Ho-Fung Chan

Auckland Bioengineering Institute,
University of Auckland,
Auckland 1142, New Zealand
e-mail: hcha184@aucklanduni.ac.nz

Alys R. Clark

Auckland Bioengineering Institute,
University of Auckland,
Auckland 1142, New Zealand
e-mail: alys.clark@auckland.ac.nz

Eric A. Hoffman

Departments of Radiology
and Biomedical Engineering,
University of Iowa,
Iowa City, IA 52242
e-mail: eric-hoffman@uiowa.edu

Duane T. K. Malcolm

Auckland Bioengineering Institute,
University of Auckland,
Auckland 1142, New Zealand
e-mail: d.malcolm@auckland.ac.nz

Merryn H. Tawhai¹

Auckland Bioengineering Institute,
University of Auckland,
Private Bag 92019,
Auckland 1142, New Zealand
e-mail: m.tawhai@auckland.ac.nz

Quantifying Normal Geometric Variation in Human Pulmonary Lobar Geometry From High Resolution Computed Tomography

Previous studies of the ex vivo lung have suggested significant intersubject variability in lung lobe geometry. A quantitative description of normal lung lobe shape would therefore have value in improving the discrimination between normal population variability in shape and pathology. To quantify normal human lobe shape variability, a principal component analysis (PCA) was performed on high resolution computed tomography (HRCT) imaging of the lung at full inspiration. Volumetric imaging from 22 never-smoking subjects (10 female and 12 male) with normal lung function was included in the analysis. For each subject, an initial finite element mesh geometry was generated from a group of manually selected nodes that were placed at distinct anatomical locations on the lung surface. Each mesh used cubic shape functions to describe the surface curvilinearity, and the mesh was fitted to surface data for each lobe. A PCA was performed on the surface meshes for each lobe. Nine principal components (PCs) were sufficient to capture >90% of the normal variation in each of the five lobes. The analysis shows that lobe size can explain between 20% and 50% of intersubject variability, depending on the lobe considered. Diaphragm shape was the next most significant intersubject difference. When the influence of lung size difference is removed, the angle of the fissures becomes the most significant shape difference, and the variability in relative lobe size becomes important. We also show how a lobe from an independent subject can be projected onto the study population's PCs, demonstrating potential for abnormalities in lobar geometry to be defined in a quantitative manner. [DOI: 10.1115/1.4029919]

Keywords: shape analysis, lung anatomy, finite elements

Introduction

The human lung is anatomically and functionally divided into five lobes. The geometry of individual lobes and fissures can now be segmented reliably from HRCT at full inspiration in healthy individuals [1]. The fissures are considered important landmarks in assessing pulmonary HRCT imaging, as knowledge of their normal anatomy allows one to distinguish normal from abnormal lungs and to assess disease progression [2–5]. However, cadaver and specimen studies of lungs from large populations have suggested significant intersubject variability in lobar geometry [2–4], including the lobar fissure orientation and absent or incomplete fissures; this complicates the definition of “abnormal” lobar geometry. A quantitative definition of what constitutes normal lobar shape would provide a means to distinguish between normal population shape variability and pathology that affects the lobe shape [6]. For example, normal aging and emphysema (a component of chronic obstructive pulmonary disease) both reduce lung tissue elastic recoil and can result in lung hyperinflation. Distinguishing early emphysema from normal aging is challenging because of their similarity on imaging and pulmonary function tests (PFTs). Emphysema tends to develop predominantly in the apical regions, whereas tissue changes due to normal aging occur throughout the lung; lobe shape distortion due to localized disease could therefore be a useful additional biomarker for diagnosis.

One approach to defining normality was proposed by Li et al. [7], who used a landmark and intensity-based consistent image registration technique to compute the average whole lung shape of a population of normal subjects. Population subjects were registered to an ideal “normal” template subject and shape variations of the population from this template were subsequently determined. A limitation of this approach is the requirement of a “perfectly normal” template subject with complete, normal airway tree and pulmonary blood vessel tree segmentations available. Average population shape variations were shown to be slightly dependent on the template subject selection.

Statistical shape modeling provides an alternative method to describe normal geometric variability [8] and does not rely on a registration-based approach. As has previously been demonstrated in analysis of breast shape [9], the method allows assessment of the variation in shape across a cohort of anatomically customized finite element models. Specifically, PCA can be used to decompose the shape of an organ across a population into the superposition of a series of shape modes that explain the variability in the population. The main goal of PCA is to reduce the dimensionality of a data set that consists of a large number of interrelated variables, whilst still representing as much of the variation that is present in the data set as possible [10]. This involves transforming the data set into a new set of variables called the PCs, which are uncorrelated and are ordered such that the first few components account for most of the variation present in the original data set [10].

