Automatic airway-artery analysis on lung CT to quantify airway wall thickening and bronchiectasis

Adria Perez-Rovira
Biomedical Imaging Group Rotterdam, Departments of Medical Informatics, Radiology, and Paediatric Pulmonology, Erasmus MC-Sophia Childrens Hospital, 3015 CE, Rotterdam, The Netherlands.

Wieying Kuo
Department of Paediatric Pulmonology and Department of Radiology, Erasmus MC-Sophia Childrens Hospital, 3015 CE, Rotterdam, The Netherlands.

Jens Petersen
Department of Computer Science, University of Copenhagen, DK-2100, Copenhagen, Denmark.

Harm A.W.M. Tiddens
Department of Paediatric Pulmonology and Department of Radiology, Erasmus MC-Sophia Childrens Hospital, 3015 CE, Rotterdam, The Netherlands.

Marleen de Bruijne
Biomedical Imaging Group Rotterdam, Departments of Medical Informatics and Radiology, Erasmus MC, 3015 CE, Rotterdam, The Netherlands.
Department of Computer Science, University of Copenhagen, DK-2100, Copenhagen, Denmark.
Author to whom correspondence should be addressed (marleen.debruijne@erasmusmc.nl).

(Dated: 24 October 2016)

Purpose: Bronchiectasis and airway wall thickening are commonly assessed in computed tomography (CT) by comparing the airway size with the size of the accompanying artery. Thus, in order to automate the quantification of bronchiectasis and wall thickening following a similar principle, there is a need for methods that automatically segment the airway and vascular trees, measure their size, and pair each airway branch with its accompanying artery.
Methods: This paper combines and extends existing techniques to present a fully automated pipeline that, given a thoracic chest CT, segments, measures, and pairs airway branches with the accompanying artery. Then quantifies airway wall thickening and bronchiectasis by measuring the Wall-Artery Ratio (WAR) and lumen and outer wall Airway-Artery Ratio (AAR). Measurements that do not use the artery size for normalisation are also extracted, including Wall Area Percentage (WAP), Wall Thickness Ratio (WTR) and airway diameters.
Results: The method was thoroughly evaluated using 8,000 manual annotations of airway-artery pairs from 24 full-inspiration paediatric CT scans (12 diseased and 12 controls). Limits of agreement between the automatically and manually measured diameters were comparable to inter-observer limits of agreement. Differences in automatically obtained WAR, AAR, WAP and WTR between bronchiectatic subjects and controls were similar as when manual annotations were used: WAR and outer AAR were significantly higher in the bronchiectatic subjects ($p < 0.05$), but lumen AAR, WAP and WTR were not. Only measurements that use artery size for normalisation led to significant differences between groups, highlighting the importance of airway-artery pairing.
Conclusions: The fully automatic method presented in this paper could replace time-consuming manual annotations and visual scoring methods to quantify abnormal widening and thickening of airways.

Key words: CT, airway, artery, quantification, bronchiectasis.

I. INTRODUCTION

Computed Tomography (CT) has been widely used to monitor airway disease as it can detect structural changes of the airways even before symptoms occur,[1] pulmonary function decline can be measured,[2] and structural changes can be observed in other imaging modalities, such as X-ray and MRI.[3,4]

CT can be used to asses several diseases, including Chronic Obstructive Pulmonary Disease (COPD),[5] Cystic Fibrosis (CF),[1,6,7] and Common Variable Immune Deficiency (CVID).[8] Such diseases lead to structural abnormalities that can be observed on CT, including bronchiectasis (an abnormal widening of the airway) and airway wall thickening.[6,8,9]

Paper published in Medical Physics, http://dx.doi.org/10.1118/1.4963214

Several visual scoring systems exist[9–13] that semiquantitatively describe wall thickening and bronchiectasis. However these systems are prone to large inter- and intra-observer variations and are likely to ignore subtle local abnormalities.[11,13,14]

Alternatively, it is possible to manually quantify bronchiectasis and airway wall thickening by measuring each individual airway branch size. However, manual annotation is a very laborious task that might require several hours per subject, even when state-of-the-art semi-automatic tools are used. There is therefore a need to automate extraction of airway measurements from volumetric chest CT.

Several publications have reported automatic extraction of measurements related to wall thickening and bronchiectasis, including: wall area percentage,[15,16] wall thickness to lumen diameter ratio,[17,18] and outer airway diameter and area.[5,17] A disadvantage of such measurements is that they depend on the size of the airway, and therefore are not comparable across the lung: an abnormally wide airway with an abnormally thick wall will look like a healthy large airway in terms of outer and lumen diameter and wall-to-lumen ratio. Pi10 (the square root of wall area estimated for an airway with 10 mm lumen perimeter)[19] has also been proposed, and while it does not suffer of such limitation, it does not allow to quantify individual airway branches.

The most common solution to deal with measurements that depend on airway size is to group airways per generation, or group of generations, and obtain a summarising measurement (e.g. the mean) for each group, exploiting the fact that airways from the same generation have similar properties and sizes. This approach has been applied to manually[5,20,21] and automatically extracted measurements.[16,17,22–24] However, airway generation labelling is a non-trivial problem that remains unsolved and that is prone to introduce errors in the system. State-of-the-art methods mislabel airways in up to 20% of cases when the airway segmentation is accurate and complete, but the error percentage is likely to be higher when dealing with incomplete segmentations.[25,26]

An alternative approach is to normalise airway properties using the diameter (or the area) of the accompanying artery. This idea relies on the fact that, in a healthy subject, airways and their accompanying artery are expected to be of similar size.[27] This approach is more in line with standard clinical practice, where radiologists define bronchiectasis as an airway-artery ratio larger than 1,[27] and it has been widely used to characterise bronchiectasis using manual annotations.[6,11,28,29]

Few publications, however, automatically extract the accompanying artery size. Mumcuoglu *et al.*[30] developed a method to quantify airway-artery ratios in 2D CT slices, but their method requires the user to locate the centre of the airway and artery for each measurement. Odry *et al.*[31,32] automatically paired airway branches with their accompanying artery to extract wall thickness and bronchiectasis measurements normalised using artery dimension in 3D CT. However the validation was limited and only correlations with subjective visual inspection in 8 diseased and one healthy subject were reported. Fetita *et al.*[33] automatically quantified airway-artery ratios on asthmatic patients in 3D CT relying on simple geometric rules. Namely, each airway branch was paired with its closest artery, and all airway-artery pairs with a difference in orientation over 60 degrees or pairs where the airway was more than twice the size of the artery were discarded.

