Extended Quantitative Computed Tomography Analysis of Lung Structure and Function

In Kyu Lee

B.S. in Aerospace Engineering and Mechanics, University of Minnesota, 2020

Submitted to the graduate degree program in Bioengineering and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Science.

Dr. William Brooks, Chairperson

Committee members

Dr. Jiwoong Choi, Advisor

Dr. Suzanne Shontz

Date defended: July 15, 2022

The Thesis Committee for In Kyu Lee certifies
that this is the approved version of the following thesis :

Extended Quantitative Computed Tomography Analysis of Lung Structure and Function

Dr. William Brooks, Chairperson

Date approved: July 15, 2022

Abstract

Computed tomography (CT) imaging and quantitative CT (QCT) analysis for the study of lung health and disease have been rapidly advanced during the past decades, along with the employment of CT-based computational fluid dynamics (CFD) and machine learning approaches. The work presented in this thesis was devoted to extending the QCT analysis framework from three different perspectives.

First, to extend the advanced QCT analysis to more data with undesirably protocolized CT scans, we developed a new deep learning-based automated segmentation of pulmonary lobes, incorporating z-axis information into the conventional UNet segmentation. The proposed deep learning segmentation, named ZUNet, was successfully applied for QCT analysis of silicosis patients with thick (5 or 10 mm) slices, which used to be excluded in QCT analysis since three-dimensional (3D) volumetric segmentation of the lungs and lobes were hardly successful or not automated. ZUNet outperformed UNet in lobe segmentation of human lungs.

In addition, we extended the application of the QCT framework, combining CFD simulations for the entire subjects of the QCT analysis. One-dimensional (1D) CFD simulations of tidal breathing have been added to the inspiratory-expiratory CT image matching analysis of 66 asthma patients (M:F=23:43, age=64.4$\pm$10.7) for pre- and post-bronchodilator comparison. We aimed to characterize comprehensive airway and lung structure and function relationship in the entire group response and patient-specific response to the bronchodilator. Along with the evidence of large airway dilatation in the entire asthmatics, the CFD analysis revealed that improvements in regional flow rate fraction, particularly in the right lower lobe (RLL), airway pressure drop, airway resistance, and workload of breathing were significantly associated with the degree of large airway dilatation.

Finally, we extended the approach using machine learning analysis to integrate numerous QCT

variables with clinical features and additional information such as environmental exposure. In pursuit of investigating the effects of particulate matter (PM) exposure on human lung structure and function alteration, principal component analysis (PCA) and k-means clustering identified low, mid, and high exposure groups from directly measured air pollution exposure data of 270 healthy (age=68±10, M:F=15:51), asthma (age=60±12, M:F=39:56), chronic obstructive pulmonary disease (COPD) (age=69±7, M:F=66:10), and idiopathic pulmonary fibrosis (IPF) (age=72±7, M:F=43:10) subjects. Based on the exposure clusters, the RLL segmental airway narrowing was observed in the high exposure group. Various associations were found between the exposure data and about 200 multiscale lung features, from quantitative inspiratory and expiratory CT image matching and 1D CFD tidal breathing simulations. To highlight, small PM increases small airway disease in asthma. PM at all sizes decreases inspiratory low attenuation area in COPD and diseases luminal diameter of the RLL segmental airways in IPF.

Acknowledgements

The presented work would not have been possible without many people's incredible support and help. First of all, my most sincere gratitude and appreciation go to my M.S. advisor Dr. Jiwoong Choi. Meeting him was my life turning point. I am grateful for giving me a great opportunity to participate in lung research. Since I came to KU, I have had the privilege of working with Dr. Choi. Also, his comprehensive knowledge of computational modeling, image processing, and medicine constantly stimulated exciting discussions and research topics. I enjoyed every discussion we had.

I am grateful for working as a member of the Airway & Asthma Translational Research Unit (AATRU). I would like to sincerely thank Dr. Mario Castro for his support and guidance. Also, I appreciate other faculty members of ATTRU, especially Dr. Jonathan Boomer, Dr. Peter Niedbalski, and Dr. Isaac Sundar, with whom I had insightful discussions, and their feedback broadened my horizon. Also, it has been my pleasure to work with past and current members of ATTRU: Taewon Kim, Rohit Biswas, Suha Cho, Melanie Bassham, Fred Shi, Alexander Alsup, and Huiqing Yin-DeClue. Also, I would like to thank my committee members, Dr. William Brooks and Dr. Suzanne Shontz, and Dr. Ken Fischer, for their contributions and insightful feedback.

In addition, I am thankful for the help of the Seoul National University Hospital (SNUH) team in Korea: Dr. Chang Hyun Lee, Dr. Chang-Hoon Lee, Junwoong Park, WonCheol Chung, Dongha Kang, Sara Park, Dongkyu Sung, and Minseob Lee. Especially, I would like to acknowledge Dr. Chang Hyun Lee, who has provided invaluable input and supported me since my undergraduate studies.

I would like to appreciate all support from my friends, especially my best neighbor Yuna and my best roommate Tae. Also, I thank YoungJin for preparing the best airway illustration. Lastly, I would like to thank the unconditional support of my family to whom I dedicate my thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Advancements of medical imaging have enabled to analyze the structure of human physiology and provide a novel perspective on the pathophysiology of different diseases. Medical imaging plays an essential role in diagnosing diseases, monitoring disease progression, and planning for surgery. Among many medical imaging modalities, computed tomography (CT) is one of the most commonly used techniques for investigating pulmonary structure. Although CT imaging can noninvasively capture the detailed status of lung anatomy, CT itself does not provide functional information. Radiologists and pulmonologists qualitatively perform structural and functional assessments. This qualitative assessment is subject to intra- or inter-observer variability and insensitive to longitudinal changes (Ostridge & Wilkinson, 2016). In the past decade, quantitative CT (QCT) analysis has become a promising technique to link the structure and function relationship of the lung.

Unlike other clinical variables, such as pulmonary function tests (PFT) results, QCT can provide regional features of the lungs. In diseased lungs, its severity can be regionally quantified, and its progression can be monitored (Zagers et al., 1996; Dowson et al., 2001; Dirksen et al., 1997, 1999). As many QCT features are associated with clinical markers and predictors of clinical outcomes (Tanabe et al., 2012; Hoesein et al., 2012), QCT analysis can also contribute to disease diagnosis by providing objective information in a clinical setting. In addition, QCT-based structural and functional assessment in healthy and diseased lungs has improved understanding of preclinical pathophysiology and clinical outcomes (Hoffman et al., 2014; Podolanczuk et al.,

2016). Therefore, QCT analysis makes CT scans more meaningful and abundant by fully utilizing its information in both clinical and research environments.

One crucial step for QCT analysis is the segmentation of anatomical parts. Different segmentation masks identify different regions of the lung from the gray scale CT image data prior to measuring and computing QCT features. For example, a lung mask can be used for extracting tissue fraction of the lung, while airway branch angles can be computed from an airway mask. Therefore, the accuracy of QCT features heavily depends on the quality of the lung segmentation masks. While these masks can be acquired from manual annotation, it is incredibly time-consuming because human lung CT consists of three-dimensional (3D) image data with several hundred slices. An alternative method is to use algorithm-based semi-automatic software with consequent manual editing. Although this kind of pipeline has been successful and dominant for over decades, there are three main issues with this semi-automatic pipeline. First, it is still time consuming. As CT technology advances, the demand for CT scans has grown exponentially (Sentinel-Event-Alert, 2011). There are more CT scans to be analyzed. The second disadvantage of the semi-automatic pipeline is intra- or inter-observer variability. Manual editing can be done by multiple people with different backgrounds, and there is no clear guideline for manual editing. Qualities of segmentation can vary, and it will affect QCT analysis. Third, CT imaging quality affects the performance of semi-automatic segmentation software. Segmentation results can be inaccurate if a patient is severely diseased, or the method may completely fail if imaging protocols do not comply with software. For example, chest CT scans with slice thickness greater than 2 mm may not be suitable in the current quantitative CT imaging cohort studies (Sieren et al., 2016). In this case, QCT analysis of CT scans cannot be completed.

Recently, QCT analysis has advanced functional understanding of the lung regions related to structural alteration. One of the primary functional roles of the lung is ventilation through approximately 60,000 conducting airway branches linked with 30,000 acini in the lung parenchyma. In the past decades, CT-based computational fluid dynamics (CFD) analysis has been advanced to resolve physiologically consistent regional ventilation distribution under various breathing con-

ditions (Lin et al., 2007; Choi et al., 2009; Yin et al., 2010; Choi et al., 2010; Yin et al., 2013; Choi et al., 2019b). Tawhai et al. (2000, 2004) introduced a fractal-based volume filling technique to model CT-unresolved conducting airways from the distal end of CT-resolved airways to the terminal bronchioles. With the volume-filling technique, an entire conducting airway model was established, which links functional features of acinar scale parenchymal units and the CT-resolved airways (Yin et al., 2010; Choi, 2011; Lin et al., 2013). However, a full 3D CFD simulation is computationally expensive and time consuming. Since human airway geometry is complex and different from person to person, generating computational mesh requires lots of manual effort. Instead, a one-dimensional (1D) CFD simulation takes significantly reduced computational effort and can be automated. Therefore, it can be applied to a more extensive dataset instead of demonstrating the breathing simulations only in representative subjects.

Despite the advances in QCT and CFD analysis in lung health and diseases, previous studies have been limited in integration with other clinical variables. However, interest in lung health and disease associated with environmental risk factors such as ambient air pollution is growing. While the need for utilizing the newly developed QCT analysis to investigate the effects of the non-clinical variables on the alteration of regional lung structure and function grows, it is challenging to integrate additional sets of data into the combination of QCT outcomes and clinical outcomes. Recent advances in machine learning applications elucidate the possibility of clustering and classification in resolving the challenges with various data sets for QCT applications.

1.2 Objectives

The goal of this work is to extend the state-of-the-art QCT lung structure and function analysis in technological and application aspects. To achieve this goal, the following specific aims were established.

1. Extend the QCT application by improving segmentation using a deep learning approach.

2. Extend the application of QCT and CFD analysis to the assessment of treatment response.

3. Extend QCT analysis by integrating QCT features with various clinical and environmental variables using machine learning analysis.

The established QCT framework enables a more comprehensive analysis of multiscale lung structural and functional characteristics, seeking better physiological and pathophysiological relationships.

1.3 Thesis Overview

To achieve this goal, the QCT framework is elaborated in Chapter 2. Basics of lung anatomy and function are introduced, and QCT features are explained with image registration. Finally, machine learning and statistics that extend the QCT analysis are introduced as post analyses.

The framework is applied to three datasets. In Chapter 3, the deep learning-based segmentation approach is used to acquire pulmonary lobe masks for silicosis patients. Features of silicosis are quantified through the QCT framework and compared with features of healthy lungs and IPF patients. The QCT framework is used for asthmatic patients in Chapter 4. Inspiratory and expiratory CT scans before and after beta-2 agonist bronchodilator are acquired. QCT analysis, including 1D CFD, is performed to investigate the effects of bronchodilators in asthma. A comprehensive set of QCT features before and after the bronchodilator inhalation is analyzed statistically. Chapter 5 introduces machine learning methods to integrate air pollution exposure data, particularly particulate matter (PM), with QCT analysis. PM exposure-based clusters are determined, and the effects of PM exposure are compared using QCT features. Finally, Chapter 6 presents a summary and future works of this extensive study.

Chapter 2

Quantitative Compute Tomography Analysis

2.1 Lung Anatomy and Function

The structure of human respiratory system is highly complicated. It comprises multiple structural components, such as lungs, lobes, airways, and vessels. In this section, anatomy of the respiratory system, focusing on lobes and airway, is briefly reviewed.

Human lungs have five compartments called pulmonary lobes. The left lung has upper and lower lobes (LUL and LLL), and the right lung has upper, middle, and lower lobes (RUL, RML, and RLL). Within the lung, lobes are separated by fissures. The left lung is divided by an oblique fissure (between LUL and LLL), and the right lung is divided by horizontal (between RUL and RML) and oblique (between RML and RLL) fissures. Branches of airways and vessels occupy different regions of lobes. Because the fissures physically divide lungs, spread of diseases can be prevented by fissures (Sofranik et al., 1992). Although the lung is one continuous compartment, disease features and severity vary between lobes (Jeffery, 1998; Leung, 1999; Morgan, 1979; American Thoracic Society & European Respiratory Society, 2002). Therefore, the pulmonary lobe segmentation is a crucial part for quantification of regional lung characterization.

Another key structure of the lung for quantitative analysis is the airway. The airway can be divided into the upper airway and the lower airway. The upper airway includes the nasal cavity, oral cavity, pharynx, and larynx. In this thesis, the structure of the lower airway will be focused on since the scope of this work is quantification from chest CT that does not include the upper airway. The lower airway starts from the trachea and has a tree-like structure. The trachea is first divided into the right and the left main bronchi, respectively. The main bronchi further branch into lobar

bronchi, which are subdivided into segmental bronchi and subsegmental bronchi. The structure of airways up to subsegmental branches is illustrated in Figure 2.1. In the RUL, there are RB1, RB2, and RB3. RB4 and RB5 are located in the RML, and RB6, RB7, RB8, RB9, and RB10 are found in the RLL. Unlike the left lung, the right lung has an intermediate bronchus that connects the right middle bronchus with the middle lobar bronchus. In the left lung, there are only two lobes. LB1, LB2, LB3, LB4, and LB5 exist in the LUL. LB6, LB7, LB8, LB9, and LB10 are found in the LLL. However, LB7 is rarely found airway variant (Smith et al., 2018). Excluding LB7, there are 19 segmental airway branches. The segmental bronchi continue to separate into bronchioles and terminal bronchioles. Not only these branches become shorter and narrower, but also the proportion of cartilage decreases as the airways progress distally. Airways from the trachea to the terminal bronchioles make up the conducting airways. No gas exchange occurs in the conducting zone. Terminal bronchioles are divided into respiratory bronchioles, which are subdivided into alveolar ducts. Gas exchange occurs in the respiratory bronchioles and alveolar sacs. This region is called a respiratory zone.

Accurate airway masks play important roles not only in providing structural metrics, such as luminal area and branching angle, but also in preparing functional connections from large airways to small airways in lobar and sublobar regions. Both are crucial prerequisites of geometry and boundary condition preparation for lung CFD simulation.

2.2 Quantitative Computed Tomography Features

Since CT scans can be taken at different inspiratory levels, different QCT features can be acquired from CT scans at different lung volumes. When there are more than one CT scan for a patient, CT images can be matched either cross volumes or cross time points, and this process is called image registration. More functional QCT features, which are not visible from a single static image, can be acquired through image registration. In this section, QCT variables computed from a single inspiratory or expiratory CT will be reviewed, and QCT variables through image registration will be introduced consequently.

2.2.1 QCT from a single CT scan

Segmentation masks, which are requirements for extracting QCT variables, allow capturing different components of the lung in CT images. One of the basic QCT variables is lung volumes. Since the size of a voxel is stored in the Digital Imaging and Communications in Medicine (DICOM), lung volumes can be calculated by summing up every voxel of the lung masks. If the left and right lungs are annotated in the lung masks, each lung volume can be computed. The pulmonary lobes mask is finer and provides more regional information than the lung mask. Likewise, lobar volume can be computed using the lobe mask. Since the lobe mask is compatible with the lung mask yet provides richer information, the lobe mask is preferably used for extracting QCT features by lobe. In the following discussion, QCT features extracted from the lobes mask will be introduced.

Hounsfield units (HU) are a unit of CT. It measures the radio density of tissue during a CT scan. High density regions, such as bone and tissue, attenuate more radiation and have high HU, while low density regions, such as air, have low HU. High attenuation regions appear white or bright when images are reconstructed, and vice versa. Assuming only air and tissue are present in the lung, the tissue fraction (β_{tissue}) of a voxel can be calculated from the voxel CT density because CT densities of tissue and air are known to be 55 HU and -1000 HU, respectively (Yin et al., 2009).

$$\beta_{tissue} = \frac{CT_{density} - (-1000)}{55 - (-1000)} \tag{2.1}$$

Likewise, air fraction (β_{air}) of the voxel can be computed:

$$\beta_{air} = \frac{55 - CT_{density}}{55 - (-1000)} \tag{2.2}$$

The lobes mask can be used to calculate tissue fractions and air fractions of the lobes by taking the average of the voxels.

Instead of separating a voxel into either tissue or air, a threshold-based method classifies an entire voxel for its abnormality. Volumetric percentages of high attenuation areas (HAA%) and

low attenuation areas (LAA%) in the lung are two examples of the threshold-based QCT features. HAA generally represents fibrotic regions in the lung and is more commonly used for interstitial lung disease (ILD) patients. The threshold range varies from study to study. Tanizawa et al. (2015) defined HAA% as CT densities greater than -200 HU and found this threshold of HAA% correlated with the decrease in percent predicted forced vital capacity (FVC%pred) and carbon monoxide diffusion capacity (DLCO) in ILD patients. On the other hand, in the Multi-Ethnic Study of Atherosclerosis (MESA) study, HAA% is defined as the percentage of CT attenuation between -600 HU and -250 HU (Podolanczuk et al., 2016). This range of HAA% was associated with ILD hospitalization and ILD-specific death (Podolanczuk et al., 2017). Matsuoka et al. (2016) has more extensively analyzed different threshold values in patients with pulmonary fibrosis and found a significant correlation between HAA% ranging from -700 HU to 0 HU and the overall extent of interstitial lung lesions.

While HAA% is significantly associated with interstitial lung diseases (ILD), LAA% is more related to obstructive diseases. LAA% can be computed on either inspiratory or expiratory CT. Inspiratory LAA% ($LAA_{IN}\%$), also known as percent emphysema (Emph%), is defined as the percentage of lung voxels below -950 HU. This threshold was shown to have the strongest associations with both microscopic and macroscopic emphysema (Gevenois et al., 1995, 1996). As its name suggests, it quantifies emphysematous regions from CT. Emph% is widely adopted as a primary outcome for chronic obstructive pulmonary disease (COPD) studies (Wang et al., 2019; Pistenmaa et al., 2021; Paulin et al., 2018; Altınsoy et al., 2016; Smith et al., 2018). Expiratory LAA% ($LAA_{EX}\%$), also known as percentages of air trapping (AirT%), in chest CT is defined as the percentages of lung voxels less than -856 HU on expiratory CT. This threshold is commonly used because the expected CT density of a normal lung at inspiratory CT is -856 HU (Chen et al., 2020). In COPDGene study, the association between AirT% and spirometry measures (FEV1 and FEV1/FVC) is stronger than the association between Emph% and the spirometry measures (Schroeder et al., 2013). Also, the association between AirT% and asthma related hospitalization is significant in asthmatic patients (Busacker et al., 2009).

The only requirement of the threshold-based QCT variables is the lobes or lungs masks. It can be easily applied to various datasets, such as rat CT data. Also, since it is easily applicable and has fewer limitations, threshold-based QCT analysis can be automated if masks of either lungs or lobes can be automatically acquired. Thus, as one of the specific aims of this study, extracting threshold-based QCT features has been automated and applied to thick section CT data (More in chapter 3).

Other than using lungs or lobes masks, segmentation of the airway provides structural QCT features of the airway. Airway related QCT structural variables can be acquired at both inspiratory and expiratory CTs. When both CTs and corresponding airways masks are present, airway structural changes between inspiratory and expiratory CTs can be computed. One of the airway structural QCT features is the hydraulic diameter (D_h).

$$D_h = \frac{4 \times LA}{P_e} \tag{2.3}$$

where LA is a luminal area, and P_e is the perimeter of the luminal area. D_h is a significant parameter that affects flow characteristics, especially flow resistance (Choi et al., 2017b). Since gender, age, and height affect the size of the airway, normalized hydraulic diameter (D_h^*) was introduced for population-based analysis (Choi et al., 2015).

