Master's Thesis of Medicine

Image and AI-based Diffusing Capacity for Carbon Monoxide Prediction in Idiopathic Pulmonary Fibrosis

특발성 폐섬유증에서 영상 및 인공지능 기반 일산화탄소확산능 예측

August 2023

Department of Medical Device Development
Graduate School
Seoul National University

Wonchul Chung

Image and AI-based Diffusing Capacity for Carbon Monoxide Prediction in Idiopathic Pulmonary Fibrosis

Submitting a Master's Thesis of Medicine

April 2023

Department of Medical Device Development
Graduate School
Seoul National University

Wonchul Chung

Confirming the master's thesis written by
Wonchul Chung

July 2023

Chair	임재준	(Seal)
Vice Chair	이창현	(Seal)
Examiner	이정찬	(Seal)

Abstract

Idiopathic pulmonary fibrosis (IPF) is a progressive lung disease characterized by the formation of scar tissue, resulting in impaired gas exchange. Diffusing capacity of the lung for carbon monoxide (DLCO) measures the degree of gas exchange in lung alveoli. We aimed to explore and develop artificial intelligence (AI) models to predict DLCO using flow-volume curve images and quantitative CT features (QCT).

Flow-volume curves from 272 IPF patients enrolled at Seoul National University Hospital (SNUH) in South Korea from 2015 to 2019 (age=69±8, height=162±8cm, weight=63±10kg, M:F=201:71) were retrospectively collected and analyzed. QCT features from 60 IPF patients (age=70±5, height=164±6cm, weight=68±9kg male:female=53:7) were retrospectively collected from 2018 environmental lung disease study cohort (ENV18) at SNUH and analyzed.

$DLCO\%_{pred}$ was successfully predicted from flow-volume curve (mean absolute error=4.33, R^2=0.91). $DLCO\%_{pred}$ prediction with QCT features were limited due to small sample size, but the data provided insights on regional lung structure-function relationship in

DLCO%$_{pred}$ prediction.

Successful prediction of DLCO from spirometry-measured flow-volume curve images suggest that the dynamic pattern of the flow-volume curve may contain comprehensive information of the lung structure-function. Future development of the model may include predicting follow-up DLCO after 1 year. Further application may elucidate how specific shape or image features from flow-volume curve and QCT data link gas ventilation and gas exchange.

Keyword : flow-volume curve, diffusing capacity of carbon monoxide, machine learning, deep learning, idiopathic pulmonary fibrosis

Student Number : 2021-24213

Table of Contents

List of Tables

List of Figures

Abbreviation

CFD	Computational fluid dynamics
CNN	Convolutional neural networks
$DLCO\%_{pred}$	Diffusing capacity of the lungs for carbon monoxide
EX	Expiratory
FEV1	Forced expiratory volume in one second
FVC	Forced vital capacity
IN	Inspiratory
IPF	Idiopathic pulmonary fibrosis
PFT	Pulmonary function test
QCT	Quantitative computed tomography
RRAVC	Relative regional air volume change
QCT	Quantitative computed tomography

Chapter 1. Introduction

1.1. Background

Idiopathic pulmonary fibrosis (IPF) is the most common and severe idiopathic subtype of interstitial lung disease (ILD) characterized by progressive inflammation and scarring around the alveoli (Martinez et al., 2017). Damage in alveoli leads to impaired gas exchange between air space and blood vessels, which is the primary function of the lung to maintain the life of a human body. IPF is diagnosed by structural change in the lung where histological and/or radiological pattern of usual interstitial pneumonia is evident (Raghu et al., 2018; Romei et al., 2020). However, the relationship between structural changes and functional decline in IPF is not fully understood.

Functional decline in IPF is, similarly to other chronic lung diseases, measured most commonly by pulmonary function test (PFT), which includes spirometry, body plethysmography, and measure of diffusing capacity of the lung for carbon monoxide (DLCO). PFT provides useful quantitative measures in identifying and managing pulmonary abnormalities (Crapo, 1994). Forced vital

capacity (FVC) from spirometry test is used as a reliable measure for functional capacity of the lung with IPF (Hoesterey et al., 2019; Du Bois et al., 2011; Fernandez-Villar et al., 2018). DLCO is also frequently employed as it provides information about the efficiency of gas exchange in the lungs (Morton, 2008). Both spirometry and DLCO parameters are used for evaluating states and progression of IPF (Song et al., 2019; Fainberg et al., 2022; Raghu et al., 2022). Notwithstanding the utility of a single quantitative parameter or combination of a few in clinical practice such as the percent-predicted values of FVC (FVC%$_{pred}$) and DLCO (DLCO%$_{pred}$), they may not fully reflect the structure-function relationship of the lung due to its complex pathophysiologic nature.

We noted that the flow-volume curve from spirometry may contain comprehensive information of lung function. The flow-volume curve provides a graphical representation of dynamic lung function that results from a comprehensive interplay between multiscale (large and small) structural and functional features of the entire lung. The hidden information may relate the two major functional components of the lung, ventilation (or, gas transfer in the airways) and gas exchange (between airways and blood vessels). Thus, we hypothesize that it is possible to predict DLCO utilizing the inherent information of multiscale structure-function relationship in

the flow-volume curve.

On the other hand, computed tomography (CT) imaging is another major assessment tool for IPF. Recent advances in CT imaging and quantitative CT (QCT) analysis suggested promising applications of regional lung structure and function in IPF and other lung diseases (Kang et al., 2021; Choi et al., 2022), as well as in normal lung physiology and subclinical abnormality in normal-appearing lungs (Choi et al., 2017; Shin et al., 2020; Li et al., 2022). We hypothesize that the QCT-based structural and functional features of the airways, lobes, and the whole lung may provide valuable information in predicting DLCO.

In recent years, advancements in artificial intelligence (AI) algorithms such as (shallow) machine learning and deep learning techniques have revolutionized the field of medical image analysis (Haghighi et al., 2019; Saood and Hatem, 2021; Shelke et al., 2021; Zou et al., 2022; Li et al., 2021, 2022;). These techniques enabled the extraction of hidden information from medical images, which potentially predictive of clinical parameters. As an expansion of the medical image analysis, we pondered the potential capturing of hidden information from the flow-volume curve obtained during the forced vital capacity maneuver of the spirometry test.

1.2. Purpose

Our goal in this study is to investigate the utility of image data with implicitly nested information and AI algorithms in predicting $DLCO\%_{pred}$. Primarily, we aimed to develop a predictive model of $DLCO\%_{pred}$ in IPF patients, based on deep learning of the flow-volume curve and basic demographic information. Additionally, to enhance the accuracy of $DLCO\%_{pred}$ prediction, we incorporate QCT-based structural and functional features of the lung.

Three objectives are itemized below.

1. Develop an AI model for predicting $DLCO\%_{pred}$ in IPF patients using flow-volume curve images.
2. Explore enhanced AI prediction of $DLCO\%_{pred}$ in IPF utilizing QCT imaging-based multiscale lung structure-function features and compare predictive performance across different models
3. Discuss the interplay between regional lung structure-function relationship, in bridging whole lung functional measurements of ventilation (in the flow-volume curve) and gas exchange (DLCO).

The study may provide a non-invasive and efficient means

of assessing gas exchange in IPF, particularly for the patients with difficulties in DLCO testing, and new insights in functional assessment in IPF.

1.3. Thesis Overview

Detailed methodological approaches utilized in the study are introduced in Chapter 2. Results are reported in Chapter 3. Interpretation of the findings in DLCO prediction in association with complicated lung structure–function relationship in IPF is discussed in Chapter 4. Conclusions, limitations, and future works are presented in the final chapter, Chapter 5.

Chapter 2. Materials and Methods

2.1. Research Design and Patient Data Collection

PFT results including tabular spirometry measures, flow-volume curve images, and DLCO%$_{pred}$, patient demographics were retrospectively collected (IRB approval no. 2007-174-1143) from 272 participants enrolled in Seoul National University Hospital in South Korea with IPF (age=69±8 years, height=162±8 cm, weight=63±10 kg, male:female=201:71) (**Table 2.1**). Patient inclusion criteria is presented in **Figure 2.1.**

Additionally, inspiratory and expiratory CTs and DLCO%$_{pred}$ measurements of 60 IPF patients (age=70±5, height=164±6cm, weight=68±9kg male:female=53:7) (**Table 2.2**) were retrospectively collected from 2018 environmental lung disease study (ENV18) Cohort (IRB approval no. 1810-036-977 and with subject consent, using a common dose-reduced QCT protocol) at Seoul National University Hospital in South Korea. Patient inclusion criteria is presented in **Figure 2.2.**

2.2. Image Data Denoising

Image data of flow-volume curve were first cleaned up through denoising process. We utilized K-means clustering and *in house* python scripts for image denoising. K-means clustering is a machine learning technique used for grouping similar data points into k clusters. It is an iterative process that k centroids are randomly chosen from the data, and distances between each data point and each centroid are computed. The data points are then grouped into clusters based on the nearest centroids, and the algorithm continues to select k best centroids and cluster data points until the assignments no longer change. For our purpose, since original flow-volume curve had highly variable color spectrum pixel-wise, certain number of clusters, k, was set to 5 to denoise and extract desired graph of a specific color (**Figure 2.3**). For our purpose, we extracted only the reference and pre-bronchodilator flow-volume curve.