PC models have been shown to improve speed and accuracy in developing finite element meshes within which to predict the mechanics of deformable tissues and to provide guidance in

¹Corresponding author.

Manuscript received August 10, 2014; final manuscript received February 16, 2015; published online March 18, 2015. Assoc. Editor: Ender A. Finol.

Fig. 1 Schematic depictions of the left and right lungs with anatomical landmarks labeled. The entrance to the lungs of the arteries (Art), airways (Air), and veins (V) is labeled along with other key anatomical features. Left lung landmarks are: (1) the highest point on fissure, (2) the top of aorta impression, (3) the intersection between cardiac notch, anterior surface, and diaphragmatic surface, (4) the lowest point on fissure /base of the lower lobe, and (5) the base of aorta impression. Right lung landmarks are: (1) the apex of arch of azygos vein impression, (2) the apex of the oblique fissure, (3) the apex of esophageal groove (lower lobe), (4) the apex of the intersection between horizontal and oblique fissures, (5) the base of the esophageal groove (lower lobe), (6) the intersection between the cardiac impression, anterior surface, and diaphragmatic surface, (7) the point on horizontal fissure at anterior surface boundary, and (8) the base of venous insertion.

segmentation from imaging. In the pulmonary lobes, this prior knowledge of normal lobar shape could potentially improve algorithms to detect lobar fissures in imaging [11,12], which present difficulties in dealing with gaps in observable fissures between image slices and regions of pathology [13]. A PCA guided approach has already proved useful in segmenting lung shape from chest radiographs where lung boundaries are often difficult to distinguish [14], but the approach has not been used to provide insight into normal lung shape. Quantifying the average lobe geometry and its normal variation is important for defining a baseline against which geometric variation due to disease or congenital abnormalities can be assessed.

Here, we use PCA on pulmonary HRCT data obtained from normal, healthy volunteers, to quantify the modes of shape variation in a normal population aged <50 years.

Methodology

A semi-automated mesh fitting procedure and PCA were performed on each of the five lobes [left upper lobe (LUL), left lower lobe (LLL), right upper lobe (RUL), right middle lobe (RML), and right lower lobe (RLL)] of the lungs of 22 human subjects with normal lung physiology. The mesh fitting procedure was performed using a desktop PC with a 2.4 GHz processor and 12 GB RAM. Each lobe took less than 1 min to mesh after manual identification of landmark nodes.

Data Acquisition. Subjects were selected for retrospective analysis from the Human Lung Atlas [15], which is a database of normal pulmonary structure and function comprising volumetric multidetector computed tomography (MDCT) imaging acquired at two lung volumes and PFTs. All study and imaging protocols were approved by the University of Iowa Review Board and Radiation Safety Committees. MDCT scan parameters were 120 kV, 100 mAs, and a pitch of 1.2. Subject data were acquired during the years 2004–2011; all data were acquired in the same center (The University of Iowa Comprehensive Lung Imaging Center (I-Clic)) under the same protocols with regular scanning calibration. Each volumetric image set contained 500–800 images, with

a distance between slices of 0.50 mm and a reconstruction matrix of 512×512 pixels.

The criteria for study inclusion were: normal radiological appearance of the lung as assessed by a radiologist; less than 50 years in age; body mass index (BMI) of $19\text{--}31 \text{ kg/m}^2$; normal PFTs; and no history of smoking. Of the 22 subjects, 10 were female and 12 male. End-inspiratory images were used in this study, with subjects volume-controlled to 85% of vital capacity.

Initial Lobar Mesh Definition. Lobe volumes were segmented automatically from each set of images using custom-written software (PASS, the Pulmonary Analysis Software Suite, University of Iowa, Iowa City, IA) [1]. Segmented lobar volumes were rendered individually in the visualization software CMGUI,¹ and a cloud of surface data was generated for each lobe.

The goal of the PCA was to quantify the variation in lobar shape across the population using methods that allow minimal user input in defining pulmonary lobar shape in an individual. Each template finite element mesh had topology appropriate to the characteristic shape of the lobe. Each lobar mesh was consistent between subjects (the same mesh connectivity); however, the meshes for the different lobes were structured to capture the size and shape of that lobe as closely as possible so meshes differed between lobes.