Another important QCT variable is circularity (Cr). The circularity of an airway branch can be calculated using the inner area and the inner perimeter of the branch (Choi et al., 2015).

$$Cr = \frac{\pi D_{ave}}{P_e} \tag{2.4}$$

where D_{ave} is an average luminal diameter and can be calculated as:

$$D_{ave} = \sqrt{\frac{4 \times LA}{\pi}} \tag{2.5}$$

Maximum value of Cr is 1, meaning the shape is circular. As Cr decreases, the shape becomes

more elliptical.

In addition, Tschirren et al. (2002) and Palágyi et al. (2006) introduced skeletonization of 3D structures. The proposed method is applied to convert a 3D airway mask to a 1D skeleton of a tracheobronchial tree. The bifurcation angle can be computed using the 1D skeleton airway.

$$Bifurcation\ angle(\theta) = \cos^{-1}\left(\frac{d1 \cdot d2}{|d1||d2|}\right) \tag{2.6}$$

where $d1$ and $d2$ are directional vectors of daughter branches, and $\cdot$ and $||$ denote the inner product and magnitude of the vector. These structural QCT variables have been associated with pulmonary function tests in asthma and COPD patients (Choi et al., 2015, 2017b). Structural alterations of D_h at lower lobes and branching angles of central airways are found in cement dust exposed subjects (Kim et al., 2020). More recently, these structural alterations of the airway are observed in response to bronchodilator (Lee et al., 2022), and 1D CFD airflow simulation is used to explain these effects (More in Chapter 4).

2.2.2 QCT from Image Registration

Image registration is a method to match different images at different time points. In chest CT, inspiratory and expiratory CT scans can be matched to extract more features. The objective of image registration is to determine a spatial transformation from the expiratory scan to the inspiratory scan. There are three main components in image registration. First, a transformative model is constructed. The model determines how a floating image can be transformed or deformed to match a reference image. Because motions of the lung are nonrigid, a free-form deformation (FFD) method based on cubic B-spline is used (Rueckert et al., 1999; Kybic & Unser, 2003). Second, a similarity between two images is measured. The similarity can be measured by calculating the squared sum difference (SSD) between two images. However, overall CT density increases as the lung inflates, and this discrepancy may lead to inaccurate matching. Therefore, squared sum tissue volume difference (SSTVD), which is less susceptible to changes due to inflation, is used in inspiratory and

expiratory image registration (Yin et al., 2009). Finally, the SSTVD between inspiratory and expiratory scans is optimized. A limited-memory quasi-Newton minimization method with bounds on the variables (L-BFGS-B) is used to optimize the high dimensional parameter space efficiently (Byrd et al., 1995). To make this optimization faster, two images are rigidly matched first, and nonrigid image registration is done subsequently. This overall process is done at six different resolutions. The original image is resampled, and the transformation model is determined at each resolution. Finally, with the spatial transformation mapping, more functional QCT features can be computed.

One example of a functional QCT feature after image registration is functional small airway disease (fSAD). Galbán et al. (2012) proposed fSAD and found that it is particularly important in monitoring COPD progression. Emph% is computed with inspiratory CT, while AirT% is computed with expiratory CT. Their regions may be overlapped. fSAD is regions of non-emphysematous air trapping, representing the air-trapped regions by small air obstruction but not by tissue destruction. Regions with inspiratory voxel greater than the emphysema threshold but image registered voxel value less than airtrapping threshold are considered as fSAD.

With image registration, the determinant of the Jacobian matrix of the voxel deformation, J, can be calculated.

$$J = \lambda_1 \lambda_2 \lambda_3 \tag{2.7}$$

where λ_i is an eigenvalue of a stretch tensor from image registration. It represents inspiratory-to-expiratory local lung volume expansion ratio (Amelon et al., 2011). Higher J means more volumetric expansion.

We can also compute the local displacement field (s^*). s^* is the local displacement between inspiratory and expiratory images normalized by the lung volume change.

$$s^* = \frac{s}{\left(V^{insp} - V^{exp}\right)^{\frac{1}{3}}} \tag{2.8}$$

where s is a displacement magnitude from image registration. s^* depicted the difference in lung

motionography between supine and prone positions due to the change in gravitational force (Kang et al., 2021).

After a mapping from expiratory image to inspiratory image is acquired, the air volume change map between inspiratory and expiratory can be computed. Using the air volume map, regional ventilation defect areas can be identified, and lobar-wise ventilation can be quantified. More recently, Chae et al. (2020) proposed a relative regional air volume change (RRAVC), which is a standardized regional lung ventilation capacity between the inspiratory and expiratory CT pair. It can be calculated by:

$$RRAVC = \frac{(v_{air}^{insp} - v_{air}^{exp})/v_{air}^{TLC}}{(V_{air}^{insp} - V_{air}^{exp})/TLC} \tag{2.9}$$

where v_{air} and V_{air} are local and entire air volume, respectively. Superscripts $insp$ and exp mean inspiration and expiration CTs. Negative RRAVC represents regions that have more air volume at expiration. RRAVC with greater than 1 are regions with above the average air volume change. 3D distributions of the regional ventilation and deformation are connected to the entire conducting airway model. The air volume change map serves to determine flow boundary conditions for CFD airflow simulations.

2.3 Machine Learning Analysis

Machine learning is a subfield of artificial intelligence (AI) that identifies patterns in data and develops a model without being explicitly programmed. There are three main types of machine learning: reinforcement learning, supervised learning, and unsupervised learning. Reinforcement learning, in general, means science and framework of learning to make decisions from interactions. A way of learning is based on the reward hypothesis. Supervised learning uses both labeled and unlabeled data for training. An example of supervised learning is semantic segmentation. The semantic segmentation model uses both images and masks to train. More details about semantic segmentation will be introduced in chapter 3. Unsupervised learning only uses unlabeled data for training. In this chapter, unsupervised machine learning methods that are applied to QCT analysis

will be introduced.

2.3.1 Principal Component Analysis

Principal component analysis (PCA) was first introduced by Karl Pearson in 1901 (Abdi & Williams, 2010). It is the most common linear dimension reduction method. Since analyzing data in high dimensional space requires a tremendous amount of computation, referred to as the curse of dimensionality, reducing the dimension is often the first step in analyzing high dimensional data. There are two ways to reduce dimensionality: feature selection and feature extraction. In feature selection, a number of variables are discarded, and thus a subset of variables is chosen. In feature extraction, a new set of variables are computed using the original variables. PCA is an example of a feature selection method.

PCA selects a new set of variables by maximizing the variance of the data. New variables are computed by a projection of variables on the weights, which are called the principal components. The first principal component is the direction of maximum variance, and the second principal component is the second largest and orthogonal to the first principal component. These principal components are the eigenvectors of the covariance matrix of the original input. Each principal component corresponds to each eigenvector, and they are in order of associated eigenvalues. In other words, the first principal component is the eigenvector with the largest eigenvalue.

Since a number of covariance eigenvectors and the dimensionality of the input are the same, n-dimensional data gives n principal components. In this way, the purpose of PCA, which is to reduce dimensionality, can not be reached. There are several ways to select the number of principal components. First, either 2 or 3 principal components can be selected for visualization. Since humans are incapable of seeing higher than 3D, 2D or 3D plots can provide the overall distribution of a high dimension dataset. The other way of determining the number of principal components is to use a scree plot. It displays eigenvalues of corresponding principal components. Since PCs are ordered by eigenvalues, there is a trend that eigenvalues decrease as the principal component decreases. The point at which the eigenvalues seem to level off can be adopted for

dimension reduction. The other similar approach is to use the proportion of variance explained. As the role of PCA is to maximize the variance of the data, the proportion of variance explained can be computed for each eigenvector. Instead of finding the "elbow" of the scree graph manually, this method can be automated by selecting a threshold value. Either 90% or 95% of the variance explained is commonly used. Finally, Horn's parallel analysis can be used to determine the number of principal components to keep (Horn, 1965). A random dataset with the same size as the original data set is created from a Monte-Carlo simulation. A correlation matrix and eigenvalues of the correlation matrix are computed from the randomly generated dataset. Eigenvalues of the randomly generated dataset are compared with the eigenvalues from the PCA. When the eigenvalue from the random data is greater than the eigenvalue from the PCA, the parallel analysis stops and suggests the number of components.

As discussed in the previous sections, there are more than 100 QCT variables. Combining QCT variables with PFT or other clinical data, there are more than 200 variables for one patient. Analyzing such a high-dimensional dataset is complicated and time-consuming. One simple application is to make subgroups based on the principal components.

2.3.2 K-means Clustering

K-means clustering is an iterative method to group similar data points into k clusters (Lloyd, 1982). It is another unsupervised machine learning method, and no label is required. Initially, the k-means algorithm randomly selects k centroids from the data. Distances between each data point and each centroid are computed. Consequently, data points are classified into different clusters based on the nearest centroids (the least squared Euclidean distance). Iteratively, k best centroids are selected, and data points are clustered into k clusters. The algorithm stops when the cluster assignments no longer change. Because the K-means clustering algorithm is not guaranteed to find the global optimum (Hartigan & Wong, 1979), multiple implementations with different initialization are needed. Using the K-means clustering, different subgroups can be identified, and subgroup analysis can be done.

Although the k-means clustering algorithm is automatically optimized, it requires the number of clusters as an input. The selection of k can be done by a visual inspection. For high dimensional data, PCA is done to extract the first two principal components, and distribution is plotted to determine an optimal k. To automate this process, the silhouette coefficient is adopted to determine the optimal number of k. The silhouette of a sample can be calculated by:

$$s(x) = \frac{b(x) - a(x)}{max(a(x), b(x))} \tag{2.10}$$

where a(x) and b(x) are average intra-cluster distance and nearest inter-cluster distances. It is a combination of how data points are dense within a cluster and how data points are separated between clusters. It ranges from -1 to 1, and a higher score means a better fit. The silhouette coefficient can be computed by taking the average:

$$SC = \frac{1}{N} \sum_{i=1}^{N} s(x) \tag{2.11}$$

where N is the number of data points. For a set of k clustering, the silhouette coefficients for each k can be computed, and the k with the best silhouette coefficient can be selected for automation. One disadvantage of K-means clustering is its computational inefficiency due to its iterative process. Therefore, K-means clustering is often used with PCA because PCA can reduce computational dimensions significantly. In Chapter 5, how a combination of PCA and k-means clustering is used to identify particulate matter exposure-based subgroups.

2.4 Statistical Analysis

Statistical analysis is extensively used with the QCT pipeline. Different tests can be used to compare different groups. Where there is only one group for longitudinal comparison, a paired t-test is used. When there are two groups to compare, parametric Welch's t-test or nonparametric Wilcoxon test are used. For more than three groups, analysis of variance (ANOVA) is first performed. If the significance is found from ANOVA, post hoc tests are performed to identify which

group is significantly different from other groups. To find associations between variables, both Pearson's and Spearman's correlations can be used, and the correlation coefficient is computed to measure the strength of the relationship between two variables.

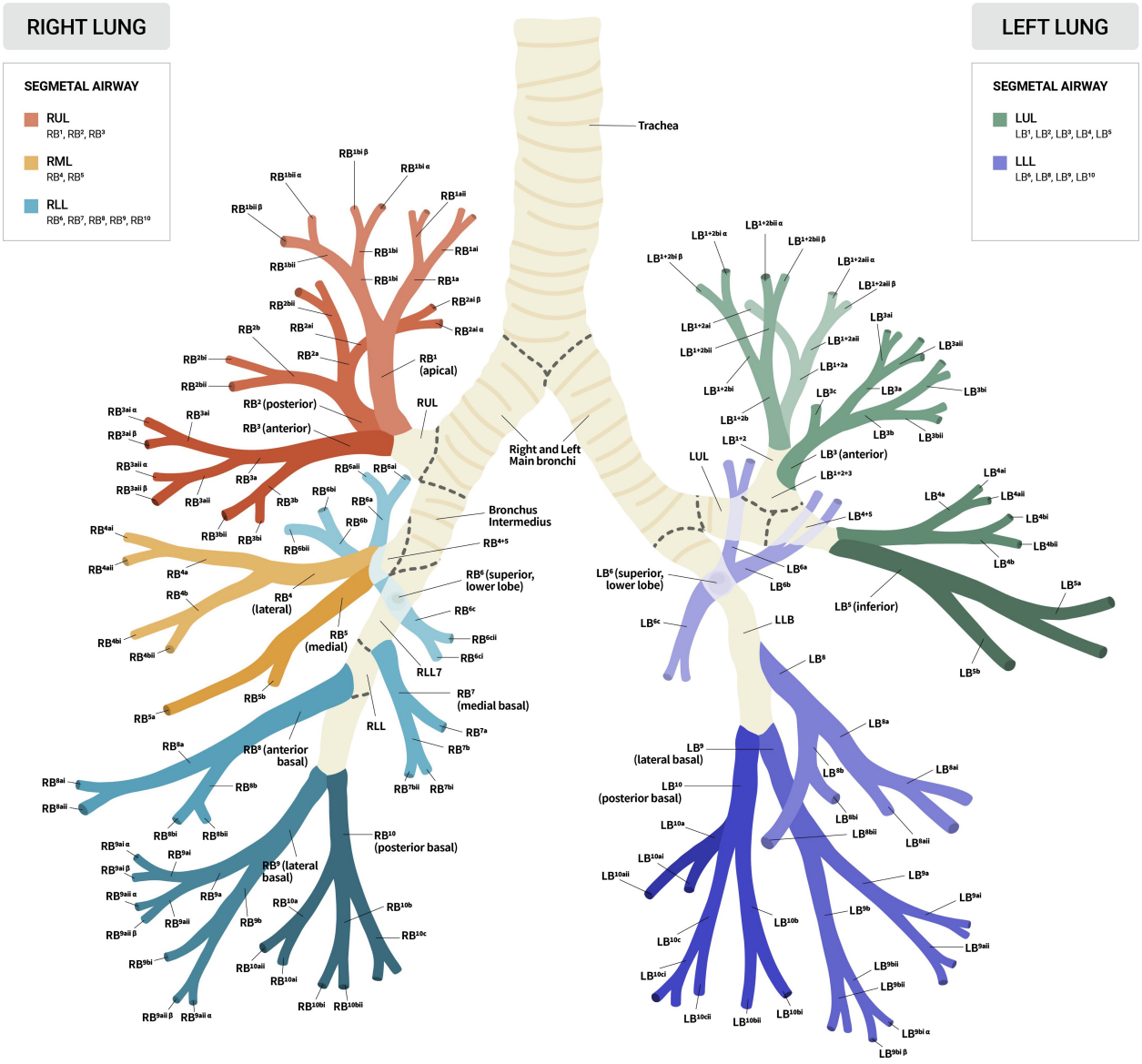

Figure 2.1: Airway branches from trachea to subsegmental branches. Each color tone of branch represents different lobes, and segmental branches are colored and labeled. Right lung and left lung are shown in left and right, respectively.

Chapter 3

A Slice Thickness Independent Segmentation Approach

In this chapter, we have sought to develop a slice thickness independence approach for pulmonary lobe segmentation. The model was trained on thin section CT images and tested on thick section CT images without transfer learning. Our proposed approach achieved dice similarity coefficients of 0.982, 0.981, 0.949, 0.896, and 0.962 for left upper, left lower, right upper, right middle, and right lower lobes, respectively. Then, we used the lobe masks to perform the whole lung quantitative analysis of chest CT images with thin and thick slice thickness (1mm & 5mm) and compared regional characteristics of silicosis, IPF, and healthy lung.

3.1 Deep Learning Segmentation

Advancements in computer hardware and enormously accumulated data have made a breakthrough in machine learning, especially in deep learning. Deep learning models have been successful in every subfield of computer vision, including classification, detection, and segmentation. It has also influenced medical imaging by outperforming previous algorithm-based models (Chan et al., 2020). There was no exception in the medical imaging segmentation task. Especially, UNet based models have dominated this field (Ronneberger et al., 2015). In this section, each component of a deep learning model, focusing on the UNet, is introduced first. Based on these components, the structure of a deep learning segmentation model, training strategies, and the model deployment are elaborated.

3.1.1 Convolutional Neural Network

There are several branches of deep learning models. Artificial neural network (ANN) is the most basic types of neural network. The building block of the ANN is a perceptron, which accepts input and computes output. Multiple perceptrons can be connected with each other, forming a hidden layer. The first layer of perceptrons is called input layer, and the last layer is called output layer. Recurrent neural network (RNN) is a variant of ANN. While computations of ANN can only move forward (to next layer), intermediate computations of RNN can return to its current layer. This recurrent connection enables the RNN to capture sequential information, and thus the RNN models have been successful in natural language processing (Hochreiter & Schmidhuber, 1997).

Convolutional neural network (CNN) is another variant of ANN that is specifically designed for image recognition task (LeCun et al., 2015). Unlike other networks, it is used for processing high dimensional data (2D or 3D). A convolution layer is composed of multiple kernels that learn features of input. Kernels, which are matrices, are slid and multiplied across the input. Dimensions of kernels are same with input dimension. Sizes of 2D inputs (images) are at least several hundreds by several hundreds, and sizes of kernels are usually 3x3, 5x5, or 7x7. Although larger kernels can capture more global features, computation cost increase exponentially. To overcome this, a pooling layer is introduced. Instead of increasing the size of kernels, the pooling layer reduces size of the inputs. There are two main types of pooling layers: maxpooling and average pooling. In max pooling, only maximum values within the size of the pool are stored. In other words, the most signaled regions are highlighted. On the other hand, average pooling takes the average within the size of the pool. It allows gradients to backpropagate through every pixel during the training. The pooling layer is added after the convolutional layer. A multiple combination of the convolutional layer and the pooling layer converts local features into global features. In this way, the CNN can recognize abstract images.

Another contributing factor of success of CNN is its non-linearity(Xu et al., 2015). Activation function is used with the convolution layer to increase non-linearity. 'Deep' in deep neural network means that it has multiple layers. However, multiple layers become useless if the model is

linear because all linear layers can be multiplied to one single layer. The most common activation function is rectified linear unit (ReLU).

$$f(x) = max(0,x) \tag{3.1}$$

The slope of ReLU for negative values is zero. Although it can make computations fast, when a neuron encounters a negative value, its gradient will not be updated and possibly never recovered. These untrained neurons become useless. To overcome this issue, leaky rectified linear unit (leakyReLU) can be used. Instead of making gradients zero, leakyReLU allows small gradient when values are negative.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x \leq 0 \end{cases} \tag{3.2}$$

, where α is a negative slope hyperparameter.

The value of α can be adjusted for a better performance. Decreasing α has a stronger rectification effect on negative values, and it becomes ReLU when α is zero.

In machine learning and statistics, it is common to observe that input data is normalized as a preprocessing step. Normalization makes sure the input falls in a reasonable range and increases numerical stability. Deep convolutional neural network is composed of multiple convolution layers, and numerical stability is not guaranteed after a single convolution layer. The unstable distribution can be accumulated and increases variations. This phenomenon, known as an internal covariate shift, can be minimized by batch normalization (Ioffe & Szegedy, 2015). Due to tremendous amount of data that do not fit into memory, deep neural networks are trained with mini batch. Batch normalization normalizes distribution of each layer over mini batch of each dimension.

$$\hat{x_i} = \frac{x_i - \mu_B}{\sqrt{\sigma_B + \varepsilon}} \tag{3.3}$$

where x_i represents one input instance. Mean (μ_B) and variance (σ_B) of mini batch can be calcu-

lated, and x_i is normalized to $\hat{x}_i$. ε is a noise to prevent denominator becomes zero. Each layer is normalized by mean and variance of the mini batch. In other words, distribution can depend on how data are grouped. Therefore, two learnable parameters, γ and β, are introduced.

$$y_i = \gamma \hat{x}_i + \beta \tag{3.4}$$

During training over the entire dataset, γ and β learn mean and bias of the layer.

