

Chapter 1

Quantitative computerized analysis of idiopathic pulmonary fibrosis

As introduced in Chapter 2, the natural history of IPF is poorly understood, and the clinical course for a given patient is unpredictable. Currently, there is a shortage of accepted bio-markers that can indicate the likely progression of IPF (Bartholmai et al., 2013). A successful quantification scheme that allows for recognition of disease across radiology, pulmonary and pathology disciplines still remains difficult. Development of accurate and automatic tools for quantitative assessment of alterations in the lung with IPF will be essential for a rapid patient-specific diagnosis and treatment. To date, the shape of the lungs and lobes in IPF has not been quantified, and nor has the spatial distribution of tissue abnormalities. This chapter describes a study of quantitative analysis of IPF disease based on HRCT scans, including both assessment of tissue abnormalities and lung lobe shape analysis.

1.1 Background

1.1.1 Challenges of IPF diagnosis

Managing patients with IPF presents a substantial health-care burden, due to short survival time and lack of effective treatments (with associated morbidity) (Olson et al., 2007; Raghunath et al., 2014). Accurate assessment and diagnosis of IPF is very challenging, since there is significant individual radiological and physiological variability among patients (Devaraj, 2014). The progression of disease varies considerably, ranging from rapid worsening of symptoms to relatively slow deterioration over several years (King Jr et al., 2011; Richeldi et al., 2017). The American Thoracic Society (ATS) and European Respiratory Society (ERS) has developed a diagnostic criteria and schema for adult patients with IPF, and this criteria strongly recommends a multidisciplinary discussion between pulmonologists, radiologists and pathologists for an accurate diagnosis (Raghu et al., 2011; Travis et al., 2013). However, a successful classification and quantification tool that allows recognition of disease consistently across radiology, pulmonary and pathology disciplines still remains difficult.

The complex appearances of IPF abnormalities that keep changing in extent over time is difficult to assess by traditional methods. Traditional radiological observation to distinguish disease patterns is tedious and not reproducible, and this manual evaluation is not consistent due to variation of inter- and intra- assessment (Flaherty et al., 2007; Watadani et al., 2013). Specifically, the difference in perception and interpretation of visual features of disease, which is associated with the experience and skills of clinical doctors, may lead to variable description of the same patient or even cause "reader error". However, this "error" can not be fully solved by training or improvement of imaging technologies (Kundel, 2006; Bartholmai et al., 2013). More importantly,

the final decision of clinical diagnosis is often based on independent evaluation from the radiologist, clinician and pathologist, which makes it hard to ensure consistency and dependability of results (Flaherty et al., 2004; Sverzellati et al., 2011). Another challenge of diagnosing IPF is the clinical problem of how to consistently detect and discriminate IPF from other idiopathic interstitial pneumonias (IIPs). These distinctive diseases usually have similar clinical presentations or indeterminate pathologic and radiographic appearances. Some cases may even have mixed restrictive/fibrotic and destructive/obstructive processes (Bartholmai et al., 2013). For example, non-specific interstitial pneumonia (NSIP), a pathological subtype of IIPs, appears to behave similarly to those with IPF/UIP patterns, especially for the cases with coexisting UIP and fibrotic NSIP patterns (Monaghan et al., 2004; Flaherty et al., 2001). All of these various IIPs have distinctly different prognosis, and specific therapy that is targeted to a particular pathological process is becoming necessary (Lynch et al., 2005). Therefore, non-IPF IIPs must be discriminated from IPF (Bjoraker et al., 1998).

1.1.2 Advantages of quantitative analysis using HRCT

Recent development in radiological imaging techniques offers exciting opportunities to develop radiological patient-specific biomarkers as important indicators of specific phenotypes (Devaraj, 2014; Gotway et al., 2007). HRCT has played an essential role in evaluating lung disease through recognizing visual patterns and features of disease regions such as ground-glass opacities, reticular patterns and honeycombing (Mueller-Mang et al., 2007). HRCT is also a useful diagnostic tool to differentiate between IPF and other pathologies. It is generally believed that the extent of visual lesion present on HRCT strongly relates to the severity of pathological abnormalities, and therefore can

be used to monitor the progression of disease and then response to therapy (Kazerooni et al., 1997; Kim et al., 1999; Wells et al., 2003; Saketkoo et al., 2011). In addition, it has been noted that the use of HRCT can actually decrease the need for surgical lung biopsy which is risky for older patients with comorbidities (Bartholmai et al., 2013). As a non-invasive tool for visualizing abnormal parenchymal densities, HRCT has its own advantage in IPF diagnosis even for the cases where HRCT fails to show enough specific features to reflect typical UIP pattern, since HRCT can provide guidance for optimizing the site to obtain a surgical lung biopsy (Kazerooni, 2001; Diette et al., 2005; Misumi and Lynch, 2006; Costabel et al., 2007).

Currently, how the changes of disease can be consistently characterized and quantified over time and how these changes can predict disease progression are still challenging tasks. Manual classification and subjective evaluation are usually complicated and not accurate enough. Image-based quantitative analysis is therefore strongly needed for developing a robust and consistent IPF assessment system (Gotway et al., 2007; Lynch et al., 2005). A number of studies have indicated that the quantification of abnormalities on thoracic HRCT has potential to determine the extent of disease. Moreover, it can help with stratifying different types of disease with numerous imaging features in variable distribution (Best et al., 2008; Wells et al., 2003; Sumikawa et al., 2008; Bartholmai et al., 2013). The hypothesis is that through quantifying abnormalities from radiological images, robust objective bio-markers can be developed to drive a patient-specific prediction based on specific phenotypes. This will be important for facilitating individualized clinical management and identifying specific phenotypes linked to clinical disease presentations or therapeutic responses (Raghunath et al., 2014).

1.1.3 Review of current published methods of quantitative analysis of lung disease

In the past few years, there has been considerable effort to provide quantitative analysis of lung parenchymal abnormalities on HRCT scans. Fortunately, quantitative methods to analyse disease patterns of chronic obstructive pulmonary disease (COPD) such as emphysema have been well developed over the last 20 years, and have been able to provide reproducible bio-markers used in sub-clinical diagnosis and assessment (da Silva Felix et al., 2008; Gietema et al., 2011; Galbán et al., 2012; Wang et al., 2013; Castaldi et al., 2013). However, quantification of disease patterns of lung fibrosis or other interstitial lung disease is more challenging, since the appearances and changes of these abnormalities are even more complicated than the characteristics seen with emphysema (Lynch, 2007; Delorme et al., 1997; Galbán et al., 2012; Depeursinge et al., 2010). In the early days, first-order textural analysis methods such as mean lung density (MLD) and histogram analysis (HIST) were used to analyse radiological lung imaging (Gilman et al., 1983; Gould et al., 1988; Müller et al., 1988; Kinsella et al., 1990; Knudson et al., 1991; Behr et al., 1992). These approaches to assess CT data are automatic and objective, but they simply examined a single parameter for the whole lung. The measurement of attenuation (which relates to density) is highly dependent upon lung volume and can be significantly affected by beam-hardening effects, scatter and drifts in scanner calibration. Therefore, these methods are difficult to apply in the presence of mixed disease.

Uppaluri et al. (1999b,a) was an early group to present a multiple feature based method to automatically quantify and classify pulmonary parenchyma of interstitial lung disease based upon HRCT. An adaptive multiple feature method (AMFM) which combined statistical texture measures with a fractal measure was developed to initially

assess the emphysema regions of the lung, and subsequently extended to study subjects with IPF or sarcoidosis. AMFM was further improved to assess as many as 22 independent texture features which enabled classification of pulmonary parenchymas into six tissue patterns, including: honeycombing, ground-glass, broncho-vascular, nodular, emphysema-like, and normal. These 22 texture features consist of statistical features (grey level distribution features, run-length features and co-occurrence matrix features) and fractal features (geometric fractal dimension and stochastic fractal dimension). The lung slices were divided regionally into 31×31 pixel regions of interest (ROI). In each ROI, an optimal subset of texture features was evaluated to determine which of the six patterns in the region could be characterized. A non-linear statistical classifier, the Bayesian classifier, was built to do the classification through calculating the probability that the ROI belongs to each tissue pattern. The whole algorithm contains two stages: the first stage involves training AMFM to recognize different HRCT based tissue patterns using a preselected dataset; the second stage involves a new CT data to be analysed using the above introduced method. This multiple textural feature based method established the foundation and framework for subsequent studies on quantitative assessment of lung imaging. Xu et al. (2006b,a) enhanced the ability of AMFM based on the work of Uppaluri et al. (1999b,a). In this study, the 2D textural feature-based tissue analysis method was extended to 3D space for quantifying emphysema and early smoking related lung pathologies. The extracted features involved both first-order features (mean, variance, kurtosis and entropy of grey level distribution) and second-order features (run-length and co-occurrence measurements). The results showed that 3D AMFM analysis of lung parenchyma improved discrimination compared to 2D AMFM of the same volumes of interest (VOIs).

In the past 15 years, AMFM has been widely used to quantify and classify lung dis-

ease from CT imaging (Van Ginneken et al., 2002; Chabat et al., 2003; Best et al., 2003; Uchiyama et al., 2003; Kim et al., 2005; Zavaletta et al., 2006; Arzhaeva et al., 2007; Best et al., 2008; Kim et al., 2010, 2011, 2015). A general workflow of texture analysis is: 1. preset several types of tissue patterns; 2. divide lung slices into a set of ROIs with a square shape; 3. use texture analysis to extract multiple features for each ROI; 4. select representative expert-labelled ROIs as a training set for a classifier; 5. use the features of ROIs from a new set as input to the classifier to find the corresponding tissue pattern. Currently, most of the published CT based lung tissue analysis methods were developed based on this framework. Different textural features, ROI sizes or classifier types were specifically selected to target different lung disease or clinical requirement. Best et al. (2003, 2008) used mean lung attenuation (MLA), skewness (asymmetry) and kurtosis (peakedness) as features combined with a further univariate and multiple correlation and regression statistical analysis to determine relationships between histogram signals and results of PFTs in patients with IPF. The result showed that CT histograms of the lungs were correlated with PFT measurements, therefore the visual disease extent on CT images can be used as a strong independent predictor of mortality in IPF. Chabat et al. (2003) used similar textural features and classifier to the method of Best et al. (2003, 2008) for differentiating centrilobular emphysema, panlobular emphysema, constrictive obliterative bronchiolitis and normal lung tissue. Zavaletta et al. (2007) trained and tested three classifiers which included 10 Nearest Neighbour Classifier, Fisher Linear Discriminant Analysis, and Parzen Window to classify normal and abnormal structures in lungs with IPF.

Uchiyama et al. (2003) and Kim et al. (2005) employed multi-layered artificial neural networks (ANNs) with a back-propagation algorithm as a classifier to distinguish between different tissue patterns which includes both normal and diffuse lung disease

slices. ANNs, using their simplest definition, are the modelling of the human brain, and their building blocks are neurons. They are excellent tools for finding patterns which are too complex or numerous for a human programmer to extract and teach the classifier to recognize, therefore will increase the accuracy of classification. In addition, Kim et al. (2010, 2011, 2015) published a series of papers presenting a texture-based computer-aided diagnosis (CAD) scoring system to assess quantitative lung fibrosis (QLF) as a measurement of lung disease severity and as a surrogate imaging marker. The QLF score (a texture feature-based measure) was compared to the CT histogram metric (a global statistical measure), and the baseline severity and early change within 7 months in patients with IPF was assessed. The result concluded that classifier-model-derived scores (QLF scores) were associated with baseline disease extent and were also a sensitive measure of change over time, and a QLF score could be used for measuring the extent of disease severity and longitudinal changes.

Most published methods focus on texture-based classification of lung parenchyma or the severity and volumetric quantification of disease for the whole lung (such as the QLF score). Currently, very few studies attempt to quantitatively characterize the spatial distribution of each disease pattern or describe how each tissue pattern changes or converts over time. Moreover, many of the existing methods are time consuming and sometimes even require several hours. These real-world limitations present difficulties when translating these techniques into clinical applications (Bartholmai et al., 2013).

In addition, it has been generally believed that there is a decrease in lung volume (both FRC and TLC) in patients with IPF, but up to now, few studies have been presented to explore the lung and lobe shape alteration in IPF lungs compared to normal ones. It is a reasonable assumption that shape changes will be observed in IPF lungs due to physiologic alterations and disease progression over time.

1.1.4 Computer-Aided Lung Informatics for Pathology Evaluation and Ratings

CALIPER (Computer-Aided Lung Informatics for Pathology Evaluation and Ratings) is a computational image analysis platform developed by the Biomedical Imaging Resource Laboratory at the Mayo Clinic (Rochester, MN, USA) for the characterization and classification of lung parenchymal findings on HRCT (Maldonado et al., 2013; Bartholmai et al., 2013; Raghunath et al., 2014). The data processing step includes lung segmentation and classification of the remaining pulmonary parenchyma on the CT dataset. Briefly, CALIPER isolates the lung parenchyma by extracting central airways and vascular structures, and then classifies every parenchymal voxel into the following characteristic CT patterns: normal (N), reticular (R), honeycomb (HC), ground-glass (GG), mild low attenuation areas (LAA), moderate LAA, and severe LAA (including emphysema). This novel computer-aided method for analysing pulmonary tissue features provides a consistent and reproducible quantification of lung disease that relates to the semi-quantitative assessment from radiologists(Maldonado et al., 2013). Figure 1.1 shows visual appearance of each characteristic CT pattern.

Pre-processing is conducted before the eventual classification of the pulmonary parenchyma. This required segmentation of anatomic lung regions. The lungs are initially segmented using an adaptive density-based morphology (thresholding) method (Hu et al., 2001). Airways are segmented by thresholding combined with 3D region growing algorithm and vessels are segmented by an enhancement filter based on Hessian matrix (Sato et al., 2000), then the final lung segments are extracted.

The volumetric detection and classification of pulmonary parenchyma by CALIPER uses a sliding window supervised classification scheme based on histogram signature

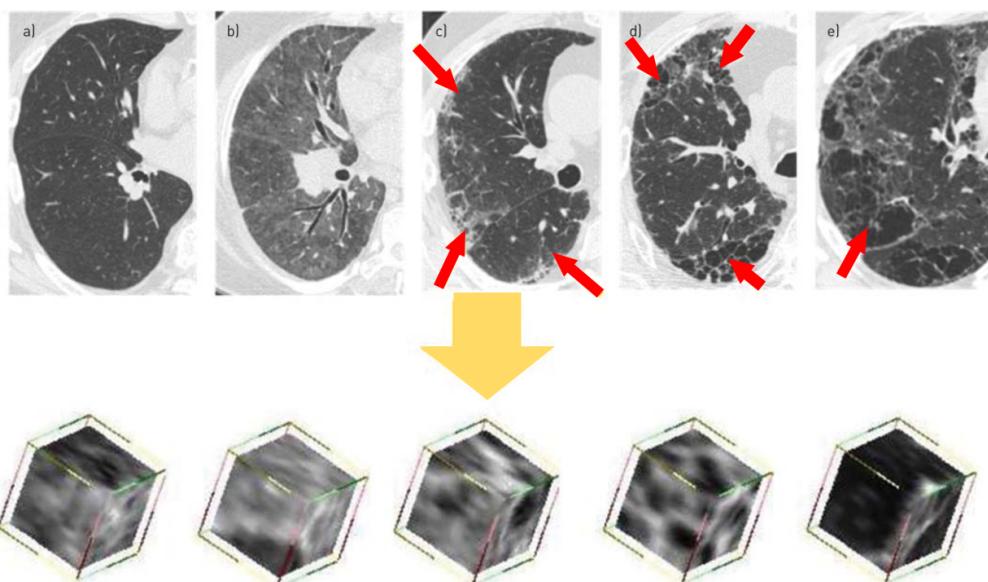


Figure 1.1: Computed tomography images demonstrating appearance and various visual manifestations of idiopathic pulmonary fibrosis: a) normal, b) ground glass, c) reticular changes (arrows), d) honeycombing (arrows) and e) emphysema (arrow). In training datasets, the consensus of four thoracic radiologists was used to identify multiple volumes of interest (VOIs) corresponding to normal, ground-glass density, reticular abnormalities, honeycombing and emphysema. Reproduced from (Maldonado et al., 2013).

mapping techniques (Zavaletta et al., 2007). This classification technique was trained by expert radiologists via consensus assessment of pathologically confirmed datasets, which were obtained from the Lung Tissue Research Consortium (LTRC). LTRC is a resource program sponsored by the NIH/NHLBI that provides clinical and physiologic data of human lung tissues to qualified investigators for use in their research and to help investigators develop a better understanding of lung disease. The central part of the classification scheme is the selection of a set of expert-labelled volumes of interest (VOIs) as the training data for a classifier. The training data used in CALIPER comprises $15 \times 15 \times 15$ -voxel VOIs acquired from HRCT scans of subjects with proven pathological diagnosis of interstitial lung disease (ILD) or emphysema from the LTRC repository. These VOIs were selected from CT scans through independent analysis by four experienced thoracic radiologists, with instructions and criterion to determine if the visual appearance should represent normal, emphysema or one of the characteristic lung fibrosis CT patterns: honeycomb, reticular or ground-glass (Maldonado et al., 2013; Bartholmai et al., 2013).