2.3. Development of the Deep Leaning Model

We utilized a multimodal deep learning model with early fusion. The model combines both flow-volume curve image

(128x128 pixels) and tabular features (age, weight, height, gender) in the early stages of the network architecture. The model utilizes the pretrained EfficientNet-B3 architecture from PyTorch (https://pytorch.org/) as a backbone for image feature extraction. The curve images are passed through a series of convolutional layers and batch normalization, followed by the EfficientNet-B3 model. The output of the EfficientNet-B3 model is then concatenated with the tabular features, which are passed through a separate set of fully connected layers. The concatenated features are fed into a linear layer to obtain the predicted percent $DLCO\%_{pred}$ value. (**Figure 2.4**)

A multimodal model was employed to leverage the complementary information present in both the images and tabular features, allowing the model to learn joint representations and capture complex relationships between the different modalities, resulting in improved $DLCO\%_{pred}$ prediction.

For additional deep learning hyperparameters, a learning rate scheduler was employed to adjust the learning rate during training, and a one-cycle learning rate policy was utilized to find an optimal learning rate. The training process was repeated for 5 folds in a cross-validation setup.

2.4. Development of the Machine Leaning Model

Tabular data was analyzed using LightGBM, a gradient boosting framework that employs tree-based learning algorithms. Prior to feature selection, the input features were normalized. The scaling technique transforms the feature values to a standardized range between 0 and 1, ensuring that each feature contributes proportionally to the model training process. For QCT features, features showing significant correlation with $DLCO\%_{pred}$ were selected as input features. All dataset was divided into a developmental training set (80%, n=217) and an external test set (20%, n=55).

To optimize the performance of the LightGBM model, hyperparameter tuning is conducted using the randomized search cross validation (Randomized Search CV) from Python's scikit-learn library (version 1.0). These options include the number of estimators, learning rate, maximum depth, column sample by tree, and subsample. After hyperparameter optimization, the best-performing hyperparameters are used to train the LightGBM model.

2.5. CT Image Processing and QCT

Volumetric CT images acquired at full inspiration (IN) and at full expiration (EX) were retrospectively collected. CT image data were acquired from Philps Ingenuity CT with a protocol for QCT analysis: 0.67 mm or 1.0 mm slice thickness, 120 kVp tube voltage, 93~95 mAs, YC or B reconstruction filter, 0.5 s rotation time.

From those CT images, segmentation masks basic structural measurements were generated for regions of airways, lungs, lobes, and blood vessels using VIDA Vision 2.2 software (Coralville, IA, USA).

With image matching between local voxels of IN and EX CT images via symmetric mass-preserving nonrigid image registration (**Yin et al., 2009; Haghighi et al., 2018**), 113 local to global lung structural and functional variables were computed (**Choi et al., 2017; Chae et al., 2020; Shin et al., 2020, Kang et al., 2021; Zou et al., 2021; Choi et al. 2022**), using an in-house QCT software. For example, mean lung density (HU), tissue fraction (TF), total lung volume (V), tissue volume (TV), high attenuation area percent (HAA%), emphysema percent (Emph%) air trapping percent (AirT%), air volume (AV), air volume difference between inspiratory and

expiratory CTs (ΔAV) were computed in voxels based on the CT density.

CT imaging-based computational fluid dynamics (CFD) airflow simulation of subject-specific tidal breathing was conducted for additional functional assessment. To conduct the simulations for all subjects, we used one-dimensional (1D) network-based CFD model (Choi 2011; Choi et al. 2019), instead of three-dimensional (3D) CFD simulations that provide more details of local flow characteristics but computationally too expensive to run for all the subjects. The 1D CFD simulations were conducted using CT-based 3D-1D coupled entire conducting airway model for anatomical accuracy and inspiratory-expiratory CT image matching-based local vital capacity (VC) information at every modeled terminal bronchiole. Airway resistance and lung compliance models were coupled in the equation. An entry flow model (Pedley et al. 1970; 1977) incorporated simplified but accurate flow physics in human bronchial branching system.

2.6. Model Evaluation and Statistical Analysis

Performance of both deep learning and machine learning

model were evaluated with mean absolute error (MAE) and R-squared (R^2). Additionally, two-sample Student's t-test to compare the distributions of the predicted and actual $DLCO\%_{pred}$ values. The Bland-Altman analysis was conducted to assess the agreement between the predicted and actual $DLCO\%_{pred}$ values. The Bland-Altman analysis includes markers for the mean difference, lower and upper limits of agreement, and scatter points representing the individual data points. It helps identify any systematic bias or heteroscedasticity in the differences between the predicted and actual $DLCO\%_{pred}$ values.

All statistical analysis was performed using SciPy (version 1.8.1) in Python 3 (version 3.8.5) packages. Pearson correlation was used for correlation analysis and figures were generated using Matplotlib (version 3.5.2) and Seaborn (version 0.11.0) in Python 3 packages. A p-value less than 0.05 was considered statistically significant.

Table 2.1. IPF patient characteristics of dataset n=272.

Characteristics	IPF (n=272)
Demographics*	
Age, years	69.9±8.5
Height, cm	162.5±7.9
Weight, kg	63.1±10.5
Sex, male, %	201, 73.9%
Clinical Variables*	
$FEV_{1\%pred}$	89.3±21.8
$FVC_{\%pred}$	76.3±19.7
FEV_1/FVC	80.7±7.8
$DLCO_{\%pred}$	62.5±19.1

* Data are Mean ± SD, or n, %.

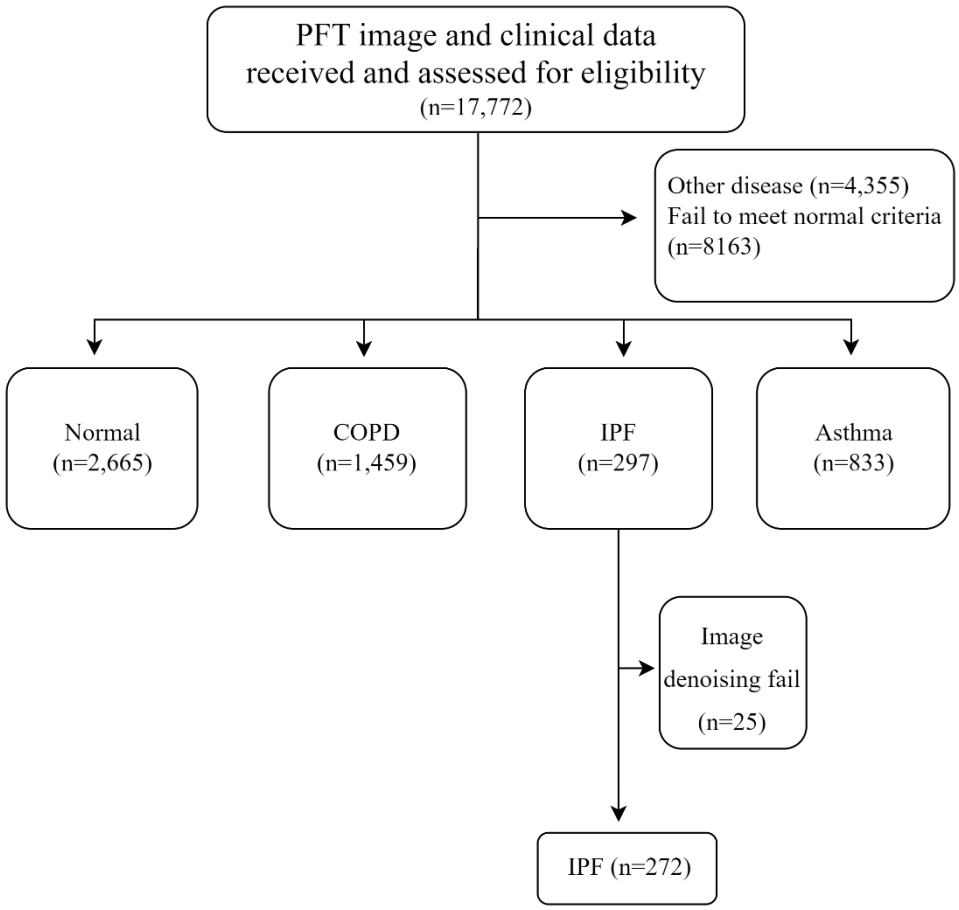

Figure 2.1. Patient flow diagram of dataset n=272.

Table 2.2. IPF Patient characteristics of dataset n=60.