3.2 Introduction

As introduced in the previous chapters, current applications of CT imaging and QCT analysis enable structural and functional assessment of various lung diseases. Disease severity can be regionally quantified, and its progression can be monitored with QCT. One of the most critical components of QCT analysis is segmentation masks, such as airway, vessel, or lobes. Regions of interest are extracted from the masks for QCT analysis. These masks can usually be acquired semi-automatically using modern software with manual editing. However, the modern software only works for CT images with well-controlled imaging protocols, such as thin slice thickness (less than 2mm), for successful lung segmentation and measurement. This requirement limits the use of advanced QCT analysis for more lung disease CT scans with thick slices (2mm or greater) since lobe segmentation is an essential initial step for quantitative analyses. Chest CT images are three-dimensional data with several hundreds of slices. Manual segmentation is time consuming and cannot be done in a large-scale dataset.

Human lungs have five compartments called pulmonary lobes: LUL, LLL, RUL, RML, and RLL. Within the lung, lobes are separated by fissures. The left lung is divided by one oblique fissure, and the right lung is divided by one oblique and horizontal fissures. Although lobes can be segmented using fissures, fissures can be either invisible radiographically or incomplete (Raasch et al., 1982). This incomplete fissure is more common in diseased lungs. Also, because the lobes are 3D structures with local textures, both global and local information is needed for lobe segmen-

tation. To tackle these problems, there have been lots of deep learning-based works proposed.

However, most of previous pulmonary lobe segmentation works either rely on other pulmonary segmentations, such as vessel, airway, and lung, or use multiple CNN models for different anatomical regions. Bragman et al. (2017) proposed an unsupervised lobe segmentation algorithm based on a probabilistic fissure segmentation, but it uses both airway and vessel masks as inputs. Gerard & Reinhardt (2019) combined four different 3D CNNs (two for fissure and two for lobes) to segment lobes, and the model requires masks of left and right lungs. Park et al. (2020) trained two separate 3D UNet models (left lung and right lung) using CT images with slice thickness less than 1mm. While 2D CNN cannot capture global features, 3D CNN is computationally expensive to capture local texture features. George et al. (2017) introduced a 2D CNNs model with 3 window channels and applied a 3D random walker for refinement. This model achieved a high accuracy, but this method requires a lung mask, which is acquired from another deep learning model. In addition, most of the previous methods for pulmonary lobe segmentation use thin slice chest CT data. Gu et al. (2021) first attempted deep learning lobe segmentation using thick section (5 mm) CT images. They used a combination of a 3D CNN for extracting global information and a 2D CNN for extracting local texture information. Their combined model was trained and tested on CT images with 5 mm slice thickness. However, both the 3D models and the hybrid 2D-3D model cannot be generalized for CT images with different slice thickness.

To overcome needs for prior masks and constant slice thickness of CT, we propose a method to automatically segment pulmonary lobes. Our method only uses CT images as an input and can segment lobes regardless of slice thickness. We call our method ZUNet, which is a UNet based segmentation model (Ronneberger et al., 2015) with additional axial location of the CT slice to capture a 3D structure. We trained ZUNet using thin section CTs of healthy subjects, IPF patients, COPD patients, and asthmatic patients. The trained model was tested to segment lobes of silicosis patients for QCT analysis. Silicosis is a fibrotic and occupational lung disease caused by inhalation of crystalline silica (Rose et al., 2019). Consequently, lobe masks were used for QCT analysis. QCT features of each group were statistically compared to characterize lobar features of

silicosis. Based on our knowledge, this is the first successful work of deep learning-based lobe segmentation using CT images with different slice thicknesses (1 - 10 mm) and the first work to conduct the whole lung and lobar quantification of QCT features on silicosis patients.

3.3 Methods

We first present our dataset breakdown for deep learning training. We then introduce our proposed method to segment lobes from CT images with different slice thicknesses. The method includes model architecture, postprocessing, and training strategy. The segmentation model is validated on thin section CT scans and tested on thick section CT scans with postprocessing. We conducted regional QCT analysis on a lobar basis using the lobe masks from the model. Finally, we performed statistical analysis to characterize and compared QCT features of silicosis with those of IPF and healthy lungs.

3.3.1 Datasets

Two different datasets are used in this study: thin section chest CT scans with 1mm or less slice thickness and thick section CT scans with 5mm or 10mm slice thickness. Thin section CT images and corresponding lobe segmentation masks are retrospectively collected from a QCT analysis study at Jeonbuk National University Hospital (n=351). For data diversity, four different disease phenotypes are included: healthy lungs (n=67), COPD (n=99), and IPF (n=77), and asthma (n=98). This dataset was split into a training dataset (n=293) and a validation dataset (n=48). Inspiratory thick-section CT scans of silicosis patients were retrospectively collected from Ankara Hospital for Occupational Diseases, Ankara, Turkey (n=30). The model was trained on thin section CT images and tested on thick section silicosis CT images without transfer learning. A breakdown of datasets is summarized in Table 3.1.

3.3.2 ZUNet Architecture

The model architecture of multi-channel ZUNet is illustrated in fig 3.2. Similar to UNet (Ronneberger et al., 2015), ZUNet is composed of a contracting path, a bottleneck layer, and an expansive path. Each layer consists of a 2D convolutional layer (transpose convolutional layer in the expansive path), leaky ReLU, and instance normalization. Zero padding is applied before each convolutional layer. Inputs to the ZUNet are a 4-channel image and a normalized axial location of the slice, called a z vector. CT images are preprocessed volumetrically first. Each voxel was clipped at -1024 HU because some CT scans use -3024 HU as a background value. After 3D CT images are normalized by minimum and maximum voxel values, one axial slice of the normalized CT images is extracted, and locations (x, y, and z) of each voxel construct positional maps. These positional maps are concatenated to the slice of CT and form a multi-channel CT image, which is fed to the contracting path. Axial position of the slice is normalized by the number of slices and one hot encoded to form the z vector. The one hot encoded z concatenated to the bottleneck layer instead of being fed to the first layer with the image. Concatenated features are upsampled by 2D transpose convolution. Finally, softmax function is used to get probability maps from the output of the last expansive layer. Unlike other common segmentation models, which output discrete values of segmentation masks, ZUNet outputs probability maps of each class. In lobe segmentation, there are six classes: LUL, LLL, RUL, RML, RLL, and background. Therefore, each CT slice outputs six probability maps, which is combined with probability maps of other CT slices. Postprocessing is performed on the 3D probability maps to make the final lobe mask.

3.3.3 Postprocessing

Two postprocessing steps are done to acquire less noise and more axially consistent lobe masks. Chest is segmented from CT slices using image processing algorithms. Then, 3D normalized convolution is done on the probability maps from the ZUNet.

Normalized CT image is first binarized using Otsu's method. Morphological opening, which is an erosion followed by a dilation, is performed to remove noises outside the chest. A Contour line

with maximum area from the opened binary image is acquired, and the contour line is converted to a chest mask. The chest of each slice is segmented, and a 3D chest mask is formed by combining each slice of the mask. The chest mask is used to remove noises outside chest, such as clothes and CT scanners. Figure 3.3 illustrates how chest segmentation is done with intermediate results.

CT images are three dimensions, but output of the ZUNet is two dimensions. While the 2D deep learning model can deal with more features and higher resolution thanks to less memory usage, segmentation results of the 2D model are axially inconsistent, especially with thicker slices. To overcome this challenge, mask probability map smoothing is used. After the chest mask is used to remove noises outside the chest, probability maps of lobes are converted to lung mask by binarizing each voxel into either lung or background. Because operating the Gaussian filter directly on the lung masked probability maps can make boundary of the lung blur and thus peripheral regions inaccurate, channel-wise 3D normalized convolution is applied (Knutsson & Westin, 1993). After Gaussian filter is applied on the masked region, the effect of blurred mask is calculated and is reversed. In this way, outside masks regions do not affect the masked region. After the normalized convolution, the smooth lobe mask is acquired. Figure 3.4 shows effects of the mask probability map smoothing.

3.3.4 Implementation Details

We train and test three models: UNet, single channel ZUNet, and multi-channel ZUNet. All the models are trained by a NVIDIA Quadro RTX 8000 GPU. We use PyTorch for training the models and "Weights and Biases" for monitoring the training. Input images are not resized, and input resolution is 512×512. Other than normalization after clipping the background value to -1024 HU, image augmentations are done weakly during the training. Because morphology of the lobes is highly relevant to positions, only a rotation with a limit of 15 degrees and blurring are done for data augmentation. Training and validation batch sizes are 16 and 32, respectively. Adam optimizer with a learning rate of 0.0001 is used to minimize a combo of cross entropy loss and

dice coefficient loss.

$$L_{CE}(y,\hat{y}) = -\sum_{c=1}^{6} y * log(\hat{y}) \tag{3.5}$$

$$L_{Dice}(y,\hat{y}) = 1 - \frac{2y\hat{y}+\varepsilon}{y+\hat{y}+\varepsilon} \tag{3.6}$$

$$L_{combo}(y,\hat{y}) = \lambda L_{CE} + (1-\lambda)L_{Dice} \tag{3.7}$$

where y and $\hat{y}$ are labels and predicted probability. ε, which was 10^{-7} in this case, is added to avoid loss function becomes undefined. Weights of cross entropy loss and dice loss are adjusted with *lambda*, which is 0.5. Models are trained for 50 epochs, and the ones with the lowest validation loss are used for testing.

3.3.5 QCT Analysis

QCT analysis is done to evaluate features of silicosis. Within the lobe masks, mean lung CT density, tissue fraction (TF), tissue volume (TV), total volume (V), HAA% between -700 HU and 0 HU, and LAA%, which is known as emphysema percentage in COPD, are computed. Mean CT density are calculated by taking the average CT density of each lobe. HAA% and LAA% are threshold-based measures. These measures are computed by counting number of voxels above or below thresholds within each lobe. Tissue fraction is calculated using the equation 2.1. More details about QCT calculations are in chapter 2.

QCT variables of the silicosis group (S) were compared with those from healthy 5-mm thick slice images (H5) and IPF 5-mm thick slice images (I5) adapted from 1-mm CTs of healthy (H1) and IPF (I1) subjects.

Paired t-test is done to compare QCT variables with ground truth masks and ZUNet masks. For disease wise QCT comparisons, ANOVA test is performed. Post hoc Tukey test is used to find which groups are significantly different when p-value from ANOVA is less than or equal to 0.05.

3.4 Results

In this section, we aim to evaluate the performance of multi-channel ZUNet in lobe segmentation and present results of QCT analysis. Robustness of the model is tested in two different ways. Segmentation accuracy is measured in thick slice thickness datasets, and QCTs extracted from predicted lobe masks are compared with those from ground truth masks. Four different models were investigated, UNet, ZUNet, multi-channel ZUNet, amd multi-channel ZUNet with postprocessing. Each model is trained on thin slice dataset with healthy lung, IPF, COPD, and asthma. We test models in the thick section silicosis dataset. Lobe segmentation masks are obtained for 30 thick section images. CT images with 5mm and 10mm slice thickness segmentation results are visualized in Figure 3.5. There are lots of false positives outside the chest, these false positives are completely removed with the chest mask. Quantitative results are shown in table 3.2. Because the objective of this study is to quantitatively analyze chest CT features of silicosis patients, we compared QCT variables resulting from ground truth and multi-channel ZUNet with postprocessing. MeanHU, TF, TV, total volume, HAA%, and LAA% are computed and plotted in Figure 3.6. Among all QCT variables, no significant difference is found between features from ground truth masks and ZUNet masks. From the lobe masks of the silicosis images, mean lung CT density, mean TF, TV, total volume, HAA%, and LAA% of lungs are -803±50 HU, 18.7±4.8%, 809±175 ml, and 16.2±13.3%, respectively. QCTs of healthy lung, IPF, and silicosis are compared. For QCT analysis of the whole lung, differences between IPF and silicosis are not significant ($p>0.05$), although there are significant differences between healthy and silicosis in meanHU, TF, and HAA ($p<0.05$). There is a significant difference between TF of IPF and silicosis in RUL but not in the other lobes. Also, HAA% of silicosis is significantly greater than that of IPF in RUL. No significant differences are found in the other lobes. A full comparison of QCT variables for each lobe and lung are presented in Figure 3.7.

3.5 Discussion

Advances in QCT-based structural and functional assessment in healthy and diseased lungs has improved understanding of preclinical pathophysiology and clinical outcomes (Hoffman et al., 2014; Podolanczuk et al., 2016). A great portion of lung diseases such as silicosis uses thick section CTs, which has limited the use of advanced QCT analysis, while visual assessment of the chest CT is one of the main tools to diagnose silicosis, reviewing CT patterns of impairment such as confluence and coalescence (dos Santos Antao et al., 2005; de Castro et al., 2014). The main barrier is requirement of well-controlled imaging protocols with thin section full volumetric images for 3D whole lung segmentation. In this study, we proposed the slice thickness independent segmentation approach and evaluated on CT images with various disease phenotypes (healthy lung, COPD, IPF, asthma, and silicosis) and a range of slice thickness (1 mm to 10 mm). We also checked reliability of the proposed method by comparing QCT features extracted from the ground truth and ZUNet with postprocessing.

Most of previous lobe segmentation works rely on prior masks, such as lung, airway, and vessel. It not only requires more computations and time but also cannot be used without successfully segmented prior masks. Our proposed method only uses slices of CT images as inputs. Lung masks are used to remove false positive regions outside the lungs. After having tested the lobe segmentation model on silicosis dataset, we found that lung segmentation is unnecessary. All the noises occurred outside of the chest regions. There are only two noise regions: around clothes and CT scanner bed. Therefore, chest mask is enough to remove those noises. Also, the proposed chest segmentation method is reliable, as it successfully segments every slice of all 30 silicosis CT scans. The second postprocessing was 3D smoothing. 3D volumetric construction of 2D masks can be done by concatenation of 2D slices, but in this way, the combined 3D mask is axially inconsistent. As addressed in other literature, this global inconsistency is the main limitation of 2D CNN models. Most of previous lobe segmentation works, thus, employ either 3D or 2D-3D hybrid models to capture global context. However, one of disadvantages of 3D models is that it cannot be generalized for different slice thickness CT scans, especially when the difference is large. Instead

of using 3D CNN models, we stick to the 2D CNN model. We used two steps to make lobe masks more globally consistent. First, we used positional location as an input to the 2D model: multi-channel and z-vector. Second, we applied 3D normalized convolution for near slice smoothing. Because values of mask are discrete, the gaussian filter cannot be directly applied to the mask. Instead, 3D gaussian blurring is applied on lung masked probability maps. Effects of postprocessing are compared by measuring accuracy before and after the postprocessing. According to Table 3.2, dice scores and sensitivity are improved remarkably in thick section CT images after the postprocessing. One contributing factor is that the model is trained on different dataset. Not only slice thickness, but also other image parameters and diseases phenotypes are different between train and test datasets. As Hofmanninger (2020) investigated clinical applicability of deep learning segmentation approaches, accuracy of lung segmentation models relies on diversity of the training data. When the model is tested on the thick section silicosis CT images, the model encounters unseen features and noises, especially outside chest. In this case, chest mask helps the model ignore the noises. In addition, thick section CT images usually have greater slice-wise discrepancies. As slice thickness increases, changes in lung anatomy become more remarkable. This discrepancy results in more axially inconsistent segmentation masks. Although the z-vector of the ZUNet can alleviate this effect globally, it can't fix near the lobe boundaries where values of z-vector would be similar. Therefore, 3D normalized convolution corrects those regions and increase the performance significantly. Finally, our approach successfully segments pulmonary lobes in unseen dataset by combining invariant positional information and robust postprocessing.

We have shown our deep-learning based approach provides robust lobe segmentation masks allowing successful QCT analysis from 5-mm and 10-mm thick-section CTs of silicosis patients. QCT features of silicosis are compared statistically with those of healthy lungs and IPF. Compared to healthy lungs, silicosis and IPF had higher HAA% in all lungs. As HAA% corresponds to fibrotic component of the lung and negatively correlates to physiological parameters such as force vital capacity and carbon monoxide diffusion capacity (Tanizawa et al., 2015), the increase in HAA% well represent the characteristics of fibrotic lung diseases. Although HAA% of silicosis

was higher than that of IPF, difference was not significant in all lungs. However, their difference was significant in right upper lobe. Although upper lobes dominant fibrosis is known as a pattern of silicosis and qualitatively confirmed (Mlika et al., 2022; Ferri, 2021), it is first quantitatively found and compared with other fibrotic disease. In addition, mean CT density and tissue fraction showed the same pattern. Means of silicosis are significantly higher than IPF only in the right upper lobe.

Although our proposed method showed robust results in segmenting lobes from thick section CT images, there are some limitations that should be considered. Pulmonary lobe segmentation is not clearly defined. Most of previous pulmonary lobe segmentation works and public datasets include vessels and airway. However, to what extent of vessel and airway can be included in the lobe is not clearly mentioned in previous works. In works of Park et al. (2020); Bragman et al. (2017), gold standards of lobe segmentation excluded portions of vessels and airways, even though they are in the lung. Because airway and vessel boundaries can be unclear when included in the lobes, airway and vessels should be excluded from lobe masks to have more accurate lobe segmentation. Then, excluding vessel and airway can make lobe segmentation much harder task. Depending on the purpose of lobe segmentation, these variances may affect analysis. Because our goal of using lobe segmentation is to perform QCT analysis, vessels and segmental airways are already excluded from the ground truth lobe masks. There are multiple holes, which indicate airways or vessels, in the predicted lobe masks. Although these holes can be filled with morphological closing, this post-processing may result in inaccurate QCT analysis. Evaluation of the effects of including airway and vessels by segmentation metrics and QCT analysis remains for future studies.

3.6 Conclusion

Our deep learning-based pulmonary lobe segmentation approach enabled QCT analysis of thick section HRCT lung images of silicosis patients. Increased mean lung CT density, tissue fraction, and HAA% in the RUL quantified degrees of fibrotic changes of the lungs with silicosis. This quantitative analysis may provide precise assessment of longitudinal lung structural and functional

changes along with disease progression and treatment responses of the patients. The proposed approach may expand the capability of QCT analysis applications to more clinically obtained thick section CTs.

Dataset	$N_{Healthy}$	N_{COPD}	N_{IPF}	N_{Asthma}	$N_{silicosis\_5mm}$	$N_{silicosis\_10mm}$	N_{Total}
Train	67	83	61	82	0	0	293
Validation	0	16	16	16	0	0	48
Test	0	0	0	0	16	14	30

Table 3.1: Slice thickness and disease distributions of lobe segmentation datasets.

Models	Metrics	LUL	LLL	RUL	RML	RLL
UNet	Dice	0.937	0.940	0.790	0.652	0.812
	Sensitivity	0.960	0.943	0.895	0.817	0.850
	Specificity	0.998	0.999	0.994	0.992	0.995
ZUNet	Dice	0.946	0.908	0.853	0.762	0.886
	Sensitivity	0.962	0.949	0.918	0.802	0.897
	Specificity	0.998	0.997	0.996	0.998	0.997
Multi-channel ZUNet	Dice	0.974	0.968	0.933	0.877	0.956
	Sensitivity	0.978	0.973	0.936	0.926	0.948
	Specificity	0.999	0.999	0.999	**0.999**	0.999
Multi-channel ZUNet With PP	Dice	**0.982**	**0.981**	**0.949**	**0.896**	**0.962**
	Sensitivity	**0.985**	**0.979**	**0.949**	**0.934**	**0.947**
	Specificity	**1.00**	**1.00**	**0.999**	**0.999**	**1.00**

Table 3.2: Metrics of different models when tested on the silicosis dataset. The highest metric for each lobe is made bold.