This paper presents a fully-automated method that combines, extends and validates pre-existing techniques to produce a fully automated pipeline that, given a 3D thoracic chest CT, segments the airway and vascular trees, measures their diameters, pairs airway branches with their accompanying artery branch, and quantifies airway wall thickening and bronchiectasis.

The contributions of this paper are: 1) Adaptation of the optimal front method,[16] originally presented to segment airways, to segment the vascular tree; 2) Introduction of a novel airway-artery pairing method; 3) Combination of different techniques to produce a fully automated pipeline that, given a 3D chest CT, quantifies wall thickening and bronchiectasis in a fully automated manner; and 4) Extensive evaluation of the methodology using 8,000 manual annotations of airway-artery pairs.

II. MATERIALS AND METHODS

II.A. Data

Inspiratory scans of 24 paediatric subjects (12 diseased and 12 controls) were retrospectively included from the Sophia-Erasmus MC cohort. The diseased subjects included 11 Cystic Fibrosis (CF) patients and 1 patient with Common Variable Immune Deficiency (CVID). Both diseases lead to recurrent lung infections that induce airway structural abnormalities including bronchiectasis and airway wall thickening.[8,34] The 12 controls were patients who underwent thorax CT scanning for diagnostic purpose and with a normal lung assessment from CT reported by the attending radiologist and a second independent radiologist blinded to patient information. Both groups were gender and age matched: age range 6 to 17 yo, 5 females. Scanning was performed using spirometry-guidance in a Siemens SOMATOM Definition Flash scanner. Similar kernel reconstructions were used for all scans: I70f/3, B75f, and B70f. Slice thickness ranged between 0.75 and 1 mm with slice spacing from 0.3 to 1 mm.

A trained observer (W. Kuo) manually annotated the 5024 visible airway-artery pairs in the dataset. First, she traced approximated centrelines of the entire airway tree using specialised software (Myrian, Intrasense, France). Second, for each branch, a cross-sectional plane perpendicular to the extracted centreline was reconstructed mid-way between bifurcations, and 3 ellipses were traced around the airway lumen, outer airway wall, and adjacent

artery (See Figure 1) and the diameter of each structure was computed. Airway generation was also annotated. This process took ∼8 hours per subject. A second observer (H. Ozturk) fully annotated a random subset of 15 subjects (2996 airway-artery measurements). HO also used the CFCT-score system (cystic fibrosis in computed tomography)[14] to semiquantitatively asses the severity and extent of bronchiectasis and wall thickening in all images.

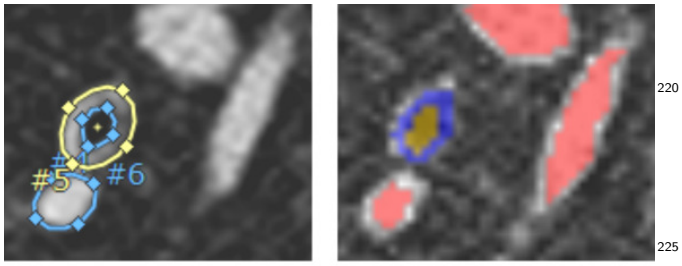

FIG. 1 a) Manual and b) automatic delineation of airway lumen, outer airway and artery.

II.B. Methodology

II.B.1. Overview of the complete pipeline

First, lung segmentation was performed and a multi-scale Hessian Eigen analysis approach was used to obtain an initial segmentation of the vascular tree (Section II.B.2). Initial segmentation of the airway centrelines was obtained using a classifier approach (Section II.B.3). Airway and vessel segmentations were refined using an optimal surface graph-cut method to obtain the boundary of the airway lumen, airway outer wall, and vessel with subvoxel precision (Section II.B.4). Centrelines were extracted and diameter measurements, orientation, and location were used to pair airway branches with their accompanying artery branch (Section II.B.5). Measurements for bronchiectasis and wall thickening were then extracted for all airway-artery pairs (Section II.B.6).

II.B.2. Lung and vasculature extraction

To confine the algorithms to the lung region, we extracted the lung fields, trachea and main bronchi using thresholding and morphological smoothing.[35] Vessels in the lungs were segmented using a multi-scale (7 scales; 1 to 12 mm; exponential scale) Hessian Eigen analysis approach,[35] that detects elongated tubular structures. We set the tubeness parameters $T_1 = T_2 = 0.4$, and the contrast parameter $T_\omega = 0.5$. These parameters are slightly less restrictive than the ones in the original work, as false positives will likely not be paired with an airway and thus they will be discarded. In order to simplify the topology, a region growing approach was applied to add connected bright voxels (> -400 HU, Hounsfield units) that are within 1 mm from the resulting segmentation. Isolated regions smaller than 25 voxels were discarded.

II.B.3. Initial airway lumen centreline extraction

An initial estimate of airway lumen centreline regions was obtained in two steps. First, for each voxel, a classifier was used to estimate its probability of being airway lumen centreline. Second, the probability maps were binarised using a vessel-guided region growing.

This work is based on Lo *et al.*[35] where lumen and non-lumen voxels were separated using a classifier approach, but we define the classes as lumen-centreline and non-lumen-centreline instead. Using the centrelines instead of the whole lumen has two main advantages: 1) All airways contribute to the classifier proportionally to their length, avoiding a bias towards larger airways. 2) Lumen voxels located close the airway wall might look similar to non-lumen voxels located immediately outside the airway wall. By not using such voxels as positive samples, it is easier for the classifier to separate the two classes.