A consistent mesh definition that is robust to interuser variability (as assessed later) was required in order to conduct the PCA. This requires the definition of landmarks that define the shape of the object in question. In the lung, there are typically few distinct anatomical landmarks on the lung surface. In this study, five anatomical landmarks were defined in the left lung, and seven in the right lung as shown schematically in Fig. 1 (note that some anatomical points lie on the surface of both lobes). In the left lung, nodes were selected from the cloud of surface data to represent the most apical and basal points in the impression of the ascending aorta in the LLL, the most apical and basal points on the fissure surface, and the point of intersection of the cardiac notch and anterior and

¹www.cmiss.org/cmgui

Fig. 2 Initial LUL surface mesh nodes. (a) Manually selected nodes. (b) Generated linear Lagrange mesh structure.

diaphragmatic surfaces of the lung (see [Supplemental](#) material and Fig. 2(a)) ([Supplemental](#) figures are available under the “[Supplemental Data](#)” tab for this paper on the ASME Digital Collection). In the right lung, the anatomical nodes were the apex and base of the esophageal groove in the lower lobe, the base of venous insertion into the lung, the apex of the intersection between horizontal and oblique fissures, the apex of the arch of the azygos vein impression, the point on the horizontal fissure at the anterior surface boundary, and the point of intersection of the cardiac impression and anterior and diaphragmatic surfaces of the lung. Additional pseudolandmarks were manually placed to define “ridges” or edges on the lung surface. That is, the aortic impression, the ridge between anterior and posterior surfaces, fissure edges and the diaphragmatic edge. These edges are identified in Fig. 2 and Figs. S1–S4 ([Supplemental](#) figures are available under the “[Supplemental Data](#)” tab for this paper on the ASME Digital Collection). Linear Lagrange surface meshes (comprising triangular and quadrilateral elements) were created by combining the manually selected anatomical landmarks with automatically generated additional pseudolandmarks that were evenly distributed across the lobar surface (Fig. 2(b)), with the aim of pseudolandmark insertion being to generate a mesh of evenly sized elements. Initial mesh structures for each of the other lobes are shown in Figs. S1–S4 ([Supplemental](#) figures are available under the “[Supplemental Data](#)” tab for this paper on the ASME Digital Collection).

Cubic Surface Mesh Fitting. A surface mesh fitting algorithm was developed in a Python script environment. The initial mesh had the same coordinate system as the computed tomography (CT) images and was transformed to obtain a local lobe-specific x_L , y_L , z_L such that each axis was consistent between subjects. This was attained by a rotation about an arbitrary z_L -axis, where this axis and rotation about this axis were defined by utilizing nodes in the initial mesh. Each lobe had a different set of initial mesh nodes from which the z_L -axis was derived. For example, the z_L -axis for the LUL was the average of two vectors created by nodes 1, 7, and nodes 8, 6 (refer to Fig. 2). The initial mesh nodes used to define the z_L -axis for the other lobes are summarized in Table 1. The origin of each mesh, defined as the center of mass, was translated to the global origin point (0,0,0).

The initial surface mesh was then converted to a cubic Lagrange mesh using the process summarized in Fig. 3. First, the new node points were inserted on the surface of the initial linear mesh to create a refined mesh with equispaced node points to define the cubic Lagrange mesh. The structure of our initial mesh allows grouping of node points into “rings” of nodes at the same z_L -height (Fig. 3(a)). Data points with a z_L -coordinate of ± 0.5

Table 1 Initial mesh nodes used to define the z_L -axis for each lobe ([Supplemental](#) figures are available under the “[Supplemental Data](#)” tab for this paper on the ASME Digital Collection)

Lobe	Initial mesh nodes	Refer to figure
LUL	1, 7 and 8, 6	Fig. 2
LLL	1, 13	Fig. S1
RLL	2, 11	Fig. S2
RML	2, 5	Fig. S3
RUL	7, 9	Fig. S4

mm of the node points were grouped, to represent a curvilinear path around the lung. The 1 mm thickness of the data point grouping was chosen to be larger than the image slice thickness, but sufficiently small that it would not include curvature in the z -direction, avoids significant vertical movement of nodes, and retains elements of approximately equal size. The coordinates of these planes were calculated by utilizing the z_L -axis as the normal vector describing the planes. Any two nodes that lie on the edges of the planes create a vector and are pushed to the plane surface via a function that pushes the nodes the maximum distance apart from each other while being constrained to the plane ring. Other (intermediate nodes) between these edge nodes were placed within the ring of data points using a data crawl function (Figs. 3(b)–3(d)).

The data crawl utilizes the two previously placed edge nodes and creates a line that covers ring data in between the edge nodes. This line was then divided into the required number of sections to reflect the number/position of intermediate nodes. To ensure that the data crawl function forms a line in the correct direction, the data plane ring is split into two parts. The split plane function uses the information of two vectors from three nodes to filter out points that are above or below the two edge nodes. In some cases, the data planes were modified gradually to align with another axis, allowing for a better distribution of mesh nodes around the surface data.

