COPD and HS patients, were the study subjects, and these subjects were included to verify the observed features. The reasons for this were as follows: First, previous studies have shown that lung aging is closely correlated to the onset and development of chronic respiratory diseases [1]. Second, the risk factors of COPD included age, suggesting that COPD has an association with lung aging. Third, declining lung function is a significant feature for the development of COPD, which increases with age,
especially in smoking individuals \[29\]. Furthermore, an injured lung function cannot recover after smoking cessation \[1\]. To the date, the mechanism of injury in lung function in COPD patients remains unclear. However, given the strong similarities between elderly and COPD lungs, the mechanism that involves aging has been investigated, and it has proven that accelerated lung aging occurs in COPD patients \[1\]. Similar to other organs, lung demonstrates physiological and structural changes associated with aging, resulting in a progressive decrease in lung function among healthy populations. In terms of smoking, this can accelerate the decline in lung function over time, suggesting that this has an association with lung aging \[30\]. This may be explained, as follows: (1) smoking could accelerate the aging of the airway epithelium \[31\], (2) changes occurred in the micro-structure of the lung between HSs and HNSs, and these are similar to that found in the COPD cohort.

Interestingly, in the present study, it was found that the discriminative performance of radiomic features between HSs and HNSs in the training cohort were slightly better than those between COPD and non-COPD patients. The possible reason for this is that the changes between COPD and non-COPD patients may be more obvious. Thus, radiomic features that reveal these micro-changes may be insensitive in distinguishing between these. In a word, aging would make an effect on the manifestations of lung disease and even lower the response to treatment, and the present findings may have an ancillary role in conventional diagnostic methods and clinical examinations.

Although the pathology of the lung is distinct, in terms of PF, all aging lungs can result in impairment of lung function, which can be measured using FEV1 and FVC \[30, 31\]. The present study revealed that some features of lung aging can indirectly evaluate the pulmonary function. Furthermore, some features were closely correlated to FEV1/FVC, such as \textit{HighGreyLevelRunEmphasis\_AllDirection\_offset8\_SD} (HGLRE), which has the strongest association with FEV1/FVC. Indeed, \textit{HGLRE} is a kind of a combination of image intensity and homogeneity. Therefore, this feature quantifies the joint-relationship between CT attenuation and HU homogeneity. A high HGLRE value indicates a dense and homogenous tissue, while a low value indicates a low attenuating and heterogeneous tissue. The relationship between HGLRE and FEV1 observed in the present study means that: (1) patients with normal spirometry may have a dense and homogenous lung; and (2) patients with abnormal spirometry may also have a low attenuating and heterogeneous lung \[29, 32\].

The present study revealed that two coarse texture features, \textit{LowGreyLevelRunEmphasis\_AllDirection\_offset2\_SD} (\(\rho = -0.63\)) and \textit{ShortRunLowGreyLevelEmphasis\_AllDirection\_offset9\_SD} (\(\rho = -0.70\)), were found to be negatively correlated with DLCO. Similar results were reported by Barjaktarevic \textit{et al.} \[29\]. \textit{Low Gray Level Run Emphasis} (LGLRE) is only a measure of CT attenuation, while \textit{Short Run Low Gray Level Emphasis} (SRLGLE) is a joint-measure of CT attenuation and HU heterogeneity. A high SRLGLE value indicates a low attenuating and heterogeneous tissue. A stronger correlation coefficient was found for SRLGLE, suggesting that spatial information may be more useful than intensity information in the identification. In addition to texture features, radiomic analysis often requires the analysis of a combination of intensity, morphology, fractal geometry and higher-order features\[33\]. This information is integrated, and may
thereby provide novel insights and a better detail of the lung using CT images. Eventually, radiomic analysis may serve as a useful tool that could help to improve the clinical management of patients[34].

Although these present findings provide meaningful implications, the present study has several limitations. First, CT examinations are usually performed without spirometric control of lung volume. However, all subjects in the present study received education on how to perform the respiratory maneuvers prior to the CT scan. Second, the present study has a relatively small sample size. Thus, the results may be biased. Hence, further analysis using a larger sample size is required. Fourth, the relationship between the severity of COPD and radiomics features was not investigated. Hence, another trial is needed to specifically address this issue.