A k-nearest neighbour classifier (KNN)[36] was used with the following features: spatial derivatives up to and including second order; Eigenvalues of the Hessian matrix; determinant and trace of the Hessian matrix; combinations of Hessian eigenvalues that measure tube, plate, and blobness. Features were extracted at 7 scales (1 mm to 12 mm; exponential scale) and standardised to zero mean and unit variance. The optimal set of image descriptors was determined using sequential floating forward feature selection.[37]

A vessel-guided region growing[35] was then performed to obtain a segmentation of the airway lumen centreline region (Figure 2d). This method relies on a vessel orientation similarity measure, which indicates how parallel an airway is to its neighbouring vessel, to overcome regions in the airway tree that have a low classifier posterior probability. Starting from the trachea, any neighbouring voxel with a posterior probability $p(x) > t_u$, or $p(x) > t_l$ and a vessel vessel orientation similarity score $s > t_s$ is added to the segmentation. Implementation details and parameter settings are further explained in Section II.C.

II.B.4. Boundary detection, diameters measurement and centreline extraction

A surface graph-cut approach[16] was applied to the initial airway segmentation to obtain lumen wall and outer airway wall with subvoxel accuracy. The method, originally developed only for airways, was adapted to also detect artery wall boundaries. This method, when applied to airways, converts the initial segmentations of the airway centreline regions (the binarisation produced in Section II.B.3, Figure 2d) into a 3D graph with image intensity information from the original CT, smoothness and topology constraints encoded in their edges. Two 3D surfaces (corresponding to the lumen wall and outer

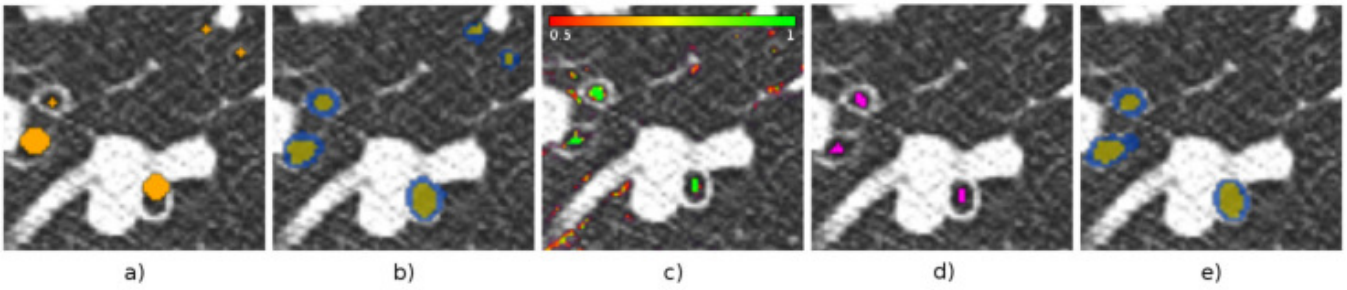

FIG. 2 **a)** Manual lumen segmentation. **b)** Automatically corrected Lumen (M_l) and outer wall (M_w) segmentations derived from the manual segmentation. **c)** Posterior probability output by the classifier. **d)** Binarised posterior probability maps (S). **e)** Final automatic lumen and outer airway segmentation.

wall) that cut the graph edges with minimal cost are then computed simultaneously.

Flow-lines, paths that depart perpendicular from the initial segmentation but do not intersect each other, are computed and used to initialise the topology of the graph. We used the first derivative of the image along the flow-lines as the edge data term to obtain the lumen surface (the method estimates a surface at the point of maximum gradient between the lumen and the wall), and the negative of the first derivative for the outer airway wall surface. Once the graph is created, an optimal surface is obtained using a min-cut max-flow algorithm[38] (see Figure 2e). We refer the reader to Petersen *et al.*[16] for a detailed explanation of the method, and its implementation.[39]

The graph-cut approach was adapted to obtain vessel boundaries based on the the initial binary segmentation obtained in Section II.B.2. The negative of the first derivative was chosen as the data term in order to find the point with maximum gradient between the bright vessel and the darker parenchyma.

The graph-cut produced 3D mesh surfaces that were converted back to isotropic 3D voxelised volumes. Subsequently airway and vessel centrelines and airway generations (being 0 at the trachea and increased by one at each bifurcation) were extracted with a front propagation method.[24] The front propagation for airways started at the trachea (generation 0). Because the vessel segmentation contained disconnected regions, the starting point for the front propagation for each region was set as the closest point to the centre of mass of the vessel segmentation, which roughly corresponded to the heart. The radii of airway lumen, outer airway wall, and artery were computed for each centreline point (every 0.5 mm) by averaging the distance from each vertex of the corresponding segmented mesh to the closest centreline point. To reduce the effect of spurious measurements, radii values were smoothed along the centrelines using a Gaussian function of $\sigma = 1$ mm. The local orientation was computed for each centreline point corresponding to the orientation of a centreline section of 5 mm centred on the point being evaluated. Airways and vessels under 5 mm were discarded.

In order to cope with differences in gender, age, weight, and height, all centrelines and associated diameter measurements were isotropically rescaled according to the predicted FVC (forced vital capacity) based on the patients gender, height and age.[40] Consequently, all reported measurements correspond to a lung with a predicted FVC of 2.859 litres, the average predicted FVC of the used dataset, roughly equivalent to a 10 year old male subject 128 cm tall.

II.B.5. Airway-artery pairing

Once the airway and artery branches were segmented and quantified, the method detected airway-artery pairs based on similarity in orientation, proximity and size. Equation (1) was used to estimate airway-artery pairing scores $s(b_p, a_q)$ at each centreline point b_p of each airway branch b with the closest artery centreline point a_q of each artery branch a. The range of the score was $s(b_p, a_q) = [0, 1]$.

$$s(b_p, a_q) = s_o(b_p, a_q) \cdot s_d(b_p, a_q) \cdot s_s(b_p, a_q) \tag{1}$$

where an orientation score $s_o = [0, 1]$ was computed given the angle $\theta(b_p, a_q) = [0, \pi/2]$, between the local orientations of the airway point b_p and the artery point a_q, as:

$$s_o(b_p, a_q) = 1 - \frac{\theta(b_p, a_q)}{\pi/2} \tag{2}$$

A distance score $s_d(b_p, a_q)$ was used to favour airway-artery pairs that were close together using Equation (3):

$$s_d(b_p, a_q) = \frac{m - (\|b_p - a_q\| - r_o(b_p) - r(a_q))}{m} \tag{3}$$

where the locally estimated radii of airway outer wall, $r_o(b_p)$, and artery, $r(a_q)$, were subtracted from the distance between the candidate pair points in order to estimate the distance between borders. The constant m is the maximum distance allowed between borders in millimetres. s_d was clipped between 0 and 1. Finally, size

similarity was also taken into consideration using Equation (4):

$$s_s(b_p, a_q) = \begin{cases} \left(\frac{r_o(b_p)}{r(a_q)}\right)^{\frac{1}{c}}, \text{ if } r_o(b_p) < r(a_q) \\ \left(\frac{r(a_q)}{r_o(b_p)}\right)^{\frac{1}{c}}, \text{ otherwise} \end{cases} \quad (4)$$

where airway-artery pairs with similar diameter were favoured. The scalar c limits the influence of the size similarity. A larger c allows to match airways and arteries despite having very different diameters due to disease.