The data crawl approach was used for all mesh nodes that could be located on data plane rings. For other nodes, such as ones located at the ends of lobes or around the diaphragm, a different approach was implemented. For the nodes at the end of lobes, new data planes were defined, which were orthogonal to the standard data planes, which allowed these nodes to be pushed either to the top or bottom of the lobe. The outer diaphragm nodes were all placed using a function that pushed the required node to the edge

Fig. 3 Example of cubic surface mesh fitting process at the top of LUL. (a) Data plane ring with mesh nodes in their original positions after conversion from linear to cubic mesh. (b) Edge nodes pushed to the edge and data plane ring split into two parts. Intermediate node placement via data crawl function for front (c) and back (d) regions of data plane ring.

of the diaphragm by creating a plane from the middle (initial) diaphragm node and the nodes directly above the required node. The inner nodes were also placed using a similar approach as the outer diaphragm nodes. The inner nodes utilized the middle diaphragm node, an outer node, and the apex node to create a node plane to place the required inner diaphragm nodes.

The cubic Lagrange mesh was fitted to the data cloud using least squares minimization. The accuracy of the surface mesh fits was assessed via the root mean square (RMS) error between each mesh and its surface data cloud.

Mesh Fitting Validation. An interuser analysis was performed to determine the influence of user variability in selection of initial nodes on the subsequent fitted surface mesh. 13 volunteers were recruited to fit the same two lobes (LUL and LLL) of a single subject. Volunteers had knowledge of lung anatomy and were supplied with a detailed set of instructions regarding the mesh fitting process. The differences between each volunteer were quantified by examining the fitting error of the final surface mesh, and the variability in node placement of the initial mesh nodes.

PCA. Surface meshes were generated for each lobe of each of the 22 subjects. For each lobe, the meshes were translated by moving the centroid of the mesh to the global origin. The combination of consistent surface mesh orientation and the translation of the centroid removed possible orientation and positional effects on the subsequent PCs.

Using the Python Modular toolkit for Data Processing module (MDP),² a PCA was performed to quantify the variation in lung lobe shape across the population. The PCA was executed in a similar approach by Rajagopal et al. [9]. A single matrix representing the datasets of the population was constructed with rows representing mesh nodes and columns representing the number of population subjects. Singular value decomposition was used in the PCA to decompose the population matrix into the mean shape vector ($\bar{x}$) and the number (p) of singular values (w) and associated vectors (N) that quantify the orthonormal directions and magnitudes of the principal modes of variation. The resultant PCA mesh (X) is shown in the following equation:

$$X = \bar{x} + w_j N_j, \quad j = 1 \dots p \quad (1)$$

where w_j is the weight value, and N_j is the associated vector for the j th PC.

The number of PCs (p) chosen affects the amount of explained variance in the PCA. The relationship between number of PCs and explained variance (seen in Fig. 4) shows that after 90% explained variance, additional PCs contribute very little toward the total explained variance. Therefore in this study, nine PCs were chosen, corresponding to around 90% explained variance for each lobe PCA.

PCA allows a subject to be approximated by computing a set of corresponding weight values for the p number of PCs. To compute the weight values, the desired subject dataset is projected onto the population PCs. This was achieved by using the dot product of the PCA variance matrix and the desired subject shape vector

$$s = (X_i - \bar{x}) \cdot V \quad (2)$$

where s is the array of PC scores, X_i is the desired subject dataset, $\bar{x}$ is the mean shape vector, and V is the population PCA variance matrix. The scores (s) were then converted into relative standard deviation (SD) weight values. The resultant weight values were then used to approximate the shape of the desired subject dataset by modifying the population PCA mesh, as shown in Eq. (1).

Fig. 4 The percentage of explained variance for each lobe versus the number of PCs for the study population PCA

Results

Mesh Fitting. Surface meshes were fitted to each lobe of the 22 subjects. The accuracy of the surface mesh fits was assessed via the RMS error between each mesh and its surface data cloud. As shown in Table 2, the RMS errors for all lobes were between 3 and 4 mm, while the SD was between 0.3 and 0.7 mm. This is a similar error estimate to fits of lobar surfaces to the same data clouds as obtained using the established methods of fitting cubic Hermite elements described by Fernandez et al. [16].

Fitted surface meshes and initial user-selected mesh nodes for a single subject were used to determine interuser variability in the accuracy of the fitted meshes across 13 users. Table 3 illustrates the variation in user-placed nodes for the LLL and LUL initial surface meshes. The mean and SD Euclidean distance values were calculated for each node, however, only SD values are shown in Table 3 because they represent the variation in user placement. The values shown in bold are nodes that were defined at anatomical locations. It can be seen that these nodes have smaller SD Euclidean distances, therefore exhibit lower variance in comparison to the other nodes that were placed in nonanatomical locations.