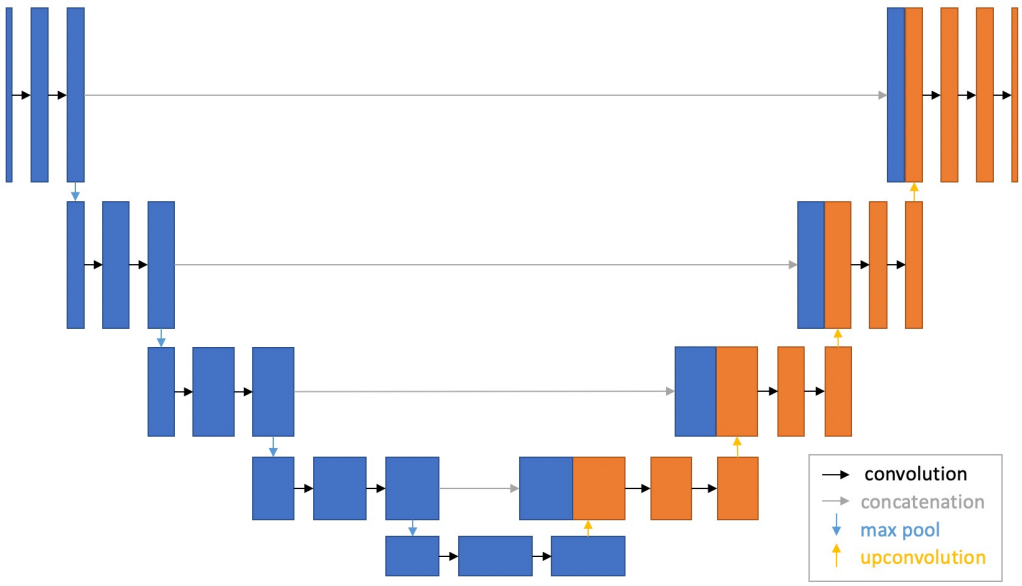

Figure 3.1: Architecture of UNet for medical imaging segmentation.

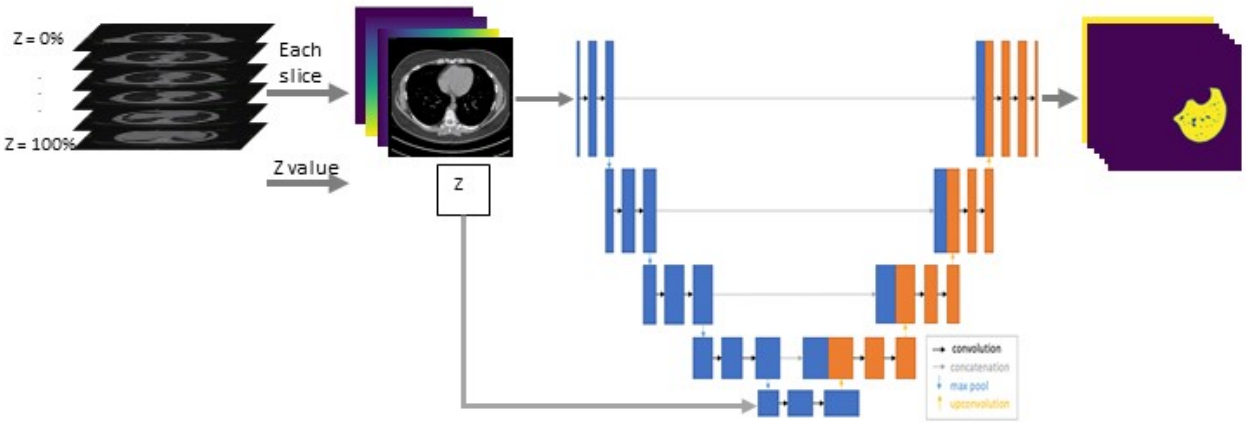

Figure 3.2: ZUNet architecture with slices of images and position of the slice in the craniocaudal (z) axis.

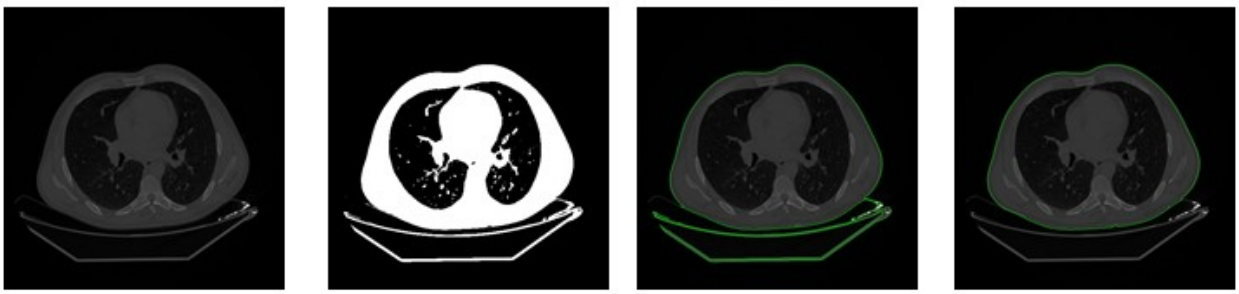

Figure 3.3: Steps to segment a chest. Axial view of a chest CT slice is shown in (a). (b) is the binarized image after Otsu's method. (c) shows a contour line with noises is drawn on (a). (d) shows a contour line after the morphological opening.

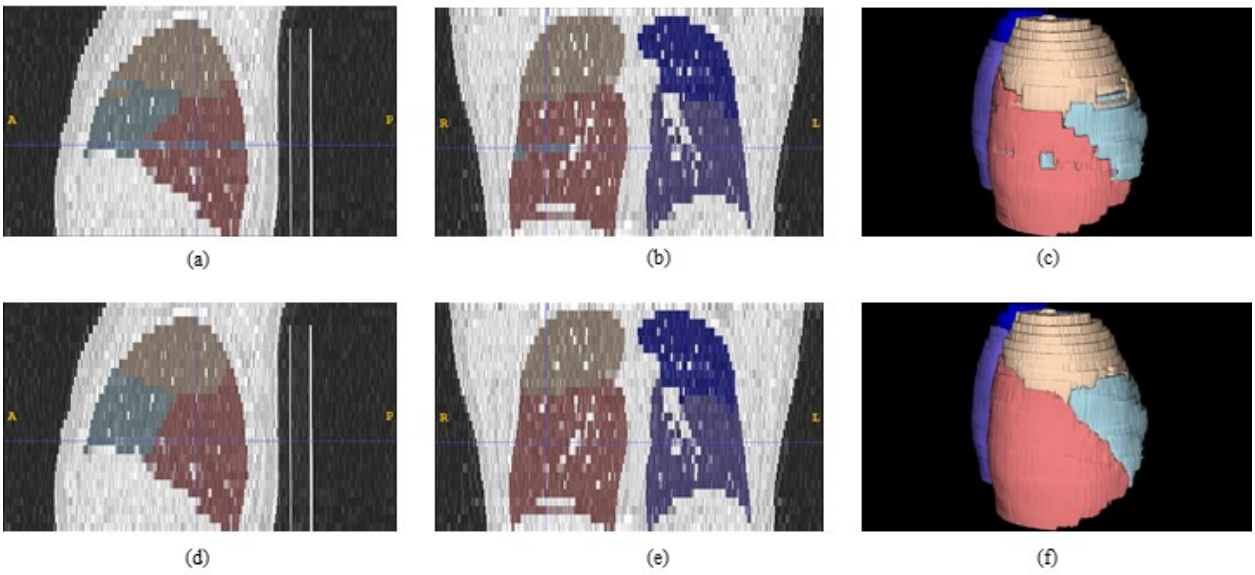

Figure 3.4: Comparisons before, (a), (b), (c), and after, (d), (e), (f), the mask probability map smoothing. (a) and (d): sagittal view, (b) and (e): coronal view, and (c) and (f): 3D shape.

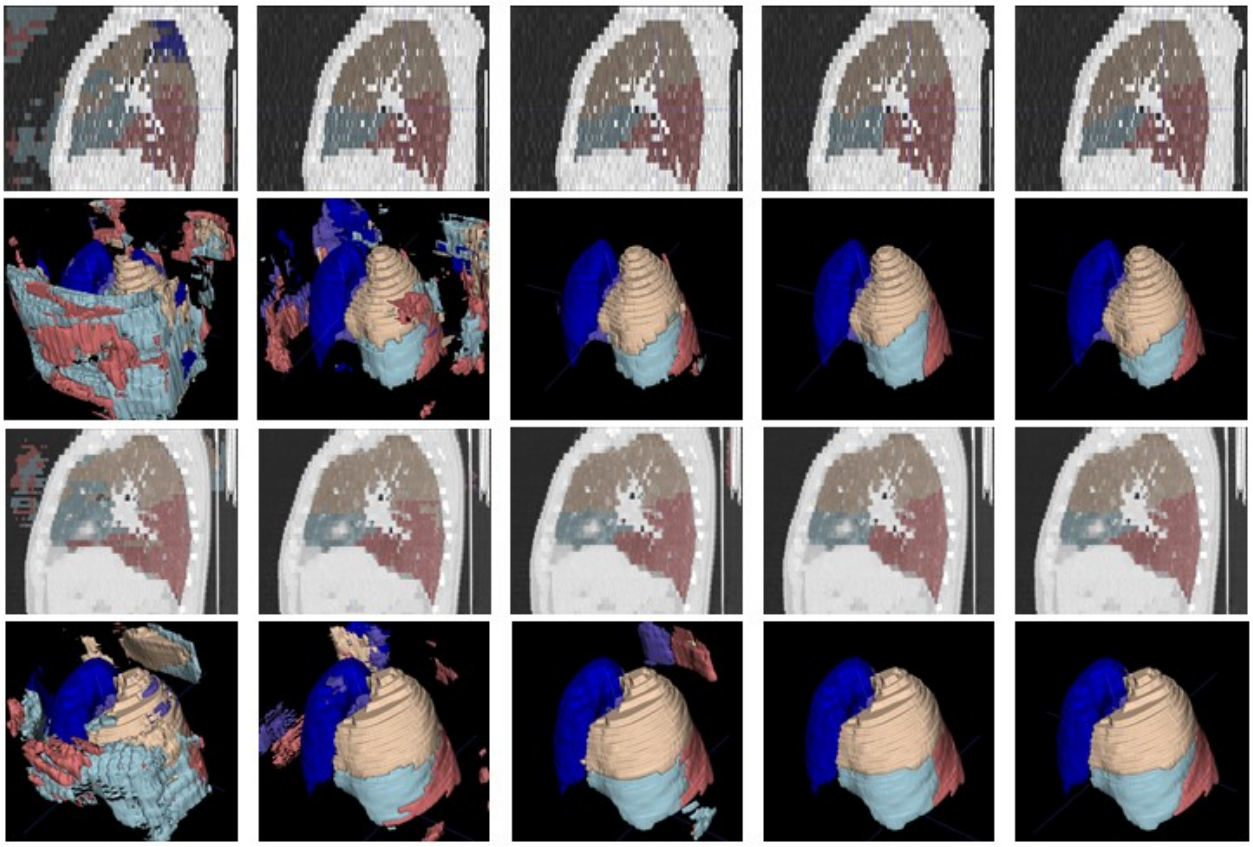

Figure 3.5: Qualitative lobe segmentation results of two representative silicosis subjects. First and third rows show sagittal view. Second and the last row show 3D view. Each column represents different segmentation model: (a) UNet, (b) ZUNet, (c) Multi-channel ZUNet, (d) Multi-channel ZUNet with postprocessing, and (e) ground truth.

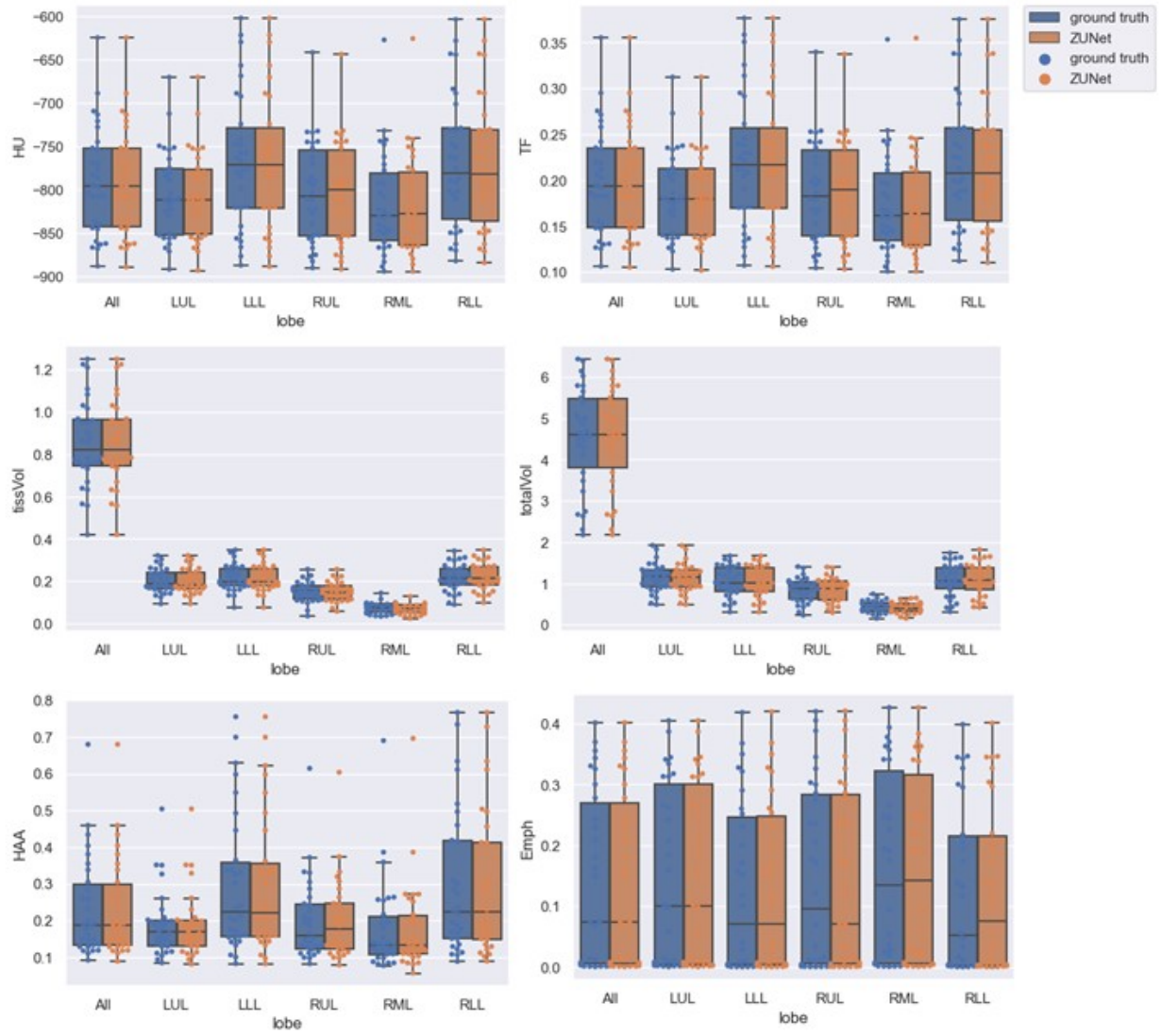

Figure 3.6: QCT comparisons of lobe masks of ground truth and ZUNet with postprocessing. Each data point represents one patient.

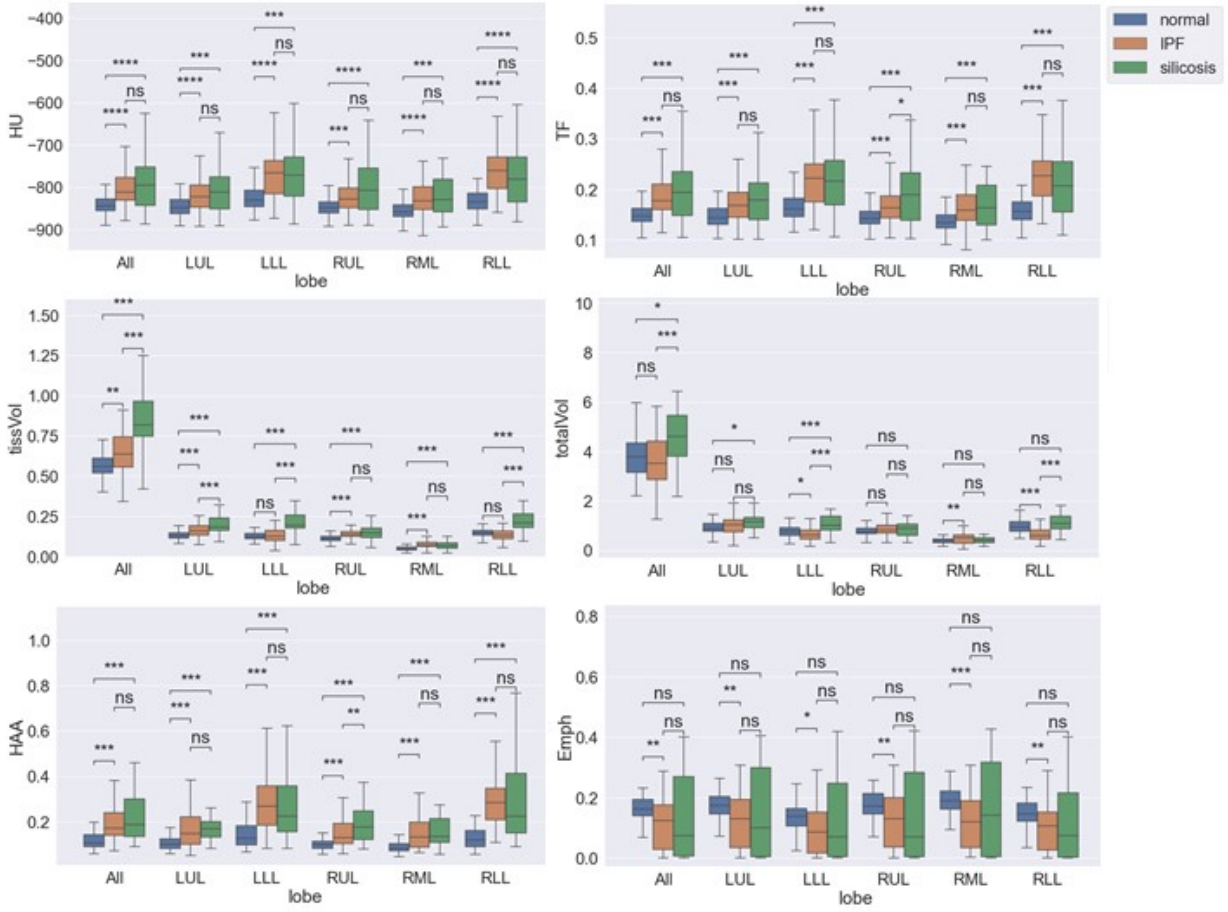

Figure 3.7: Bar plot for each QCT variables with 5 lobes. *,**, and *** denote that p values are less than 0.05, 0.01, and 0.001, repectively.

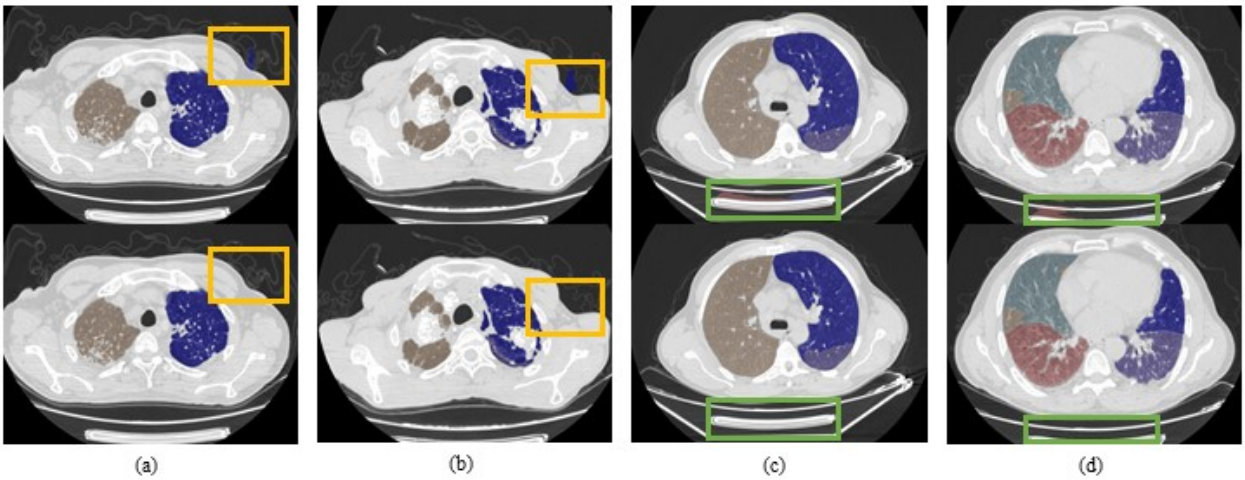

Figure 3.8: Comparisons before (first row) and after (second row) the chest segmentation. (a) and (b): clothes noises, (c) and (d): Noises between CT scanner and patient. Yellow bounding boxes highlight noises around clothes. Green bounding boxes highlight noises around CT scanner bed.