The most reliable location to find the accompanying artery of an airway branch is away from its branchings. Thus, for each airway branch b with P centreline points, the $P/2$-th point was the first to be attempted to pair. If a (b_p, a_q) pair with $p = P/2$ and $s(b_p, a_q) \geq s_t$ was found, it was stored as the accompanying artery, precluding the evaluation of other pairing points for the airway branch. If $s(b_p, a_q) < s_t$ for all candidate a_q, $P/2-1$ and $P/2+1$ were then evaluated. All possible b_p points were evaluated in this middle-to-the-border fashion until a pair was found or no more points were left to explore in the airway branch.

II.B.6. Quantification of airway structural properties

Using the diameter (twice the computed radius) for airway lumen (d_l), outer airway (d_o) and the accompanying artery (d_a) at the airway-artery pairing points, the following structural airway properties related to wall thickening and bronchiectasis were computed for each airway-artery pair:

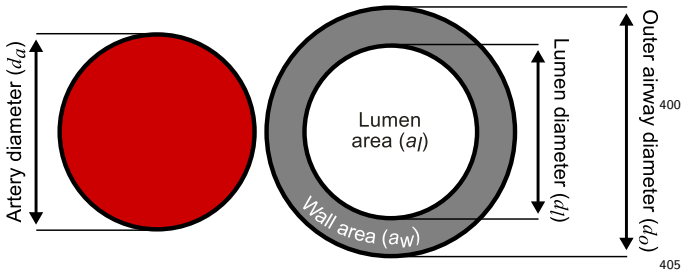

FIG. 3 Diameter and area measurements used to extract the airway structural measurements.

Wall Thickness (WT): The difference between outer airway diameter and airway lumen, divided by 2.

$$WT = \frac{d_o - d_l}{2} \quad (5)$$

Wall Thickness Ratio (WTR): Wall thickness divided by the outer diameter.

$$WTR = \frac{WT}{d_o} = \frac{(d_o - d_l)/2}{d_o} \quad (6)$$

Wall Area Percentage (WAP): The percentage of the total airway area that is airway wall.

$$WAP = \frac{a_w}{a_w + a_l} \cdot 100 \quad (7)$$

Wall-Artery Ratio (WAR): The difference between outer airway diameter and airway lumen, divided by the diameter of its accompanying artery.

$$WAR = \frac{d_o - d_l}{d_a} \quad (8)$$

Airway-Artery Ratio (AAR): The ratio between airway diameter and artery diameter, computed separately for airway lumen and outer airway as:

$$AAR_l = \frac{d_l}{d_a}, \quad AAR_o = \frac{d_o}{d_a} \quad (9)$$

II.C. Implementation details and parameters setting

This section details the most relevant implementation choices and setting of parameters for the presented methodology. For the remainder of the paper, we will consider that an automatically extracted airway (or vessel) branch corresponds to a manual measurement if the distance between the airway (or vessel) centreline and the centre of the manual annotation is smaller than the manually annotated radius.

II.C.1. Initial airway lumen extraction (II.B.3)

Approximate lumen centreline segmentations were available for the full dataset. However, these were only accurate in orientation and not in position, as they were intended to be used only to create the cross-sectional images for the manual annotations (Figure 2a). To use such data to train the classifier, the surface graph-cut approach described in Section II.B.4 was used to obtain more accurate segmentations of the airway lumen (M_l) and airway wall (M_w), as seen in Figure 2b. Airway centrelines were automatically obtained using a front propagation method[22] on M_l. All voxels traversed by this centreline path formed the centreline mask (M_c).

All voxels from M_c were used for the lumen-centreline class. For the non-lumen-centreline class, voxels located in the airway wall and immediately outside the airway were considered. The airway wall (M_w) was dilated outward by 5 mm and a random selection of voxels were used as the non-lumen-centreline samples, thus using equal number of positive and negative samples.

A leave-4-out cross-validation approach was implemented to avoid using the same data for training and testing: Sets of 4 scans (2 scans from diseased patients, 2 from controls) were separated from the other 20. A KNN classifier, with 21 nearest neighbours (Section II.B.3) was then trained using these 20 scans and evaluated on the other 4 not previously seen by the algorithm. The process was repeated 6 times to evaluate the 24 scans.

Different values of parameters t_u, t_l, and t_s were tested for lumen centreline extraction. Resulting segmentations (S) were evaluated for airway centreline completeness c and volume leakage (l). Completeness was computed as the percentage of manually extracted centreline voxels

M_c found inside the segmented lumen region S, and volume leakage as the percentage of voxels from S found outside the corrected manual segmentations of lumen (M_l) and airway wall (M_w). The trachea and the two main bronchi were excluded from this evaluation. In general, with lower threshold values the resulting segmentations were more complete but contained a larger portion of false positives. We opted for the set of parameters that provide the most complete tree while keeping l under 2.5%. These parameters were $t_u = 18/21$, $t_l = 16/21$, and $t_s = 0.5$, resulting in $c = 60.5\%$ and $l = 2.4\%$.