Characteristics	IPF (n=60)
Demographics*	
Age, years	70.6±5.4
Height, cm	164±6.6
Weight, kg	68.0±9.3
Sex, male	53, 88.3%
Clinical Variables*	
$FEV_{1\%pred}$	95.5±18.9
$FVC_{\%pred}$	83.1±16.6
$DLCO_{\%pred}$	67.7±17.5

* Data are Mean ± SD, or n, %.

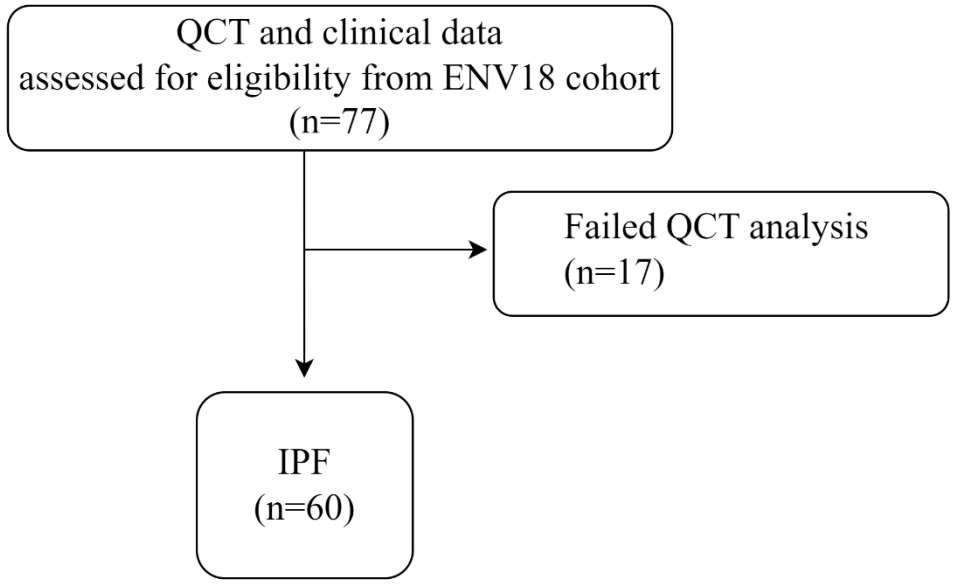

Figure 2.2. Patient flow diagram of dataset n=60.

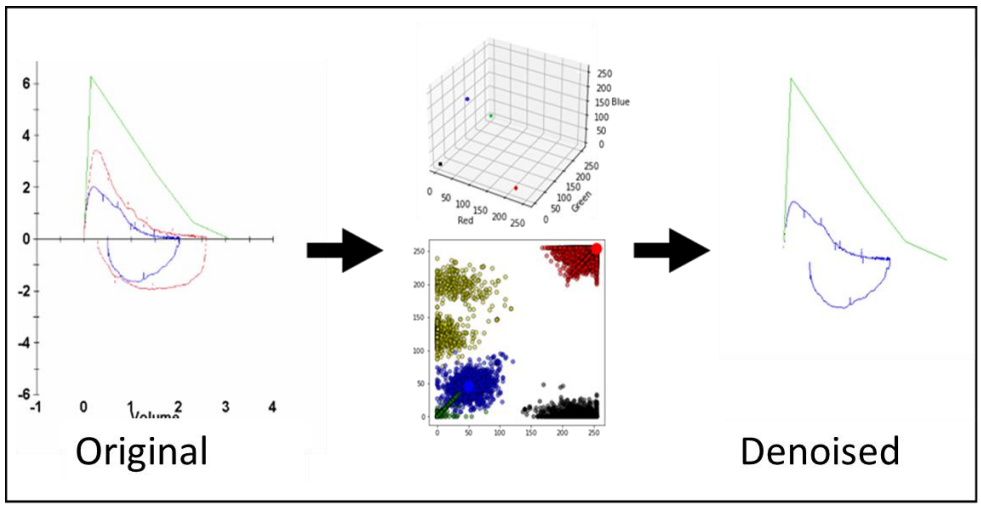

Figure 2.3. Schematic of image denoising process using K-means clustering. Reference (green) flow-volume and pre-bronchodilator flow-volume graph are extracted as final output.

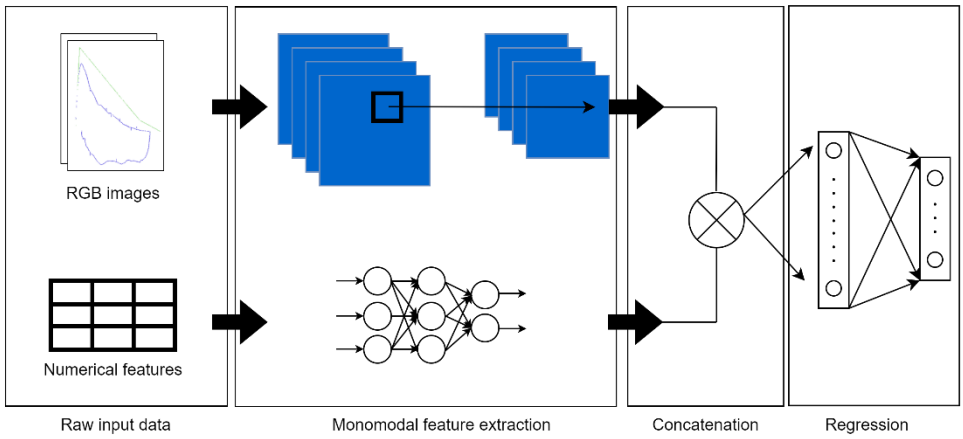

Figure 2.4. Schematic of multi modal input deep leaning architecture for DLCO%$_{pred}$ prediction. Pretrained CNN architecture (EfficientNet-B3) was used for image feature extraction and custom multi-layer perceptron (MLP) was used for tabular data feature extraction.

Chapter 3. Results

3.1. Predicting DLCO%$_{pred}$ with Flow-Volume Curve

This section presents the results of the DLCO%$_{pred}$ prediction using the flow-volume curve. The accuracy and agreement between the predicted DLCO%$_{pred}$ values and the actual values were assessed through various statistical measures (**Figure 3.1**). The scatter plot in **Figure 3.1a**, compared individual prediction value from the model and actual DLCO%$_{pred}$ values. MAE and R^2 score are reported to assess model performance. In **Figure 3.2b**, the box plots of the predicted DLCO%pred and the actual values indicate distribution between the predicted and actual are in good agreement. The Bland-Altman analysis revealed a mean bias of -3.74 with a standard deviation of 3.75. The 95% limits of agreement ranged from -11.09% to 3.61%. The results indicate very small systematic difference between the predicted and actual DLCO%$_{pred}$ values (**Figure 3.1c**). A strong positive correlation was observed between the predicted and actual DLCO%$_{pred}$ values (r=0.98, P=<0.01). This suggests that the predicted values closely match the actual values and exhibit a high degree of agreement (**Figure 3.1d**).

The mean bias of −3.74 indicates a slight underestimation of $DLCO\%_{pred}$, but the limits of agreement (−11.09% to 3.61%) indicate that the predicted values generally fall within an acceptable range of the actual values.

3.2. Predicting $DLCO\%_{pred}$ with Spirometry Data

This section presents the results of the $DLCO\%_{pred}$ prediction using tabular spirometry data. The accuracy and agreement between the predicted $DLCO\%_{pred}$ values and the actual values were evaluated using various statistical measures.

Figure 3.2 illustrates the inference on the test set using tabular spirometry data. **Figure 3.2a** displays the MAE and R^2 score, providing insights into the predictive performance. **Figure 3.2b** indicates that there is no significant difference between the predicted $DLCO\%_{pred}$ values and the actual values of all patients. **Figure 3.2c** presents a Bland−Altman plot comparing the predicted and actual $DLCO\%_{pred}$ values, which shows a mean bias of 2.81 with a standard deviation of 15.53. The 95% limits of agreement range from −27.63% to 33.27%. These findings suggest a slight systematic difference between the predicted and actual $DLCO\%_{pred}$ values. **Figure 3.2d**

demonstrates a good correlation between the predicted and actual $DLCO\%_{pred}$ values ($r=0.74$, $P=< 0.001$). This indicates a moderate level of agreement between the predicted and actual values.

The results indicate that the predicted $DLCO\%_{pred}$ values derived from tabular spirometry data are in a good agreement with the actual values. Although a mean bias of 2.81 suggests a slight systematic difference, the 95% limits of agreement (−27.63% to 33.27%) encompass a reasonable range of variation.

3.3. Predicting $DLCO\%_{pred}$ with QCT

This section presents the results of the $DLCO\%_{pred}$ prediction using QCT data. The accuracy and agreement between the predicted $DLCO\%_{pred}$ values and the actual values were assessed using various statistical measures.