Chapter 4

Explain Bronchodilator-induced Regional Ventilation Improvement in Asthma using QCT and CFD

Despite recent advancements in QCT and CFD analysis in lung health and diseases, the application of QCT-CFD combined analysis has been rarely applied to assess lung response to treatments. In general, comprehensive full three-dimensional CFD studies for lung use only a few representative subjects, lacking the statistical evidence for the physiological and pathophysiological interpretation of the findings. In this section, our QCT pipeline is applied to analyze the effects of bronchodilator inhalation in 66 asthmatic patients. Structural changes in the airway induced by bronchodilator are linked with 1D CFD airflow simulations. First, a brief introduction to asthma, a bronchodilator, and a 1D CFD simulation are presented. Then, detailed methods of extensive QCT analysis linking with 1D CFD simulation are introduced. Finally, the results of the QCT-CFD analysis are presented.

4.1 Introduction

Asthma is a chronic condition that causes the airways narrow and inflamed, and this condition causes airway obstruction. Despite there are more than three hundred million people worldwide suffering from asthma, there is no cure for asthma (Masoli et al., 2004; Vos et al., 2020). However, it is manageable with medications. One of the most common types of asthma medication is a bronchodilator. It helps make breathing easier by relaxing the muscles around the airways. Despite its common use to improve lung functions, the underlying regional lung structure-function relationship in response to the bronchodilator is not fully understood.

Previous application of the QCT framework for analysis of bronchodilator response showed a tendency to reduce regional discordance of ventilation distribution. Bronchodilator inhalation induced patchy pairs of hyperinflated and hypoinflated regions compared to the pre-bronchodilator lung and improved expiratory lung deflation (Choi et al., 2017a). Statistical analysis of QCT-based airway metrics between pre- and post-bronchodilator showed a significant increase in diameter and circularity of proximal airways, such as the trachea, LMB, RMB, and bronchus intermedius (BI), and decreased end-tracheal branching angle (Lee et al., 2022). In this extended study, we used 1D CFD airflow simulations to further characterize bronchodilator (BD)-induced improvement in regional air flow dynamics and to associate the flow features with other QCT features and clinical data. Hereafter, the terms "bronchodilator" and "BD" are both used according to specific needs for natural and concise expressions, respectively.

4.2 Methods

The current analysis used retrospectively collected data from 72 patients with asthma enrolled in a bronchodilator response study conducted at Seoul National University Hospital, Seoul, Korea in 2013. Among the 72 patients, 6 patients were excluded by the exclusion criteria in Figure 4.1, and the other 66 patients (M:F=23:43, age=64.4 $\pm$ 10.7, severe:non-severe=37:29) were included in the analysis. The overall pipeline is illustrated in Figure 4.2.

Full inspiratory CT at total lung capacity (TLC) and full expiratory CT at residual volume (RV) acquired both before and after bronchodilator inhalation were collected. Segmentation and structural measurement of airway, lungs, lobes, and vessels from each of four CT scans were done using VIDA Apollo 2.0 (Coralville, IA) and in-house software. Airway structural metrics include D_h^*, Cr, branching angle and more. Pulmonary lobe masks were used to acquire threshold based QCT features, such as HAA% and LAA%. Symmetric mass-preserving nonrigid image registration (Yin et al., 2009; Haghighi et al., 2018) between TLC and RV images were performed for before and after bronchodilator inhalation. Functional QCT features, such as J, RRAVC and s^*, were extracted after image registration.

Based on CT-resolved airways and lobe masks, entire conducting airway models were made using a fractal-based volume filling method (Fig 4.3) (Tawhai et al., 2000, 2004; Choi, 2011), which links airways with approximately 30,000 parenchymal units at an acinar scale. For CFD simulation, regional air volume change distribution at those parenchymal units provided the initial flow rate distribution in the entire conducting airway model, which serves as boundary conditions for CFD air flow simulations. Subject-specific tidal breathing flow rate patterns were generated by sinusoidal waveform with an amplitude of tidal volume determined by individual subject's weight (kg) multiplied by 6 ml/kg. To analyze all 66 patients, we used 1D CFD simulation instead of 3D CFD simulation. We used 1D CFD simulations that solved energy balance equation for pressure drop due to unsteady effect, total effect, kinetic energy, and viscous dissipation and a lung compliance model applied to pleural region that link acinar and pleural region. For viscous pressure drop entry flow model (Pedley & TJ, 1977; Choi, 2011) was employed through the entire conducting airway. Pre- and post-bronchodilator whole lung and lobar distributions of local air flowrate (FR) and flowrate fraction (FF), airway pressure drop (P) at peak inspiratory (PI) and peak expiratory (PE) flow rates were used for comparison between pre- and post-BD tidal breathing characteristics and associations between variables. Furthermore, acinar pressure work (W_p) and airway resistance (R) were computed for the whole lung and each lobe.

$$W_P = (P)(FR) \tag{4.1}$$

$$R = \frac{P}{FR} \tag{4.2}$$

All QCT features were integrated with CFD features for analysis. Finally, the CFD variables were associated with other QCT and clinical features.

4.3 Analysis

First, effects of bronchodilator inhalation were visually inspected. A demonstrative result of CFD simulation of 56-year-old severe asthmatic patient with 15% increase in FEV1 is shown in

Figure 4.4. In this demonstrative case, regional discordance in distributions of flowrate fraction and airway pressure was substantially reduced. More large airways defined by the inner diameters greater than 2 mm appeared after bronchodilator inhalation. Also, D_h^* and circularity at the trachea as well as end-tracheal branching angle were improved after bronchodilator inhalation. Airway resistance in the entire conducting airway decreased by 36%. It is shown that heterogeneity of airwasy resistance was reduced.

The demonstrated BD-induced functional improvement was not clearly shown in all 66 patients. Thus, effects of bronchodilator were compared statistically in all 66 subjects. Significant increase in large airway diameters were found in the entire 66 patients, in the trachea (6.7%, p<0.001), LMB (5.6%, p<0.001), RMB (5.6%, p<0.001), BI (2.8%, p=0.0016), LLB (11.1%, p<0.001), and average over the segmental airways in each of the five lobes (LUL, LLL, RUL, RML, RLL: 8.6%, 16.3%, 11.0%, 10.7%, 9.5%, respectively; p<0.001, all). However, CFD-derived functional outcomes were not significantly improved in all subjects. To find important functional imaging biomarkers, associations between CFD, QCT, and clinical features were found. We found the increase in large airway diameters were significantly associated (p<0.05) with CFD features. Specifically, decreases in WP at PI and PE were associated with increase in D_h^* of the trachea (r=0.46, 0.51), the BI (r=0.43, 0.47), and the LLB (r=0.30, 0.33). Also, FR increases in the RLL at PI and PE were associated with increases in D_h^* of the trachea (r=0.41, 0.45), the RMB (r=0.37, 0.36), and the BI (r=0.49, 0.48). It is worth noting that the trachea, the RMB, and the BI are an airway path to the RLL. In this finding, it is implied that the degree of bronchial dilatation indicative in the large airways improves regional breathing, by reducing regional pressure-driven workload in the whole lung and particularly recovering air flow in the RLL.

Not only the diameters of the airways, but also the shape of the airway cross section affected flow characteristics. In the LUL, WP at PI and PE was moderately associated with increase in Cr (r=0.50, 0.45). In the LLL, decrease in Cr was associated with FF at PI (r=-0.297), and FF at the LLL was associated with percent predicted value of FEV (FEV1$\%_{pred}$), FVC, and the FEV1 to FVC ratio (FEV1/FVC) (r=0.336, 0.293, 0.271). This finding is linked with our recent finding

that pre-BD D_h^* at the LLB was mildly associated with FEV1%$_{pred}$ and FEV1/FVC and support implication of the impact of the LLB constriction in bronchodilator inhalation in the distal airways (Lee et al., 2022). Constriction of LLB was also found to differentiate severe asthmatics into two distinct groups, one with airway constriction and air trapping but no airway wall thickening, and the other with wall thickening without airway constriction or air trapping, multiscale imaging-based cluster analysis of 248 asthmatic patients enrolled in Severe Asthma Research Program (SARP) (Choi et al., 2017c) and have a significant impact on forming high speed air flow and particle deposition hot spots that increases proximal airway particle deposition fraction (Choi et al., 2019a) in the same cohort data. The LLB is the entrance to segmental and further distal airways of the three basal regions in the LLL.

Not only airway structural variables but also other QCT features were associated with CFD features. In the RML, FR decreases at PI and PE were associated with LAA%$_{IN}$ (r=-0.44, -0.42), which increased by bronchodilator inhalation in all asthmatics (13.8%, p=0.024), and mean ventilation fraction decrease was associated with FEV1%$_{pred}$ increase (r=-0.49).

4.4 Conclusion

Our QCT-CFD framework was successfully applied for the characterization of multiscale structure and function relationship in bronchodilator response of 66 asthma patients. Integrated QCT-CFD lung analysis in bronchodilator response in asthma-associated changes in regional air flow features with improvement in proximal airway features and clinical data. Results elucidated how bronchodilation by β 2-agonists reduced CFD simulated workload of breathing with the interplay between various regional features of airway and lung structure and function. An increase in airway diameter by bronchodilator was associated with a decrease in CFD-simulated workload.

The QCT and CFD-derived features could identify regional differences in bronchodilator response and provide specific structure-function association. Such subject-specific characterization of structure-function response to treatment may be necessary if one was trying to modify the airways, such as by bronchial thermoplasty and inhaled biologic.

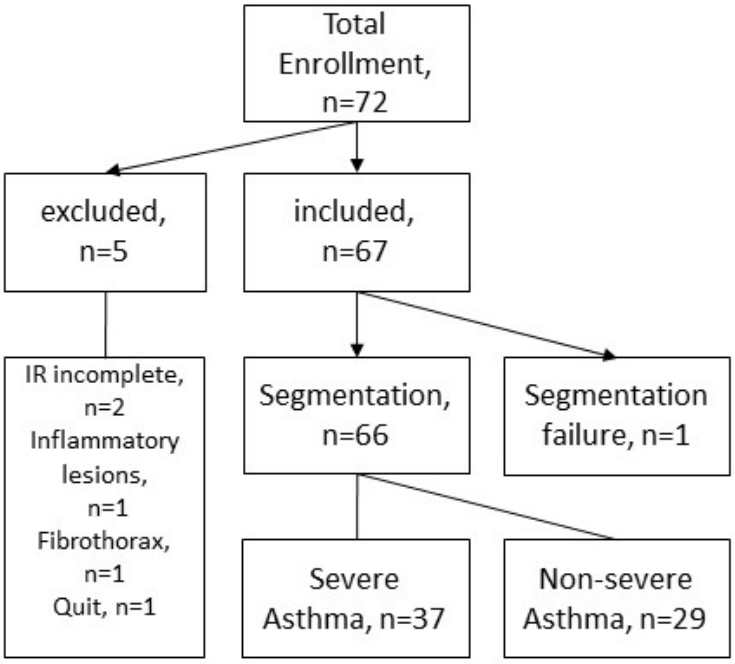

Figure 4.1: An enrollment flowchart of asthma bronchodilator response analysis. Among 72 patients, CTs of 66 participants were analyzed.

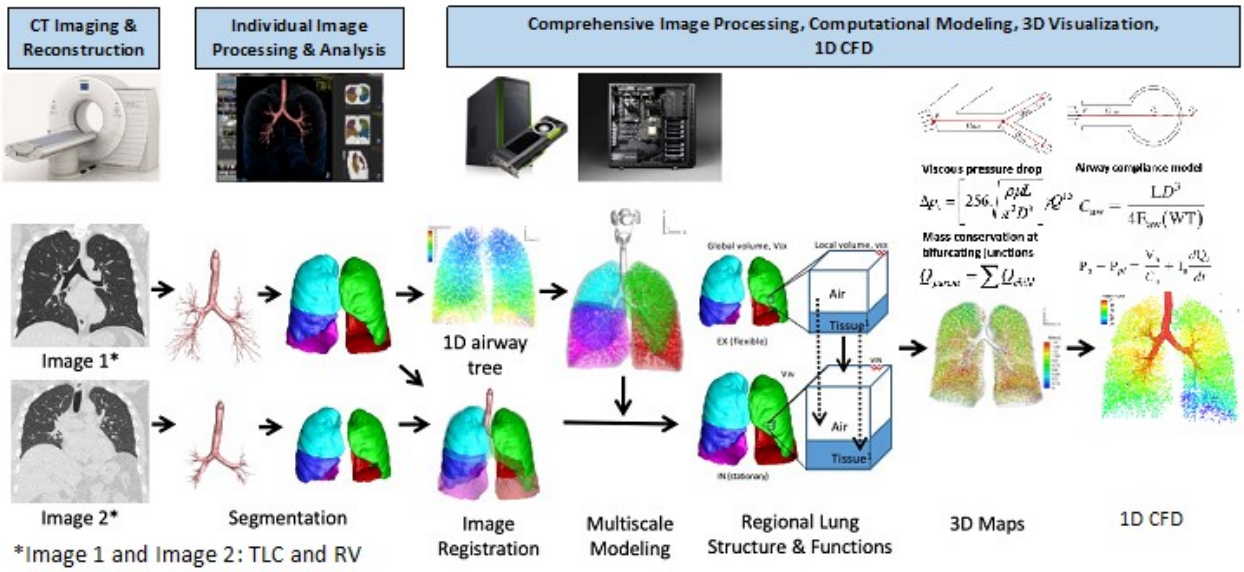

Figure 4.2: A workflow of inspiratory and expiratory CT image matching and 1D CFD airflow simulation for lung structure-function assessment (adapted from Chae et al. (2020))

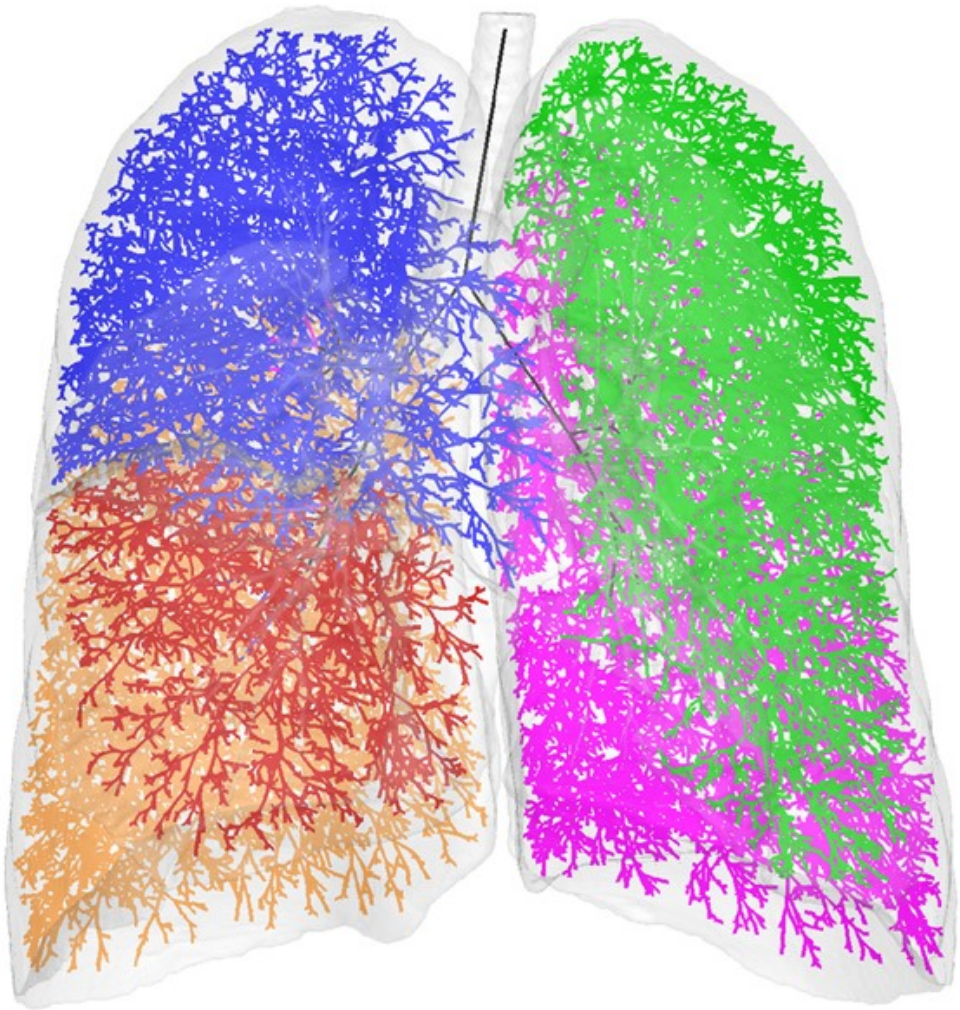

Figure 4.3: A visualization of the modeled skeletonized 1D airway tree overlayed with shaded CT-resolved 3D geometry surfaces of airway and lobes. Green, light purple, blue, red, and orange indicate the five lobes (LUL, LLL, RUL, RML, and RLL), respectively.

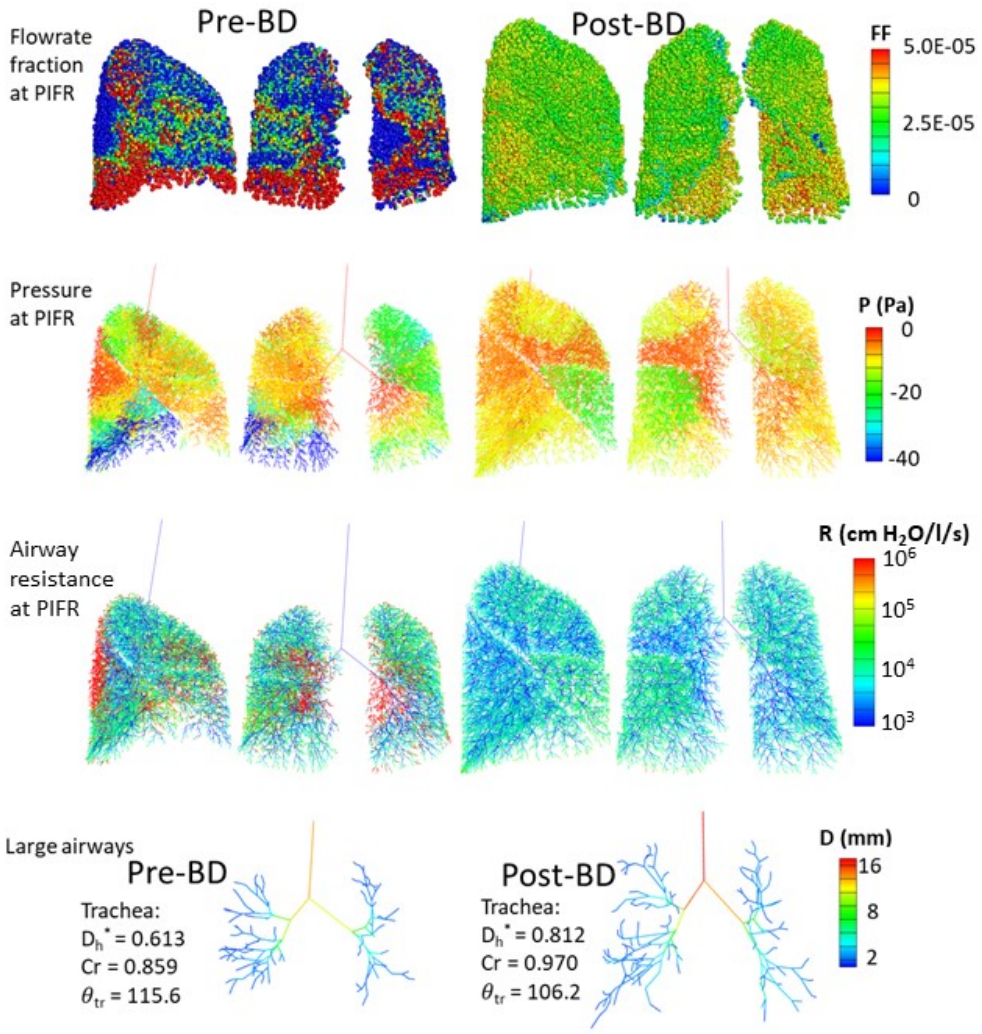

Figure 4.4: Pre- (left) and post-BD (right) comparison of (1st row) flowrate fraction (FF) distribution, (2nd row) air pressure (P) distribution, and (3rd row) airway resistance at peak inspiratory flowrate, and (4th row) large airway branches with inner diameter > 2 mm.