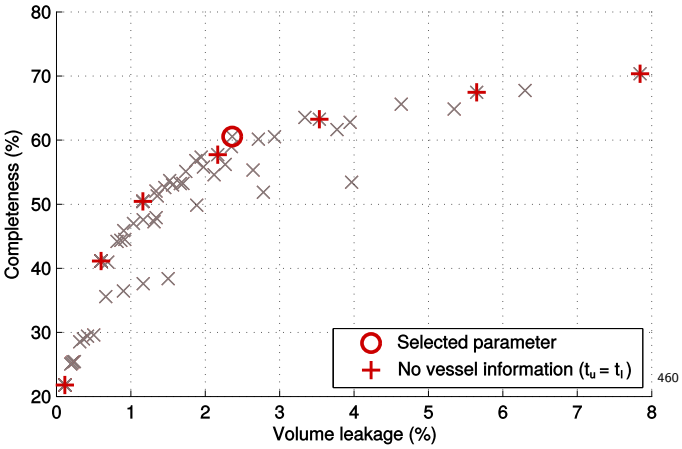

FIG. 4 Vessel-guided segmentation results with different parameters.

II.C.2. Boundary detection (II.B.4)

Flow-line length were restricted to 1 mm inwards and 10 mm outwards, with 0.3 mm spacing between nodes for airways, and 1 mm inward and 5 outward, with 0.4 mm spacing between nodes (due to memory restrictions) for arteries. We refer the reader to Petersen *et al.*[16] for a full explanation of the parameters.

Different smoothness penalty and smoothness constraint values (denoted as p_m and δ)[16] were tested independently for lumen, outer airway and artery segmentations. Spearman correlation coefficient was used to assess agreement between the manual diameter measurements and the the diameters derived from the graph-cut segmentation with the different parameters.

Figure 5 shows that the different values had a very limited influence on the correlation coefficients. $p_m = 64$ and $\delta = 3$ were selected for the airway lumen wall ($p_m = 64$ and $\delta = 1$ for outer airways, and $p_m = 4$ and $\delta = 3$ for vessel wall.

II.C.3. Airway-artery pairing threshold (II.B.5)

In order to limit the influence of similarity in size between airway and artery in Eq. 4, $c = 3$ was used. The value of m in Eq. 3 was set to 20 (mm).

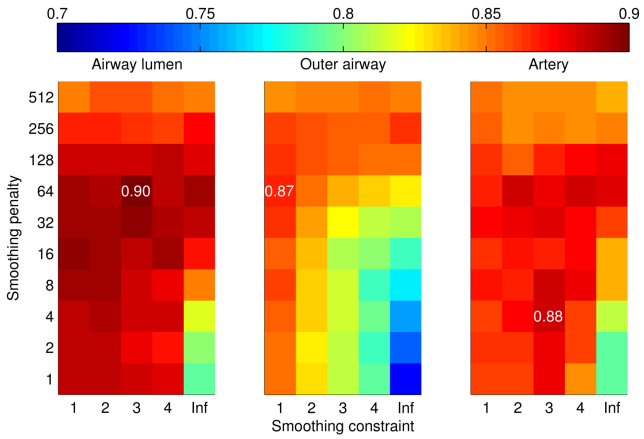

FIG. 5 Spearman correlation coefficients between automatic and manual annotations of diameters with different smoothing constraints and penalties. Selected parameters are marked with their corresponding coefficient.

To assess the quality of the airway-artery pairs we measured the Spearman correlation coefficients (SCC) between manually obtained and automatically obtained diameters at the automatically obtained pairing points b_p and a_q.

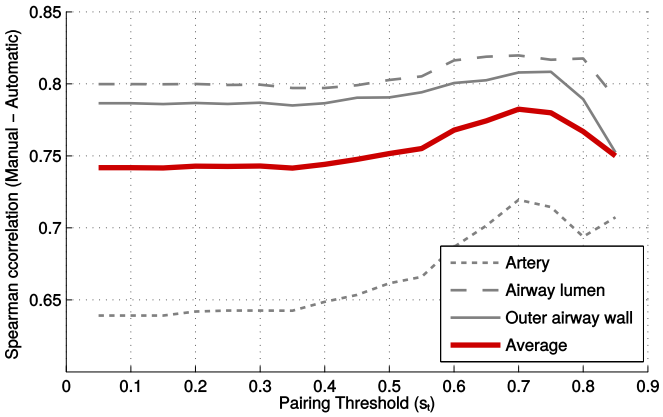

FIG. 6 Spearman correlation coefficient between automatic and manual annotations for different s_t.

While there is little variation on SCC for different s_t values (Figure 6), $s_t = 0.7$ produced the best results in terms of SCC, and thus it was the value selected for the remainder of the paper.

III. RESULTS

III.A. Agreement with manual annotations

To validate diameter measurements and airway-artery ratios (WAR and AAR), the non-parametric *bias* (median difference) and 95% *limits of agreement* (2.5^{th} to 97.5^{th} percentiles) between observers and between the automatic method and each observer was computed on the 15 subjects that had annotations from both observers. Because annotations from each observer were

obtained at different locations, only automatically extracted branches that were associated with one manual annotation for each observer were used, thus guaranteeing that both manual annotations corresponded to the same branch. 1330 automatically extracted airway branches and 1444 artery branches fulfilled that criteria.

From the 1330 automatically extracted airway branches with ground truth from both observers, the proposed method paired 1039 (78.9%) with an artery and failed to pair 291 (21.9%).

TABLE I Bias and limits of agreement for diameter measurements (in mm)

	Bias			*Limits of agreement*		
	d_l	d_o	d_a	d_l	d_o	d_a
Obs. 1 vs. Obs. 2	-0.040	0.166	0.013	0.678	1.039	0.853
Auto vs. Obs. 1	-0.181	-0.348	-1.025	0.501	0.836	0.884
Auto vs. Obs. 2	-0.144	-0.520	-1.039	0.545	0.808	0.929
Avg. Auto vs. Obs.	-0.162	-0.434	-1.032	0.523	0.822	0.906

Table I shows a small bias on diameter measurements between observers (-0.04 to 0.17 mm), and relatively large limits of agreement ranging between 0.68 and 1.04 mm. For automatic diameters compared with both observers, we observed an increased bias of up to 1.04 mm but similar limits of agreement (0.52 to 0.91 mm). The automatic method systematically measures smaller diameters than human observers. This might be explained by the observers delineating the artery and airway boundaries at the point where the wall is no longer visible (immediately outside the wall), while the surface graph-cut approach aimed to find the highest intensity gradient and thus measuring smaller diameters (see Figure 1). However, the limits of agreement between automatic and manual are comparable to between observers, indicating that the quality of the automatic measurements is similar to manual measurements.