Table 4 shows the variation in surface mesh error for the LLL and LUL in each of the 13 volunteers. The variation in mesh RMS error is small, this is in contrast to the larger variation seen in volunteer node placement. The difference in initial node placement between volunteers did not markedly affect the RMS error of surface meshes; therefore implying the fitting process was robust to interuser variation.

PCA. Figure 4 shows the proportion of variation in lobe shape that is explained with the addition of each PC. The variation in the population shape of the RML can be explained with the fewest PCs; this is in contrast to the RLL, where explaining any given proportion of variation requires the most modes. Each lobe shows a similar trend in their explained variance per mode, and >90% of the variance can be explained with nine PCs in each.

Figure 5 illustrates how the first three PCs for the LLL change the shape of the lobe. The mean (μ) population lobe shape is depicted in column two of Fig. 5, while columns one and three show the lobe shape at three SDs either side of the mean (-3σ , $+3\sigma$). Column four describes the shape change to the lobe

Table 2 RMS error for fitted lung lobe surface meshes

Lobe	Mean RMS (mm)	SD RMS (mm)
LLL	3.08	0.33
LUL	3.19	0.50
RLL	3.15	0.42
RML	3.59	0.48
RUL	3.37	0.65

²<http://mdp-toolkit.sourceforge.net/>

Table 3 SD Euclidean distance values of user-placed nodes in LLL and LUL interuser analysis

LLL node variability		LUL node variability	
Node number	Euclidean distance SD (mm)	Node number	Euclidean distance SD (mm)
1 ^a	6.03	1 ^a	9.76
2 ^a	3.52	6 ^a	2.19
3	9.81	7 ^a	7.53
5	14.28	8	11.81
6	11.93	11	7.48
7 ^a	3.87	12	9.48
8	7.23	—	—
9 ^a	5.25	—	—
12	20.01	—	—
13	9.32	—	—
Mean Euclidean distance SD	9.13		8.04

Note: Rows 7–10 are blank for LUL columns, because the LUL initial mesh only required six user-placed nodes, this is in contrast to the ten required for LLL.

^aNodes that were placed at anatomical locations. See Fig. 1 (LUL) and Fig. S1 (LLL) for initial mesh node locations (Supplemental figures are available under the “Supplemental Data” tab for this paper on the ASME Digital Collection).

associated with a particular PC. The shape changes were determined visually by inspection of the surface renderings of lobe mesh transformations over a range of SD values for each PC. Extra figures illustrating the shape changes from the first three PCs in each of the other lobes can be found in Supplemental Material (Supplemental sections are available under the “Supplemental Data” tab for this paper on the ASME Digital Collection).

The descriptions of lobe shape changes associated with the first three PCs are summarized in Table 5. For the first mode of variation, the majority of the lung lobes (except RUL) have shape changes associated with scaling of the lobe size (larger or smaller lungs). The second and third PCs show different shape changes for each lobe. However, most changes can be associated with the diaphragm for lower lobes and the lobe apex for the upper lobes.

The effect of age, BMI, and total lung capacity (TLC, as defined from PFT in the upright seated posture) on the first PC of this population of normal subjects was assessed. As the first PC visually represents a volume scaling of the lobe in most cases, a negative weight would imply a smaller than average lobe and a positive weight a larger than average lobe. Female subjects tended

Table 4 Surface mesh RMS errors for interuser variability analysis

Volunteer	LLL RMS (mm)	LUL RMS (mm)
1	2.94	3.31
2	3.04	4.18
3	2.88	3.55
4	2.72	3.47
5	2.97	4.20
6	2.92	3.41
7	2.82	3.19
8	2.90	3.58
9	2.89	3.30
10	2.95	3.29
11	2.84	3.16
12	3.03	3.70
13	2.77	3.35
Mean	2.90	3.51
SD	0.09	0.32

to have negative weighting compared to positive weighting for the males, reflecting the smaller on-average female lung size. Figure 6 illustrates this trend. Correlation values ($R^2 = 0.51$, $p < 0.001$) indicate a reasonable linear fit and a strong correlation between the first PC of the LLL and the subjects' TLC. In contrast, no significant relationship was found between subject age or BMI and the weighting of the first PC.

Normalized PCA. The first PC for each lobe has been shown here, in most cases, to represent the volume scaling of the lobe. This volume scaling PC was removed by scaling each lung to a common size to remove the effect of size on the study population. The scaling factor for each subject in the population was calculated by scaling the subject's TLC volume to a 1 l volume. PCA was then implemented again with the new normalized surface meshes.

The effect of the first three PCs on lobe shape is summarized in Table 6. When compared with Table 5 (lobe shape change including lung size), the first PC, in most cases, now represents the different fissure surface angles of the study population. The third PC for each lobe now represents the volume scaling of the lobe. That is, while total volume was scaled to a common value, the relative volume of the lobes can differ between subjects and this is reflected in the third PC.