The VOIs with agreement by all four radiologists on the class of abnormality were used as exemplars to determine canonical histogram signatures of the CT patterns of visual abnormality by automatic cluster affinity techniques. Quantitative discriminability of a series of pairwise dissimilarity metrics based on the VOI histograms was tested using multi-dimensional scaling (MDS). The Cramer Von Mises Distance (CVM), which was found to be most consistent with the expert groupings, was selected as the dissimilarity metric to train CALIPER. For each of the parenchymal voxels needing to be classified, the local histograms of its neighbouring $15 \times 15 \times 15$ voxels were compared against the histograms of the exemplars identified in the training phase. A CVM dissimilarity measure was used in the comparison and the fundamental type of the exemplar

(N,R,H,G or emphysema) with the lowest CVM was assigned as the parenchymal CT pattern to this classified voxel. The parenchymal voxels identified as vessel structures were classified as normal pattern. Figure 1.2 shows a representative dataset with axial, coronal and sagittal sections of a CT lung volume where every voxel of the parenchyma is characterized and colour coded into one of the parenchymal patterns (N, R, H, G, and mild, moderate and severe LAA).

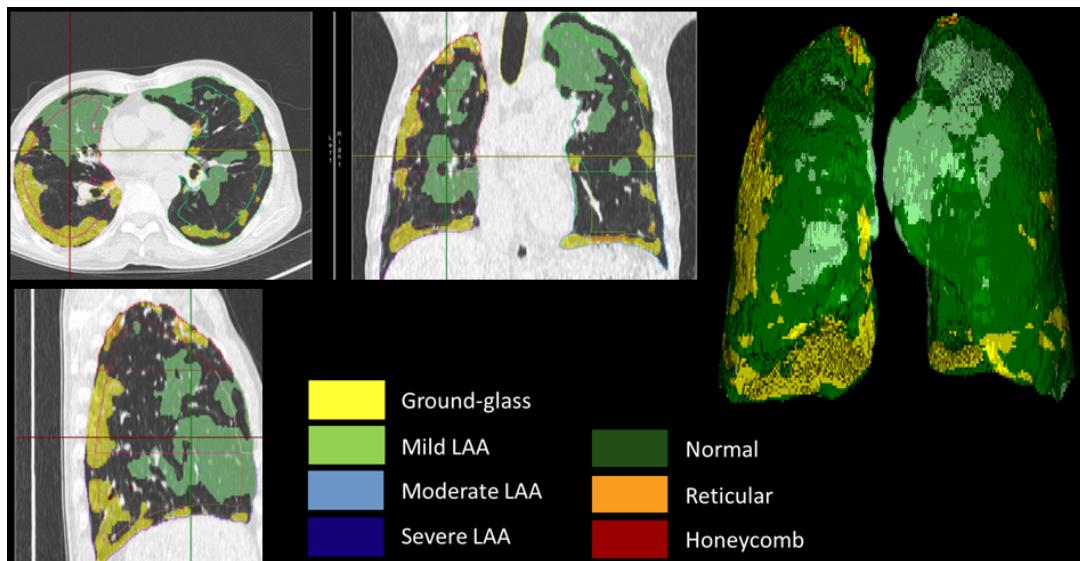


Figure 1.2: Color labelled classification result from CALIPER of one subject diagnosed with IPF. (a) Transverse plane. (b) Coronal plane. (c) Sagittal plane. (d) 3D color labelled lung.

1.2 Methods: quantitative analysis of IPF lungs

This section describes the quantitative methods used in this chapter to analyse and characterize IPF tissue abnormalities and lung lobe shape, longitudinally and in comparison to normal subjects. In summary, HRCT imaging was classified by pattern using a validated image analysis process (CALIPER) (Maldonado et al., 2013; Bartholmai et al.,

2013; Raghunath et al., 2014). The tissue classification data was mapped to a mean statistical shape model (SSM) allowing a quantitative approach to analyse tissue pattern density, tissue pattern volume, spatial distribution of abnormalities, and regional changes in tissue abnormalities over time. In the shape analysis, lobar finite element meshes for both IPF and normal subjects were projected to the SSM, and lobe shape differences in IPF were then quantitatively characterized.

1.2.1 Tissue classification of IPF lungs

Imaging and clinical data

Data used in this study was acquired as part of routine clinical diagnosis or follow up. Data use was approved by the Southern Health and Disability Ethics Committee. Data include volumetric HRCT and PFTs from 13 patients who were diagnosed with or suspected to have IPF. All patients were under clinical care at Auckland City Hospital, Auckland, New Zealand. Volumetric HRCT images (slice thickness 1.25-5.00 mm) were acquired at the end of inspiration during routine diagnostic inspection and/or monitoring for IPF disease. Eleven of the subjects had more than one serial CT scan within a 5-79 month interval, representing different time points (7 subjects had 3 time points, 4 subjects had 2 time points). The population demographics for these subjects is shown in Table 1.1.

Table 1.1: Demographic data.

Description	
Age (years)	43-83
Females/Males	3/10
Slice thickness	1.25-5.00 mm
Scan month interval	5-79 month
Slice resolution	512 × 512 mm
Number of slices	68-227

Normalization of classified data

Tissue CT patterns for each patient at each time point were classified using the CALIPER software introduced in Section 1.1.4. Then, lung surface data and fissure surface data were acquired using the lobe segmentation method introduced in Chapter 3. A bi-cubic Hermite finite element surface mesh was fitted to the shape of the lung and its fissures via a least squares fit using the CMISS software package (<https://www.cmiss.org>). The details for the generation of lobe data and the lobe mesh are given in Chapter 3, Section ??.

There is lung shape variation between different subjects and often between clinical images obtained at different times, as well as variation in the extent to which a patient inhales during imaging, even with careful training. Thus, the classified volumetric lung data was then mapped to a statistical shape model (SSM) of the "normal" older human lung to provide a consistent mapping of tissue abnormalities in individuals to a consistent lung shape. The steps for the construction of the SSM were introduced in Chapter 3, Section ??, which described an SSM for a cohort aged 21-83. In this chapter, data from 35 normal subjects aged 50 years and older were used to derive an SSM because this is consistent with the typical age of onset of IPF. The SSM used for mapping data is

the average mesh of the lung lobes which was derived from these 35 training subjects, and it provides a description of a statistical mean lung and fissure surface shape for a cohort of normal adults aged > 50 years.

In order to map the individual classified data to the SSM mesh, the bi-cubic Hermite finite element mesh of the lung surface was converted to a tri-cubic Hermite volumetric mesh which describes not only the lung surface but also the internal anatomy (Tawhai and Burrowes, 2003). The volumetric mesh has 40 nodes and 30 elements for left lung, and 56 nodes and 38 elements for right lung. Each node has 24 DOFs which store the global coordinates (x , y and z) and the first, second and third nodal derivatives ($\frac{\partial n}{\partial \xi_1}$, $\frac{\partial n}{\partial \xi_2}$, $\frac{\partial n}{\partial \xi_3}$, $\frac{\partial^2 n}{\partial \xi_1 \xi_2}$, $\frac{\partial^2 n}{\partial \xi_2 \xi_3}$, $\frac{\partial^2 n}{\partial \xi_1 \xi_3}$ and $\frac{\partial^3 n}{\partial \xi_1 \xi_2 \xi_3}$), where n is x , y and z , and ξ is the local element coordinate.

During the data mapping, all of the classified data should be completely enclosed inside its lobe volume mesh. The position of each point within the finite element mesh was defined locally in each element of the mesh by ξ_i , for $i=1,\dots,3$ with $0 < \xi_i < 1$. The ξ_i location denotes the local coordinates of the data point with respect to its element. The local coordinate ξ_i was then used to calculate the global coordinates of the mapped data points using

$$u(\xi_i) = \sum_{n=1}^N \psi_n(\xi_i) u_n, \quad (1.1)$$

where u_n is a vector of N element nodal parameters of the SSM lobe mesh associated with the interpolation functions ψ_n .

In order to force a uniform data point distribution throughout each lung, the gaps in the mapped data caused by the mapping deformation from individual shape to SSM (shown in Figure 1.3a) were "filled" by matching the classification of their closest

neighbour point among the classified data. Uniform data point distribution facilitates further density and spatial distribution analysis of abnormalities. Briefly, the gaps in the mapped data were filled using the following steps:

1. Mapped data were cut into a series of axial slices (shown in Figure 1.3a).
2. The lung mesh of average SSM was used as a mask to define the lung boundary. (shown in Figure 1.3b).
3. Morphological operations were applied to smooth the lung boundary, then a lung mask was generated (shown in Figure 1.3c).
4. The gaps enclosed inside the lung mask were filled with the CT pattern color of its closest point among the classified data (shown in Figure 1.3d).

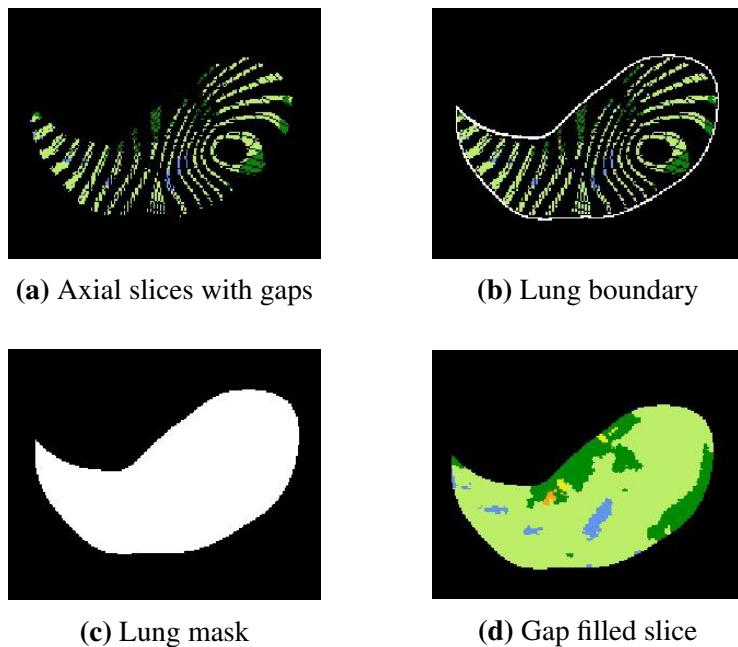


Figure 1.3: Diagram of gap filling steps within the mapped data. (a) Get axial slices of mapped data (with gaps). (b) Define lung boundary (SSM defined). (c) Get lung mask. (d) Fill the gaps within lung mask.

1.2.2 Tissue quantification of IPF lungs

Density analysis

The average density value of each classified CT pattern was calculated. In a typical CT image, intensity is measured in Hounsfield units (HU) which corresponds linearly to the actual density of the imaged tissue. HU was calculated with the segmentation software PTK calibrated to values of approximately -1024 for air density, zero for water density, and over 40 for blood, bone, and other non-parenchymal tissue. The tissue density (ρ , g/cm^3) was then acquired at each voxel using

$$\rho = \frac{HU}{1024} + 1. \quad (1.2)$$

The average density of each CT pattern was then calculated from individual voxel density.

Spatial distribution analysis

Based on the criteria of IPF defined by the ATS and ERS, the diagnosis of IPF is usually associated with the presence of a UIP pattern in HRCT (see details in Chapter 2, Section ??). The distribution of UIP on HRCT is characteristically basal and peripheral (subpleural), though often patchy. Therefore, in order to quantitatively analyse the spatial distribution of IPF abnormalities, the percentage of honeycomb, reticular, emphysema and ground-glass which represent typical UIP disease patterns on HRCT were calculated in basal-to apical sections, dorso-to-ventral sections, from subpleural to internal, and by lobe:

Basal-to-apical: In the direction from base to apex, the volume percentage of each disease region was averaged in bins representing 5% of lung height (along the crano-caudal axis).

Dorso-to-ventral: In the direction from posterior to anterior, the volume percentage of each disease region was averaged in bins representing 5% of distance along the crano-caudal axis.

Subpleural to internal: The distance from the abnormalities to the boundary of the lung and to the centre of the lung were measured to analyse the location of disease with respect to the lung surface. To be specific, the centre location of each connected cluster of disease area was firstly extracted, and the subpleural to internal distance percentage $R_{subpleural}$ of each connected cluster, which described how far the connected cluster was from the lung surface, was then calculated as Figure 1.4

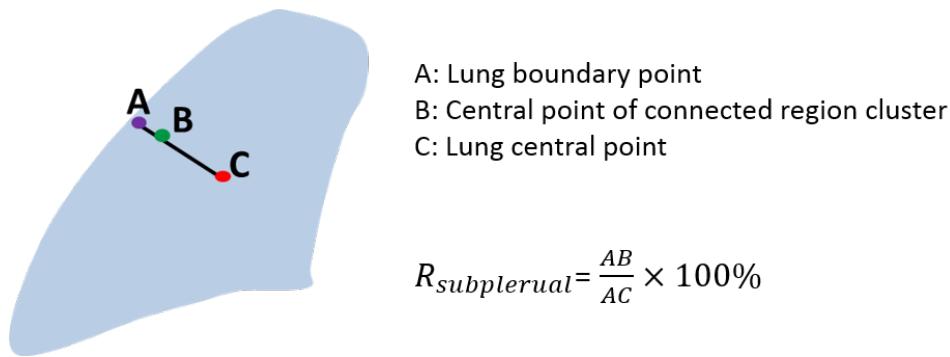


Figure 1.4: Diagram of subpleural-to-internal percentage measurement.

The coordinate of the lung central point $P_{centre}(x, y, z)$ was calculated as

$$P_{centre}(x, y, z) = \frac{\sum_{i=1}^N P_{surface}(x, y, z)}{N}, \quad (1.3)$$

where $P_{surface}(x, y, z)$ is the coordinate of the lung surface data point, and N is the number of lung surface data points.

Lobar distribution: In order to analyse the lobar distribution of disease, the volume percentage of each disease CT pattern located in each lobe was calculated.

Change in classification of tissue pattern over time

Median survival time of patients with IPF is generally from 3 to 5 years from time of diagnosis. However, individual progression of disease is variable and how the characteristic disease pattern changes over time (e.g. whether a disease region deteriorates and changes to other tissue patterns or stays the same over time) still remains elusive. In this study, the classified data of all subjects and time points have been normalized to a standard lung shape (SSM, as introduced in Section 1.2.1), thus making it possible to detect the disease change in specific regions over time, by extracting the CT pattern of each voxel at different time points.