In conclusion, some radiomics features that reveal the micro-changes of lung structure were found to be associated with lung aging, and radiomic signatures, which are constructed using LASSO regression, could be used to identify a lung aging-related population of COPD patients and a population of smoking patients. However, a larger study is still needed to verify this finding, and further analysis are required to assess its prognostic value in an aging population.

**Abbreviations**

CT: computed tomographic; HS: healthy smoker; HNS: healthy non-smoker; COPD: chronic obstructive pulmonary disease; DL\textsubscript{CO}: carbon monoxide diffusing capacity; FEV1: forced expiratory volume in one second; FVC: forced vital capacity; LASSO: least absolute shrinkage and selection operator; SVM: support vector machine; KNN: K-Near Neighbor; ROC: receiver operating characteristic

**Declarations**

**Ethics approval and consent to participate**

The study was approved by the Ethics Committee of Zhejiang Provincial People's Hospital. The informed consent was waived due to the retrospective nature of the study.(Additional file 2)

**Consent for publication**

Not applicable.

**Availability of data and materials**

The datasets generated and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests
Funding

National Key Research and Development Project(2017YFC0114103);
Key Research and Development Project of Zhejiang Province(2020C01058)

Authors’ contributions

SJC. Data analysis and manuscript writing;

ZYS, YS. Image collection and rating system design;

XYG. Study design and manuscript editing;

PPP. Data training and modelling;

XYG. Study design, manuscript writing and editing;

Acknowledgements

Not applicable.

References

1. Miller, M.R. Structural and physiological age-associated changes in aging lungs. Semin Respir Crit Care Med, 2010. 31(5): 521-527. doi:10.1055/s-0030-1265893

2. Janssens, J.P. Aging of the respiratory system: impact on pulmonary function tests and adaptation to exertion. Clin Chest Med, 2005. 26(3): 469-484.doi:10.1016/j.ccm.2005.05.004

3. Bodini R,Santus P, Di Marco F, Aliberti S, Centanni S, Blasi F, et al. Epidemiology of clinical trials of medicines in respiratory diseases in Europe and Italy. Respir Med, 2017. 125: 8-11. doi:10.1016/j.rmed.2017.02.009

4. Lowery, E.M, Brubaker AL,Kuhlmann E, Kovacs EJ. The aging lung. Clin Interv Aging, 2013. 8: 1489-1496. doi:10.2147/CIA.S51152

5. Hanania NA, King MJ, Braman SS, Saltoun C, Wise RA, Enright P, et al. Asthma in the elderly: Current understanding and future research needs—a report of a National Institute on Aging (NIA) workshop. J Allergy Clin Immunol, 2011. 128(3 Suppl): S4-24. doi: 10.1016/j.jaci.2011.06.048

6. Kelley, R.C. and L.F. Ferreira. Diaphragm abnormalities in heart failure and aging: mechanisms and integration of cardiovascular and respiratory pathophysiology. Heart Fail Rev, 2017. 22(2): 191-207. doi: 10.1007/s10741-016-9549-4

7. Skloot, G.S. The Effects of Aging on Lung Structure and Function. Clin Geriatr Med, 2017. 33(4): 447-457. doi: 10.1016/j.cger.2017.06.001
8. Miller, M. Structural and Physiological Age-Associated Changes in Aging Lungs. Seminars in Respiratory and Critical Care Medicine, 2010. 31(05): 521-527. doi: 10.1055/s-0030-1265893

9. Vaz Fragoso, C.A. and T.M. Gill. Respiratory impairment and the aging lung: a novel paradigm for assessing pulmonary function. J Gerontol A Biol Sci Med Sci, 2012. 67(3): 264-75. doi: 10.1093/gerona/grl198

10. Sayiner, A. Hague C, Ajlan A, Leipsic J, Wierenga L, Krowchuk NM, et al., Bronchiolitis in young female smokers. Respir Med, 2013. 107(5): p. 732-8. doi: 10.1016/j.rmed.2012.12.023

11. Copley SJ, Wells AU, Hawtin KE, Gibson DJ, Hodson JM, Jacques AE, et al. Lung Morphology in the Elderly: Comparative CT Study of Subjects over 75 Years Old versus Those under 55 Years Old. Radiology, 2009.251:566-573. doi:10.1148/radiol.2512081242

12. Chiesa, A.M, Ciccarese F, Gardelli G, Regina UM, Feletti F, Bacchi Reggiani ML, et al. Sonography of the normal lung: Comparison between young and elderly subjects. Journal of Clinical Ultrasound, 2015. 43(4): p. 230-234. doi: 10.1002/jcu.22225