Independent Subject Projection. The PCA can be used to decompose an individual's lobe shape into the weighted sum of these PCs. Table 7 shows the weight values in the LUL of an independent subject that was not included in the original study population. The weight values for the first eight PCs for the independent subject were all within ± 1 SD except for the second PC (-1.53 SD). This indicates, with respect to the study population, that this subject's LUL shape is quite similar to the average LUL population shape, assuming that the population exhibits a Gaussian distribution. After the application of the first eight PCs, which explains 92% of the variance in the LUL population, the deformed mean PCA mesh (RMS error = 4.68 mm) provides a good representation of the independent subject's surface mesh (RMS error = 3.58 mm).

Fig. 5 Description of shape changes to the LLL for the first three PCs

Table 5 Summary of shape changes for each lobe associated with the first three PCs

Lobe	PC 1	PC 2	PC 3
LLL	Scaling of the lobe with extension of the diaphragm front	Change in the width of the diaphragm and shape of lobe apex area	Change in the shape of the diaphragm
LUL	Scaling of the lung lobe	Rotation of the lobe and change in the pulmonary artery and vein entry point	Change in the bottom edge of the lobe
RLL	Slight scaling of the lobe along with fissure surface changes	Lobe height scaling	Change and scaling of the lobe apex region
RML	Scaling of the lobe and change in bottom edge nodes location	Rotation of the lobe and change in top edge nodes	Scaling of the fissure surface and the diaphragm section
RUL	Angle of fissure surface and change in anterior-posterior lobe distance	Scaling of the bottom half of lobe	Scaling of the top half of lobe

Discussion

We have presented a new image-based analysis of human lung lobe shape, using a multivariate PCA to determine the normal supine end-inspiratory shape of a cohort of 22 subjects with normal lung function, and the principal modes of variation, which describe the shape across the cohort. The PCA conducted here quantifies variation in lobar geometry in a normal population, with more than 60% of the population variation being explained by the first three PCs and 90% by the first nine PCs. The shape analysis has potential application in radiological assessment of the lung.

The benefit of PCA over previous more qualitative analyses of pulmonary lobar structure [1–3] is that it allows a quantitative description of shape variance in a population. Here, we have selected a normal population that excluded smokers, subjects over 50 years of age, and BMI >31 kg/m². Variation in our study is in agreement with previous qualitative studies, although these studies did not always distinguish between normal and abnormal subjects [1–3]. The first PC in each lobe, contributing to 26–52% of the explained variance, can be attributed at least in part to normal variation in the size of the lung. This was also found in a previous PCA study that analyzed pulmonary function measures (lung volumes and capacities) rather than 3D shape [17], where the first PC was attributed to a “general measure of lung function,” therefore representing lung size.

The second PCs for the lower lobes in our study corresponded to differences in diaphragm shape. This is consistent with a similar study by Li et al. [7], where they used a landmark and intensity-based consistent image registration technique to compute the average shape of a population of normal subjects. Subjects were registered to a common template subject and the average transformation was calculated. SD of displacement was used as a measure of shape variation in the population and indicated a significant amount of displacement in the diaphragm region of the lung. This was also seen in the current study, where

the main variation in shape, after size scaling, of the LLL and RLL was associated with the diaphragm. The diaphragm tends to displace cranially when in the supine posture, with respect to upright. The modes of variation associated with diaphragm shape could therefore potentially be less significant in the upright posture, but are of relevance to radiological analysis (via HRCT) where the lung is typically imaged supine. The normal shape variations in Li et al. [7] were measured in an origin-dependent way that can influence the magnitude of the standard displacement. Origins for each lung were located at the end of the corresponding mainstem bronchus; this was in contrast to the centroid of each lobe in our study.

The second PCs for the upper lobes described significant variation in the angle of the fissures and the location of the hilum. Aziz et al. [2] have previously reported variation in lobar fissure orientation, and our finding is consistent with this previous study. However, this study provides a quantitative assessment of the variability. The upper lobes also exhibited variation in the hilum region that may be related to the segmentation of the lung. The automated segmentation software that was used in the current study does not include the large pulmonary vessels that enter the lungs, and their “cut-off” point between intra- and extrapulmonary tissue is somewhat user-dependent. This can result in noisy segmentations around the hilum area. There is a trade-off in this region between increased user interaction to smooth the data cloud at the hilum, and maintaining as automated a procedure as possible. This must be accounted for when interpreting variation in these subsequent second PCs.

The trend seen in SD values for the first PCs between male and female is expected. The first PC is related to the size of the lobe; therefore as males are statistically larger than females, one would expect to see more positive weighting values in males and more negative weighting values in females. This trend is also seen when comparing the first PC and the TLC volume of each subject from PFTs. The weak relationship (*R*-squared value of 0.51) indicates that larger SD values correlate to larger measured TLC volumes in subjects.