1.2.3 Volume analysis of IPF lungs

The change of whole lung volume over time

The volume of each classified tissue pattern in both left and right lung was calculated using the following equation:

$$V_{Region} = N \times R_x \times R_y \times R_z \quad (1.4)$$

where N is the number of voxels of each tissue pattern, R_x , R_y are the x, y resolution of the CT scan, and R_z is the thickness of the CT scan. Then the whole volume of left and

right lung was calculated as the sum of the volume of each CT pattern.

Lobe volume difference between old normal lungs and IPF lungs

The lobe volume of the IPF cohort was compared with the older normal cohort described in Section 1.2.1. In order to quantitatively analyse the lobe volume difference between the two groups, the volume proportion of each lobe was calculated as:

$$P_i = \frac{V_i}{\sum_{i=1}^5 V_i}, \quad (1.5)$$

where V_i is the volume of each lobe with $i=1$ corresponding to left lower lobe, $i=2$ corresponding to left upper lobe, $i=3$ corresponding to right lower lobe, $i=4$ corresponding to right middle lobe and $i=5$ corresponding to right upper lobe. The differences in lobe volume proportions between IPF subjects and normal subjects were compared statistically using a t-test. In order to further compare lobe volumes between these two groups, the average lobe volume proportion among IPF subjects and among the normal subjects were then calculated.

1.2.4 SSM based shape analysis of IPF lungs

The SSM was used to quantitatively analyse the alterations in lung lobe shapes of patients with IPF. As previously described in Section 1.2.1, 35 normal subjects aged > 50 years were used as training data to construct the SSM which contained both lung surface and fissure surface. PCA was used to decompose the shape variation of the lung lobe into a set of modes, and each mode represented one type of lung and fissure surface shape variation. Thus, each lung lobe shape was described by a linear combination of the mode vector and its corresponding weight by

$$S_{new} = S_{mean} + \sum_{i=1}^L \mathbf{u}_i w_i, \quad (1.6)$$

where S_{mean} is the average lobe shape model across all the training subjects, $\mathbf{u}_i (i = 1, 2, \dots, L, L = 34 \text{ in this study})$ is the mode vector of shape variation which corresponds to the i^{th} largest principal component from PCA, and w_i is a weight factor given to each mode of variation.

The lung lobe FE mesh of each IPF subject was then procrustes projected on to the average SSM after alignment to the reference model (details can be seen in Chapter 3, Section ??) The new weight values of all shape modes $w_{new} = [w_{new1}, w_{new2}, \dots, w_{newL}]$ ($L = 35$) were calculated from the projection. These mode weights can be used as quantitative indexes to analyze and compare the shape variation and difference between IPF and the control group.

Shape difference between IPF lungs and normal control lungs

For the SSM of the control group, the first three shape modes explained over 30% of the total variation in the training set. Therefore, the weight values of the first three modes were used as the measurement to compare the shape difference of lung lobe between controls and IPF. The p-values of the first three shape modes between the two groups were calculated using a t-test.

Relationship between lung lobe shape and fibrosis and low attenuation area extent

In order to quantitatively investigate the association between fibrosis extent and lung shape variation, the association of the first three mode weightings with the overall volume percentage of fibrosis and low attenuation area (LAA) was estimated using linear

regression. Total fibrosis extent was estimated as the sum of reticular, honeycomb and ground-glass opacification of both lungs, and total LAA extent was estimated as the sum of mildLAA, moderateLAA and severeLAA of both lungs. The behaviour of the first three modes with respect to overall fibrosis percentage and LAA percentage was analysed using linear regression.

1.3 Results

1.3.1 Normalization of classified data

The data classified using the CALIPER software were mapped to the SSM using the method introduced in 1.2.1. Figure 1.5 shows the mapped classification data for a single subject. The mapped data for the other subjects can be found in Appendix ??.

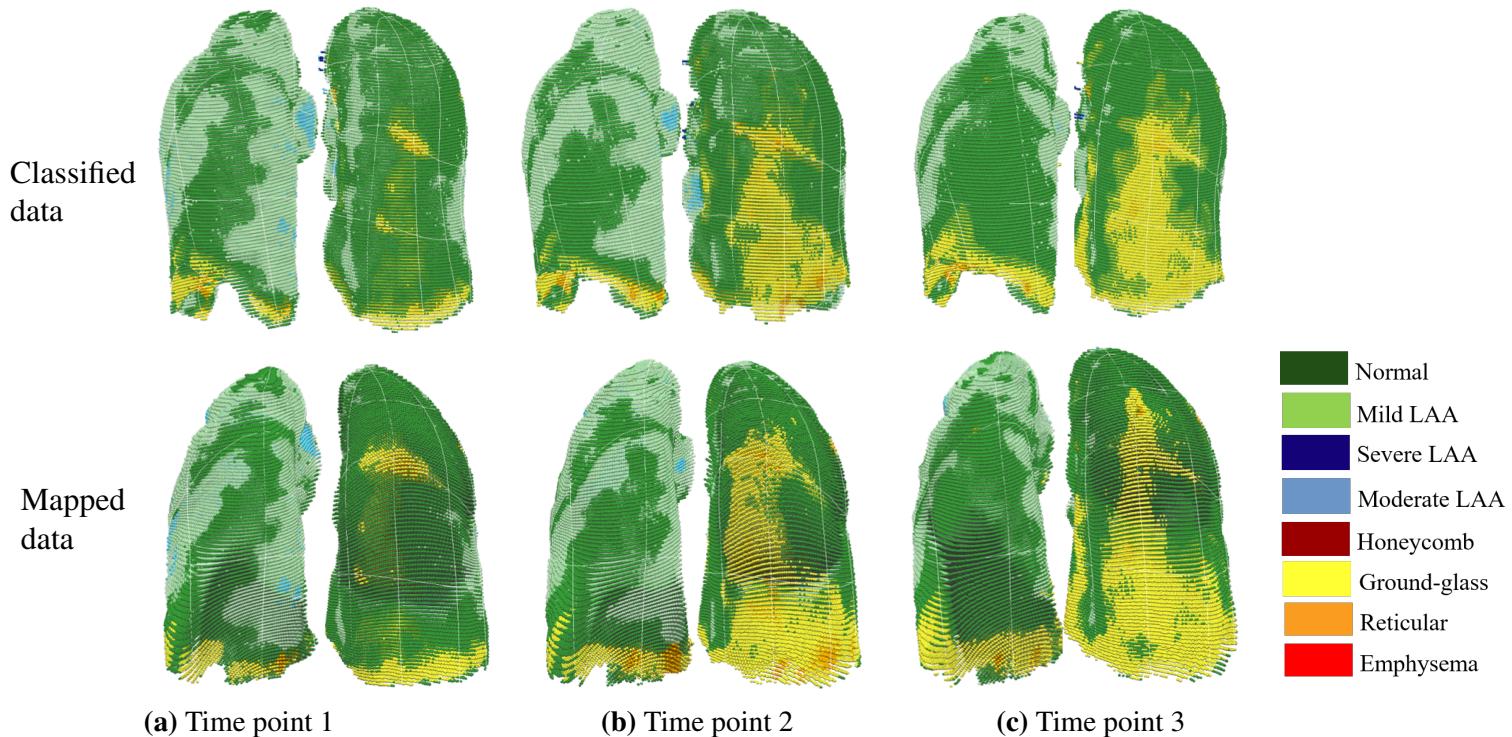


Figure 1.5: Classified data (top row) and mapped data (bottom row) for three time points from one subject diagnosed with IPF. (a) The first time point, scan time: 0 months. (b) The second time point, scan time: 15 months. (c) The third time point, scan time: 20 months.

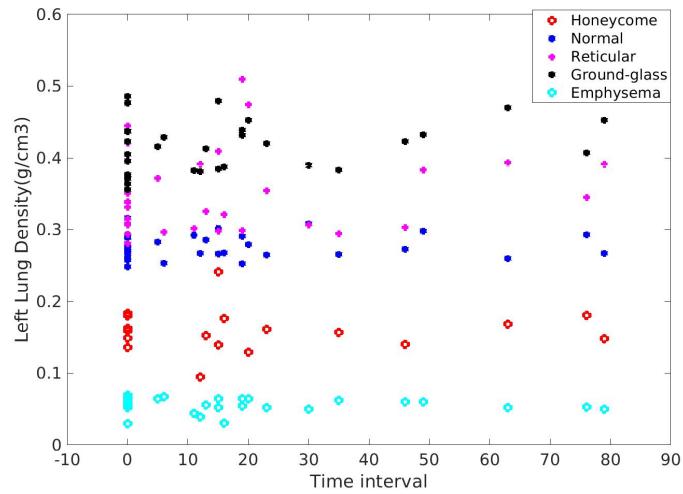
1.3.2 Tissue quantification of IPF lungs

Density analysis

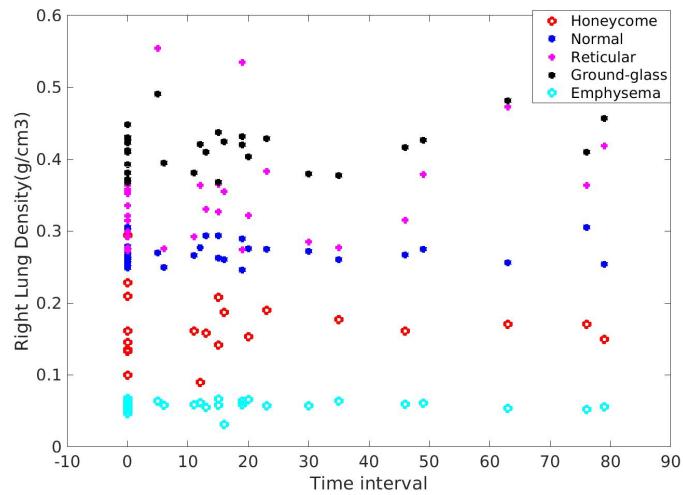
The density analysis result is shown in Figure 1.6 and Table 1.2. It can be seen that the average density of each CT pattern remains consistent and fluctuates only within a specific range over time in both left and right lung. The ground-glass region has the highest average tissue density and emphysema has the lowest average tissue density.

Table 1.2: Mean tissue density of each CT pattern for left and right lung (mean value \pm standard deviation)

	Mean tissue density (left lung)(g/cm³)	Mean tissue density (right lung)(g/cm³)
Normal	0.280 \pm 0.022	0.271 \pm 0.015
Honeycomb	0.148 \pm 0.039	0.155 \pm 0.027
Reticular	0.355 \pm 0.054	0.321 \pm 0.039
Ground-glass	0.424 \pm 0.051	0.404 \pm 0.018
Emphysema	0.080 \pm 0.014	0.082 \pm 0.016



(a) Left lung tissue density



(b) Right lung tissue density

Figure 1.6: Average tissue density (g/cm^3) of each CT pattern in IPF lungs. Each data point represents the tissue density of one time point from one patient. X axis represents the month interval of scan time for each patient, and "0" represents the first scan for this patient. (a) Tissue density of left lung. (b) Tissue density of right lung.

Spatial distribution analysis

Basal-to-apical distribution Figure 1.7 shows the percentage distribution (the average volume percentage of all time points from all patients) against lung height (cranio-caudal axis) of four characteristic CT patterns: ground-glass, reticular, honeycomb and emphysema for left and right lung. It can be seen from the result that ground-glass mainly locates in the basal part of the lung. The percentage of ground-glass decreases gradually with increasing lung height, and the distribution of ground-glass is quite similar in left and right lung. In contrast, the percentage of emphysema trends toward increasing from lung base to apex. The reticular tissue is mainly located in the most basal and apical areas: it is much lower in the middle part of the lung. The distribution of honeycomb does not have a relationship with lung height.

Figure 1.8 shows the percentage distribution at month 0 (the average volume percentage of the first time point from all patients) against lung height (cranio-caudal axis) of four characteristic CT patterns (ground-glass, reticular, honeycomb and emphysema) in left and right lung at month 0. From the result, the cranio-caudal distribution of the four CT patterns at month 0 are quite similar to the total average distribution shown in Figure 1.7, but the absolute percentages of the abnormalities at month 0 are lower than the total average values. That is, as expected, the volume of disease increases over time as a whole in the IPF lung.

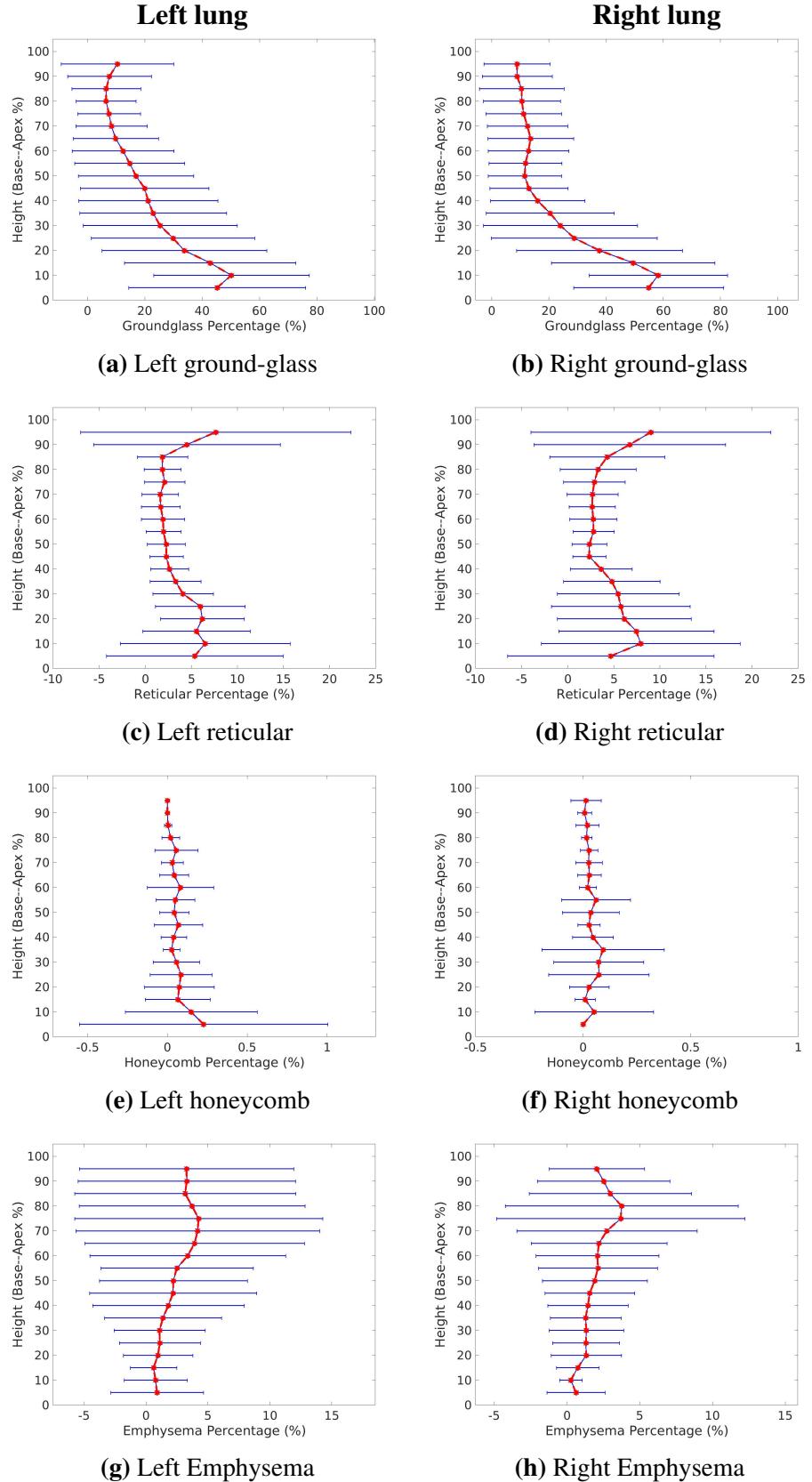


Figure 1.7: Volume percentage of each tissue classification plotted against lung height (cranio-caudal axis) in IPF left and right lungs. The average percentage was calculated within 5% sections of the lung height from the base to apex. Red shows the average value at each position across all patients, and blue shows the standard deviation. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