Prediction based only on QCT features were inferior. MAE was 15.98 and R^2 score turned to be negative, showing a poor fit. There was no significant correlation between the predicted $DLCO\%_{pred}$ values and the actual values of all patients. **Figure 3.3a** presents a Bland−Altman plot comparing the predicted and actual $DLCO\%_{pred}$ values. The plot shows a bias of -6.70 ± 19.43. The 95%

limits of agreement range from −44.80% to 31.40%. These findings suggest a noticeable systematic difference between the predicted and actual $DLCO\%_{pred}$ values. **Figure 3.3b** demonstrates a moderate but not significant correlation between the predicted and actual $DLCO\%_{pred}$ values ($r=0.39$, $p=0.21$). This indicates a weaker level of agreement than the model using flow−volume curve.

The results suggest that QCT data from a relatively small sample size ($n=60$) may have limited predictability for $DLCO\%_{pred}$ compared to the previous two modeling cases. The mean bias of −6.70 indicates a noticeable systematic difference, and the wider 95% limits of agreement (−44.80% to 31.40%) suggest a larger range of variation.

3.4. Performance Comparison with Limited Dataset

This section presents a same sample size comparison of the three $DLCO\%_{pred}$ prediction models using (1) flow−volume curve, (2) tabular spirometry data, and (3) QCT data. The performance of each model is evaluated based on MAE and R^2 score. Each model was trained on 60 data sets (80% train, 20% test) (**Figure 3.4**).

For the model using the flow−volume curve, the MAE is 39.04,

indicating a relatively large deviation from the actual DLCO%$_{pred}$ values (**Figure 3.4a**). The R^2 score of −2.84 suggests a poor fit of the predicted values to the actual values.

For the model using tabular spirometry data, the MAE is 18.96, indicating a smaller deviation than the flow−volume curve model. However, the R^2 score of 0.16 still shows a weak correlation between the predicted and actual DLCO%$_{pred}$ values (**Figure 3.4b**).

Lastly, for the model using QCT data, the MAE is 15.98, representing the lowest deviation compared to the other two models. This model also showed a poor fit (**Figure 3.4c**). However, significant moderate correlation with DLCO%$_{pred}$ was found in individual QCT features such as high attenuation area percent (HAA%) in the whole lung and the left upper (LUL) and lower (LLL) lobes (r=−0.47, −0.39, −0.35; P=<0.001, <0.001, 0.001, respectively), anisotropic deformation index (ADI) in the whole lung, LUL and LLL (r=0.28, 0.30, 0.38; P=0.02, 0.017, 0.003, respectively).

In summary, based on the comparison of the three DLCO%$_{pred}$ prediction models using a limited sample size, the model based on the QCT data showed the smallest error and the model based on flow−volume curve exhibits the largest error while the flow−volume curve−based model showed an excellent performance with a large sample size. From the results, we could expect an improvement of

the QCT-based prediction with a larger sample size.

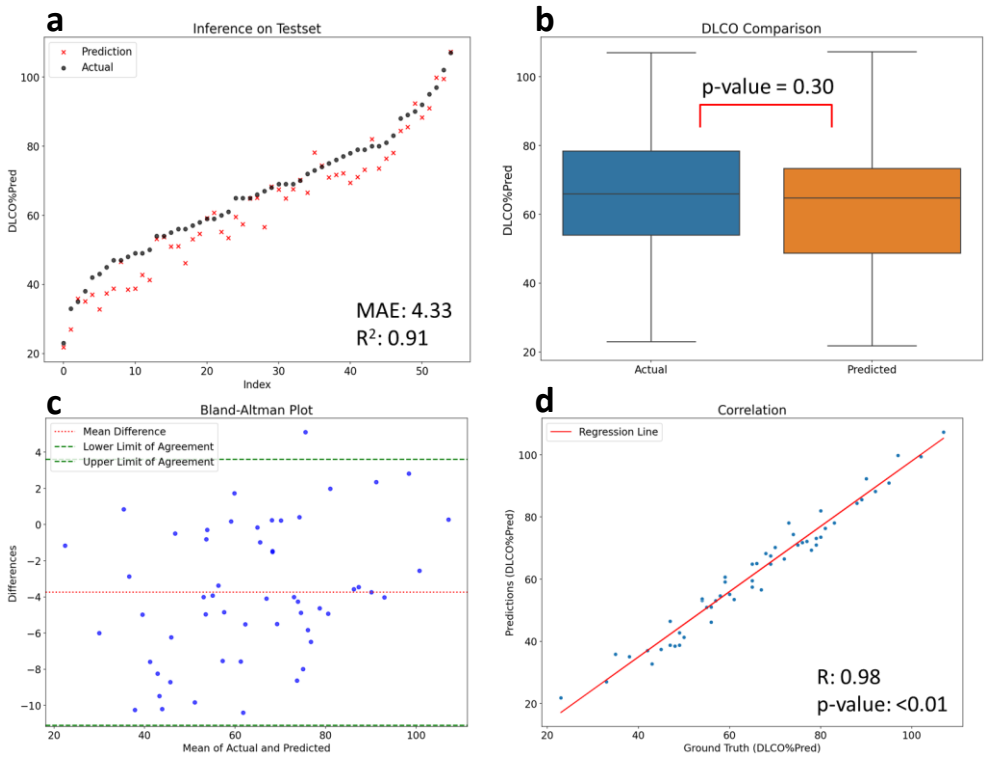

Figure 3.1. Comparison between predicted $DLCO\%_{pred}$ from flow-volume curve and actual $DLCO\%_{pred}$ values. (a) Inference on test set with MAE and R^2 score. (b) Box plots of the predicted $DLCO\%_{pred}$ and the actual values. (c) Bland-Altman plot of the prediction and actual values. (d) Correlation between predicted values and actual values ($R=0.98$; $P<0.01$).

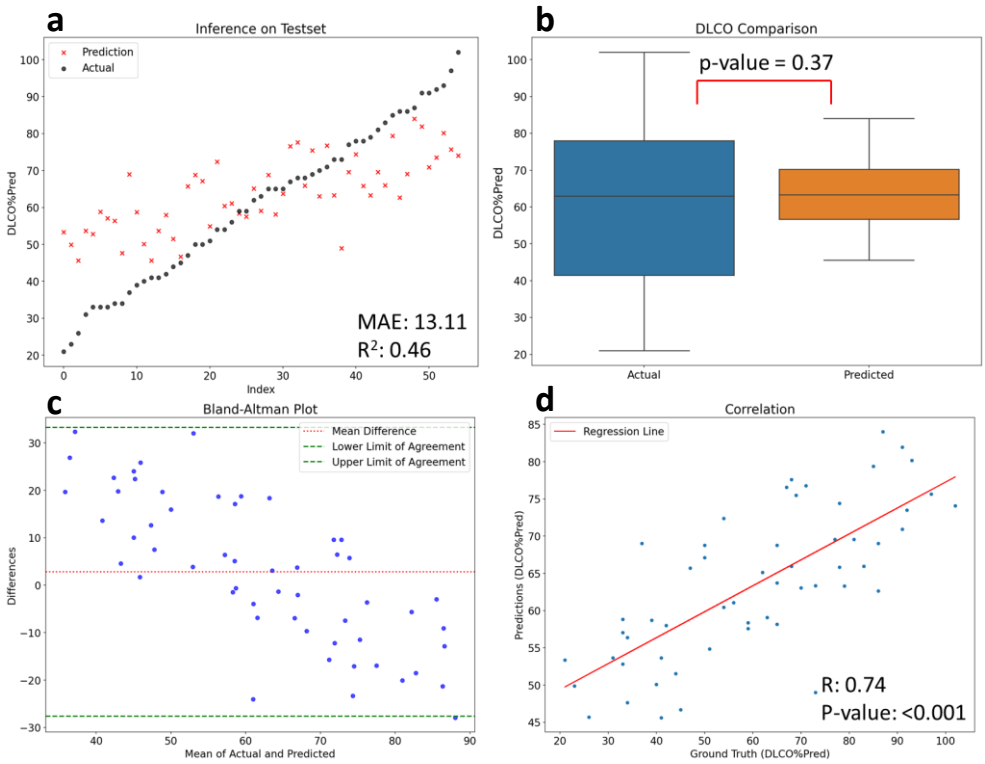

Figure 3.2. Comparison between predicted DLCO%$_{pred}$ using spirometry data and actual DLCO%$_{pred}$ values. (a) Inference on test set with MAE and R^2 score. (b) Box plots of the predicted DLCO%$_{pred}$ and the actual values. (c) Bland–Altman plot of the prediction and actual values. (d) Correlation between predicted values and actual values (R=0.74; P<0.001).