To check whether removing the effect of lung size would influence the principal modes of shape variation, the PCA was repeated using lobe meshes that were scaled by the ratio of 1 l to the subject’s TLC. In these normalized lungs, the lobe fissure location and orientation became the major contributor to lobe shape. With the effect of lung size removed via lobe scaling, the first PC now represents the different lobe fissure angles of the study population. This is consistent with previous cadaver and specimen studies of lungs which suggested significant intersubject variability in lobar geometry [1–4], showing that this variability is consistent between in vivo and in vitro lung anatomy. It was found that the third PC now represented lobar volume scaling, reflecting a relative change in lobe volume between the lobes of a subject.

The method for surface mesh generation was semi-automated; hence there may be user error in selection of initial node locations. This would be significant if it altered the final fitted lobe mesh

Fig. 6 Relationship between the weighting of the first PC of the LLL and the TLC (obtained from PFTs) of each subject in the study population

Table 6 Summary of shape changes for each lobe associated with the first PCs of the volume normalized PCA

Lobe	PC 1	PC 2	PC 3
LLL	Change in the angle of the lobe fissure surface and shape of lobe apex	Change in diaphragm shape	Scaling of the lobe
LUL	Rotation and change in the lobe fissure surface	Change in the hilum area. Rotation of the lobe	Scaling of the lobe
RLL	Change in the angle of the lobe fissure surface and width of the diaphragm	Lobe height scaling	Scaling of the lobe
RML	Change in the size and angle of the lobe fissure surface	Rotation of the lobe	Scaling of the lobe
RUL	Rotation of the lobe	Scaling of the bottom half of lobe	Scaling of the top half of lobe

Table 7 The PC weights (SD values) for the LUL of an independent subject

PC	1	2	3	4	5	6	7	8
SD	0.79	−1.53	0.98	−0.17	0.46	−0.73	−0.32	0.58

such that the definition of the principal shape modes was affected. Analysis of intersubject node placement (Table 3) showed that nodes that were placed in distinct anatomical locations did not suffer from large variation in user placement, when the user was familiar with pulmonary anatomy. However, the lung has relatively few well-defined anatomical landmarks on its surface, and we noted variation in the placement of initial nodes that were not associated with well-defined anatomical locations. Anatomical landmark nodes had SD Euclidean distance values of less than 8 mm, while nonanatomical nodes had values up to 20 mm. This was to be expected as the description of node locations allowed for some flexibility in placement for nonanatomical nodes. This large variation, however, did not translate to RMS error variation in the final surface meshes, indicating a robust fitting algorithm. Interuser error could be further minimized in future studies now that an initial PCA data set has been established. That is, new subjects could have surface meshes fitted using a PCA-based fitting procedure [9], which would remove the need for initial manual placement of nodes at landmark locations. This would allow new subjects to be added to the cohort in a relative straightforward way and the removal of manual node placement would greatly decrease analysis time, allowing the PCA to be used in a practical or clinical setting.

Study Limitations. The subjects considered here were normal, never smokers, whose imaging was obtained under carefully controlled conditions. This limits the absolute number of normal subjects that can be considered, as HRCT is not routinely acquired in human subjects without clinical indication of the need for diagnostic CT. Data can be added to the PCA over time as more becomes available and this will allow, for example, a distinction to be made between normal males and normal females, between BMI ranges or between other population groups.

Imaging in this study was obtained at close to full inspiration. The set of PCs obtained from the PCA therefore are only applicable to this one particular imaged volume.

Conclusions

We have demonstrated a semi-automated procedure for quantifying normal variability in pulmonary lobar geometry via a PCA. We have shown that much of the variability across the study population was because of the size of the lung and the shape/displacement of the diaphragm. This is the first quantitative assessment of pulmonary lobar geometry that serves as a basis from which one can consider how pathologies can contribute to abnormalities in pulmonary geometry and function. The quantitative description of normal variability in the lung could contribute to improved clinical diagnostics, as comparison of imaging data to a normal shape could indicate regional tissue abnormalities. As

it is usually possible to automatically segment lobar structures from HRCT, our semi-automated procedure—that is robust to user variability—could provide a means to compare patient imaging to the population mean, and hence to identify substantial deviations from that mean. Future studies in abnormal populations could identify how anatomical features vary between groups.

Acknowledgment

This work was supported by the New Zealand Health Research Council via Project No. 12/668 (Clark PI), Project No. 09/143 (Tawhai PI), and NIH Grant No. HL064368.