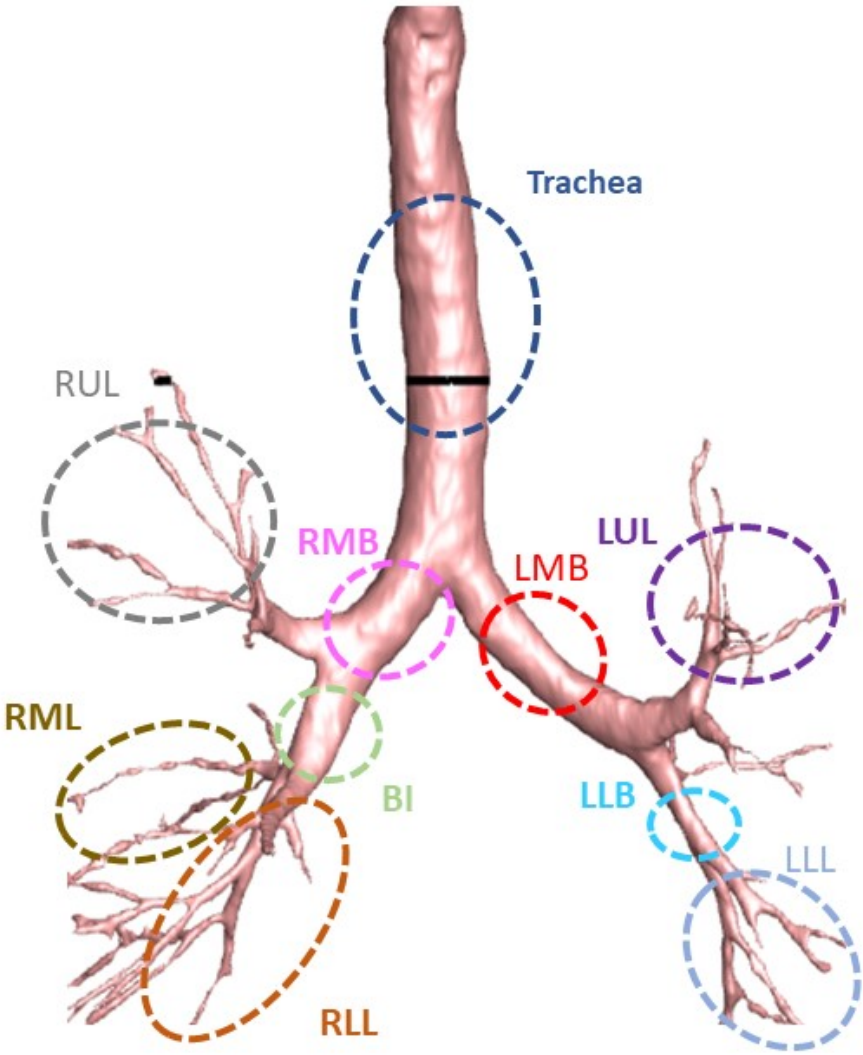

Figure 4.5: Airway branches with anatomical labels.

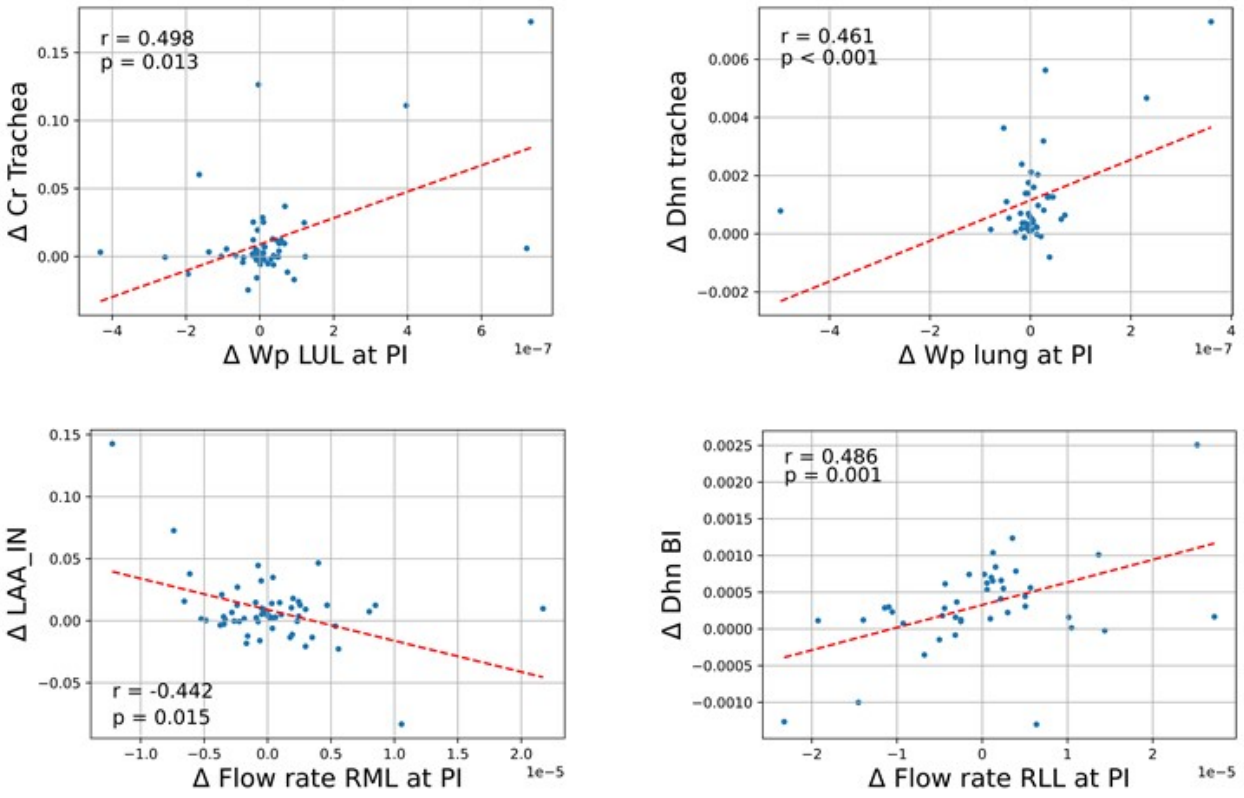

Figure 4.6: Statistical associations between 1D CFD features at peak inspiratory flow rate (PIFR) and QCT features. The red dotted line is a correlation trend line.

Chapter 5

QCT Analysis of Ambient Particulate Matter Exposure-associated Multiscale Structure And Function Alteration

Various QCT features have been well associated with clinical outcomes through statistics. Recently, machine learning has been identified as useful for linking imaging features with clinical characteristics (Choi et al., 2017a; Haghighi et al., 2018; Zou et al., 2021; Trivedi et al., 2022). For example, imaging-based subgroup analysis was used to predict clinical characteristics longitudinally. However, an integrated association of additional information, such as environmental exposure, has not been investigated through machine learning. In this chapter, our aim is to extend QCT-CFD analysis integrating with the third dataset that plays a vital role in the analysis in addition to clinical variables, using unsupervised machine learning approaches. In this study, the specific target information in the third dataset is air pollution exposure.

5.1 Introduction

Exposure to air pollution has been identified as a risk factor that affects human health severely. There are two main types of air pollutants: gaseous pollutants and atmospheric particulate matter (PM). Gaseous pollutants are accumulated in urban regions mainly due to the combustion from vehicles. Examples of gaseous pollutants are nitrogen dioxide (NO_2), ozone (O_3), and carbon monoxide (CO). On the other hand, PM is a mixture of liquid and solid particles suspended in the atmosphere. Due to the high heterogeneity of PM components, there are several ways to classify

them. One of the most common ways of classification is based on size because the physical size of PM affects how deeply PM can be deposited in the human body. PM can be classified into: PM_{10}, PM_4, $PM_{2.5}$, PM_1, and total suspended particles (TSP). PM_{10}, PM_4, $PM_{2.5}$, and PM_1 are particles with an aerodynamic diameter smaller than 10 μm, 4 μm, 2.5 μm, and 1 μm, respectively. The two major classes are PM_{10} and $PM_{2.5}$, which are referred to as fine dust with a diameter of 10 μm or smaller and ultrafine dust with a diameter of 2.5 μm or smaller. PM enters the human body through either nose or mouth and spreads through the respiratory system. Despite the fact that a large amount of PM is trapped by mucus and expelled, some PM remains or even moves around the body through the blood, damaging different parts of the body (Manojkumar & Srimuruganandam, 2021). Recently, there have been a growing amount of reports that investigated the effects of PM on the cardiovascular system (Cao et al., 2011; Gondalia et al., 2020), nervous system (Younan et al., 2020; Costa et al., 2020), and immune system (Gao et al., 2019). However, its effects on the primary site of PM invasion, i.e., the respiratory systems, may remain as an important topic.

Adamkiewicz et al. (2020) investigated clinical aspects of ambient air pollution and identified outdoor air pollution as a risk factor for exacerbation of respiratory diseases. However, its long-term exposure and many associated confounding factors make the pathophysiological effects of PM exposure unclear. Therefore, there have been several studies utilizing QCT analysis to assess its effects because QCT analysis can evaluate regional disease severity. Tung et al. (2021) found associations between $PM_{2.5}$ and emphysema severity in 86 COPD patients. Wang et al. (2019) investigated 7,071 participants without clinical cardiovascular disease longitudinally and found ambient concentrations of $PM_{2.5}$ and NO_x were significantly associated with an increase in percent emphysema. Kwon et al. (2020) included more QCT features for analysis and found associations of PM_{10} and NO_2 with emphysema and airway wall area percentage.

Most QCT and air pollution studies have been limited to looking for associations between specific exposure and specific QCT variables. Also, disease specific study of the association between QCT features and air pollution exposure has been only found in the COPD population. In this study, we extended the application of our established QCT framework for QCT-exposure

combined multi-modal dataset to investigate the effects of ambient air pollution in healthy lungs, asthma, COPD, and IPF patients, not only seeking associations between exposure and QCT data but also establishing exposure-based PCA and clustering as machine learning applications.

5.2 Methods

Inspiratory and expiratory CTs, direct measurements of PM_1, $PM_{2.5}$, PM_4, PM_{10}, TSP, and NO_2, and PFT results were prospectively collected (under IRB approval and with subject consent, using a common dose-reduced QCT protocol (Sieren et al., 2016)) from 270 participants enrolled in five tertiary medical centers collected in Seoul National University Hospital in South Korea. The population includes adults with healthy lungs (age=68$\pm$10, M:F=15:51), asthma (age=60$\pm$12, M:F=39:56), COPD (age=69$\pm$7, M:F=66:10), and IPF (age=72$\pm$7, M:F=43:10). 113 QCT and 72 CFD features of multiscale structure and function were derived from VIDA Vision (Coralville, IA) and *in-house* QCT-CFD pipeline (from Chapter 2).

Dimension of air pollution exposure data was reduced using PCA. Four principal components were determined by Horn's parallel analysis. Consequently, K-means clustering was used to generate exposure based subgroups using the principal components. The number of clusters for the K-means clustering was determined by the best Silhouette coefficient. Both exposure clusters and disease types were considered in the analysis. We first performed ANOVA test to find any difference among groups, and post-hoc Tukey-Kramer test was used to identify significantly difference between groups. Then, associations between exposure variables and QCT-CFD features were found using Pearson's correlation. P-values less than 0.05 were considered significant.

5.3 Results and Discussion

We derived three exposure-based clusters and found significant differences in QCT and PFTs between high (cluster 3) and low (cluster 1) exposure clusters (Fig 5.1). Also, many associations between QCT features and exposure variables within diseases and within exposure clusters were

found.

First, subjects with all disease states were statistically compared. Associations between air pollution exposure and two major pathological processes of airflow limitations were confirmed: airway narrowing and small airway disease (Wang et al., 2020). Narrowing of D_h^*(sRLL) was significantly associated with PM_1 and $PM_{2.5}$ (r=-0.144, -0.141), across all disease types. Also, D_h^*(sRLL) was significantly smaller in cluster 3 (0.34±0.05) than in cluster 1 (0.31±0.07) (Fig 5.2). fSAD% was increased with high exposure of PM_1, $PM_{2.5}$, PM_4, and NO_2 (r=0.189, 0.178, 0.132, 0.178) (Fig 5.3) but not with large particles (PM_{10}, TSP). Bhatt et al. (2016) found that both fSAD% and $LAA\%_{IN}$ were associated with FEV_1 decline in COPD patients but only fSAD% was associated with FEV_1 decline for subjects without airflow obstruction, meaning functional small airway disease may precede emphysema development (Galbán et al., 2012). Exposure to small particles (PM_1, $PM_{2.5}$, and PM_4) and NO_2 plays a significant role in developing obstructive lung diseases. Findings in the D_h^* and fSAD analyses may suggest PM induces in airway constriction in both proximal and distal regions. It is speculated that the greatest decrease in D_h^*(sRLL) is attributable to hypoinflation of lungs with IPF as the characteristic of a restrictive lung disease.

Analysis in each disease was also conducted. In COPD, PM_1, $PM_{2.5}$, PM_4 and PM_{10} were negatively correlated with $LAA\%_{IN}$ (r=-0.265, -0.257, -0.227, -0.258). This result is consistent with previous findings (Tung et al., 2021; Kwon et al., 2020). We also found that J was associated with PM_1, $PM_{2.5}$, and PM_4 (r=-0.257, -0.283, -0.266). Figure. 5.4 shows $LAA\%_{IN}$ and J vs $PM_{2.5}$ in COPD patients. Interestingly, $LAA\%_{IN}$ is negatively correlated with the exposure, which may indicate that $PM_{2.5}$ is associated with hypoinflation, presumably due to inflammation of the local lung response as a response to the toxic particle deposition. This is a first QCT-based suggestion on hypoinflaion associated with local lung inflammation. There have been only few studies investigating effects of air pollution exposure using QCT in diseases other than COPD. In asthma, associations of $LAA\%_{IN}$ with PM_{10} and TSP (r=0.307, 0.316) but not with small particles were found. Also, NO_2 was found to be significant in asthma, as it was associated with fSAD%, AirT%, and Cr of the tracheal lumen cross section (r=0.310, 0.318, -412) (Fig 5.5).

1D CFD analysis showed overall higher associations between the CFD-derived functional features and the exposure. Across all disease types, directly measured PM_1, $PM_{2.5}$, and PM_4 exposures correlated with CFD-derived airway resistance of the whole lung (r=0.416, 0.485, 0.436) (hereafter: at the peak inspiratory flow rate; p<0.05). D_h^*(sRLL) was significantly smaller (-8.2%, p=0.017) in the high PM exposure group than the low exposure group, and D_h^*(sRLL) moderately associated with acinar pressure work Wp computed by $\Delta P \times FR$ (ΔP is pressure drop through the entire conducting airway, and FR is flow rate) in the three lobes of the right lung (RUL, RML, and RLL; r=0.409,0.402, 0.439) and the whole lung. In asthma, residential exposures to PM2.5, PM10, and ozone associated with ΔP in the whole lung (r=0.425, 0.441, 0.419) and greater in the lower lobes (r=0.590, 0.613, 0.589 in LLL). Directly measured PM_1, $PM_{2.5}$, and PM_4 associated with transpulmonary pressure (Ptp) in COPD (r=-448, -0.492, -0.478) and IPF (r=-0.500, -0.536, -0.526), with R (r=-534, -0.628, -0.587), and with ΔP (r=0.500, 0.536, 0.526) in COPD. In COPD, work address-based CO exposure highly associated with ΔP (r=0.689, whole lung; 0.793, RLL) and Wp (r=0.632), and moderately with R (r=0.554). PM2.5 at work associated with Ptp (r=0.587) and R (r=0.604). Residential exposure to PM2.5 and PM10 associated with Ptp in COPD (r=0.587) and ΔP in IPF (r=0.462, whole lung; 0.672, RLL). The results were submitted to the 2022 Radiological Society of North America (RSNA) Annual Meeting, Chicago, IL, USA, and has been accepted for oral presentation.

5.4 Conclusion

Machine learning algorithms, such as PCA and K-means clustering, were used to derive three exposure-based cluster membership out of 21 directly measured exposure variables. The resulting cluster memberships were successfully used for QCT analysis of air pollution's impact on lung health and disease. Inspiratory and expiratory QCT image matching-based multiscale airway and lung structure-function analysis suggested regional structural and functional impairments associated with directly measured PM and NO_2 exposure in health and diseases (asthma, COPD, and IPF). PCA and K-means clustering were used to derive three exposure-based clusters. QCT

features were compared within diseases and within clusters. In the low exposure cluster, QCT values were highly variable between patients due to various disease states. However, correlations are more noticeable at higher exposure levels. Furthermore, different air pollutants, depending on their phase and physical size, alter structure-function relationships differently in different disease states. In all disease states, small particles and NO2 were associated with obstructive lung diseases. Small PM increases small airway disease in asthma, PM at all sizes decreases inspiratory low attenuation area in COPD and diseases luminal diameter of RLL segmental airways in IPF.