13. Xu, Y, Lu L, E LN, Lian W, Yang H, Schwartz LH, et al. Application of Radiomics in Predicting the Malignancy of Pulmonary Nodules in Different Sizes. American Journal of Roentgenology, 2019:1-8. doi: 10.2214/AJR.19.21490

14. Ji, G.W. Zhang YD, Zhang H, Zhu FP, Wang K, Xia YX, et al. Biliary Tract Cancer at CT: A Radiomics-based Model to Predict Lymph Node Metastasis and Survival Outcomes. Radiology, 2019. 290(1): 90-98. doi: 10.1148/radiol.2018181408

15. Vogelmeier, C.F. Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. Am J Respir Crit Care Med, 2017. 195(5): 557-582. doi: 10.1016/j.arbres.2017.06.001

16. Celli, B.R, W. MacNee, and A.E.T. Force, Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. Eur Respir J, 2004. 23(6):932-946.doi:10.1183/09031936.04.00014304

17. Roca J, Burgos F, Sunyer J, Saez M, Chinn S, Antó JM, et al. References values for forced spirometry. Group of the European Community Respiratory Health Survey. Eur Respir J, 1998. 11(6): 1354-1362.doi:10.1183/09031936.98.11061534

18. Xue, X. Yang Y, Huang Q, Cui F, Lian Y, Zhang S, et al. Use of a Radiomics Model to Predict Tumor Invasiveness of Pulmonary Adenocarcinomas Appearing as Pulmonary Ground-Glass Nodules. Biomed Res Int, 2018. 6803971. doi: 10.1155/2018/6803971

19. Acharya U.R,Dua S,Du X, SRee SV,Chua CK. Automated Diagnosis of Glaucoma Using Texture and Higher Order Spectra Features. IEEE Transactions on Information Technology in Biomedicine, 2011. 15(3): 449-455. doi: 10.1109/TITB.2011.2119322

20. Parmar, C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning methods for Quantitative Radiomic Biomarkers. Sci Rep, 2015. 5:13087. doi: 10.1038/srep13087
21. Boehmer, E.D. Grossmann P, Bussink J, Lambin P, Aerts HJWL. Aging negatively skews macrophage TLR2- and TLR4-mediated pro-inflammatory responses without affecting the IL-2-stimulated pathway. Mechanisms of Ageing and Development, 2005. 126(12):1305-1313. doi:10.1016/j.mad.2005.07.009

22. Maciel-Baron, L.A. Morales-Rosales SL, Aquino-Cruz AA, Triana-Martínez F, Galván-Arzate S, Luna-López A. Senescence associated secretory phenotype profile from primary lung mice fibroblasts depends on the senescence induction stimuli. Age (Dordr), 2016. 38(1): p. 26. doi: 10.1007/s11357-016-9886-1

23. El Assar, M., J. Angulo, and L. Rodriguez-Manas, Oxidative stress and vascular inflammation in aging. Free Radic Biol Med, 2013. 65: 380-401. doi: 10.1016/j.freeradbiomed

24. Calabresi, C. Arosio B, Galimberti L, et al., Natural aging, expression of fibrosis-related genes and collagen deposition in rat lung. Experimental Gerontology, 2007. 42(10): 1003-1011. doi:10.1016/j.exger.2007.06.016

25. Mohamed Hoesein, F.A. de Hoop B, Zanen P, Gietema H, et al., CT-quantified emphysema in male heavy smokers: association with lung function decline. Thorax, 2011. 66(9): 782-7. doi: 10.1136/thx.2010.145995

26. Vidic, I. Egnell L, Jerome NP, Teruel JR, Sjøbakk TE, Østlie A, Support vector machine for breast cancer classification using diffusion-weighted MRI histogram features: Preliminary study. J Magn Reson Imaging, 2018. 47(5): 1205-1216. doi: 10.1002/jmri.25873

27. de Jong, P.A, Müller NL, Paré PD, Coxson HO. Computed tomographic imaging of the airways: relationship to structure and function. Eur Respir J, 2005. 26(1):140-52. doi:10.1183/09031936.05.00007105

28. Yuan, R., Hogg JC, Paré PD, Sin DD, Wong JC, Nakano Y. Prediction of the rate of decline in FEV(1) in smokers using quantitative Computed Tomography. Thorax, 2009. 64(11):944-9. doi: 10.1136/thx.2008.112433