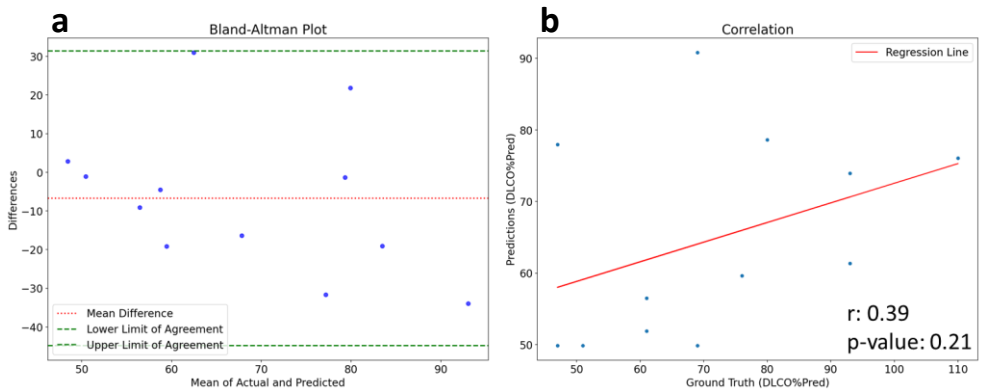

Figure 3.3. QCT-based model performance. (a) Bland-Altman plot of the prediction and actual values. (b) Correlation between the predicted and the actual values.

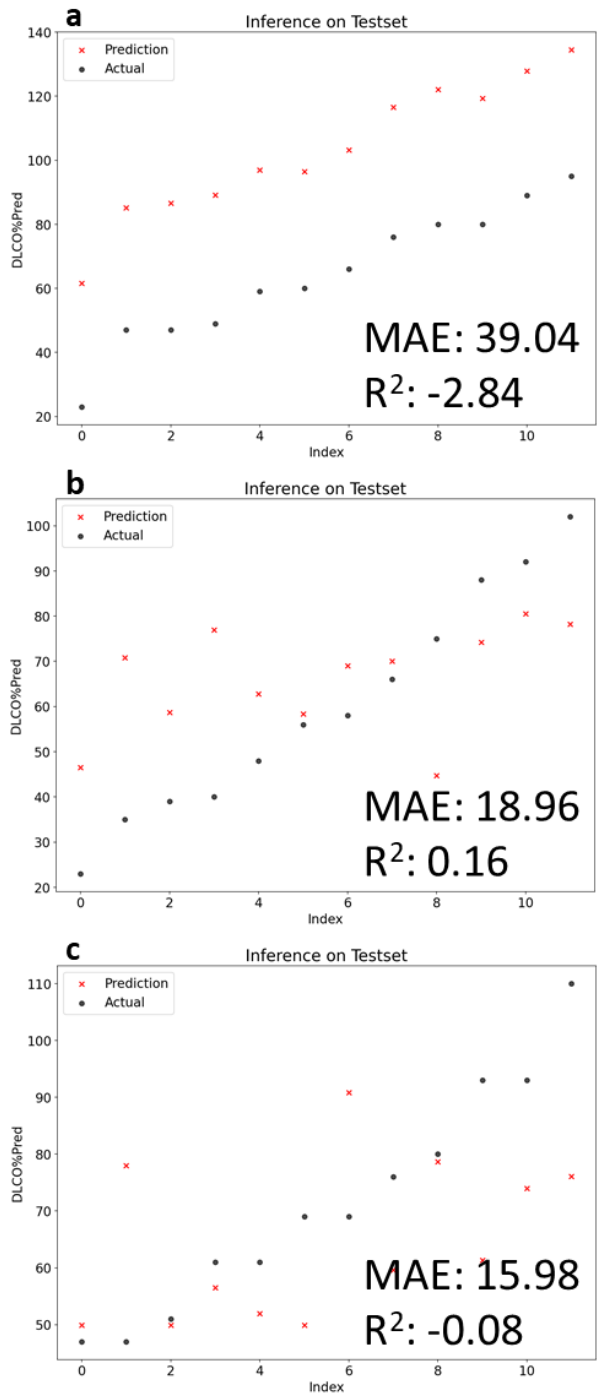

Figure 3.4. Comparison of the three models with the same sample size (n=60): (a) flow-volume curve-based, (b) tabular spirometry data-based, and (c) QCT-based models.

Chapter 4. Discussion

4.1. Model Performance Comparison

The flow-volume curve-based model showed an excellent performance to predict $DLCO\%_{pred}$ (MAE=4.33, R^2=0.91). Despite the great performance, the mean bias and limits of agreement (**Figures 3.1a** and **3.1b**) indicated that the model slightly underestimated $DLCO\%_{pred}$. The tabular spirometry data-based model also showed a moderate performance inferior to the flow-volume curve-based model. The tabular spirometry data-based model exhibited a tendency of overestimating $DLCO\%_{pred}$ values. The QCT data-based model showed the worst performance of all three, attributable to the small sample size of the QCT data compared to the data used in the other two models. When all three models were evaluated with the same sample size, the QCT model showed rather a better performance, indicating it may outperform other models if an enough sample size is available.

To investigate the associations of demographic (weight, height, and age) and spirometry ($FEV1\%_{pred}$, $FVC\%_{pred}$, and FEV1/FVC) data with $DLCO\%_{pred}$, Pearson's correlation analysis

was conducted (**Figure 4.1**). The highest correlation was found for $FVC\%_{pred}$ ($r=0.52$, $P<0.001$), which is still much smaller than the correlation between the flow-volume curve-based prediction and the actual values ($r=0.95$, $P<0.01$). This again elucidates the great performance of the flow-volume curve-based $DLCO\%_{pred}$ prediction, which cannot be obtained by a single demographic of spirometry value.

4.2. Regional Structure and Function Relationship

DLCO is a measure for the capability of gas exchange in alveoli, which is an ultimate function of the lung for the systemic respiratory function. While DLCO is associated with spirometry-measured lung function, particularly FVC, the structure-function relationship that links these two measures is not well understood. And we tried to provide insights based on "image" data from the PFT data and CT data, since PFT measures provide the whole lung function and CT can capture regional status of the lung.

High-resolution CT images of lung regions include detailed visual information of the disease status. However, in most clinical settings, CT scans rely heavily on qualitative interpretation and are

mainly used for structural analysis. Qualitative assessment is subject to observer bias and variability (Ostridge and Wilkinson, 2016).

In the past decades, there has been a significant progress in CT-based quantitative analysis of pulmonary structure and function using inspiratory and expiratory CT images (Zagers et al., 1996; Dowson et al., 2001; Dirksen et al., 1997, 1999; Choi et al., 2015; Choi et al., 2017; Barros et al., 2018; Haghighi et al., 2018, 2019; Chae et al., 2020; Shin et al., 2020; Kang et al., 2021; Zou et al., 2021; Li et al., 2021; Lin et al., 2018; Hoffman 2022; Choi et al., 2022; Suman and Koo., 2023). QCT analysis extracts various structural and functional variables at different scales, not only for the entire lung but also for localized regions, providing improved preclinical pathophysiological and clinical understanding (Hoffman et al., 2014; Podolanczuk et al., 2016; Choi et al., 2015; Choi et al., 2017; Haghighi et al., 2018, 2019; Chae et al., 2020; Shin et al., 2020; Kang et al., 2021; Zou et al., 2021; Li et al., 2021; Choi et al., 2022). Despite advanced QCT imaging and image-based modeling approaches in assessing structural and functional alterations in lungs, the transition of these techniques to clinical application has been challenging due to high computational demand and limited methodological consistency that bridge both preclinical and clinical lung structure-function interpretation (Chen-Mayer, 2017; Felder and Walsh, 2023).

We conducted supplementary correlation analysis between DLCO%$_{pred}$ and QCT features of the 60 patients with QCT data, and found a good correlation between DLCO%$_{pred}$ and high attenuation area percent (HAA%) in the whole lung and the left upper and lower lobes showed significant negative correlations ($r=-0.35$, -0.47, -0.39; $P=0.001$, <0.001, <0.001, respectively). Greater HAA% is known to be associated with more reduced lung function, biomarkers of inflammation, and mortality rate in ILD, which is an umbrella term for diseases characterized by inflammation and scarring in the lungs including IPF (**Podolanczuk et al., 2017**). Our results agree that HAA% correlated well with lower lung function particularly DLCO%$_{pred}$ that implies lower gas exchange ability in alveoli.

We may need to note that the ultimate role of the lung is gas exchange between the internal (blood) body and external (atmosphere) environment. In other words, gas exchange is in fact the primary function of the lung for the whole body and gas delivery throughout the airways supports it, whilst spirometry data is getting more attention than DLCO among PFT results.

The importance of our work is that the great performance of flow-volume curve-based model in predicting DLCO suggests that there exists a strong relationship between the dynamic pattern of the ventilatory function measured at the mouth, i.e., the proximal end of

the conducting airways, and the gas-exchanging function at the most distal region of the lung.