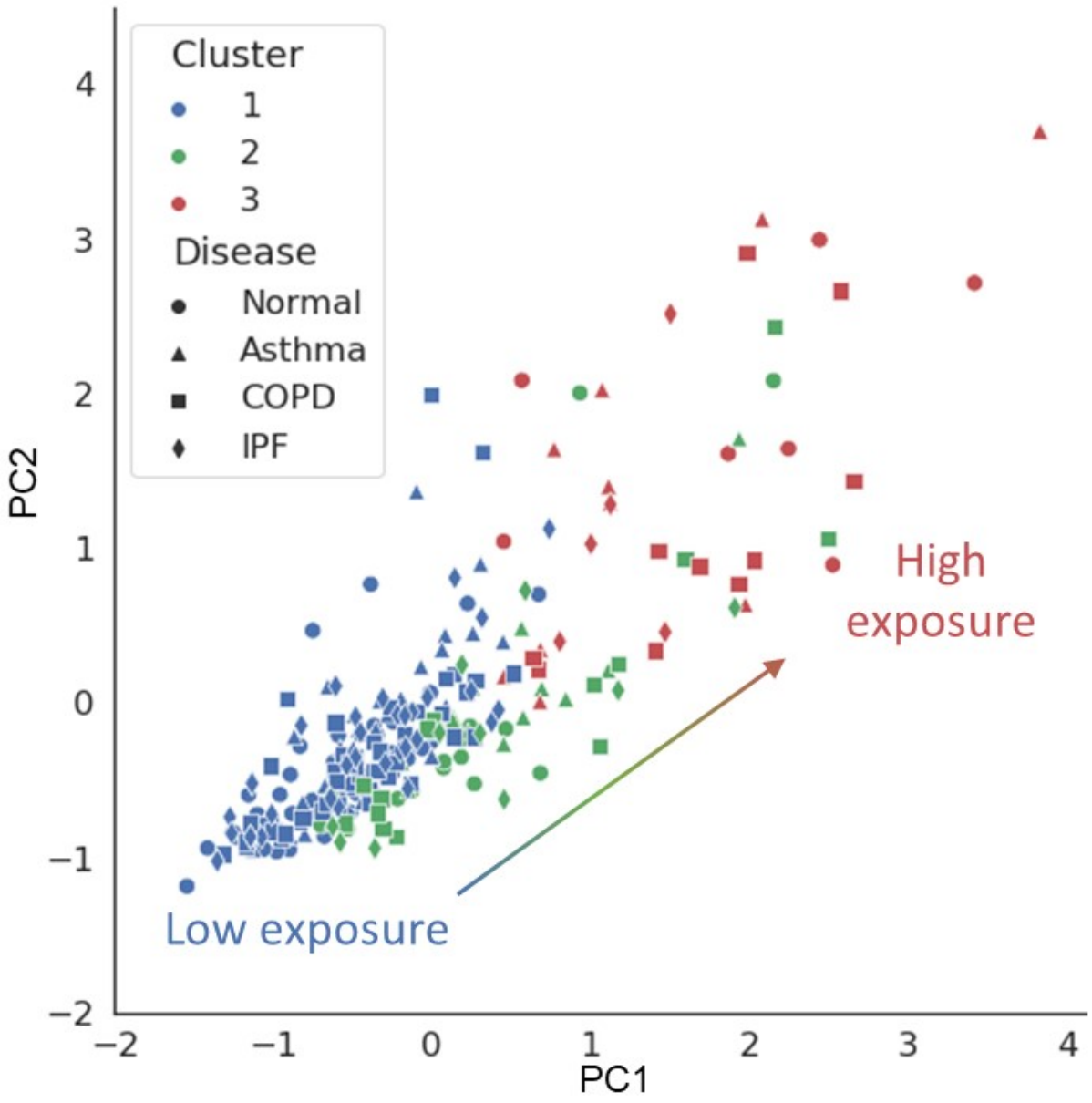

Figure 5.1: Air pollution exposure-based PCA and K-means clustering where blue, green, and red represent low exposure, moderate exposure, and high exposure respectively

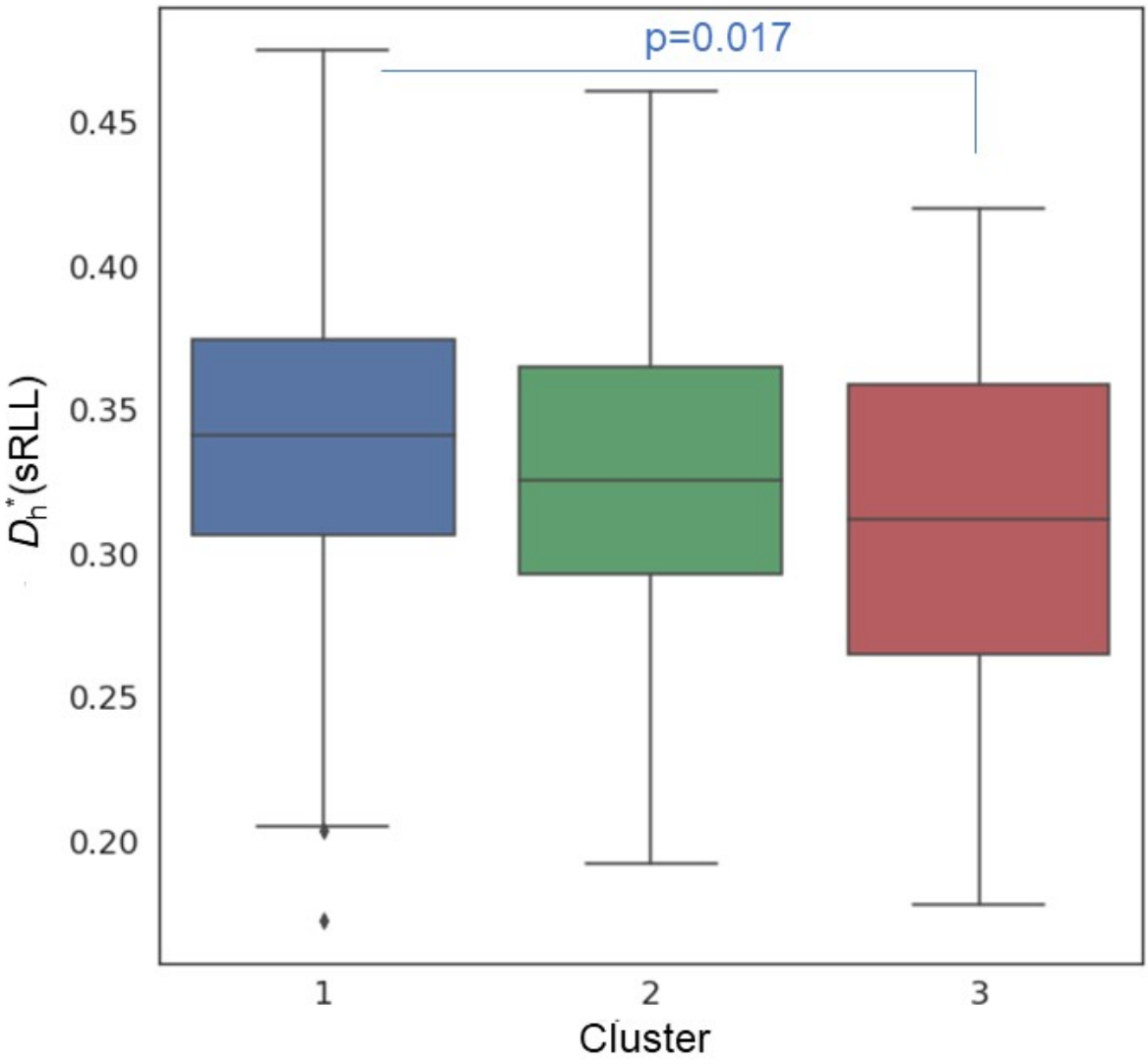

Figure 5.2: D_h^* in the right lower lobe (RLL) segmental airways in exposure clusters: significantly smaller in culster 3 (high exposure) than in cluster 1 (low exposure).

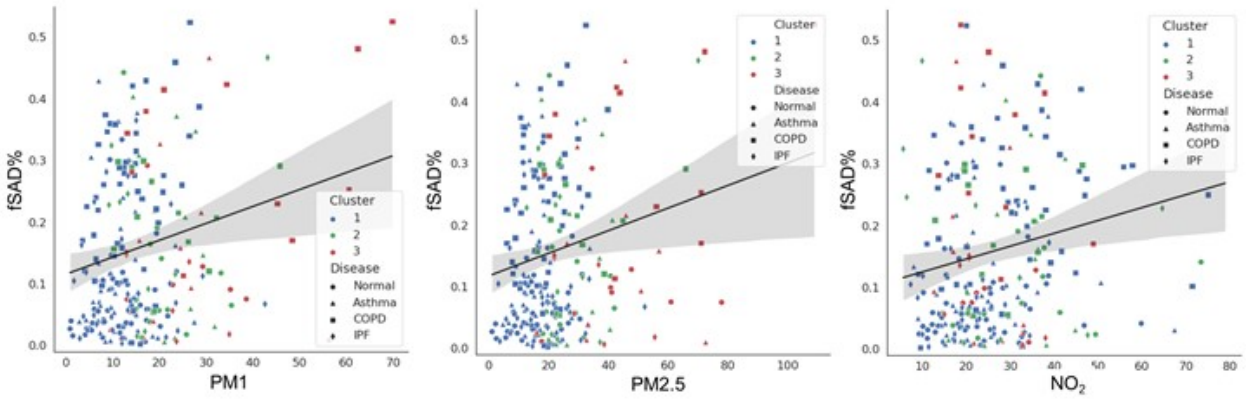

Figure 5.3: Correlations between PM_1, $PM_{2.5}$, NO_2 and fSAD% (left, middle, and right, respectively) across all disease types. The gray region represents confidence interval of 95%.

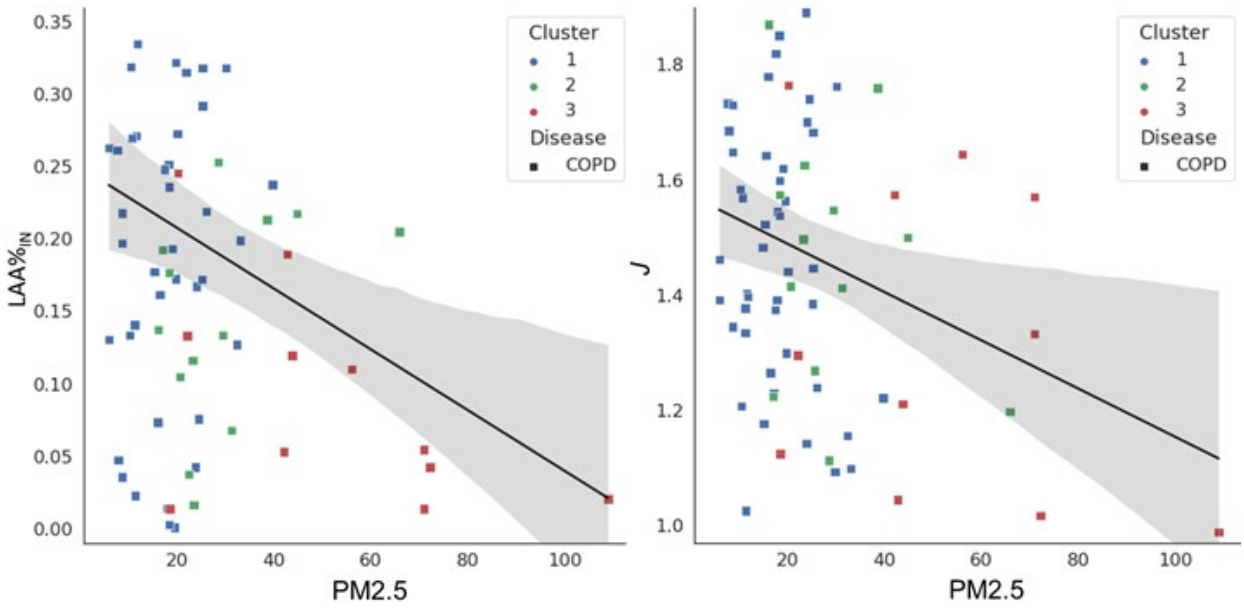

Figure 5.4: Correlations between $PM_{2.5}$ and $LAA\%_{IN}$ (left) and between $PM_{2.5}$ and J (right) in COPD. The gray region represents confidence interval of 95%.

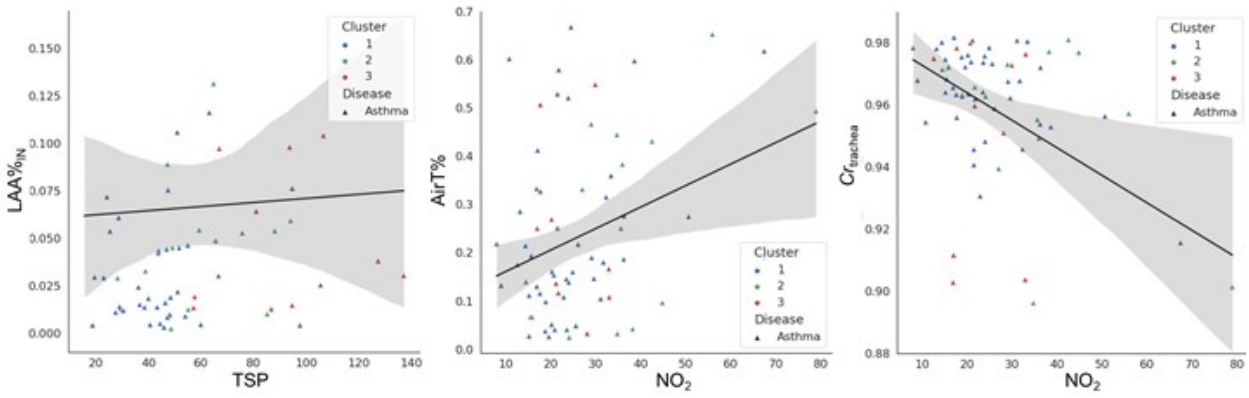

Figure 5.5: Correlations between TSP and LAA$\%_{IN}$ (left), NO_2 and AirT% (middle), and NO_2 and $Cr_{trachea}$ (right) in asthma. The gray region represents confidence interval of 95%.

Chapter 6

Conclusion

In this thesis, an extended QCT analysis framework was introduced and applied to a wide range of data with postprocessing. Each Chapter is summarized, and the conclusion of this work is made in this Chapter.

In Chapter 2, the basics of lung anatomy, focusing on the lobes and airway, were reviewed. It was emphasized that segmentation mask is a crucial part of QCT analysis and that, as well as the types of masks, the segmentation quality is important for the accuracy of QCT features. Segmentation was then connected to threshold-based QCT features. HAA, which represents overall interstitial lung lesions, can be computed with a threshold between -700 HU and 0 HU. Its lobal-wise percentages can be calculated by dividing the number of HAA voxels by the number of total voxels in each lobe. While HAA% has been widely used to analyze patients with pulmonary fibrosis and other restrictive lung diseases, LAA% values on inspiratory and expiratory CTs are more commonly used for obstructive diseases such as COPD and asthma, respectively. Airway masks are used to extract airway structural metrics such as airway circularity, diameter, and branching angle. By image registration between inspiratory and expiratory lung CTs, functional QCT variables are computed, which include regional deformation and ventilation. Airway mask and ventilation distribution are used to prepare anatomically correct airway geometry and physiologically consistent boundary conditions for CFD airflow simulations. In addition, machine learning methods were introduced as a postprocessing step of QCT analysis. PCA, an unsupervised dimensionality reduction algorithm, is used to reduce the high dimensional data space of QCT variables. Another unsupervised machine learning method, K-means clustering, was introduced to create clinical subgroups for analysis.

In Chapter 3, a new slice thickness-independent deep learning model for pulmonary lobe segmentation was developed and named "ZUNet," reflecting its core addition of z-axis information to the well-known UNet segmentation model. The new deep learning model was trained with thin slice thickness CTs and accurately segmented lobes from thick slice thickness CT of silicosis patients. From the resulting lobe masks, QCT features of silicosis patients were successfully extracted and compared with those of IPF patients and healthy controls. It was highlighted that HAA% of the RUL was significantly higher in silicosis than in IPF patients, reflecting the regional difference between the two severe diseases. To our knowledge, this is the first quantitative confirmation of regional lung characteristics of silicosis.

In Chapter 4, another application of the QCT framework linking with 1D CFD air flow simulation for assessment of treatment response was introduced. The established QCT-CFD framework was used to investigate the effects of bronchodilator inhalation in asthmatic patients. Inspiratory and expiratory CT scans before and after the bronchodilator inhalation were collected. Each inspiratory and expiratory set is image registered. Consequently, 1D CFD simulations of subject-specific tidal breathing were conducted for both pre- and post-bronchodilator conditions. On the one hand, 1D CFD simulation demonstrated functional improvement with a reduction in regional discordance. However, not all 66 asthmatic patients responded the same based on statistical comparison before and after bronchodilator, while all showed significant large airway dilatation by bronchodilator. Significant regional lung functional improvement, such as ventilation improvement in the right lower lobe along with the reduced whole lung pressure workload, was found. 1D CFD simulations successfully showed the population-based statistical analysis of regional lung functional changes after the treatment, which cannot be done with 3D CFD simulations due to high computational cost and time.

Finally, the QCT framework is applied to analyze more diverse data. In Chapter 5, PM exposure data were integrated with QCT analysis for the impact on lung structural and functional changes. PCA was first utilized to reduce the dimension of PM exposure variables, and k-means clustering was used to make three PM exposure-based subgroups (low, moderate, and high exposure).

Associations between QCT variables and exposure features were found. Also, an exposure-based cluster analysis was done. Regional airway narrowing, particularly in the RLL segmental airways, was found in the high PM exposure cluster (cluster 3) across all disease types (healthy, asthma, COPD, and IPF) and most remarkably in IPF. 1D CFD analysis suggested an association of this airway narrowing in the RLL segmental airways with a functional decline of the RLL and also the whole lung.

In summary, extended QCT analysis has been investigated in three aspects. Deep learning segmentation enabled quantitative analysis of regional lung damage in silicosis patients from thick slice CTs. QCT-CFD analysis has resolved the structure-function relationship in response to drug treatment, particularly the bronchodilator response in asthma. Moreover, machine learning approaches (PCA and K-means clustering) enlarged the QCT analysis based on the air pollution exposure dataset that was additionally collected with the CT and other clinical data.

References

Abdi, H. & Williams, L. J. (2010). Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433–459.

Adamkiewicz, G., Liddie, J., & Gaffin, J. M. (2020). The respiratory risks of ambient/outdoor air pollution. *Clinics in chest medicine*, 41(4), 809–824.

Altınsoy, B., Öz, İ. İ., Erboy, F., Tor, M., & Atalay, F. (2016). Emphysema and airflow obstruction in non-smoking coal miners with pneumoconiosis. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, 22, 4887.

Amelon, R., Cao, K., Ding, K., Christensen, G. E., Reinhardt, J. M., & Raghavan, M. L. (2011). Three-dimensional characterization of regional lung deformation. *Journal of biomechanics*, 44(13), 2489–2495.

American Thoracic Society & European Respiratory Society (2002). American thoracic Society/European respiratory society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. this joint statement of the american thoracic society (ATS), and the european respiratory society (ERS) was adopted by the ATS board of directors, june 2001 and by the ERS executive committee, june 2001. *Am. J. Respir. Crit. Care Med.*, 165(2), 277–304.

Bhatt, S. P., Soler, X., Wang, X., Murray, S., Anzueto, A. R., Beaty, T. H., Boriek, A. M., Casaburi, R., Criner, G. J., Diaz, A. A., et al. (2016). Association between functional small airway disease and fev1 decline in chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine*, 194(2), 178–184.

Bragman, F. J., McClelland, J. R., Jacob, J., Hurst, J. R., & Hawkes, D. J. (2017). Pulmonary lobe segmentation with probabilistic segmentation of the fissures and a groupwise fissure prior. *IEEE transactions on medical imaging*, 36(8), 1650–1663.

Busacker, A., Newell Jr, J. D., Keefe, T., Hoffman, E. A., Granroth, J. C., Castro, M., Fain, S., & Wenzel, S. (2009). A multivariate analysis of risk factors for the air-trapping asthmatic phenotype as measured by quantitative ct analysis. *Chest*, 135(1), 48–56.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing*, 16(5), 1190–1208.

Cao, J., Yang, C., Li, J., Chen, R., Chen, B., Gu, D., & Kan, H. (2011). Association between long-term exposure to outdoor air pollution and mortality in china: a cohort study. *Journal of hazardous materials*, 186(2-3), 1594–1600.

Chae, K. J., Choi, J., Jin, G. Y., Hoffman, E. A., Laroia, A. T., Park, M., & Lee, C. H. (2020). Relative regional air volume change maps at the acinar scale reflect variable ventilation in low lung attenuation of copd patients. *Academic Radiology*, 27(11), 1540–1548.

Chan, H.-P., Samala, R. K., Hadjiiski, L. M., & Zhou, C. (2020). Deep learning in medical image analysis. *Deep Learning in Medical Image Analysis*, (pp. 3–21).

Chen, A., Karwoski, R. A., Gierada, D. S., Bartholmai, B. J., & Koo, C. W. (2020). Quantitative ct analysis of diffuse lung disease. *Radiographics*, 40(1), 28–43.

Choi, J. (2011). *Multiscale numerical analysis of airflow in CT-based subject specific breathing human lungs*. PhD thesis, University of Iowa.

Choi, J., LeBlanc, L. J., Choi, S., Haghighi, B., Hoffman, E. A., O'Shaughnessy, P., Wenzel, S. E., Castro, M., Fain, S., Jarjour, N., et al. (2019a). Differences in particle deposition between members of imaging-based asthma clusters. *Journal of aerosol medicine and pulmonary drug delivery*, 32(4), 213–223.