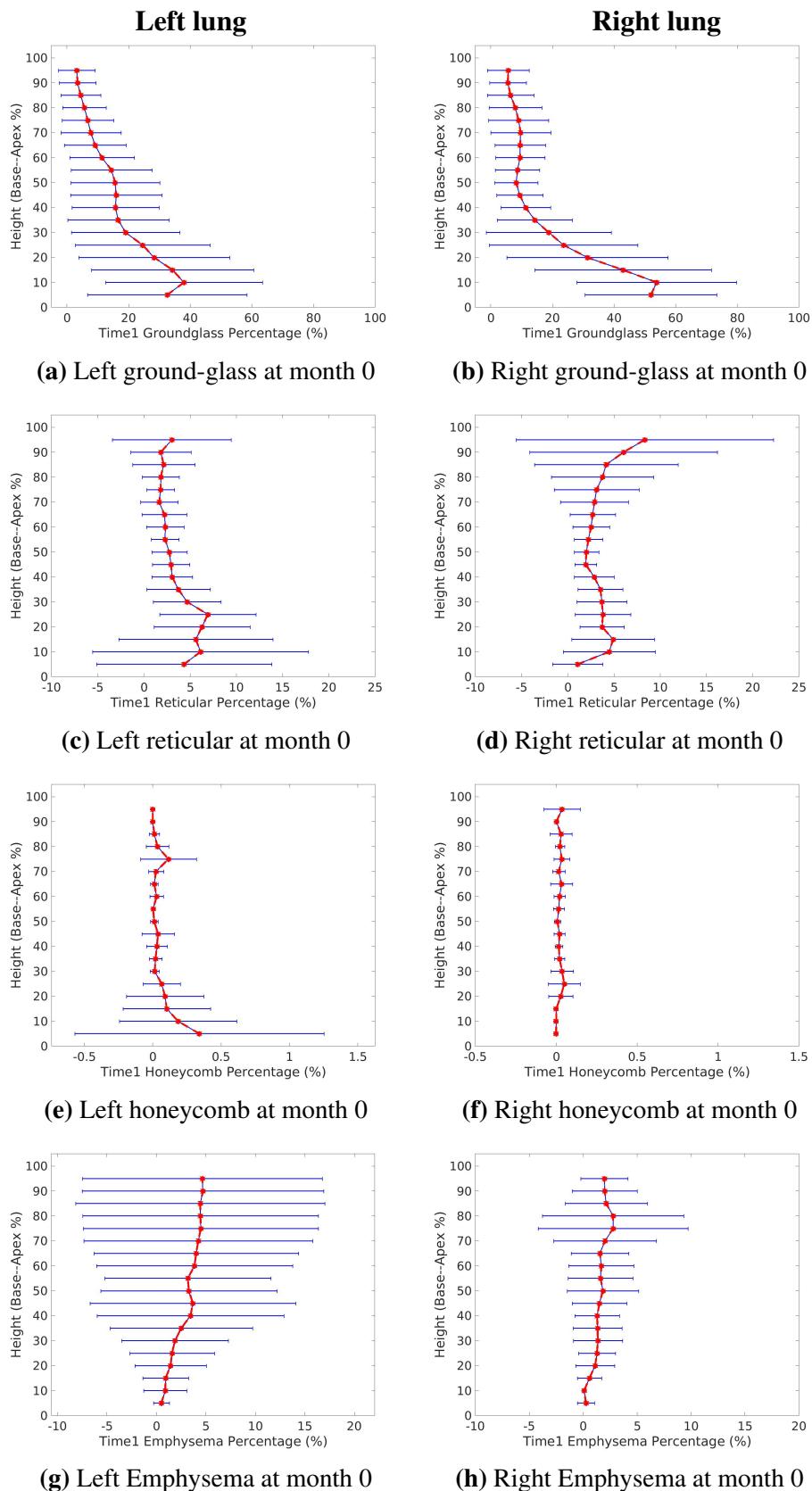


Figure 1.8: Volume percentage of each tissue classification plotted against lung height (cranio-caudal axis) in IPF left and right lungs at month 0. The average percentage was calculated within 5% sections of the lung height from the base to apex. Red shows the average value at each position across all patients, and blue shows the standard deviation. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

Figure 1.9 shows (for a single representative patient) the percentage distribution against lung height (cranio-caudal axis) of the four characteristic CT patterns in the left and right lung over time. From time point 1 to time point 3, an overall increase of the disease regions can be observed, even though the percentage of some tissue pattern may fall during this time. The distribution of ground-glass and reticular in base-to-apex does not change quantifiably over time, however, there are some fluctuations for honeycomb and emphysema (but the amount of honeycomb is very small). The percentage distribution against lung height over time for other patients can be found in Appendix ??.

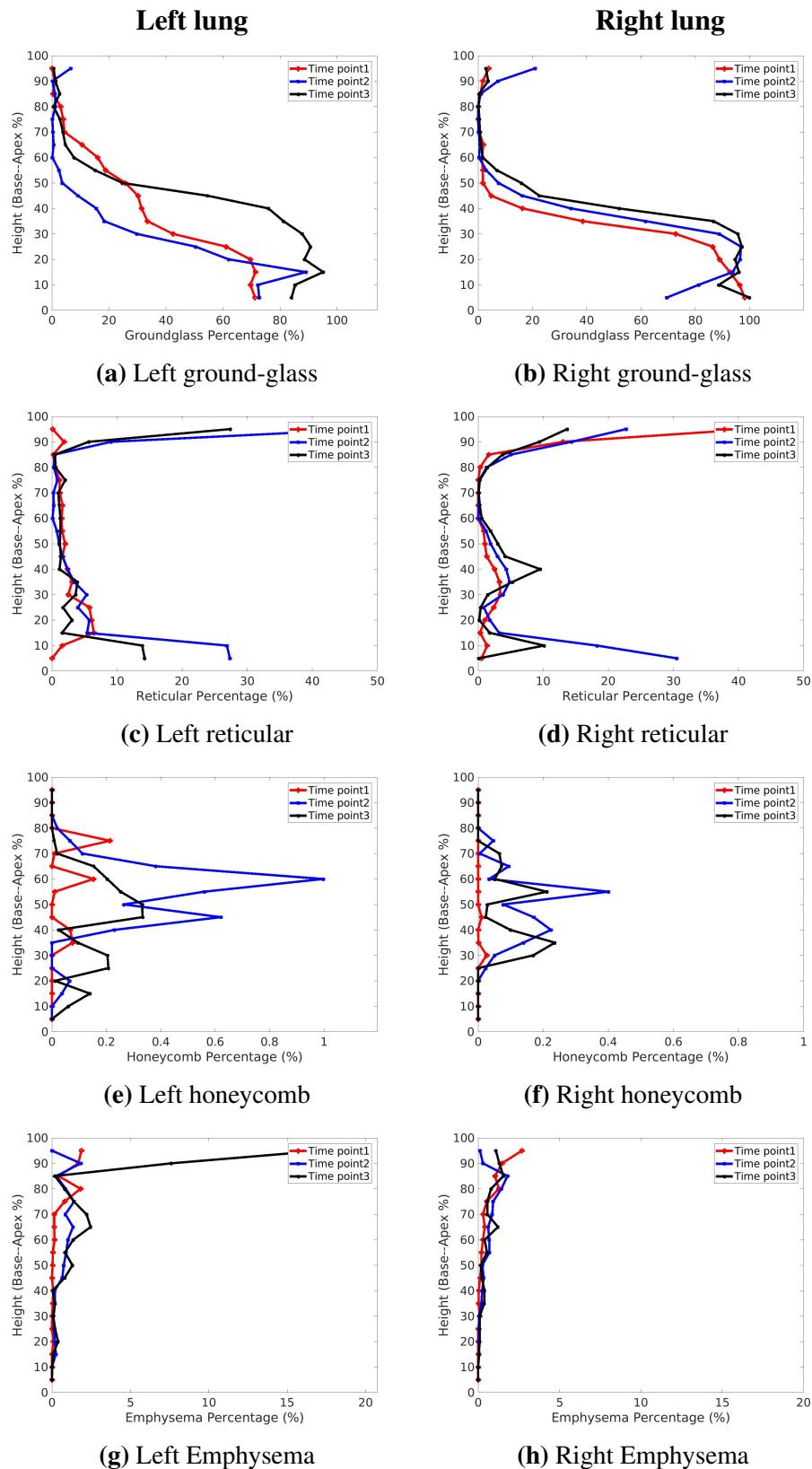


Figure 1.9: Volume percentage of each tissue classification plotted against lung height (cranio-caudal axis) of one patient diagnosed with IPF in left and right lungs over time. The average percentage was calculated within 5% sections of the lung height from the base to apex. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

Dorso-to-ventral distribution Figure 1.10 shows the percentage distribution (the average volume percentage of all time points from all patients) from posterior to anterior of the lung (dorso-ventral axis) for the four characteristic CT patterns for left and right lung. The results illustrate that the percentage of ground-glass and reticular keep decreasing from back to front, and they mostly locate in the dorsal part of the left and right lungs. Honeycomb and emphysema are both relatively constant along the dorso-ventral axis, but for the honeycomb pattern, there are relatively more abnormalities appearing in the dorsal and ventral regions compared to the middle area.

Figure 1.11 shows the percentage distribution (the average volume percentage of the first time point from all patients) from posterior to anterior of the lung (dorso-ventral axis) for the four characteristic CT patterns in left and right lung at month 0. It can be seen that, like the crano-caudal distribution, the distribution of the diseases in the dorso-ventral direction at the first time point are almost the same as the distribution in Figure 1.11, and the disease percentage at month 0 is slightly lower than the total time averaged value.

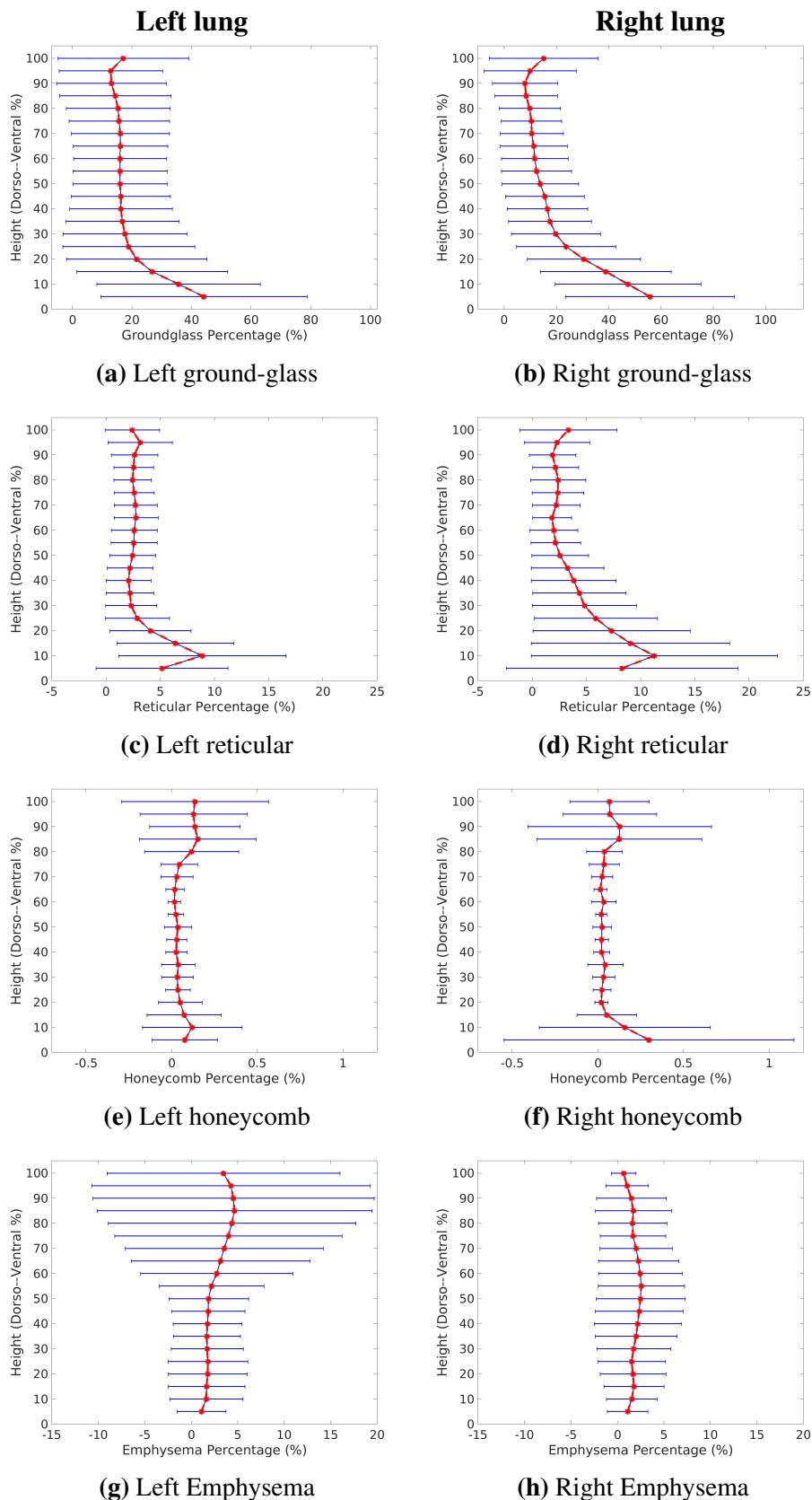


Figure 1.10: Volume percentage of each tissue classification plotted in the direction of dorso-ventral axis in IPF left and right lungs. The average percentage was calculated within 5% sections along the axis from posterior to anterior. Red shows the average value at each position across all patients, and blue shows the standard deviation. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