29. Lafata, K.J, Zhou Z, Liu JG, Hong J, Kelsey CR, Yin FF, An Exploratory Radiomics Approach to Quantifying Pulmonary Function in CT Images. Scientific Reports, 2019. 9(1). doi: 10.1038/s41598-019-48023-5

30. Kerstjens, H.A. Rijcken B, Schouten JP, Postma DS, et al., Decline of FEV1 by age and smoking status: facts, figures, and fallacies. Thorax, 1997. 52(9): 820-7. doi:10.1136/thx.52.9.820

31. McDonough, J.E. Yuan R, Suzuki M, Seyednejad N, Elliott WM, Sanchez PG, Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. N Engl J Med, 2011. 365(17): 1567-1575. doi: 10.1056/NEJMoa1106955

32. Lalley, P.M. The aging respiratory system–pulmonary structure, function and neural control. Respir Physiol Neurobiol, 2013. 187(3): 199-210. doi: 10.1016/j.resp.2013.03.012

33. Meier, A. Farrow C, Harris BE, King GG, Jones A. Application of texture analysis to ventilation SPECT/CT data. Comput Med Imaging Graph, 2011. 35(6): 438-50. doi: 10.1016/j.compmedimag.2011.01.001
Barjaktarevic I, Springmeyer S, Gonzalez X, Sirokman W, Coxson HO, Cooper CB. Diffusing capacity for carbon monoxide correlates best with tissue volume from quantitative CT scanning analysis. Chest, 2015. 147(6): 1485-1493. doi: 10.1378/chest.14-1693

Tables

**Table 1.** The baseline characteristics of Older Group and Younger Group

| Variable                        | Older Group (n=118) | Younger Group (n=141) | P-value |
|---------------------------------|--------------------|-----------------------|---------|
| Age (years)                     | 64.78 ± 4.8        | 34.5 ± 9.5            | <0.001* |
| Man-to-woman ratio              | 52 :66             | 79 :62                | NS†     |
| Ex-smoker–to-nonsmoker ratio    | 26:92               | 15:126                | NS†     |
| Body mass index (kg/m²)         | 27.2± 4.7          | 28.7± 5.2             | NS*     |

Note.—Unless otherwise specified, data are means± standard deviations. NS=not statistically significant.

* Paired t test. † 2 Test.

**Table 2.** Texture parameters after the dimensionality reduction

| Category | Feature in radiomics signature |
|----------|--------------------------------|
| Histogram| Percentile70*                  |
|          | Mean Deviation                 |
|          | Percentile70*                  |
|          | Quantile0.975                  |
|          | Range                          |
| Texture  | ClusterShade_AllDirection_offset1_SD |
|          | Inertia_AllDirection_offset2_SD |
| GLCM     | Correlation_angle45_offset9    |
|          | InverseDifferenceMoment_AllDirection_offset2_SD |
|          | InverseDifferenceMoment_AllDirection_offset6_SD* |
| RLM      | GreyLevelNonuniformity_AllDirection_offset1_SD* |
|          | LongRunEmphasis_angle0_offset7 |
|          | LongRunEmphasis_angle0_offset2 |
|          | LowGreyLevelRunEmphasis_AllDirection_offset2_SD |
|          | HighGreyLevelRunEmphasis_AllDirection_offset8_SD |
|          | RunLengthNonuniformity_AllDirection_offset8_SD* |
|          | ShortRunEmphasis_AllDirection_offset1_SD |
|          | ShortRunLowGreyLevelEmphasis_AllDirection_offset9_SD |
|          | ShortRunEmphasis_AllDirection_offset3_SD |
|          | ShortRunEmphasis_AllDirection_offset8_SD* |
| Form     | Compactness1                   |