TABLE II Bias and limits of agreement for airway-artery ratios

	Bias			*Limits of agreement*		
	AAR_l	AAR_o	WAR	AAR_l	AAR_o	WAR
Obs. 1 vs. Obs. 2	-0.010	0.045	0.056	0.197	0.335	0.250
Auto vs. Obs. 1	0.059	0.176	0.107	0.323	0.718	0.463
Auto vs. Obs. 2	0.066	0.136	0.056	0.332	0.711	0.473
Avg. Auto vs. Obs.	0.063	0.156	0.082	0.327	0.715	0.468

Table II shows that AAR and WAR limits of agreement were larger between the automatic and the manually obtained ratios than between observers. This is mainly a consequence of the systematic bias between manually and automatically obtained artery diameters: Because d_a is used as a denominator to calculate AAR and WAR, the uniform bias in d_a (as observed in Table I) leads to larger non-uniform bias and limits of agreement in AAR and WAR.

Finally, manually labelled airway generations, only available from observer 1, were compared to automatically obtained generations for the 2380 airway-artery pairs that were matched with a manual annotation. 47.69% of airways had the same generation assigned manually and automatically, 10.42% had been automatically labelled with a later generation than manually, and 41.89% with a smaller one.

III.B. Differentiation between diseased and control groups

The proposed methodology detected and quantified 3380 airway branches longer than 5 mm. From these, 2298 (68.0% of all segmented airway branches) were automatically paired with an artery, and their WAR, WAP, WT, AAR_l, and AAR_o were computed.

Statistical significance of differences in mean WAR, WAP, WT, AAR_l, and AAR_o per subject between diseased and controls were analysed using a student t-test, as subjects of each group were selected to be of similar conditions, thus Gaussianity on the means distribution was assumed.

TABLE III Wall-thickness measurements sorted by Student t-test significance with mean (and SD) for each group

	Diseased	Controls	p-value
WAR	0.71 (0.06)	0.60 (0.05)	0.000151
WAR (*Manual*)	0.62 (0.08)	0.52 (0.05)	0.002601
CFCT-score WT	9.26 (10.20)	0.00 (0.00)	0.004715
WAP	76.66 (3.28)	73.31 (6.08)	0.106996
WTR	0.26 (0.02)	0.25 (0.03)	0.136019
WT (*Manual*)	1.06 (0.11)	1.01 (0.18)	0.391342
WAP (*Manual*)	77.35 (4.10)	75.80 (5.42)	0.438866
WTR (*Manual*)	0.27 (0.02)	0.26 (0.03)	0.466496
WT	1.06 (0.09)	1.03 (0.23)	0.653551

Table III shows all measurements related to wall thickening sorted by p-value with mean values and standard deviation. Also included are values for the manual measurements and the CFCT-score for wall thickening. Only CFCT-score and mean WAR (both manual and automatic) captured statistically significant differences ($p < 0.005$) between the diseased and control populations. Measurements (automatic and manual) that are not normalised using the accompanying artery diameter (WT, WAP and WTR) did not show statistically significant differences between diseased and controls.

Similarly, Table IV lists bronchiectasis measurements per patient sorted by p-value. AAR_o (manual and automatic) and visual scores for bronchiectasis were the only measurements significantly different ($p < 0.05$) between diseased and controls.

Notice that the manual and automatic measurements lead to the same results: Only WAR and AAR_o, which rely on the airway-artery pairing step, captured significant differences between groups, while all measurements that do not rely in airway-artery pairing (WAP, WTR,

TABLE IV Bronchiectasis measurements sorted by Student t-test significance with mean (and SD) for each group

	Diseased	Controls	p-value
AAR_o (*Manual*)	1.17 (0.14)	1.02 (0.05)	0.002267
AAR_o	1.33 (0.12)	1.21 (0.10)	0.014233
CFCT-score BR	12.30 (18.60)	0.81 (1.51)	0.044432
d_l	2.07 (0.28)	2.27 (0.37)	0.154980
AAR_l (*Manual*)	0.55 (0.09)	0.49 (0.08)	0.135996
d_o (*Manual*)	4.03 (0.26)	3.95 (0.35)	0.504185
d_o	4.20 (0.36)	4.33 (0.55)	0.494872
AAR_l	0.63 (0.09)	0.61 (0.11)	0.636810
d_l (*Manual*)	1.92 (0.23)	1.94 (0.25)	0.839547

WT, d_l, d_o), plus AAR_l, did not capture any statistical differences between groups.

In subjects with bronchiectasis, some of the most peripheral airways that would be too small to be detected in CTs of controls, have become big enough to be observed and analysed. This leads to a selection bias where the number and distribution of airways might differ between diseased and controls, affecting the average measurements per patient. When airways were analysed per generation, the diseased population had more airways than controls ($p < 0.05$) for all generations ≥ 6, but not for earlier generations. When only measurements of central airways (generation ≤ 5) were averaged, no statistically significant differences were found between groups using any of the evaluated automated measurements, but manual WT, WAR, d_o, and AAR_o were significantly larger for diseased than controls ($p < 0.05$). When peripheral airways from generation ≥ 6 were averaged together, diseased patients showed a higher automatically extracted WAR ($p = 0.0077$) and AAR_o ($p = 0.0237$) than controls, but WAP, WTR, WT, d_l, d_o, AAR_l were not significantly different between groups. Similarly, for manual annotations on these airways, only WAR ($p = 0.0034$) and AAR_o ($p = 0.0013$) were significantly different between groups.

III.C. Robustness of the method to sub-optimal segmentations

To test the ability of the studied measurements to capture differences between groups when derived from poor segmentations that include numerous false positives and negatives, sub-optimal airway segmentations were generated by using 9 different thresholds ($t_u = t_l = [\frac{10}{21}..\frac{18}{21}]$, Section II.B.3). Due to the limited improvements offered by using vessel information in the region growing (Figure 4), no vessel information was used ($t_u = t_l$) for simplicity. The entire pipeline was then reran on the different initial segmentations.