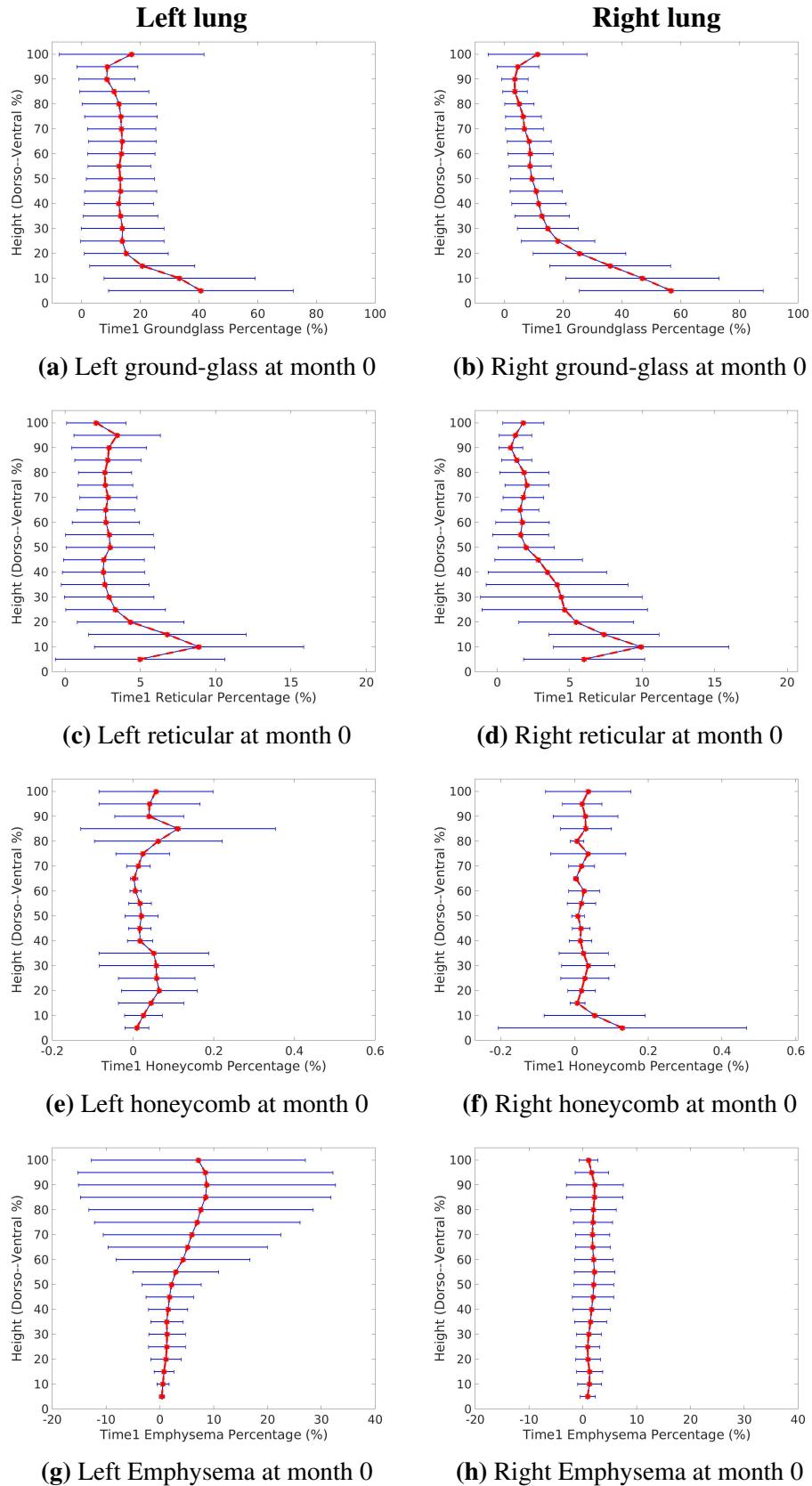


Figure 1.11: Volume percentage of each tissue classification plotted in the direction of the dorso-ventral axis in IPF left and right lungs at month 0. The average percentage was calculated within 5% sections along the axis from posterior to anterior. Red shows the average value at each position across all patients, and blue shows the standard deviation. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

Figure 1.12 shows the percentage distribution from the posterior to anterior of the lung (dorso-ventral axis) of the four characteristic CT patterns from an individual patient diagnosed with IPF in left and right lung over time. For ground-glass, emphysema and left honeycomb, an increase in the volume percentage over time can be seen and the dorso-ventral distribution does not change quantifiably during this period of time. However, the spatial location of reticular and honeycomb (in the right lungs) redistributes in the dorso-ventral direction over time, and no consistent trend is observed. The percentage distribution in this axis over time for all patients can be found in Appendix ??.

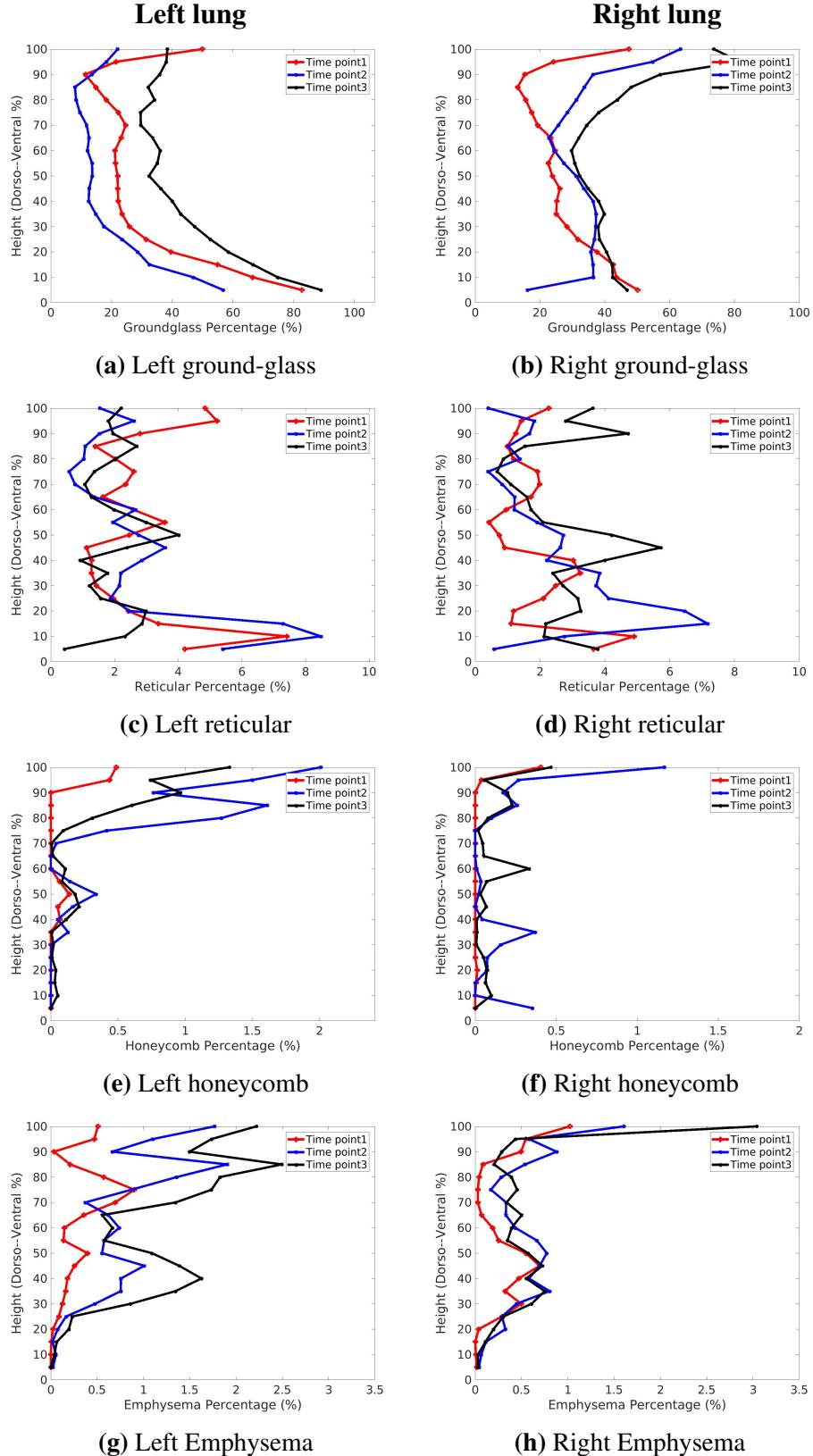


Figure 1.12: Volume percentage of each tissue classification plotted in the direction of dorso-ventral axis from one patient diagnosed with IPF in left and right lungs over time. The average percentage was calculated within 5% sections along the axis from posterior to anterior. (a) (b) is ground-glass distribution. (c) (d) is reticular distribution. (e) (f) is honeycomb distribution. (g) (h) is emphysema distribution.

Subpleural to internal distribution Figure 1.13 plots the subpleural-to-internal distance percentage of each connected cluster of three fibrosis CT patterns: ground-glass, reticular and honeycomb. The result shows that the subpleural-to-internal percentage of most connected disease clusters are < 20% for both left and right lungs. That is, the fibrosis is located preferentially within 20% of the distance from the lung surface. Reticular and ground-glass abnormalities are distributed throughout the centre to surface, whereas honeycomb is mostly not observed in the central core of the lungs in these subjects.

Lobar distribution Figure 1.14 shows the average volume percentage (across all time points from all patients) of ground-glass, reticular, honeycomb and emphysema in the five lobes (left upper, left lower, right upper, right middle, right lower). Fibrosis is located predominantly in the lower lobes (55.39%, 57.65%, 64.52% for honeycomb, reticular, ground-glass) compared to the other two lobes. For reticular pattern, the percentage in the middle lobe is much lower than the percentage in other lobes. Emphysema predominantly presents in the upper lobes (68.14%) and may also appear in the middle lobe of the right lung. Figure 1.15 shows the volume percentage of the four tissue patterns in the five lobes at each time point from an individual patient diagnosed with IPF. As shown in the figures, the lobar distribution of the disease remains almost the same at different time points, with more fibrosis appearing in the lower lobes and emphysema presenting mainly in upper lobes. The lobar distribution over time for other patients can be found in Appendix ??.

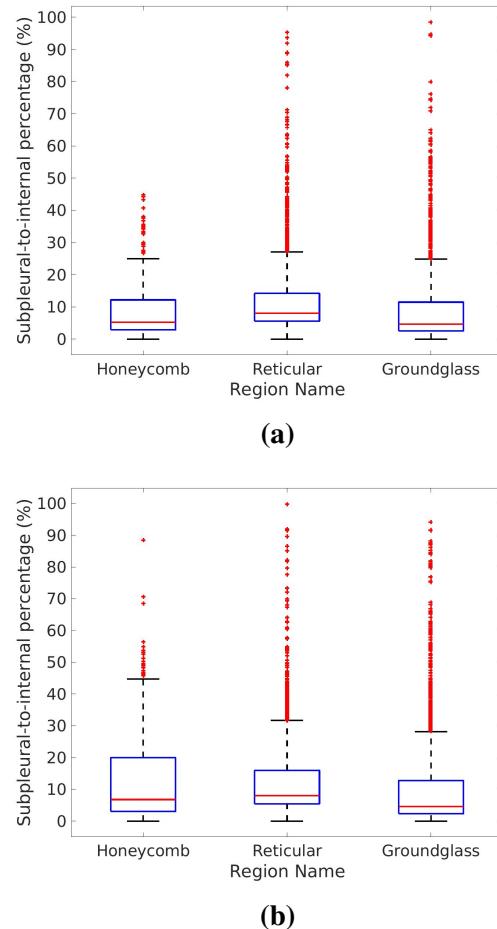


Figure 1.13: Subpleural-to-internal percentage of connected cluster of disease in IPF left and right lung. (a) Left lung. (b) Right lung.

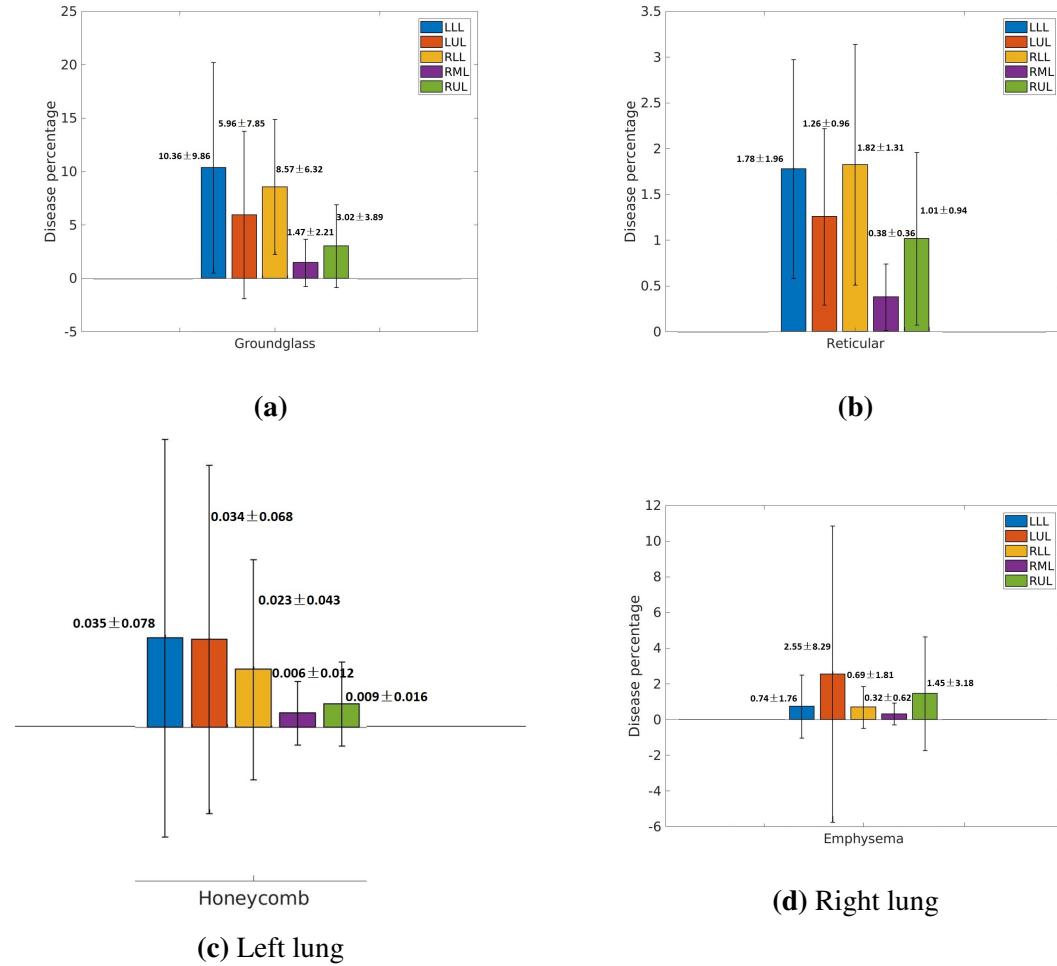


Figure 1.14: Average lobar distribution of abnormal tissue. The tissue percentage in different lobes is shown by color, and each bar represents the volume percentage of the CT pattern for one lobe across all the subjects. The black line shows the standard deviation. (a) Ground-glass lobar distribution. (b) Reticular lobar distribution. (c) Honeycomb lobar distribution. (d) Emphysema lobar distribution.

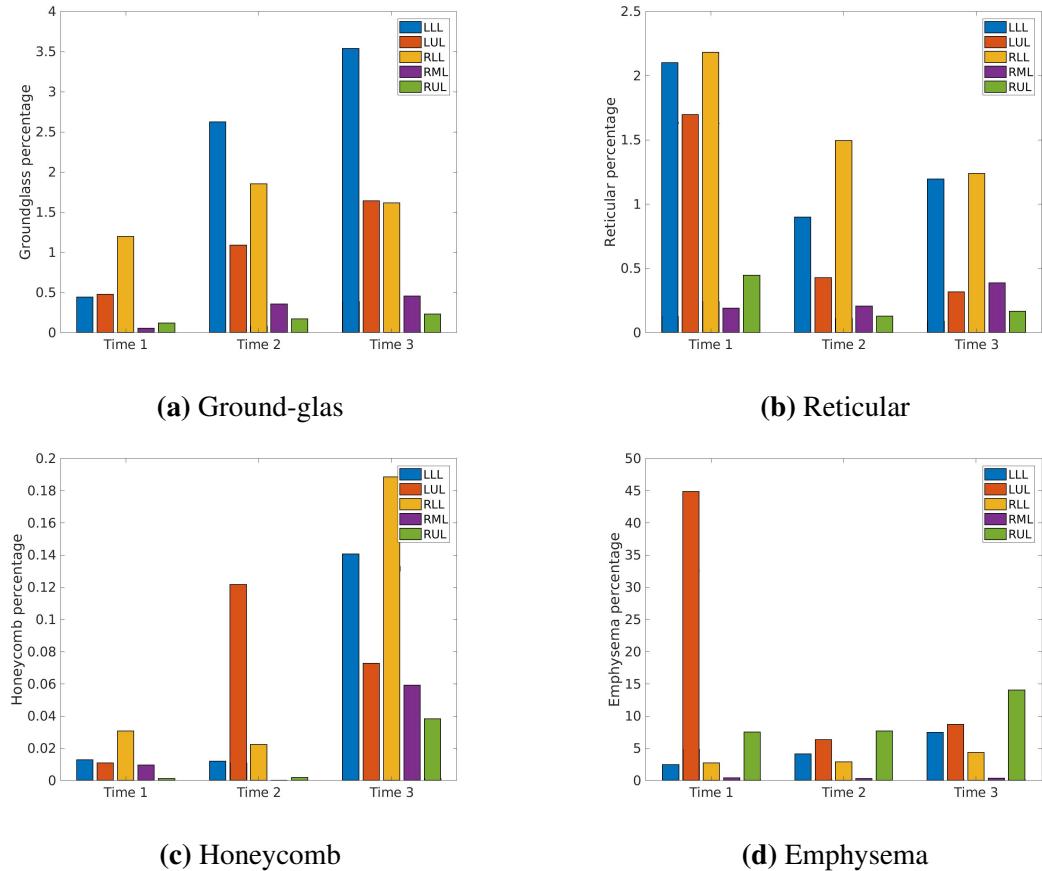


Figure 1.15: Lobar distribution of abnormal tissue for one patient diagnosed with IPF over time. The tissue percentage in different lobes is shown by colour, and each bar represents the volume percentage of the CT pattern for one lobe. (a) Ground-glass lobar distribution. (b) Reticular lobar distribution. (c) Honeycomb lobar distribution. (d) Emphysema lobar distribution.