*Show common features
| Method                | Training cohort | Validation cohort |
|----------------------|-----------------|------------------|
|                      | Accuracy | F1_score | Sensitivity | Specificity | AUC | Accuracy | F1_score | Sensitivity | Specificity | AUC |
| Logistic regression  | 0.77     | 0.43      | 0.759       | 0.797       | 0.84 | 0.84     | 0.45      | 0.893       | 0.708       | 0.85 |
| SVM                  | 0.674    | 0.061     | 0.982       | 0.301       | 0.667 | 0.78     | 0.1       | 0.996       | 0.203       | 0.64 |
| Bayes                | 0.653    | 0.411     | 0.802       | 0.359       | 0.748 | 0.598    | 0.108     | 0.746       | 0.105       | 0.578|
| KNN                  | 0.732    | 0.44      | 0.944       | 0.312       | 0.863 | 0.768    | 0.345     | 0.921       | 0.263       | 0.761|
| Decision Tree        | 0.837    | 0.627     | 0.993       | 0.45        | 0.853 | 0.634    | 0.211     | 0.873       | 0.148       | 0.538|
| Random forest        | 0.784    | 0.453     | 0.985       | 0.304       | 0.922 | 0.671    | 0.229     | 0.927       | 0.148       | 0.630|

SVM, support vector machine; KNN, K-Near Neighbor; AUC, area under the curve; CI, confidence interval
### Table 3B. Discrimination performance of the different model construction methods in the HS and HNS groups

| Method            | Training cohort | Validation cohort |  |  |  |  |  |  |  |
|-------------------|-----------------|------------------|---|---|---|---|---|---|---|
|                   | Accuracy        | F1-score         | Sensitivity | Specificity | AUC | Accuracy | F1-score | Sensitivity | Specificity | AUC |
| Logistic regression | 0.80            | 0.667            | 0.857       | 0.729       | 0.87 | 0.76     | 0.691     | 0.769       | 0.750       | 0.80 |
| SVM               | 0.742           | 0.66             | 0.915       | 0.541       | 0.793 | 0.737    | 0.694     | 0.862       | 0.607       | 0.836 |
| Bayes             | 0.727           | 0.64             | 0.831       | 0.582       | 0.800 | 0.719    | 0.68      | 0.828       | 0.735       | 0.825 |
| KNN               | 0.742           | 0.595            | 0.948       | 0.455       | 0.810 | 0.649    | 0.600     | 0.957       | 0.441       | 0.721 |
| Decision Tree     | 0.78            | 0.752            | 0.831       | 0.721       | 0.865 | 0.632    | 0.51      | 0.783       | 0.529       | 0.696 |
| Random forest     | 0.841           | 0.821            | 0.887       | 0.787       | 0.926 | 0.737    | 0.724     | 0.759       | 0.714       | 0.761 |

SVM, support vector machine; KNN, K-Near Neighbor; AUC, area under the curve; CI, confidence interval.

### Table 4. Diagnostic accuracy of the radiomics signature of the two models

| Parameter                  | Model 1 (COPD vs. non-COPD) | Model 2 (HS vs. HNS) |
|---------------------------|-----------------------------|---------------------|
|                           | Training cohort             | Validation cohort   | Training cohort | Validation cohort |
| AUC                       | 0.84                        | 0.85                | 0.87            | 0.80              |
| AUC 95% CI                | 0.79~0.90                   | 0.76~0.95           | 0.81~0.93       | 0.68~0.93         |
| Accuracy                  | 0.77                        | 0.84                | 0.80            | 0.76              |
| Accuracy 95% CI           | 0.70~0.83                   | 0.74~0.91           | 0.71~0.86       | 0.62~0.87         |
| Sensitivity (%)           | 75.94                       | 89.29               | 85.71           | 76.92             |
| Specificity (%)           | 79.66                       | 70.83               | 72.88           | 75.00             |
| PPV                       | 89.38                       | 87.72               | 77.14           | 76.92             |
| NPV                       | 59.49                       | 73.91               | 82.69           | 75.00             |

AUC, area under the curve; CI, confidence interval; PPV, predictive positive value; NPV, negative positive value

### Figures
Figure 1

The workflow of the radiomics signature construction

A: Radiomics (COPD vs non-COPD)

B: Radiomics (HS vs HNS)

Figure 2

The score diagrams of the radiomics signature in the training and validation cohorts of two populations. A: COPD (yellow), non-COPD (blue). B: HS (yellow), HNS (blue). A high score means a high risk of lung damage.
Figure 3

The ROC analysis of Model 1 for the discrimination between the COPD and non-COPD in the training (A) and validation (B) cohorts.
**Figure 4**

The ROC analysis of Model 2 for the discrimination between HSs and HNSs in the training (A) and validation (B) cohorts.

**Figure 5**

The heat map of correlation coefficients between the selected features and PF tests.
Figure 6

The heat map of correlation coefficients among the selected features.

Supplementary Files

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

- SupplementaryMaterial.docx