Figure 7 shows that differences between groups in WAR and AAR_o remain significant up to when ~30% of the initial airway segmentation were false positives.

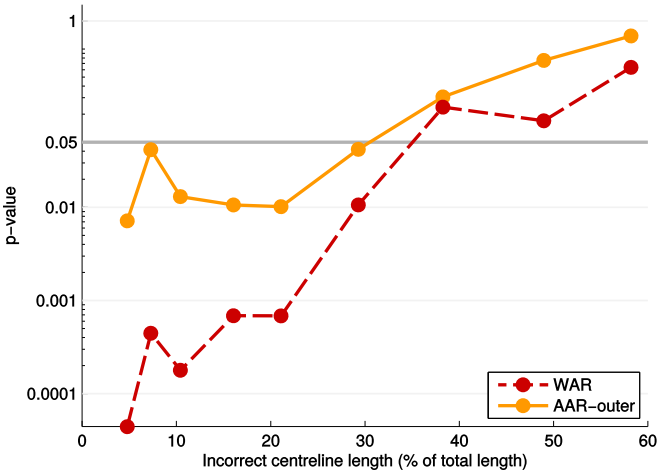

FIG. 7 Statistical significance of differences between controls and diseased using WAR and AAR_o obtained on segmentations with different percentage of incorrectly segmented airways.

IV. DISCUSSION

In this paper we have presented a fully automated pipeline that segments, measures, and pairs airway and artery branches in order to obtain measurements of bronchiectasis and wall thickening that are normalised using the artery size.

The limits of agreement between automatically and manually measured diameters were similar to these between manual observations (Table I). Only measurements normalised by artery size (AAR_o and WAR) were significantly different between diseased and controls, while measurements that are based on the airway properties alone (WT, WAP, WTR, d_l, d_o), plus AAR_l, did not show significant differences between groups. Differences in wall thickening and bronchiectasis between groups were similar when using manual measurements than the automatic ones (Tables III and IV).

The fact that AAR_l, a commonly used biomarker to asses airway bronchiectasis,[41] did not show significant differences between diseased and controls is in line with recent publications,[17] where no significant differences in lumen area were observed between controls and CF patients with visible bronchiectasis signs, while differences in outer airway area were significant up to the 9th generation in adults and in generations 4 and 5 in children. This suggests that AAR_o is better suited than AAR_l to quantify airway bronchiectasis.

The proposed method is an alternative to airway analysis per generation (or group of generations), as it computes average measurements over all airways. Airway generation labelling is a step likely to introduce errors. Methods for generation and branch labelling have been reported to achieve around 90% accuracy on healthy subjects[25] and as high as 80% on diseased subjects,[26] when the algorithm is fed with a complete and accurate airway segmentation obtained manually. The use of less accurate segmentations or less efficient methods might

reduce that accuracy even further. In this paper, where the segmentations contained numerous false positives and negatives, and the generation labelling was obtained via a simple method based on bifurcation detection, less than half the airway branches were assigned the same generation automatically than manually.

The airway segmentations obtained automatically with the proposed methodology missed 40% of the airway tree length compared to manually obtained segmentations. This compares well to other state-of-the-art automatic airway segmentation methods (Table IV in Lo *et al.*,[24] where none of the 15 methods evaluated extracted more than 60% of the tree length while keeping false positives under 2.5%).

An unexpected result from this work is that the use of vessel information did not improve airway segmentation results (Figure 4). This might be because the vessel information helps to overcome low posterior probability regions away from bifurcation points (where the vessel orientation is reliable) while the classifier used in this paper generally outputs high posterior probabilities in such regions. Most of the missed airway branches occur as a consequence of missing the bifurcation points, where the computed vessel orientation is not reliable and thus the use of vessel orientation information offers no improvement.

Airway-artery pairing remains a challenging problem. The proposed methodology failed to pair correctly segmented airways that were annotated by a manual observer in 21.9% of the cases. This compares well with the only paper known to the authors where automatic airway-artery pairing is validated.[32] In that study, two different observers paired the same airway to different arteries in 34.4% of the cases, and their automated method was reported to obtain a wrong airway-artery pair in 24.7% of the branches.

V. CONCLUSIONS

This paper presents a fully-automatic method to quantify wall thickening and bronchiectasis in individual airway branches using the artery size as a normalising factor. The automatically obtained values separate diseased from bronchiectasis-free controls as well as manual annotations or visual scoring systems. Both, manually and automatically obtained measurements lead to the same clinical conclusions: only measurements that are normalised by artery size separate disease from controls, highlighting the importance of the airway-artery pairing step.

The reported results and the the robustness of the method to poor and incomplete segmentations suggest that the proposed methodology could substitute time-consuming manual annotations and visual scoring systems to quantify airway wall thickening and bronchiectasis, allowing to obtain objective measurements on lung CTs that could not be measured manually due to time and financial limitations.

ACKNOWLEDGMENTS

The authors would like to thank Hadiye Ozturk for the additional manual annotations. This research was performed with funding from the Netherlands Organization for Scientific Research (NWO), Vertex, the European Cystic Fibrosis Society, and the Danish Council for Independent Research.

DISCLOSURE OF CONFLICTS OF INTEREST

The authors have no relevant conflicts of interest to disclose.

[1] P. D. Sly, S. Brennan, C. Gangell, N. De Klerk, C. Murray, L. Mott, S. M. Stick, P. J. Robinson, C. F. Robertson, and S. C. Ranganathan, "Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening," American Journal of Respiratory and Critical Care Medicine **180**, 146–152 (2009).

[2] P. de Jong, Y. Nakano, M. Lequin, J. Mayo, R. Woods, P. Paré, and H. Tiddens, "Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis," European Respiratory Journal **23**, 93–97 (2004).

[3] K. S. Lee, S. L. Primack, C. A. Staples, J. R. Mayo, J. E. Aldrich, and N. L. Muller, "Chronic infiltrative lung disease: comparison of diagnostic accuracies of radiography and low- and conventional-dose thin-section CT." Radiology **191**, 669–673 (1994).