Change in classification of tissue pattern over time

Table 1.3 and Table 1.4 show quantitative results for the change in classification of tissue pattern (ground-glass and reticular) for one patient with IPF between time points. The change in volume percentage of classified tissue CT patterns between time points is given (for ground-glass or reticular pattern to the other tissue CT patterns). Over 50% of areas initially identified as ground-glass remain classified the same over time in both

left and right lung, whereas other parts of ground-glass are subsequently classified as reticular or normal tissue. For the reticular region, although a high percentage doesn't change between time points, a large proportion is recognized as ground-glass at the second and third time point, especially in right lung (more than 50% on average).

Interestingly, in this subject and others in the cohort, some disease areas are classified back to normal tissue at later time points, although a decline in lung function can be observed during this period. Around 10%-30% of reticular and ground-glass areas are identified back to normal region during this time. This is clearly an artefact of analysis of IPF patient, which may be explained by the two following reasons: 1. Error caused by the lung surface fitting and classified data mapping. As introduced in Chapter 3, Section ??, the average RMS error of the fitting method for 50 subjects was 4.79 mm, and this error was inevitable. The accuracy of the classified data mapping to the mean SSM was partly relied on accuracy of lung surface fitting, as the local coordinate of within the corresponding element was used during the mapping. 2. Error caused by the CALIPER tissue classification. The distribution of abnormalities in IPF lung is mainly patchy, and the regions of fibrosis and emphysema are always not large connected areas. Figure 1.16 shows the CALIPER tissue classification overlapping on raw image. As shown in Figure 1.16, different tissue patterns in most of the lung region are mixed and interlaced, which increase the difficulty of identifying tissue pattern in these part accurately.

Table 1.3: Change in classification of tissue CT pattern (volume percentage) over time of left lung from one subject diagnosed with IPF (subject 6 with three time points) (%).

		Ground-glass	Mild-LAA	Moderate-LAA	Normal	Reticular	Honeycomb	Severe-LAA
Time1 - Time2	Ground-glass	53.65	0.36	0	37.31	8.65	0	0.03
	Reticular	29.76	3.01	0	28.41	38.82	0	0
Time2 - Time3	Ground-glass	71.61	0.13	0.03	15.41	12.82	0	0
	Reticular	39.94	0.83	0	16.59	42.56	0.07	0
Time1 - Time3	Ground-glass	70.01	0.83	0	20.80	8.35	0.01	0
	Reticular	37.97	4.73	0	15.49	41.81	0	0

The time interval from time point 1 to time point 2 is 15 months, the time interval from time point 1 to time point 2 is 5 months, the time interval from time point 1 to time point 3 is 20 months.

Table 1.4: Change in classification of tissue CT pattern (volume percentage) over time of right lung from one subject (subject 6 with three time points) diagnosed with IPF (%).

		Ground-glass	Mild-LAA	Moderate-LAA	Normal	Reticular	Honeycomb	Severe-LAA
Time1 - Time2	Ground-glass	61.76	4.15	0.92	7.14	25.92	0	0.11
	Reticular	47.61	6.88	0.01	11.10	34.37	0	0.03
Time2 - Time3	Ground-glass	78.16	0.08	0.01	18.51	3.23	0	0.05
	Reticular	55.74	0.44	0	31.46	12.35	0.01	0
Time1 - Time3	Ground-glass	79.98	1.82	1.63	12.82	3.45	0	0.30
	Reticular	65.80	3.28	0.02	19.40	11.33	0	0.16

The time interval from time point 1 to time point 2 is 15 months, the time interval from time point 1 to time point 2 is 5 months, the time interval from time point 1 to time point 3 is 20 months.

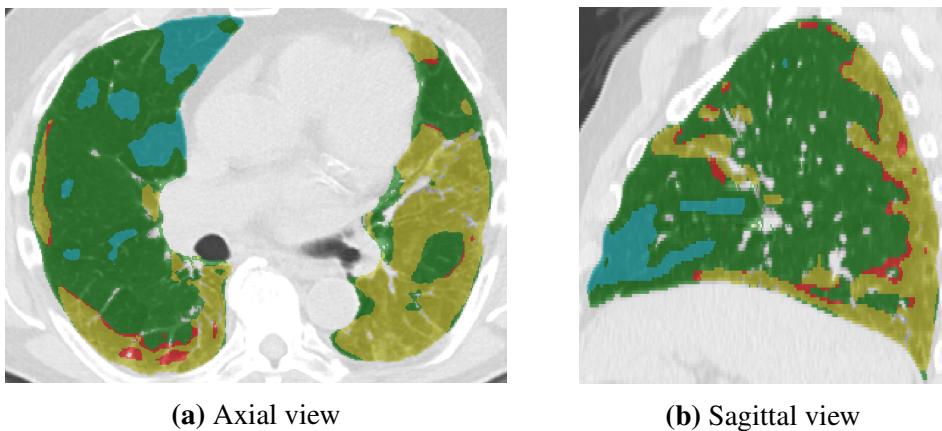


Figure 1.16: CALIPER tissue classification overlapping on raw image. Green is the normal region, red is the reticular region, yellow is the ground-glass region, and blue is the mild LAA region. (a) Axial view. (b) Sagittal view.

1.3.3 Volume analysis of IPF lungs

The change of whole lung volume over time

Table 1.5 shows the imaged volume of left and right lung (the first two columns) of all subjects for each time point. It can be seen that although the imaged lung volumes of most (9) patients decrease over time, some subjects (4 patients) do exhibit an increase in lung volume as time goes by. Interestingly, for most of the (9) patients, the changes of lung volume (decrease or increase) are consistent with the changes of the percentage of LAA over time.

Lobe volume difference between old normal lungs and IPF lungs

The average lobe volume proportion among IPF subjects and among normal subjects (the control group) are shown in Table 1.6 (the first two columns). It can be seen that the IPF group has a lower average volume proportion for left lower lobe and right lower

Table 1.5: Imaged volume of each lung (L) and the percentage of LAA area (%) in both lungs for each time point.

Sub No.	Time point	Scan time	Left lung	Right lung	LAA percentage
IPF2	Time point1	0 month	1.39	2.04	4.41
	Time point2	15 month	1.88	2.72	37.49
	Time point3	35 month	1.83	2.68	25.00
IPF3	Time point1	0 month	1.55	2.34	6.80
	Time point2	11 month	1.38	2.11	5.84
	Time point3	30 month	1.28	2.05	4.75
IPF5	Time point1	0 month	1.73	1.49	19.64
	Time point2	12 month	1.58	1.31	16.41
	Time point3	23 month	1.67	1.43	31.54
IPF6	Time point1	0 month	2.41	3.47	81.27
	Time point2	15 month	2.13	3.01	73.79
	Time point3	20 month	2.06	3.09	59.76
IPF9	Time point1	0 month	1.35	1.49	7.69
	Time point2	7 month	1.22	1.43	4.43
	Time point3	20 month	1.23	1.53	8.64
IPF10	Time point1	0 month	1.34	1.66	35.76
	Time point2	14 month	1.01	1.25	1.03
IPF13	Time point1	0 month	1.94	1.58	22.62
	Time point2	16 month	1.85	2.40	23.04
IPF14	Time point1	0 month	2.45	2.74	54.43
	Time point2	46 month	2.16	2.33	32.36
IPF15	Time point1	0 month	3.72	3.37	73.90
	Time point2	13 month	2.86	3.38	69.51
	Time point3	76 month	3.04	3.72	72.03
IPF21	Time point1	0 month	2.24	2.31	29.38
	Time point2	63 month	1.88	1.86	29.04
	Time point3	79 month	1.73	1.85	23.21

lobe compared to the value of the control group. Smaller lower lobe volume would be consistent with increase in stiffness of the lower lobes in IPF lungs, which relates to the basal appearance of fibrosis. The difference in lobe volume proportions between IPF subjects and the control group were compared statistically using a Student t-test. Table 1.6 (the third column) shows the p-values for this test when comparing volumes of the five lobes between the two groups. The result shows a significant difference in the volume proportion for both right lower lobe and right middle lobe between IPF subjects and the controls ($p < 0.001$), but not for the left lower lobe. That is, the right lower lobe is significantly smaller than in IPF, and this is at the expense of a significant increase in right middle lobe size.

Table 1.6: Lobe volume proportion comparison between IPF subjects and older normal subjects.

Lobe	IPF	Older normal	P-value
Left lower lobe	0.205	0.211	0.532
Left upper lobe	0.251	0.242	0.352
Right lower lobe	0.210	0.254	$\ll 0.001$
Right middle lobe	0.119	0.087	$\ll 0.001$
Right upper lobe	0.214	0.207	0.559

1.3.4 SSM - shape analysis of IPF lungs

Shape difference between IPF lungs and normal lungs

Figure 1.17 shows the weight value distribution of the first three principal shape modes for IPF and the control subjects.

There is a significant difference of the first mode weight between IPF and controls ($p \ll 0.001$). However, for mode 2 and mode 3, no significant shape difference was observed between the two groups ($p = 0.017$ and $p = 0.641$, respectively). Figure 1.18

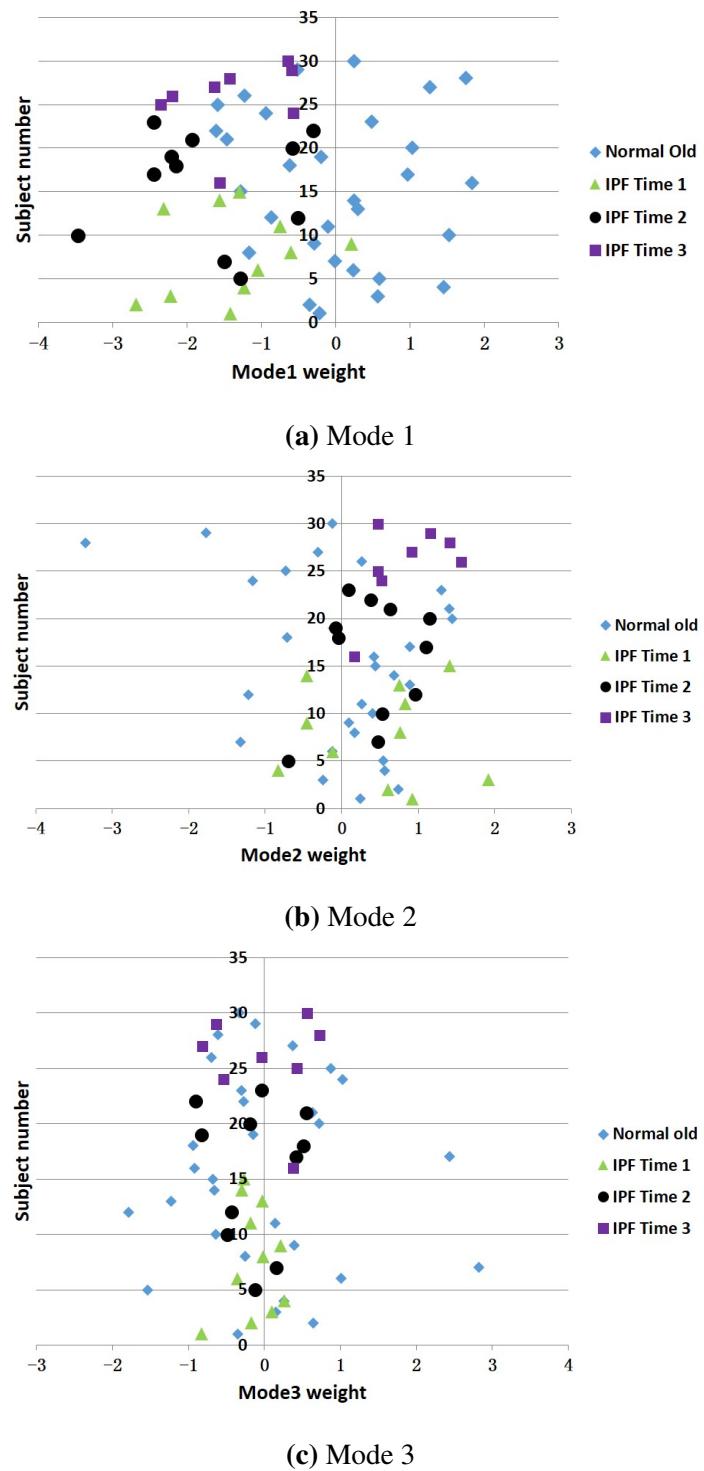


Figure 1.17: Shape differences between IPF subjects and old normal subjects of the first three modes (shown with different time point). (a) Mode 1 weight. (b) Mode 2 weight. (c) Mode 3 weight.

illustrates overall shape variation of the first mode with added different values of standard deviation to the mean shape model. The first shape mode accounts for over 20% of the entire shape variation in normal control lungs. It is seen in Figure 1.18 that the first mode relates to the largest change in the anterior-posterior diameter of the lung with a lateromedial tilt towards both apices, and is also associated with the ratio of apical to basal diameters. With the positive or negative standard deviation added to the mean shape model, there is a variation in the right anterior edge, the inferior lingular segment of the left superior lobe, the medial basal segment of the right middle lobe, and the left oblique fissure. In addition, there is a shape change in the "roundness" of the lateral surface in both left and right lung, and the variation in the distance of left and right lungs in the apex and base is also observed. As shown in Figure 1.17a and Figure 1.18, the weight values of Mode 1 for IPF subjects are mostly distributed across negative values (all but one subject at a single time point), where the negative weight corresponds to a larger ratio of anterior-posterior diameter to lung height compared to control lungs.

In Figure 1.17, the data for different time points were shown in different colors. As can be seen from the figures, the first three shape modes did not change significantly between time points (with all the p-values > 0.1 between time point 1 and time point 2, time point 1 and time point 3, time point 2 and time point 3). There were no regular trends in any of the three shape modes over time.