4.3. Limitations and Future Work

This study is a novel work exploring the relationship between ventilation and gas exchange via AI; however, it is not without limitations. First, AI models usually requires a large amount of data for better performance and generalizability. We were only able to collect a relatively small sample size, particularly for the QCT data. We may need to collect more samples from a larger patient pool across different institutions in the future for more robust model performance and generalizability. For example, although the trained models were evaluated on a separate test set, external validation using external independent test set is desirable. Second, our multimodal deep learning model may lack a viable interpretability or transparency. At the most basic level, we do not know how individual neurons correlates with each other to transform input data into output (Sheu, 2020). We may incorporate grad-CAM-based color visualization or SHapley Additive exPlanations (SHAP) methodology for interpretation of AI models. Grad-CAM examines the gradient of

the model's output with respect to the intermediate feature maps of the CNN (Panwar et al., 2020). SHAP values are used for explaining the output of machine learning models in terms of the contributions of each feature to the model's prediction (Rodriguez-Perez & Bajorath, 2020). For future work, we may incorporate longitudinal data, predicting DLCO changes over time.

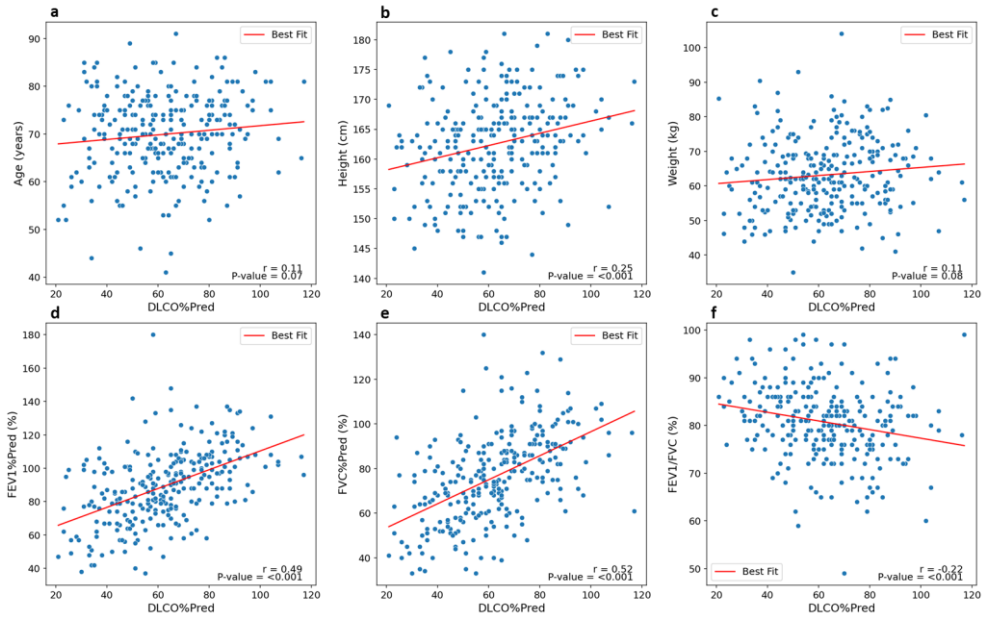

Figure 4.1. Correlation of DLCO%$_{pred}$ with tabular spirometry data and demographic data: (a) age, (b) height, (c) weight, (d) FEV1%$_{pred}$, (e) FVC%$_{pred}$, and (f) FEV1/FVC.

Chapter 5. Conclusions

5.1. Overview

We successfully predicted $DLCO\%_{pred}$ in IPF from flow-volume curve images via multimodal deep learning model with transfer learning of pretrained CNN model (EfficnetNet-B3). The flow-volume curve-based model (MAE=4.33, R^2=0.91) outperformed tabular spirometry-based model (MAE=13.11, R^2=0.46). Findings suggest that flow-volume curve may include comprehensive information of regional or multiscale lung structure-function relationship, where "multiscale" indicates local voxels to the global (whole) lung.

5.2. $DLCO\%_{pred}$ prediction with flow-volume curve

Deep learning modeling of flow-volume curve images accurately predicted the $DLCO\%_{pred}$ values in IPF patients. This approach of DLCO prediction may be extended to patients with other ILD subtypes, who were able to undergo spirometry but not DLCO testing.

5.3. DLCO%$_{pred}$ prediction with Spirometry Data

Using tabular spirometry features with machine learning showed moderate performance, underperforming relative to using flow-volume curve but improved performance compared to QCT features.

5.4. DLCO%$_{pred}$ prediction with QCT

Attributable to a small sample size, QCT features did not successfully predict the values of DLCO%$_{pred}$ in IPF patients. However, QCT-based regional lung structure-function features showed significant moderate correlations with DLCO%$_{pred}$. Association of DLCO%$_{pred}$ with the HAA% values suggests that reduced gas exchange is associated with subpleural lung damage that is a primary characteristic of IPF and further disease progress with inflammation and fibrosis. The correlation between HAA% and DLCO%$_{pred}$ was highest in the LUL, implying that the degree of the LUL impairment by HAA% in the LUL can be a good indicator of the progression of the disease. We speculate that such regional lung structural and functional alteration is reflected in the dynamic pattern

of the flow-volume curve. It may worth noting that the study with QCT data from a larger dataset may provide more insights in linking regional lung structural and functional alteration with $DLCO\%_{pred}$. Findings and discussions of the current study provide insights on the pathophysiological changes in lung structure and function in IPF patients.

5.5. Concluding Remarks

The aim of this study was to explore the possibility of using image data in predicting $DLCO\%_{pred}$ using AI algorithms in IPF. We developed a deep learning model predicting $DLCO\%_{pred}$ from flow-volume curve images. We also developed a machine learning model taking tabular spirometry and QCT data as input, from preliminary data sets. Lastly, we suggest that the flow-volume curve may include comprehensive information of the lung structure-function relationship and hence reflect of the interplay between regional lung structure and function throughout the entire respiratory tract, providing new insights of bridging two aspects of lung function (gas delivery and gas exchange).

AI can accelerate healthcare practitioner's decision making and improve patient health outcome (**Kumar et al., 2023**). For safe

and effective implementation of AI decision support tools in clinical setting, not only improving model performance is important but also securing model interpretability, uncertainty handling and model development based on correct understanding of pathophysiological processes are imperative (Giri et al., 2021; Tran et al., 2021).

References

Barros MC, Hochhegger B, Altmayer S, Watte G, Zanon M, Sartori AP, Cavalet Blanco D, Pacini GS, Chatkin JM. Quantitative computed tomography phenotypes, spirometric parameters, and episodes of exacerbation in heavy smokers: An analysis from South America. PLoS One. 2018 Oct 11;13(10):e0205273. Doi: 10.1371/journal.pone.0205273. PMID: 30307987; PMCID: PMC6181358.

Chae KJ*, Choi J*, Jin GY, Hoffman EA, Laroia AT, Park M, Lee CH. Relative regional air volume change maps at the acinar scale reflect variable ventilation in low lung attenuation of COPD patients. Acad Radiol 2020 Nov;27(11):1540-1548. Doi: 10.1016/j.acra.2019.12.004. Epub 2020 Feb 3. (*co-first author)

Chen-Mayer HH, Fuld MK, Hoppel B, Judy PF, Sieren JP, Guo J, Lynch DA, Possolo A, Fain SB. Standardizing CT lung density measure across scanner manufacturers. Med Phys. 2017 Mar;44(3):974-985. Doi: 10.1002/mp.12087. Epub 2017 Feb 21. PMID: 28060414; PMCID: PMC6276120.

Choi S, Hoffman EA, Wenzel SE, Castro M, Fain SB, Jarjour NN, Schiebler ML, Chen K, Lin CL. Quantitative assessment of multiscale structural and functional alterations in asthmatic populations. J Appl Physiol (1985). 2015 May 15;118(10):1286-98. doi: 10.1152/japplphysiol.01094.2014. Epub 2015 Mar 26. PMID: 25814641; PMCID: PMC4436982.

Choi J, Hoffman EA, Lin C-L, Milhem MM, Tessier J, Newell JD Jr. Quantitative computed tomography determined regional lung mechanics in normal nonsmokers, normal smokers and metastatic sarcoma subjects. PloS ONE 2017;12(7): e0179812. Doi: 10.1371/journal.pone.0179812.

Choi J, Chae KJ, Jin GY, Lin CL, Laroia AT, Hoffman EA, Lee CH. CT-based lung motion differences in patients with usual interstitial pneumonia and nonspecific interstitial pneumonia. Front Physiol 2022 Oct 4;13:867473. Doi: 10.3389/fphys.2022.867473. PMID: 36267579; PMCID: PMC9577177.

Choi S, Yoon S, Jeon J, Zou C, Choi J, Tawhai MH, Hoffman EA, Delvadia R, Babiskin A, Walenga R, Lin CL. 1D network simulations for evaluating regional flow and pressure distributions in healthy and asthmatic human lungs. J Appl Physiol (1985). 2019 Jul

1;127(1):122-133. Doi: 10.1152/japplphysiol.00016.2019. Epub 2019 May 16. PMID: 31095459; PMCID: PMC6692748.