References

- [1] Zhang, L., Hoffman, E. A., and Reinhardt, J. M., 2006, "Atlas-Driven Lung Lobe Segmentation in Volumetric X-Ray CT Images," *IEEE Trans. Med. Imaging*, **25**(1), pp. 1–16.
- [2] Aziz, A., Ashizawa, K., Nagaoki, K., and Hayashi, K., 2004, "High Resolution CT Anatomy of the Pulmonary Fissures," *J. Thorac. Imaging*, **19**(3), pp. 186–191.
- [3] Prakash, Bhardwaj, A. K., Shashirekha, M., Suma, H. Y., Krishna, G. G., and Singh, G., 2010, "Lung Morphology: A Cadaver Study in Indian Population," *Ital. J. Anat. Embryol.*, **115**(3), pp. 235–240.
- [4] Bhimai, D. N., Narasinga, R. B., and Sunitha, V., 2011, "Morphological Variations of Lung—A Cadaveric Study in North Coastal Andhra Pradesh," *Int. J. Biol. Med. Res.*, **2**(4), pp. 1149–1152.
- [5] Hayashi, K., Aziz, A., Ashizawa, K., Hayashi, H., Nagaoki, K., and Otsuji, H., 2001, "Radiographic Appearances of the Major Fissures," *Radiographics*, **21**(4), pp. 861–874.
- [6] Hoffman, E. A., Clough, A. V., Christensen, G. E., Lin, C. L., McLennan, G., Reinhardt, J. M., Simon, B. A., Sonka, M., Tawhai, M. H., Van Beek, E. J., and Wang, G., 2004, "The Comprehensive Imaging-Based Analysis of the Lung: A Forum for Team Science," *Acad. Radiol.*, **11**(12), pp. 1370–1380.
- [7] Li, B., Christensen, G. E., Hoffman, E. A., McLennan, G., and Reinhardt, J. M., 2012, "Establishing a Normative Atlas of the Human Lung: Computing the Average Transformation and Atlas Construction," *Acad. Radiol.*, **19**(11), pp. 1368–1381.
- [8] Coates, T. F., Taylor, C. J., Cooper, C. H., and Graham, J., 1995, "Active Shape Models—Their Training and Application," *Comput. Vision Image Understanding*, **61**(1), pp. 38–59.
- [9] Rajagopal, V., Boyes, R., Azhar, M., Gamage, T. P. B., Nielsen, P. M. F., and Nash, M. P., 2011, "Towards a Completely Automated Clinical Workflow for Breast Image Analysis Using Patient-Specific Biomechanical Models," *Proceedings of MICCAI2011 Workshop on Breast Image Analysis*, pp. 121–128.
- [10] Jolliffe, I. T., 2002, *Principal Component Analysis*, Springer-Verlag, New York.
- [11] Saita, S., Kubo, M., Kawata, Y., Niki, N., Ohmatsu, H., and Moriyama, N., 2006, "An Algorithm for the Extraction of Pulmonary Fissures From Low-Dose Multislice CT Image," *Syst. Comput. Jpn.*, **37**(9), pp. 63–76.
- [12] Zhou, X., Hayashi, T., Hara, T., Fujita, H., Yokoyama, R., Kiryu, T., and Hoshi, H., 2004, "Automatic Recognition of Lung Lobes and Fissures From Multislice CT Images," *Proc. SPIE*, **5370**, pp. 1629–1633.
- [13] Wei, Q., and Hu, Y., 2014, "A Hybrid Approach to Segmentation of Diseased Lung Lobes," *IEEE J. Biomed. Health Inf.*, **18**(5), pp. 1696–1706.
- [14] Shi, Y., Qi, F., Xue, Z., Chen, L., Ito, K., Matsuo, H., and Shen, D., 2008, "Segmenting Lung Fields in Serial Chest Radiographs Using Both Population-Based and Patient-Specific Shape Statistics," *IEEE Trans. Med. Imaging*, **27**(4), pp. 481–494.
- [15] Li, B., Christensen, G. E., Hoffman, E. A., McLennan, G., and Reinhardt, J. M., 2003, "Establishing a Normative Atlas of the Human Lung: Intersubject Warping and Registration of Volumetric CT Images," *Acad. Radiol.*, **10**(3), pp. 255–265.
- [16] Fernandez, J. W., Mithraratne, P., Thrupp, S. F., Tawhai, M. H., and Hunter, P. J., 2004, "Anatomically Based Geometric Modelling of the Musculo-Skeletal System and Other Organs," *Biomech. Model. Mechanobiol.*, **2**(3), pp. 139–155.
- [17] Cowie, H., Lloyd, M. H., and Soutar, C. A., 1985, "Study of Lung Function Data by Principal Components Analysis," *Thorax*, **40**(6), pp. 438–443.