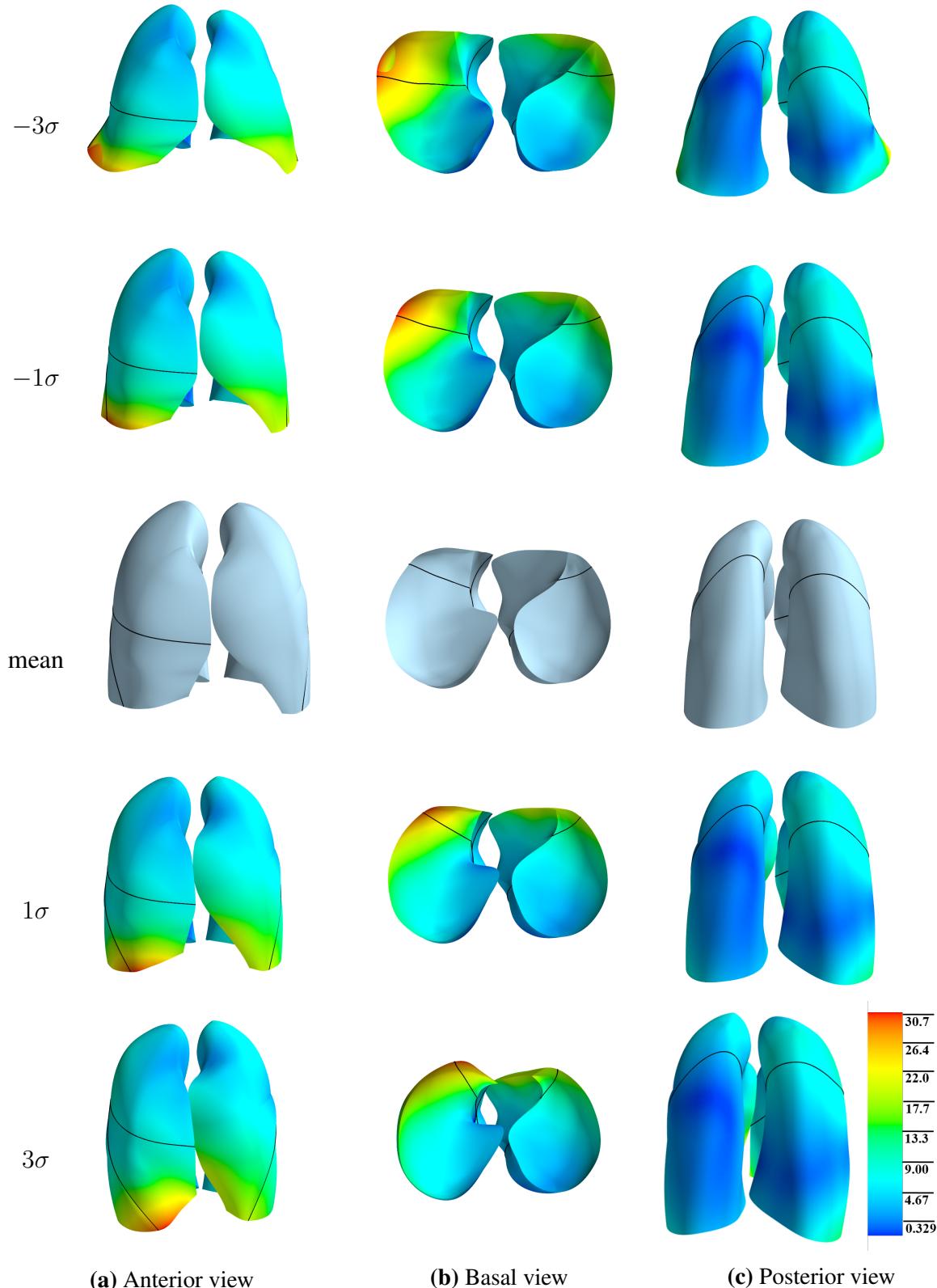


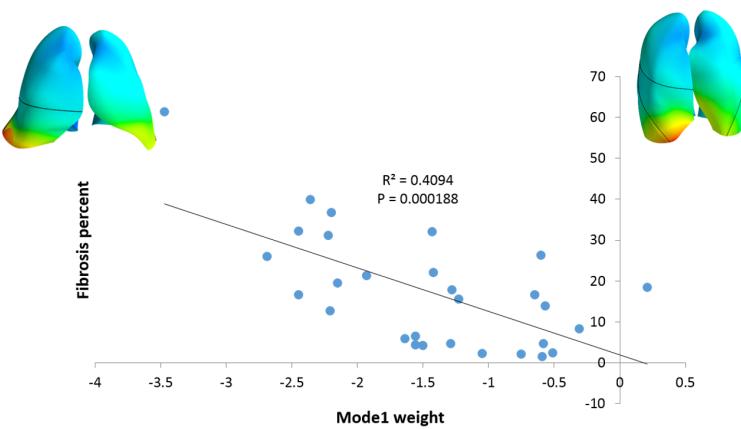
Figure 1.18: PCA-derived shape variation of the first shape mode, with different values of standard deviation (-3σ , -1σ , μ , 1σ , 3σ) added to the mean shape model. (a) Anterior view. (b) Basal view. (c) Posterior view. The colormap shows the extent of shape variation from the mean shape model.

Relationship between lung lobe shape and fibrosis and low attenuation area extent

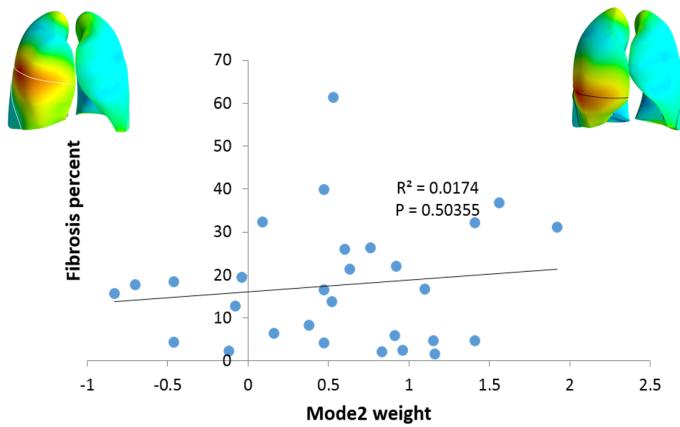
Table 1.7 lists the p-values and R values from linear regression for the correlation of the first three modes with fibrosis and LAA percentage. Figure 1.19 shows the relationship of the first three modes to overall fibrosis percentage. No correlations were found between the three shape modes and LAA extent. The first shape mode shows a significant relationship with the percentage of fibrosis ($p < 0.0001$). With an increase of the percentage of fibrosis, the weight value of mode 1 becomes increasingly negative which corresponds to a larger shape difference from the mean control lung shape. Negative weighting of mode 1 is also associated with increased ratio of anterior-posterior diameter to the height of the lung, which makes the lung appear 'fatter' and 'shorter'. Therefore in general, the lung shape change in patients with IPF is strongly associated with the fibrosis extent, and the more extensive the fibrosis, the more 'abnormal' the lung shape is.

Table 1.7: Results for linear regression of shape mode weighting against extent of fibrosis.

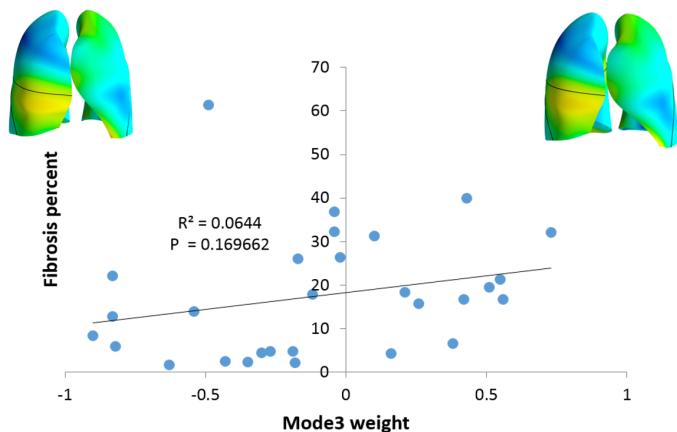
Mode	Fibrosis		LAA	
	P-value	R	P-value	R
Mode 1	<< 0.001	0.640	0.168	0.265
Mode 2	0.504	0.132	0.578	0.110
Mode 3	0.170	0.254	0.796	0.047



(a) Mode 1



(b) Mode 2



(c) Mode 3

Figure 1.19: A Linear regression of the PCA-derived first three modes with respect to overall fibrosis percentage. (a) Mode 1 weight. (b) Mode 2 weight. (c) Mode 3 weight.

1.4 Discussion

Individualized treatment strategies are urgently needed in clinical applications and are the ultimate goal of modern pulmonary medicine. A quantitative analysis of IPF based on HRCT images has the potential to build a reliable relationship between imaging tissue-level bio-markers and clinical endpoints. As shown here, tissue density, tissue volume, spatial distribution of abnormalities, and lung shape are all important indexes for describing a statistical progression of IPF disease. Through quantifying these features instead of qualitative evaluation, subjective errors could be avoided to offer a more consistent assessment of individual patients.

Average statistical shape model provides a consistent measurement to quantify disease among different people and over the whole clinical course

IPF is a progressive lung disease that has significant variable expression between different patients. In this study, the spatial distribution of CALIPER-classified data in the IPF lung were quantified along cranio-caudal, dorso-to-ventral, and subpleural-to-internal directions over time. Currently, few published studies have quantitatively analysed the spatial distribution of IPF abnormalities, especially for the progression and change of disease over time. The SSM-based quantitative method used here normalized a set of lungs of different shapes into a standard shape model, thus providing a convenient way to make a reliable comparison between different patients or within one patient at different time points. This method makes it possible to capture the disease variation across a population and describe the difference using objective indexes. Furthermore, the progression of disease over time can be analysed. The goal is that this contributes to helping describe the clinical course of disease.

It is widely acknowledged that typical symptoms occur on average 1-2 years before a definite clinical diagnosis of IPF, and some radiographic evidence of IPF may even be found before symptoms occur (Raghu et al., 2011; Devaraj, 2014). Understanding this 'sub-clinical' period of disease is very important for developing an early diagnosis of IPF disease. The SSM-based method makes it possible to compare the IPF lung with a normal cohort. While the IPF lung was shown to have significantly different shape to the normal lung (primarily with respect to mode 1), its shape did not change significantly over time. The shape differences partially reflect changes to the lung tissue, and partially to the chest wall. The results of this study suggest that considerable chest wall remodelling has occurred in these patients prior to their initial imaging.

The shape of the basal part of IPF lung is significantly different to the normal lung and correlates with fibrosis extent

The first shape mode of the SSM is significantly different between IPF and normal, and strongly correlates with the percentage of fibrosis. The first shape mode corresponds mainly to the anterior-posterior diameter of the lung which is associated with a variation of diaphragm location. This shape change of the diaphragm is potentially explained by the spatial distribution of tissue abnormalities. The basal and peripheral location of disease is very likely to increase the stiffness of the lower part of the lung, thus it could impact on the movement of the diaphragm when breathing. Furthermore, fibrosis extent in IPF is found to be related to the ratio of anterior-posterior diameter to the height of lung. A larger proportion of fibrosis is usually correlated with a larger ratio of anterior-posterior diameter to the height of lung. This "compression effect" on the lung shape may be associated with the reduction in tissue compliance caused by fibrosis. Specifically, since the IPF subject and control imaging used in this chapter were all

acquired at the end of inspiration, the lower tissue compliance caused by fibrosis will influence the expansion of the lung during inhalation if inhalation is driven by a normal muscle pressure, and therefore this may have an effect on the lung shape. In addition, it is demonstrated in the lobe volume analysis that the IPF lung has a lower average volume proportion for the left lower lobe and left upper lobe, which may be also caused by the "compression effect" of the lower lobe in IPF lungs. These features provide a potential way to indicate if the 'lung function' is deteriorating before PFTs can show any abnormalities as a shape difference of IPF patient may be observed when comparing with older normal lungs.

The quantification of combined IPF and emphysema is the difficulty and emphasis in the future work

Combined IPF and emphysema (CPFE) has been recognised and defined over the past ten years (Cottin et al., 2005; Meltzer and Noble, 2008). Some researchers suggest that CPFE should be regarded as a distinct clinical entity other than emphysema or IPF alone, since it has a characteristic pulmonary function feature different from pure emphysema or IPF. It is commonly believed that CPFE is strongly associated with heavy smokers, severe dyspnea on exertion, and impaired gas exchange. Emphysema usually appears in the upper lobe whereas IPF disease tends to appear in the lower lobes (Lin and Jiang, 2015) which is consistent with the result of our analysis. However, whether the presence of the two diseases developing in parallel is still unknown (Cottin et al., 2005; King Jr et al., 2011; Lin and Jiang, 2015). In the future work, the image based analysis result will be used as quantitative indexes in a lung functional modelling, which could help understand the relationship between the two diseases and the impact of emphysema on IPF.

1.5 Summary

In this chapter, we analysed and characterized IPF classified tissue over time using quantitative methods. The result shows that fibrosis had higher tissue density compared with normal tissue, and presented predominantly basally and peripherally. In contrast, emphysema had lower tissue density and mostly located in upper lobes. The first principal SSM mode ($> 20\%$ of the shape variation in normal lungs) was significantly different between IPF and normal and strongly correlated with fibrosis extent in IPF lungs. This quantitative analysis provides consistent potential tissue-level markers which will be used to guide the computational modelling of IPF lungs in the next chapter.

List of References

- Arzhaeva, Y., Prokop, M., Tax, D. M., De Jong, P. A., Schaefer-Prokop, C. M., and van Ginneken, B. (2007). Computer-aided detection of interstitial abnormalities in chest radiographs using a reference standard based on computed tomography. *Medical Physics*, 34(12):4798–4809. [Cited on page 7.]
- Bartholmai, B. J., Raghunath, S., Karwoski, R. A., Moua, T., Rajagopalan, S., Maldonado, F., Decker, P. A., and Robb, R. A. (2013). Quantitative ct imaging of interstitial lung diseases. *Journal of thoracic imaging*, 28(5). [Cited on pages 1, 2, 3, 4, 8, 9, 11, and 12.]
- Behr, J., Mehnert, F., Beinert, T., et al. (1992). Evaluation of interstitial lung disease by quantitative high-resolution computed tomography. *Am Rev Respir Dis*, 145(Suppl.):A191. [Cited on page 5.]
- Best, A. C., Lynch, A. M., Bozic, C. M., Miller, D., Grunwald, G. K., and Lynch, D. A. (2003). Quantitative ct indexes in idiopathic pulmonary fibrosis: relationship with physiologic impairment. *Radiology*, 228(2):407–414. [Cited on page 7.]
- Best, A. C., Meng, J., Lynch, A. M., Bozic, C. M., Miller, D., Grunwald, G. K., and Lynch, D. A. (2008). Idiopathic pulmonary fibrosis: physiologic tests, quantitative ct indexes, and ct visual scores as predictors of mortality. *Radiology*, 246(3):935–940. [Cited on pages 4 and 7.]

- Bjoraker, J. A., Ryu, J. H., Edwin, M. K., Myers, J. L., Tazelaar, H. D., Schroeder, D. R., and Offord, K. P. (1998). Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *American journal of respiratory and critical care medicine*, 157(1):199–203. [Cited on page 3.]
- Castaldi, P. J., San José Estépar, R., Mendoza, C. S., Hersh, C. P., Laird, N., Crapo, J. D., Lynch, D. A., Silverman, E. K., and Washko, G. R. (2013). Distinct quantitative computed tomography emphysema patterns are associated with physiology and function in smokers. *American journal of respiratory and critical care medicine*, 188(9):1083–1090. [Cited on page 5.]
- Chabat, F., Yang, G.-Z., and Hansell, D. M. (2003). Obstructive lung diseases: texture classification for differentiation at ct. *Radiology*, 228(3):871–877. [Cited on page 7.]
- Costabel, U., Du Bois, R. M., and Egan, J. J. (2007). Diffuse parenchymal lung disease, volume 36. Karger Medical and Scientific Publishers. [Cited on page 4.]
- Cottin, V., Nunes, H., Brillet, P., Delaval, P., Devouassoux, G., Tillie-Leblond, I., Israel-Biet, D., Valeyre, D., Cordier, J., et al. (2005). Combined pulmonary fibrosis and emphysema: a distinct underrecognised entity. *European Respiratory Journal*, 26(4):586–593. [Cited on page 52.]
- da Silva Felix, J. H., Cortez, P. C., Rebouças Filho, P. P., de Alexandria, A. R., Costa, R. C. S., and Holanda, M. A. (2008). Identification and quantification of pulmonary emphysema through pseudocolors. In Mexican International Conference on Artificial Intelligence, pages 957–964. Springer. [Cited on page 5.]
- Delorme, S., Keller-Reichenbecher, M.-A., Zuna, I., Schlegel, W., and Van Kaick, G. (1997). Usual interstitial pneumonia: quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis. *Investigative radiology*, 32(9):566–574. [Cited on page 5.]