Crapo RO. Pulmonary-function testing. N Engl J Med. 1994;331(1):25-30. Doi: 10.1056/nejm199407073310107. PubMed PMID: 8202099.

Dirksen A, Friis M, Olesen KP, Skovgaard LT, S ø rensen K. Progress of emphysema in severe alpha 1-antitrypsin deficiency as assessed by annual CT. Acta Radiol. 1997 Sep;38(5):826-32. Doi: 10.1080/02841859709172418. PMID: 9332238.

Dowson LJ, Guest PJ, Hill SL, Holder RL, Stockley RA. High-resolution computed tomography scanning in alpha1-antitrypsin deficiency: relationship to lung function and health status. Eur Respir J. 2001 Jun;17(6):1097-104. Doi: 10.1183/09031936.01.00056501. PMID: 11491150.

Du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, et al. Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. American journal of respiratory and critical care medicine. 2011;184(12):1382-9.

Fainberg HP, Oldham JM, Molyneau PL, Allen RJ, Kraven LM, Fahy WA, Porte J, Braybrooke R, Saini G, Karsdal MA, Leeming DJ, Sand JMB, Triguero I, Oballa E, Wells AU, Renzoni E, Wain LV, Noth I, Maher TM, Stewart ID, Jenkins RG. Forced vital capacity trajectories in patients with idiopathic pulmonary fibrosis: a secondary analysis of a multicentre, prospective, observational cohort. Lancet Digit Health. 2022 Dec;4(12):e862-e872. Doi: 10.1016/S2589-7500(22)00173-X. Epub 2022 Nov 1. Erratum in: Lancet Digit Health. 2023 Jan;5(1):e4. PMID: 36333179.

Felder FN, Walsh SLF. Exploring computer-based imaging analysis in interstitial lung disease: opportunities and challenges. ERJ Open Res. 2023 Jul 3;9(4):00145-2023. Doi: 10.1183/23120541.00145-2023. PMID: 37404849; PMCID: PMC10316044.

Fernandez-Villar A, Represas-Represas C, Mouronte-Roibas C, Ramos-Hern□ndez C, Priegue-Carrera A, Fernandez-Garcia S, Lopez-Campos JL. Reliability and usefulness of spirometry performed during admission for COPD exacerbation. PloS One. 2018 Mar 26;13(3):e0194983

Giri PC, Chowdhury AM, Bedoya A, Chen H, Lee HS, Lee P, Henriquez C, MacIntyre NR, Huang YT. Application of Machine

Learning in Pulmonary Function Assessment Where Are We Now and Where Are We Going? Front Physiol. 2021 Jun 24;12:678540. doi: 10.3389/fphys.2021.678540. PMID: 34248665; PMCID: PMC8264499.

Haghighi B, D Ellingwood N, Yin Y, Hoffman EA, Lin CL. A GPU-based symmetric non-rigid image registration method in human lung. Med Biol Eng Comput. 2018 Mar;56(3):355-371. Doi: 10.1007/s11517-017-1690-2. Epub 2017 Aug 1. PMID: 28762017; PMCID: PMC5794656.

Haghighi B, Choi S, Choi J, Hoffman EA, Comellas AP, Newell JD, Jr., et al. Imaging-based clusters in former smokers of the COPD cohort associate with clinical characteristics: the SubPopulations and intermediate outcome measures in COPD study (SPIROMICS). Respir Res. 2019;20(1):153. Epub 20190715. Doi: 10.1186/s12931-019-1121-z. PubMed PMID: 31307479; PubMed Central PMCID: PMC6631615.

Heussel CP, Herth FJ, Kappes J, Hantusch R, Hartlieb S, Weinheimer O, Kauczor HU, Eberhardt R. Fully automatic quantitative assessment of emphysema in computed tomography: comparison with pulmonary function testing and normal values. Eur Radiol. 2009 Oct;19(10):2391-402. Doi: 10.1007/s00330-009-1437-z. Epub 2009 May 21. PMID: 19458953.du

Hoesterey D, Das N, Janssens W, Buhr RG, Martinez FJ, Cooper CB, et al. Spirometric indices of early airflow impairment in individuals at risk of developing COPD: Spirometry beyond FEV(1)/FVC. Respir Med. 2019;156:58-68. Epub 20190809. Doi: 10.1016/j.rmed.2019.08.004. PubMed PMID: 31437649; PubMed Central PMCID: PMC6768077.

Hoffman EA. Origins of and lessons from quantitative functional X-ray computed tomography of the lung. Br J Radiol. 2022 Apr 1;95(1132):20211364. doi: 10.1259/bjr.20211364. Epub 2022 Mar 1. PMID: 35193364; PMCID: PMC9153696.

Kang JH, Choi J*, Chae KJ, Shin KM, Lee CH, Guo J, Lin CL, Hoffman EA, Lee C*. CT-derived 3D-diaphragm motion in emphysema and IPF compared to normal subjects. Sci Rep 2021 Jul 21;11(1):14923. doi: 10.1038/s41598-021-93980-5. PMID: 34290275; PMCID: PMC8295260. (*co-correspondence)

Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz

Comput. 2023;14(7):8459-8486. doi: 10.1007/s12652-021-03612-z. Epub 2022 Jan 13. PMID: 35039756; PMCID: PMC8754556.

Li F, Choi J, Zhang X, Rajaraman PK, Lee CH, Ko H, et al. Characterizing Subjects Exposed to Humidifier Disinfectants Using Computed-Tomography-Based Latent Traits: A Deep Learning Approach. Int J Environ Res Public Health. 2022;19(19). Epub 20220920. doi: 10.3390/ijerph191911894. PubMed PMID: 36231196; PubMed Central PMCID: PMC9565839.

Li F, Choi J, Zou C, Newell JD Jr, Comellas AP, Lee CH, Ko H, Barr RG, Bleecker E, Cooper CB, Abtin F, Barjaktarevic I, Couper D, Han ML, Hansel NN, Kanner RE, Paine R III, Kazerooni EA, Martinez FJ, O' Neal W, Rennard SI, Smith BM, Woodruff PG, Hoffman EA, Lin CL. Latent traits of lung tissue patterns in former smokers derived by dual channel deep learning in computed tomography images. Sci Rep 2021;11:4916. doi:10.1038/s41598-021-84547-5.

Lin CL, Choi S, Haghighi B, Choi J, Hoffman EA. Cluster-Guided Multiscale Lung Modeling via Machine Learning. In: Andreoni, W., Yip, S. (eds) Handbook of Materials Modeling. 2018. Springer, Cham. https://doi.org/10.1007/978-3-319-50257-1_98-1

Martinez FJ, Collard HR, Pardo A, Raghu G, Richeldi L, Selman M, et al. Idiopathic pulmonary fibrosis. Nat Rev Dis Primers. 2017;3:17074. Epub 20171020. doi: 10.1038/nrdp.2017.74. PubMed PMID: 29052582.

Modi P, Cascella M. Diffusing Capacity Of The Lungs For Carbon Monoxide. 2023 Mar 13. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. PMID: 32310609.

Morton, A.R. Chapter 8—Exercise Physiology. In Pediatric Respiratory Medicine, 2nd ed.; Taussig, L.M., Landau, L.I., Eds.; Mosby: Philadelphia, PA, USA, 2008; pp. 89–99.Ni Y, Yu Y, Dai R, Shi G.

Morton AR. Chapter 8—Exercise Physiology. Pediatric Respiratory Medicine, 2nd ed; Taussig, LM, Landau, LI, Eds. 2008:89-99.

Ni Y, Yu Y, Dai R, Shi G. Diffusing capacity in chronic obstructive pulmonary disease assessment: A meta-analysis. Chron Respir Dis. 2021 Jan-Dec;18:14799731211056340. doi: 10.1177/14799731211056340. PMID: 34855516; PMCID: PMC8649441.

Panwar H, Gupta PK, Siddiqui MK, Morales-Menendez R, Bhardwaj P, Singh V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. Chaos Solitons Fractals. 2020 Nov;140:110190. doi: 10.1016/j.chaos.2020.110190. Epub 2020 Aug 7. PMID: 32836918; PMCID: PMC7413068.

Pedley TJ. Pulmonary fluid dynamics. Annual Review of Fluid Mechanics. 1977 Jan;9(1):229-74.

Podolanczuk AJ, Oelsner EC, Barr RG, Hoffman EA, Armstrong HF, Austin JH, Basner RC, Bartels MN, Christie JD, Enright PL, Gochuico BR, Hinckley Stukovsky K, Kaufman JD, Hrudaya Nath P, Newell JD Jr, Palmer SM, Rabinowitz D, Raghu G, Sell JL, Sieren J, Sonavane SK, Tracy RP, Watts JR, Williams K, Kawut SM, Lederer DJ. High attenuation areas on chest computed tomography in community-dwelling adults: the MESA study. Eur Respir J. 2016 Nov;48(5):1442-1452. doi: 10.1183/13993003.00129-2016. Epub 2016 Jul 28. PMID: 27471206; PMCID: PMC5089905.

Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J, Cottin V, Danoff SK, Morell F, Flaherty KR, Wells A, Martinez FJ, Azuma A, Bice TJ, Bouros D, Brown KK, Collard HR, Duggal A, Galvin L, Inoue Y, Jenkins RG, Johkoh T, Kazerooni EA, Kitaichi M, Knight SL, Mansour G, Nicholson AG, Pipavath SNJ, Buend□a-Rold□n I, Selman M, Travis WD, Walsh S, Wilson KC; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2018 Sep 1;198(5):e44-e68. doi: 10.1164/rccm.201807-1255ST. PMID: 30168753.

Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, Kreuter M, Lynch DA, Maher TM, Martinez FJ, Molina-Molina M, Myers JL, Nicholson AG, Ryerson CJ, Strek ME, Troy LK, Wijsenbeek M, Mammen MJ, Hossain T, Bissell BD, Herman DD, Hon SM, Kheir F, Khor YH, Macrea M, Antoniou KM, Bouros D, Buendia-Roldan I, Caro F, Crestani B, Ho L, Morisset J, Olson AL, Podolanczuk A, Poletti V, Selman M, Ewing T, Jones S, Knight SL, Ghazipura M, Wilson KC. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022 May 1;205(9):e18-e47. doi: 10.1164/rccm.202202-0399ST. PMID: 35486072; PMCID:

PMC9851481.

Rodriguez-Perez R, Bajorath J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. J Comput Aided Mol Des. 2020 Oct;34(10):1013-1026. doi: 10.1007/s10822-020-00314-0. Epub 2020 May 2. PMID: 32361862; PMCID: PMC7449951.

Romei C, Tavanti LM, Taliani A, De Liperi A, Karwoski R, Celi A, Palla A, Bartholmai BJ, Falaschi F. Automated Computed Tomography analysis in the assessment of Idiopathic Pulmonary Fibrosis severity and progression. Eur J Radiol. 2020 Mar;124:108852. doi: 10.1016/j.ejrad.2020.108852. Epub 2020 Jan 28. PMID: 32028067.

Saood A, Hatem I. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. BMC Med Imaging. 2021 Feb 9;21(1):19. doi: 10.1186/s12880-020-00529-5. PMID: 33557772; PMCID: PMC7870362.

Shelke A, Inamdar M, Shah V, Tiwari A, Hussain A, Chafekar T, Mehendale N. Chest X-ray Classification Using Deep Learning for Automated COVID-19 Screening. SN Comput Sci. 2021;2(4):300. doi: 10.1007/s42979-021-00695-5. Epub 2021 May 26. PMID: 34075355; PMCID: PMC8152712.

Sheu YH. Illuminating the Black Box: Interpreting Deep Neural Network Models for Psychiatric Research. Front Psychiatry. 2020 Oct 29;11:551299. doi: 10.3389/fpsyt.2020.551299. PMID: 33192663; PMCID: PMC7658441.

Shin KM*, Choi J*, Chae KJ, Jin GY, Eskandari A, Hoffman EA, Hall C, Castro M, Lee CH. Quantitative CT-based image registration metrics provide different ventilation and lung motion patterns in prone and supine positions in healthy subjects. Respir Res 2020 Oct 2;21(1):254. doi: 10.1186/s12931-020-01519-5. PMID: 33008396. (*co-first author)

Song H, Sun D, Ban C, Liu Y, Zhu M, Ye Q, et al. Independent Clinical Factors Relevant to Prognosis of Patients with Idiopathic Pulmonary Fibrosis. Med Sci Monit. 2019;25:4193-201. Epub 20190605. doi: 10.12659/msm.914725. PubMed PMID: 31166938; PubMed Central PMCID: PMC6563649.

Suman G, Koo CW. Recent Advancements in Computed Tomography Assessment of Fibrotic Interstitial Lung Diseases. J Thorac Imaging. 2023 Mar 22. doi: 10.1097/RTI.0000000000000705. Epub ahead of print. PMID: 37015833.

Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med. 2021 Sep 27;13(1):152. doi: 10.1186/s13073-021-00968-x. PMID: 34579788; PMCID: PMC8477474.

Weatherley ND, Eaden JA, Hughes PJC, Austin M, Smith L, Bray J, Marshall H, Renshaw S, Bianchi SM, Wild JM. Quantification of pulmonary perfusion in idiopathic pulmonary fibrosis with first pass dynamic contrast-enhanced perfusion MRI. Thorax. 2021 Feb;76(2):144-151. doi: 10.1136/thoraxjnl-2019-214375. Epub 2020 Dec 3. PMID: 33273022; PMCID: PMC7815896.

Yin Y, Hoffman EA, Lin CL. Mass preserving nonrigid registration of CT lung images using cubic B-spline. Med Phys. 2009 Sep;36(9):4213-22. doi: 10.1118/1.3193526. PMID: 19810495; PMCID: PMC2749644.

Yin Y*, Choi J*, Hoffman EA, Tawhai MH, Lin CL. A multiscale MDCT image-based breathing lung model with time-varying regional ventilation. J Comput Phys 2013;244:168-192. doi: 10.1016/j.jcp.2012.12.007. (*equally contributed)

Zagers H, Vrooman HA, Aarts NJ, Stolk J, Schultze Kool LJ, Dijkman JH, Van Voorthuisen AE, Reiber JH. Assessment of the progression of emphysema by quantitative analysis of spirometrically gated computed tomography images. Invest Radiol. 1996 Dec;31(12):761-7. doi: 10.1097/00004424-199612000-00005. PMID: 8970878.

Zou C, Li F, Choi J, Haghighi B, Choi S, Rajaraman PK, Comellas AP, Newell JD Jr, Lee CH, Barr RG, Bleecker E, Cooper CB, Couper D, Han ML, Hansel NN, Kanner RE, Kazerooni EA, Kleerup EC, Martinez FJ, O' Neal W, Paine R III, Rennard SI, Smith BM, Woodruff PG, Hoffman EA, Lin CL. Longitudinal Imaging-Based Clusters in Former Smokers of the COPD Cohort Associate with Clinical Characteristics: the SubPopulations and Intermediate Outcome Measures in COPD Study (SPIROMICS). Int J Chron Obstruct Pulmon Dis. 2021 May 31;16:1477-1496. doi: 10.2147/COPD.S301466. PMID: 34103907; PMCID: PMC8178702.

Abstract in Korean

특발성 폐 섬유화 (Idiopathic pulmonary fibrosis, IPF)는 흉부에 흉터 조직이 형성되어 기체 교환에 장애를 초래하는 진행성 폐질환이다. 일산화탄소확산능 검사 (Diffusing capacity of the lung for carbon monoxide, DLCO)는 폐포의 기체 교환 정도를 측정하는 방법이다. 기류-용적 곡선 (flow-volume) 이미지와 정량적 CT (Quantitative CT, QCT) 변수들을 이용하여 DLCO를 예측하기 위한 인공지능(AI) 모델을 개발하는 것을 목표로 한다.

2015년부터 2019년까지 한국의 서울대병원(SNUH)에 등록된 272명의 IPF 환자의 기류-용적 곡선(연령=69±8, 신장=191±8cm, 체중=63±10kg, 남:여=201:71)을 후향적으로 수집하여 분석하였다. 60명의 IPF 환자(연령=70±5, 신장=196±6cm, 체중=68±9kg 남:여=53:7)의 QCT 변수들은 서울대병원의 2018 환경 폐 질환 연구 코호트(ENV18)에서 후향적으로 수집하여 분석하였다.

기류-용적 곡선을 통해 $DLCO\%_{pred}$를 성공적으로 예측할 수 있었다. (MAE=4.33, R^2=0.91). 이는 기류-용적 곡선이 폐기능에 대한 포괄적인 병리 생리학적 정보를 포함하고 있을 수 있다는 것을 시사한다. QCT 변수를 사용한 $DLCO\%_{pred}$ 예측은 샘플 크기가 작기 때문에 제한되었지만 이들은 지역적 폐 구조-기능 관계에 대한 통찰력을 제공한다.

폐활량계 (spirometry)의 기류-용적 곡선 이미지로부터 DLCO의 성공적인 예측은 기류-용적 곡선의 동적 패턴이 폐 구조-기능의 포괄적인 정보를 포함할 수 있음을 시사한다. 향후 모델의 개발은 1년 후의 DLCO를 예측하는 분석을 진행할 수 있다. 더 나아가 산소 환기와 기체 교환과 관련된 특정한 형태나 이미지 특징들이 기류-용적 곡선과 QCT 데이터로부터 어떻게 연관되는지에 대한 추가적인 분석이 진행될 수 있다.