Choi, J., Lim, H.-j., Park, M., Kang, H.-R., Lin, C.-L., Hoffman, E. A., & Lee, C. H. (2017a). Cross-volume and cross-time ct image matching assessment provides multiscale characterization of individual asthmatic lung responses to bronchodilator. In *A98. SEEING IS BELIEVING: USING NOVEL IMAGING TECHNIQUES TO UNDERSTAND THE LUNG IN HEALTH AND DISEASE* (pp. A2666–A2666). American Thoracic Society.

Choi, J., Tawhai, M. H., Hoffman, E. A., & Lin, C.-L. (2009). On intra-and intersubject variabilities of airflow in the human lungs. *Physics of Fluids*, 21(10), 101901.

Choi, J., Xia, G., Tawhai, M. H., Hoffman, E. A., & Lin, C.-L. (2010). Numerical study of high-frequency oscillatory air flow and convective mixing in a ct-based human airway model. *Annals of biomedical engineering*, 38(12), 3550–3571.

Choi, S., Haghighi, B., Choi, J., Hoffman, E. A., Comellas, A. P., Newell, J. D., Wenzel, S. E., Castro, M., Fain, S. B., Jarjour, N. N., et al. (2017b). Differentiation of quantitative ct imaging phenotypes in asthma versus copd. *BMJ open respiratory research*, 4(1), e000252.

Choi, S., Hoffman, E. A., Wenzel, S. E., Castro, M., Fain, S., Jarjour, N., Schiebler, M. L., Chen, K., Lin, C.-L., Heart, N., et al. (2017c). Quantitative computed tomographic imaging–based clustering differentiates asthmatic subgroups with distinctive clinical phenotypes. *Journal of Allergy and Clinical Immunology*, 140(3), 690–700.

Choi, S., Hoffman, E. A., Wenzel, S. E., Castro, M., Fain, S. B., Jarjour, N. N., Schiebler, M. L., Chen, K., & Lin, C.-L. (2015). Quantitative assessment of multiscale structural and functional alterations in asthmatic populations. *Journal of applied physiology*, 118(10), 1286–1298.

Choi, S., Yoon, S., Jeon, J., Zou, C., Choi, J., Tawhai, M. H., Hoffman, E. A., Delvadia, R., Babiskin, A., Walenga, R., et al. (2019b). 1d network simulations for evaluating regional flow and pressure distributions in healthy and asthmatic human lungs. *Journal of Applied Physiology*, 127(1), 122–133.

Costa, L. G., Cole, T. B., Dao, K., Chang, Y.-C., Coburn, J., & Garrick, J. M. (2020). Effects of air pollution on the nervous system and its possible role in neurodevelopmental and neurodegenerative disorders. *Pharmacology & therapeutics*, 210, 107523.

de Castro, M. C. S., Ferreira, A. S., Irion, K. L., Hochhegger, B., Lopes, A. J., Velarde, G. C., Zanetti, G., & Marchiori, E. (2014). Ct quantification of large opacities and emphysema in silicosis: correlations among clinical, functional, and radiological parameters. *Lung*, 192(4), 543–551.

Dirksen, A., Dijkman, J. H., Madsen, F., Stoel, B., Hutchison, D. C., Ulrik, C. S., Skovgaard, L. T., Kok-Jensen, A., Rudolphus, A., Seersholm, N., et al. (1999). A randomized clinical trial of α1-antitrypsin augmentation therapy. *American journal of respiratory and critical care medicine*, 160(5), 1468–1472.

Dirksen, A., Friis, M., Olesen, K., Skovgaard, L. T., & Sørensen, K. (1997). Progress of emphysema in severe α1-antitrypsin deficiency as assessed by annual ct. *Acta Radiologica*, 38(5), 826–832.

dos Santos Antao, V. C., Pinheiro, G. A., Terra-Filho, M., Kavakama, J., & Müller, N. L. (2005). High-resolution ct in silicosis: correlation with radiographic findings and functional impairment. *Journal of computer assisted tomography*, 29(3), 350–356.

Dowson, L. J., GUEST, P. J., & STOCKLEY, R. A. (2001). Longitudinal changes in physiological, radiological, and health status measurements in α1-antitrypsin deficiency and factors associated with decline. *American journal of respiratory and critical care medicine*, 164(10), 1805–1809.

Ferri, F. F. (2021). *Ferri's Clinical Advisor 2022, E-Book*. Elsevier Health Sciences.

Galbán, C. J., Han, M. K., Boes, J. L., Chughtai, K. A., Meyer, C. R., Johnson, T. D., Galbán, S., Rehemtulla, A., Kazerooni, E. A., Martinez, F. J., et al. (2012). Computed tomography–based biomarker provides unique signature for diagnosis of copd phenotypes and disease progression. *Nature medicine*, 18(11), 1711–1715.

Gao, X., Colicino, E., Shen, J., Kioumourtzoglou, M.-A., Just, A. C., Nwanaji-Enwerem, J. C., Coull, B., Lin, X., Vokonas, P., Zheng, Y., et al. (2019). Impacts of air pollution, temperature, and relative humidity on leukocyte distribution: an epigenetic perspective. *Environment international*, 126, 395–405.

George, K., Harrison, A. P., Jin, D., Xu, Z., & Mollura, D. J. (2017). Pathological pulmonary lobe segmentation from ct images using progressive holistically nested neural networks and random walker. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 195–203). Springer.

Gerard, S. E. & Reinhardt, J. M. (2019). Pulmonary lobe segmentation using a sequence of convolutional neural networks for marginal learning. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (pp. 1207–1211).: IEEE.

Gevenois, P. A., De Maertelaer, V., De Vuyst, P., Zanen, J., & Yernault, J.-C. (1995). Comparison of computed density and macroscopic morphometry in pulmonary emphysema. *American journal of respiratory and critical care medicine*, 152(2), 653–657.

Gevenois, P. A., De Vuyst, P. d., De Maertelaer, V., Zanen, J., Jacobovitz, D., Cosio, M. G., & Yernault, J.-C. (1996). Comparison of computed density and microscopic morphometry in pulmonary emphysema. *American journal of respiratory and critical care medicine*, 154(1), 187–192.

Gondalia, R., Holliday, K. M., Baldassari, A., Justice, A. E., Stewart, J. D., Liao, D., Yanosky, J. D., Engel, S. M., Jordahl, K. M., Bhatti, P., et al. (2020). Leukocyte traits and exposure to ambient particulate matter air pollution in the women's health initiative and atherosclerosis risk in communities study. *Environmental health perspectives*, 128(1), 017004.

Gu, H., Gan, W., Zhang, C., Feng, A., Wang, H., Huang, Y., Chen, H., Shao, Y., Duan, Y., & Xu, Z. (2021). A 2d–3d hybrid convolutional neural network for lung lobe auto-segmentation on

standard slice thickness computed tomography of patients receiving radiotherapy. *BioMedical Engineering OnLine*, 20(1), 1–13.

Haghighi, B., D Ellingwood, N., Yin, Y., Hoffman, E. A., & Lin, C.-L. (2018). A gpu-based symmetric non-rigid image registration method in human lung. *Medical & biological engineering & computing*, 56(3), 355–371.

Hartigan, J. A. & Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1), 100–108.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.

Hoesein, F. A. M., van Rikxoort, E., van Ginneken, B., de Jong, P. A., Prokop, M., Lammers, J.-W. J., & Zanen, P. (2012). Computed tomography-quantified emphysema distribution is associated with lung function decline. *European Respiratory Journal*, 40(4), 844–850.

Hoffman, E. A., Ahmed, F. S., Baumhauer, H., Budoff, M., Carr, J. J., Kronmal, R., Reddy, S., & Barr, R. G. (2014). Variation in the percent of emphysema-like lung in a healthy, nonsmoking multiethnic sample. the mesa lung study. *Annals of the American Thoracic Society*, 11(6), 898–907.

Hofmanninger, J. (2020). Prayer f pan j röhrich s prosch h langs g automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *Eur. Radiol. Exp*, 4(1), 1.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185.

Ioffe, S. & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448–456).: PMLR.

Jeffery, P. K. (1998). Structural and inflammatory changes in copd: a comparison with asthma. *Thorax*, 53(2), 129.

Kang, J. H., Choi, J., Chae, K. J., Shin, K. M., Lee, C.-H., Guo, J., Lin, C.-L., Hoffman, E. A., & Lee, C. (2021). Ct-derived 3d-diaphragm motion in emphysema and ipf compared to normal subjects. *Scientific reports*, 11(1), 1–12.

Kim, T., Cho, H. B., Kim, W. J., Lee, C. H., Chae, K. J., Choi, S.-H., Lee, K. E., Bak, S. H., Kwon, S. O., Jin, G. Y., et al. (2020). Quantitative ct-based structural alterations of segmental airways in cement dust-exposed subjects. *Respiratory research*, 21(1), 1–11.

Knutsson, H. & Westin, C.-F. (1993). Normalized and differential convolution. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (pp. 515–523).: IEEE.

Kwon, S. O., Hong, S. H., Han, Y.-J., Bak, S. H., Kim, J., Lee, M. K., London, S. J., Kim, W. J., & Kim, S.-Y. (2020). Long-term exposure to pm10 and no2 in relation to lung function and imaging phenotypes in a copd cohort. *Respiratory Research*, 21(1), 1–11.

Kybic, J. & Unser, M. (2003). Fast parametric elastic image registration. *IEEE transactions on image processing*, 12(11), 1427–1442.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436–444.

Lee, C. H., Choi, J., Lee, I. K., & Kang, H.-R. (2022). Quantitative ct image matching analysis of multiscale airway and lung structure-function responses to bronchodilator in asthma. In *28th Annual Scientific Meeting of European Society of Thoracic Imaging* (pp. 234–241).: ESTI.

Leung, A. N. (1999). Pulmonary tuberculosis: the essentials. *Radiology*, 210(2), 307–322.

Lin, C.-L., Tawhai, M. H., & Hoffman, E. A. (2013). Multiscale image-based modeling and simulation of gas flow and particle transport in the human lungs. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 5(5), 643–655.

Lin, C.-L., Tawhai, M. H., McLennan, G., & Hoffman, E. A. (2007). Characteristics of the turbulent laryngeal jet and its effect on airflow in the human intra-thoracic airways. *Respiratory physiology & neurobiology*, 157(2-3), 295–309.

Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2), 129–137.

Manojkumar, N. & Srimuruganandam, B. (2021). Health effects of particulate matter in major indian cities. *International journal of environmental health research*, 31(3), 258–270.

Masoli, M., Fabian, D., Holt, S., Beasley, R., & for Asthma (GINA) Program, G. I. (2004). The global burden of asthma: executive summary of the gina dissemination committee report. *Allergy*, 59(5), 469–478.

Matsuoka, S., Yamashiro, T., Matsushita, S., Fujikawa, A., Yagihashi, K., & Nakajima, Y. (2016). Objective quantitative ct evaluation using different attenuation ranges in patients with pulmonary fibrosis: correlations with visual scores. *Int J Respir Pulm Med*, 3(3), 3–7.

Mlika, M., Adigun, R., & Bhutta, B. S. (2022). Silicosis. In *StatPearls [Internet]*. StatPearls Publishing.

Morgan, E. J. (1979). Silicosis and tuberculosis. *Chest*, 75(2), 202–203.

Ostridge, K. & Wilkinson, T. M. (2016). Present and future utility of computed tomography scanning in the assessment and management of copd. *European Respiratory Journal*, 48(1), 216–228.

Palágyi, K., Tschirren, J., Hoffman, E. A., & Sonka, M. (2006). Quantitative analysis of pulmonary airway tree structures. *Computers in biology and medicine*, 36(9), 974–996.

Park, J., Yun, J., Kim, N., Park, B., Cho, Y., Park, H. J., Song, M., Lee, M., & Seo, J. B. (2020). Fully automated lung lobe segmentation in volumetric chest ct with 3d u-net: validation with intra-and extra-datasets. *Journal of digital imaging*, 33(1), 221–230.

Paulin, L. M., Smith, B. M., Koch, A., Han, M., Hoffman, E. A., Martinez, C., Ejike, C., Blanc, P. D., Rous, J., Barr, R. G., et al. (2018). Occupational exposures and computed tomographic imaging characteristics in the spiromics cohort. *Annals of the American Thoracic Society*, 15(12), 1411–1419.

Pedley, T. & TJ, P. (1977). Pulmonary fluid dynamics.

Pistenmaa, C. L., Nardelli, P., Ash, S., Come, C., Diaz, A., Rahaghi, F., Barr, R., Young, K., Kinney, G., Simmons, J., et al. (2021). Pulmonary arterial pruning and longitudinal change in percent emphysema and lung function: the genetic epidemiology of copd study. *Chest*, 160(2), 470–480.

Podolanczuk, A. J., Oelsner, E. C., Barr, R. G., Bernstein, E. J., Hoffman, E. A., Easthausen, I. J., Stukovsky, K. H., RoyChoudhury, A., Michos, E. D., Raghu, G., et al. (2017). High-attenuation areas on chest computed tomography and clinical respiratory outcomes in community-dwelling adults. *American journal of respiratory and critical care medicine*, 196(11), 1434–1442.

Podolanczuk, A. J., Oelsner, E. C., Barr, R. G., Hoffman, E. A., Armstrong, H. F., Austin, J. H., Basner, R. C., Bartels, M. N., Christie, J. D., Enright, P. L., et al. (2016). High attenuation areas on chest computed tomography in community-dwelling adults: the mesa study. *European Respiratory Journal*, 48(5), 1442–1452.

Raasch, B., Carsky, E., Lane, E., O'callaghan, J., & Heitzman, E. (1982). Radiographic anatomy of the interlobar fissures: a study of 100 specimens. *American Journal of Roentgenology*, 138(6), 1043–1049.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234–241).: Springer.

Rose, C., Heinzerling, A., Patel, K., Sack, C., Wolff, J., Zell-Baran, L., Weissman, D., Hall, E., Sooriash, R., McCarthy, R. B., et al. (2019). Severe silicosis in engineered stone fabrication

workers—california, colorado, texas, and washington, 2017–2019. *Morbidity and Mortality Weekly Report*, 68(38), 813.

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., & Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8), 712–721.

Schroeder, J. D., McKenzie, A. S., Zach, J. A., Wilson, C. G., Curran-Everett, D., Stinson, D. S., Newell Jr, J. D., & Lynch, D. A. (2013). Relationships between airflow obstruction and quantitative ct measurements of emphysema, air trapping, and airways in subjects with and without chronic obstructive pulmonary disease. *AJR. American journal of roentgenology*, 201(3), W460.

Sentinel-Event-Alert (2011). Radiation risks of diagnostic imaging and fluoroscopy. *The Joint Commission on the Accreditation of Health Organizations*, 47.

Sieren, J. P., Newell Jr, J. D., Barr, R. G., Bleecker, E. R., Burnette, N., Carretta, E. E., Couper, D., Goldin, J., Guo, J., Han, M. K., et al. (2016). Spiromics protocol for multicenter quantitative computed tomography to phenotype the lungs. *American journal of respiratory and critical care medicine*, 194(7), 794–806.

Smith, B. M., Jensen, D., Brosseau, M., Benedetti, A., Coxson, H. O., & Bourbeau, J. (2018). Impact of pulmonary emphysema on exercise capacity and its physiological determinants in chronic obstructive pulmonary disease. *Scientific reports*, 8(1), 1–10.

Sofranik, R., Gross, B., & Spizarny, D. (1992). Radiology of the pleural fissures. *Clinical imaging*, 16(4), 221–229.

Tanabe, N., Muro, S., Sato, S., Tanaka, S., Oguma, T., Kiyokawa, H., Takahashi, T., Kinose, D., Hoshino, Y., Kubo, T., et al. (2012). Longitudinal study of spatially heterogeneous emphysema progression in current smokers with chronic obstructive pulmonary disease.

Tanizawa, K., Handa, T., Nagai, S., Hirai, T., Kubo, T., Oguma, T., Ito, I., Ito, Y., Watanabe, K., Aihara, K., et al. (2015). Clinical impact of high-attenuation and cystic areas on computed tomography in fibrotic idiopathic interstitial pneumonias. *BMC Pulmonary Medicine*, 15(1), 1–10.

Tawhai, M. H., Hunter, P., Tschirren, J., Reinhardt, J., McLennan, G., & Hoffman, E. A. (2004). Ct-based geometry analysis and finite element models of the human and ovine bronchial tree. *Journal of applied physiology*, 97(6), 2310–2321.

Tawhai, M. H., Pullan, A., & Hunter, P. (2000). Generation of an anatomically based three-dimensional model of the conducting airways. *Annals of biomedical engineering*, 28(7), 793–802.

Trivedi, A. P., Hall, C., Goss, C. W., Lew, D., Krings, J. G., McGregor, M. C., Samant, M., Sieren, J. P., Li, H., Schechtman, K. B., et al. (2022). Quantitative ct characteristics of cluster phenotypes in the severe asthma research program cohorts. *Radiology*.

Tschirren, J., Palágyi, K., Reinhardt, J. M., Hoffman, E. A., & Sonka, M. (2002). Segmentation, skeletonization, and branchpoint matching—a fully automated quantitative evaluation of human intrathoracic airway trees. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 12–19).: Springer.

Tung, N. T., Ho, S.-C., Lu, Y.-H., Chen, T.-T., Lee, K.-Y., Chen, K.-Y., Wu, C.-D., Chung, K. F., Kuo, H.-P., Thao, H. N. X., et al. (2021). Association between air pollution and lung lobar emphysema in copd. *Frontiers in medicine*, 8.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., Abbasi-Kangevari, M., Abbastabar, H., Abd-Allah, F., Abdelalim, A., et al. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258), 1204–1222.

Wang, J., Li, Y., Zhao, P., Tian, Y., Liu, X., He, H., Jia, R., Oliver, B. G., & Li, J. (2020). Exposure to air pollution exacerbates inflammation in rats with preexisting copd. *Mediators of inflammation*, 2020.

Wang, M., Aaron, C. P., Madrigano, J., Hoffman, E. A., Angelini, E., Yang, J., Laine, A., Vetterli, T. M., Kinney, P. L., Sampson, P. D., et al. (2019). Association between long-term exposure to ambient air pollution and change in quantitatively assessed emphysema and lung function. *Jama*, 322(6), 546–556.

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853*.

Yin, Y., Choi, J., Hoffman, E. A., Tawhai, M. H., & Lin, C.-L. (2010). Simulation of pulmonary air flow with a subject-specific boundary condition. *Journal of biomechanics*, 43(11), 2159–2163.

Yin, Y., Choi, J., Hoffman, E. A., Tawhai, M. H., & Lin, C.-L. (2013). A multiscale mdct image-based breathing lung model with time-varying regional ventilation. *Journal of computational physics*, 244, 168–192.

Yin, Y., Hoffman, E. A., & Lin, C.-L. (2009). Mass preserving nonrigid registration of ct lung images using cubic b-spline. *Medical physics*, 36(9Part1), 4213–4222.

Younan, D., Petkus, A. J., Widaman, K. F., Wang, X., Casanova, R., Espeland, M. A., Gatz, M., Henderson, V. W., Manson, J. E., Rapp, S. R., et al. (2020). Particulate matter and episodic memory decline mediated by early neuroanatomic biomarkers of alzheimer's disease. *Brain*, 143(1), 289–302.

Zagers, H., Vrooman, H. A., Aarts, N. J., Stolk, J., KOOL, L. S. J., Dijkman, J. H., van Voorthuisen, E., & REIBER, J. H. (1996). Assessment of the progression of emphysema by quantitative analysis of spirometrically gated computed tomography images. *Investigative radiology*, 31(12), 761–767.

Zou, C., Li, F., Choi, J., Haghighi, B., Choi, S., Rajaraman, P. K., Comellas, A. P., Newell, J. D., Lee, C. H., Barr, R. G., et al. (2021). 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). *International journal of chronic obstructive pulmonary disease*, 16, 1477.