- Depeursinge, A., Iavindrasana, J., Hidki, A., Cohen, G., Geissbuhler, A., Platon, A., Poletti, P.-A., and Müller, H. (2010). Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization. *Journal of digital imaging*, 23(1):18–30. [Cited on page 5.]
- Devaraj, A. (2014). Imaging: how to recognise idiopathic pulmonary fibrosis. *European Respiratory Review*, 23(132):215–219. [Cited on pages 2, 3, and 51.]
- Diette, G. B., Scatarige, J. C., Haponik, E. F., Merriman, B., and Fishman, E. K. (2005). Do high-resolution ct findings of usual interstitial pneumonitis obviate lung biopsy? *Respiration*, 72(2):134–141. [Cited on page 4.]
- Flaherty, K. R., Andrei, A.-C., King Jr, T. E., Raghu, G., Colby, T. V., Wells, A., Bassily, N., Brown, K., Du Bois, R., Flint, A., et al. (2007). Idiopathic interstitial pneumonia: do community and academic physicians agree on diagnosis? *American journal of respiratory and critical care medicine*, 175(10):1054–1060. [Cited on page 2.]
- Flaherty, K. R., King Jr, T. E., Raghu, G., Lynch III, J. P., Colby, T. V., Travis, W. D., Gross, B. H., Kazerooni, E. A., Toews, G. B., Long, Q., et al. (2004). Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? *American journal of respiratory and critical care medicine*, 170(8):904–910. [Cited on page 3.]
- Flaherty, K. R., Travis, W. D., Colby, T. V., Toews, G. B., Kazerooni, E. A., Gross, B. H., Jain, A., STRAWDERMAN III, R. L., Flint, A., LYNCH III, J. P., et al. (2001). Histopathologic variability in usual and nonspecific interstitial pneumonias. *American journal of respiratory and critical care medicine*, 164(9):1722–1727. [Cited on page 3.]
- 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). Com-

- puted tomography–based biomarker provides unique signature for diagnosis of copd phenotypes and disease progression. *Nature medicine*, 18(11):1711. [Cited on page 5.]
- Gietema, H. A., Müller, N. L., Fauerbach, P. V. N., Sharma, S., Edwards, L. D., Camp, P. G., Coxson, H. O., of COPD Longitudinally to Identify Predictive Surrogate End-points (ECLIPSE) investigators, E., et al. (2011). Quantifying the extent of emphysema:: Factors associated with radiologists' estimations and quantitative indices of emphysema severity using the eclipse cohort. *Academic radiology*, 18(6):661–671. [Cited on page 5.]
- Gilman, M. J., Laurens, J. R., Somogyi, J. W., and Honig, E. G. (1983). Ct attenuation values of lung density in sarcoidosis. *Journal of computer assisted tomography*, 7(3):407–410. [Cited on page 5.]
- Gotway, M. B., Freemer, M. M., and King, T. E. (2007). Challenges in pulmonary fibrosis· 1: Use of high resolution ct scanning of the lung for the evaluation of patients with idiopathic interstitial pneumonias. *Thorax*, 62(6):546–553. [Cited on pages 3 and 4.]
- Gould, G., MacNee, W., McLean, A., Warren, P., Redpath, A., Best, J., Lamb, D., and Flenley, D. (1988). Ct measurements of lung density in life can quantitate distal airspace enlargement—an essential defining feature of human emphysema. *American Review of Respiratory Disease*, 137(2):380–392. [Cited on page 5.]
- Hu, S., Hoffman, E. A., and Reinhardt, J. M. (2001). Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images. *IEEE transactions on medical imaging*, 20(6):490–498. [Cited on page 9.]
- Kazerooni, E. A. (2001). High-resolution ct of the lungs. *American Journal of Roentgenology*, 177(3):501–519. [Cited on page 4.]

- Kazerooni, E. A., Martinez, F. J., Flint, A., Jamadar, D. A., Gross, B. H., Spizarny, D. L., Cascade, P. N., Whyte, R. I., Lynch 3rd, J., and Toews, G. (1997). Thin-section ct obtained at 10-mm increments versus limited three-level thin-section ct for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *AJR. American journal of roentgenology*, 169(4):977–983. [Cited on page 4.]
- Kim, E. Y., Lee, K., Chung, M., Kwon, O. J., Kim, T., and Hwang, J. (1999). Non-specific interstitial pneumonia with fibrosis: serial high-resolution ct findings with functional correlation. *AJR. American journal of roentgenology*, 173(4):949–953. [Cited on page 4.]
- Kim, H., Tashkin, D., Clements, P., Li, G., Brown, M., Elashoff, R., Gjertson, D., Abtin, F., Lynch, D., Strollo, D., et al. (2010). A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. *Clinical and experimental rheumatology*, 28(5 Suppl 62):S26. [Cited on pages 7 and 8.]
- Kim, H. J., Brown, M. S., Chong, D., Gjertson, D. W., Lu, P., Kim, H. J., Coy, H., and Goldin, J. G. (2015). Comparison of the quantitative ct imaging biomarkers of idiopathic pulmonary fibrosis at baseline and early change with an interval of 7 months. *Academic radiology*, 22(1):70–80. [Cited on pages 7 and 8.]
- Kim, H. J., Brown, M. S., Elashoff, R., Li, G., Gjertson, D. W., Lynch, D. A., Strollo, D. C., Kleerup, E., Chong, D., Shah, S. K., et al. (2011). Quantitative texture-based assessment of one-year changes in fibrotic reticular patterns on hrct in scleroderma lung disease treated with oral cyclophosphamide. *European radiology*, 21(12):2455–2465. [Cited on pages 7 and 8.]
- Kim, K. G., Goo, J. M., Kim, J. H., Lee, H. J., Min, B. G., Bae, K. T., and Im, J.-G. (2005). Computer-aided diagnosis of localized ground-glass opacity in the lung at ct: initial experience. *Radiology*, 237(2):657–661. [Cited on page 7.]

- King Jr, T. E., Pardo, A., and Selman, M. (2011). Idiopathic pulmonary fibrosis. *The Lancet*, 378(9807):1949–1961. [Cited on pages 2 and 52.]
- Kinsella, M., Müller, N. L., Abboud, R. T., Morrison, N. J., and DyBuncio, A. (1990). Quantitation of emphysema by computed tomography using a “density mask” program and correlation with pulmonary function tests. *Chest*, 97(2):315–321. [Cited on page 5.]
- Knudson, R. J., Standen, J. R., Kaltenborn, W. T., Knudson, D. E., Rehm, K., Habib, M. P., and Newell, J. D. (1991). Expiratory computed tomography for assessment of suspected pulmonary emphysema. *Chest*, 99(6):1357–1366. [Cited on page 5.]
- Kundel, H. L. (2006). History of research in medical image perception. *Journal of the American College of Radiology*, 3(6):402–408. [Cited on page 2.]
- Lin, H. and Jiang, S. (2015). Combined pulmonary fibrosis and emphysema (cpfe): an entity different from emphysema or pulmonary fibrosis alone. *Journal of thoracic disease*, 7(4):767. [Cited on page 52.]
- Lynch, D. A. (2007). Quantitative ct of fibrotic interstitial lung disease. *Chest*, 131(3):643–644. [Cited on page 5.]
- Lynch, D. A., Godwin, J. D., Safrin, S., Starko, K. M., Hormel, P., Brown, K. K., Raghu, G., King Jr, T. E., Bradford, W. Z., Schwartz, D. A., et al. (2005). High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. *American journal of respiratory and critical care medicine*, 172(4):488–493. [Cited on pages 3 and 4.]
- Maldonado, F., Moua, T., Rajagopalan, S., Karwoski, R. A., Raghunath, S., Decker, P. A., Hartman, T. E., Bartholmai, B. J., Robb, R. A., and Ryu, J. H. (2013). Automated quantification of radiologic patterns predicts survival in idiopathic pulmonary

- fibrosis. *European Respiratory Journal*, pages erj00718–2012. [Cited on pages 9, 10, 11, and 12.]
- Meltzer, E. B. and Noble, P. W. (2008). Idiopathic pulmonary fibrosis. *Orphanet journal of rare diseases*, 3(1):8. [Cited on page 52.]
- Misumi, S. and Lynch, D. A. (2006). Idiopathic pulmonary fibrosis/usual interstitial pneumonia: imaging diagnosis, spectrum of abnormalities, and temporal progression. *Proceedings of the American Thoracic Society*, 3(4):307–314. [Cited on page 4.]
- Monaghan, H., Wells, A. U., Colby, T. V., Du Bois, R. M., Hansell, D. M., and Nicholson, A. G. (2004). Prognostic implications of histologic patterns in multiple surgical lung biopsies from patients with idiopathic interstitial pneumonias. *Chest*, 125(2):522–526. [Cited on page 3.]
- Mueller-Mang, C., Grosse, C., Schmid, K., Stiebellehner, L., and Bankier, A. A. (2007). What every radiologist should know about idiopathic interstitial pneumonias. *Radiographics*, 27(3):595–615. [Cited on page 3.]
- Müller, N. L., Staples, C. A., Miller, R. R., and Abboud, R. T. (1988). “density mask”: an objective method to quantitate emphysema using computed tomography. *Chest*, 94(4):782–787. [Cited on page 5.]
- Olson, A. L., Swigris, J. J., Lezotte, D. C., Norris, J. M., Wilson, C. G., and Brown, K. K. (2007). Mortality from pulmonary fibrosis increased in the united states from 1992 to 2003. *American journal of respiratory and critical care medicine*, 176(3):277–284. [Cited on page 2.]
- Raghu, G., Collard, H. R., Egan, J. J., Martinez, F. J., Behr, J., Brown, K. K., Colby, T. V., Cordier, J.-F., Flaherty, K. R., Lasky, J. A., et al. (2011). An official ats/ers/jrs/alat statement: idiopathic pulmonary fibrosis: evidence-based guidelines

- for diagnosis and management. *American journal of respiratory and critical care medicine*, 183(6):788–824. [Cited on pages 2 and 51.]
- Raghunath, S., Rajagopalan, S., Karwoski, R. A., Maldonado, F., Peikert, T., Moua, T., Ryu, J. H., Bartholmai, B. J., and Robb, R. A. (2014). Quantitative stratification of diffuse parenchymal lung diseases. *PLoS One*, 9(3):e93229. [Cited on pages 2, 4, 9, and 13.]
- Richeldi, L., Collard, H. R., and Jones, M. G. (2017). Idiopathic pulmonary fibrosis. *The Lancet*, 389(10082):1941–1952. [Cited on page 2.]
- Saketkoo, L. A., Matteson, E. L., Brown, K. K., Seibold, J. R., and Strand, V. (2011). Developing disease activity and response criteria in connective tissue disease-related interstitial lung disease. *The Journal of rheumatology*, 38(7):1514–1518. [Cited on page 4.]
- Sato, Y., Westin, C.-F., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., and Kikinis, R. (2000). Tissue classification based on 3d local intensity structures for volume rendering. *IEEE transactions on visualization and computer graphics*, 6(2):160–180. [Cited on page 9.]
- Sumikawa, H., Johkoh, T., Colby, T. V., Ichikado, K., Suga, M., Taniguchi, H., Kondoh, Y., Ogura, T., Arakawa, H., Fujimoto, K., et al. (2008). Computed tomography findings in pathological usual interstitial pneumonia: relationship to survival. *American journal of respiratory and critical care medicine*, 177(4):433–439. [Cited on page 4.]
- Sverzellati, N., Devaraj, A., Desai, S. R., Quigley, M., Wells, A. U., and Hansell, D. M. (2011). Method for minimizing observer variation for the quantitation of high-resolution computed tomographic signs of lung disease. *Journal of computer assisted tomography*, 35(5):596–601. [Cited on page 3.]

- Tawhai, M. H. and Burrowes, K. S. (2003). Developing integrative computational models of pulmonary structure. *The Anatomical Record*, 275(1):207–218. [Cited on page 15.]
- Travis, W. D., Costabel, U., Hansell, D. M., King Jr, T. E., Lynch, D. A., Nicholson, A. G., Ryerson, C. J., Ryu, J. H., Selman, M., Wells, A. U., et al. (2013). An official american thoracic society/european respiratory society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *American journal of respiratory and critical care medicine*, 188(6):733–748. [Cited on page 2.]
- Uchiyama, Y., Katsuragawa, S., Abe, H., Shiraishi, J., Li, F., Li, Q., Zhang, C.-T., Suzuki, K., et al. (2003). Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. *Medical Physics*, 30(9):2440–2454. [Cited on page 7.]
- Uppaluri, R., Hoffman, E. A., Sonka, M., Hartley, P. G., Hunninghake, G. W., and McLennan, G. (1999a). Computer recognition of regional lung disease patterns. *American journal of respiratory and critical care medicine*, 160(2):648–654. [Cited on pages 5 and 6.]
- Uppaluri, R., Hoffman, E. A., Sonka, M., Hunninghake, G. W., and McLennan, G. (1999b). Interstitial lung disease: a quantitative study using the adaptive multiple feature method. *American journal of respiratory and critical care medicine*, 159(2):519–525. [Cited on pages 5 and 6.]
- Van Ginneken, B., Katsuragawa, S., ter Haar Romeny, B. M., Doi, K., and Viergever, M. A. (2002). Automatic detection of abnormalities in chest radiographs using local texture analysis. *IEEE transactions on medical imaging*, 21(2):139–149. [Cited on page 7.]

- Wang, J.-S., Cherg, J.-M., Perng, D.-S., Lee, H.-S., and Wang, S. (2013). High-resolution computed tomography in assessment of patients with emphysema. *Respiratory care*, 58(4):614–622. [Cited on page 5.]
- Watadani, T., Sakai, F., Johkoh, T., Noma, S., Akira, M., Fujimoto, K., Bankier, A. A., Lee, K. S., Müller, N. L., Song, J.-W., et al. (2013). Interobserver variability in the ct assessment of honeycombing in the lungs. *Radiology*, 266(3):936–944. [Cited on page 2.]
- Wells, A. U., Desai, S. R., Rubens, M. B., Goh, N. S., Cramer, D., Nicholson, A. G., Colby, T. V., Du Bois, R. M., and Hansell, D. M. (2003). Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *American journal of respiratory and critical care medicine*, 167(7):962–969. [Cited on page 4.]
- Xu, Y., Sonka, M., McLennan, G., Guo, J., and Hoffman, E. A. (2006a). Mdct-based 3-d texture classification of emphysema and early smoking related lung pathologies. *IEEE transactions on medical imaging*, 25(4):464–475. [Cited on page 6.]
- Xu, Y., van Beek, E. J., Hwanjo, Y., Guo, J., McLennan, G., and Hoffman, E. A. (2006b). Computer-aided classification of interstitial lung diseases via mdct: 3d adaptive multiple feature method (3d amfm). *Academic radiology*, 13(8):969–978. [Cited on page 6.]
- Zavaletta, V. A., Bartholmai, B. J., and Robb, R. A. (2007). High resolution multidetector ct-aided tissue analysis and quantification of lung fibrosis. *Academic radiology*, 14(7):772–787. [Cited on pages 7 and 11.]
- Zavaletta, V. A., Karwoski, R. A., Bartholmai, B., and Robb, R. A. (2006). High resolution multidetector ct aided tissue analysis and quantification of lung fibrosis. In *Medical Imaging 2006: Physiology, Function, and Structure from Medical Images*,

volume 6143, page 61432Z. International Society for Optics and Photonics. [Cited on page 7.]