[4] C. Sileo, H. Corvol, P. Y. Boelle, E. Blondiaux, A. Clement, and H. Ducou Le Pointe, "HRCT and MRI of the lung in children with cystic fibrosis: Comparison of different scoring systems," Journal of Cystic Fibrosis **13**, 198–204 (2014).

[5] M. Hackx, A. A. Bankier, and P. A. Gevenois, "Chronic Obstructive Pulmonary Disease: CT Quantification of Airways Disease," Radiology **265**, 34–48 (2012).

[6] L. S. Mott, K. G. Graniel, J. Park, N. H. de Klerk, P. D. Sly, C. P. Murray, H. A. W. M. Tiddens, and S. M. Stick, "Assessment of early bronchiectasis in young children with cystic fibrosis is dependent on lung volume," Chest **144**, 1193–1198 (2013).

[7] L. A. Tepper, E. M. W. J. Utens, D. Caudri, A. C. Bos, K. Gonzalez-Graniel, H. J. Duivenvoorden, E. C. W. Van Der Wiel, A. L. Quittner, and H. A. W. M. Tiddens, "Impact of bronchiectasis and trapped air on quality of life and exacerbations in cystic fibrosis," European Respiratory Journal **42**, 371–379 (2013).

[8] L. J. Maarschalk-Ellerbroek, P. A. de Jong, J. M. van Montfrans, J. W. J. Lammers, A. C. Bloem, A. I. M. Hoepelman, and P. M. Ellerbroek, "CT Screening for Pulmonary Pathology in Common Variable Immunodeficiency Disorders and the Correlation with Clinical and Immunological Parameters," Journal of Clinical Immunology **34**, 642–654 (2014).

[9] A. Brody, J. Klein, P. Molina, J. Quan, J. Bean, and R. Wilmott, "High-resolution computed tomography in young patients with cystic fibrosis: distribution of abnormalities and correlation with pulmonary function tests." The Journal of pediatrics **145**, 32–38 (2004).

[10] M. Bhalla, N. Turcios, V. Aponte, M. Jenkins, B. S. Leitman, D. I. McCauley, and D. P. Naidich, "Cystic fibrosis: scoring system with thin-section CT." Radiology **179**, 783–788 (1991).

[11] P. de Jong, M. Ottink, S. Robben, M. Lequin, W. Hop, J. Hendriks, P. Paré, and H. Tiddens, "Pulmonary disease assessment in cystic fibrosis: comparison of CT scoring systems and value of bronchial and arterial dimension measurements." Radiology **231**, 434–439 (2004).

[12] A. A. J. M. van de Ven, J. M. van Montfrans, S. W. J. Terheggen-Lagro, F. J. Beek, D. P. Hoytema van Konijnenburg, O. A. M. Kessels, and P. A. de Jong, "A CT scan score for the assessment of lung disease in children with common variable immunodeficiency disorders," Chest **138**, 371–379 (2010).

[13] O. M. Mets, E. J. Smit, F. a. a. Hoesein, H. a. Gietema, R. P. H. Bokkers, M. Attrach, S. van Amelsvoort-van de Vorst, E. T. Scholten, C. F. M. Buckens, M. Oudkerk, J. W. J. Lammers, M. Prokop, and P. a. de Jong, "Visual versus automated evaluation of chest computed tomography for the presence of chronic obstructive pulmonary disease," PLoS ONE **7**, 1–7 (2012).

[14] L. A. Tepper, D. Caudri, E. M. W. J. Utens, E. C. van der Wiel, A. L. Quittner, and H. A. W. M. Tiddens, "Tracking CF disease progression with CT and respiratory symptoms in a cohort of children aged 6-19 years," Pediatric Pulmonology **1189**, 1182–1189 (2014).

[15] M. Schmidt, J. Kuhnigk, S. Krass, M. Owsijewitsch, B. de Hoop, and H. Peitgen, "Reproducibility of airway wall thickness measurements," in *SPIE, Medical Imaging: Computer-Aided Diagnosis*, Vol. 7624 (2010) pp. 76241O–76241O–10.

[16] J. Petersen, M. Nielsen, P. Lo, L. H. Nordenmark, J. H. Pedersen, M. M. W. Wille, A. Dirksen, and M. de Bruijne, "Optimal surface segmentation using flow lines to quantify airway abnormalities in chronic obstructive pulmonary disease." Medical image analysis **18**, 531–41 (2014).

[17] M. O. Wielpütz, M. Eichinger, O. Weinheimer, S. Ley, M. a. Mall, M. Wiebel, A. Bischoff, H.-U. Kauczor, C. P. Heuel, and M. Puderbach, "Automatic Airway Analysis on Multidetector Computed Tomography in Cystic Fibrosis: Correlation With Pulmonary Function Testing." Journal of thoracic imaging **28**, 104–113 (2013).

[18] C. Fetita, P.-Y. Brillet, R. Hartley, P. A. Grenier, and C. Brightling, "3D mapping of airway wall thickening in asthma with MSCT: A level set approach," in *SPIE, Medical Imaging: Computer-Aided Diagnosis*, Vol. 9035 (2014).

[19] Y. Nakano, J. C. Wong, P. A. De Jong, L. Buzatu, T. Nagao, H. O. Coxson, W. M. Elliott, J. C. Hogg, and P. D. Paré, "The prediction of small airway dimensions using computed tomography," American Journal of Respiratory and Critical Care Medicine **171**, 142–146 (2005).

[20] R. Lambert, R. Castile, and R. Tepper, "Model of forced expiratory flows and airway geometry in infants." Journal of applied physiology **96**, 688–92 (2004).

[21] A. A. Diaz, C. E. Come, J. C. Ross, R. San José Estépar, M. K. Han, S. H. Loring, E. K. Silverman, and G. R. Washko, "Association between airway caliber changes with lung inflation and emphysema assessed by volumetric CT scan in subjects with COPD." Chest **141**, 736–44 (2012).

[22] T. Schlathölter, C. Lorenz, I. C. Carlsen, and S. Renisch, "Simultaneous Segmentation and Tree Reconstruction of the Airways for Virtual Bronchoscopy," in *SPIE Medical Imaging*, Vol. 4684 (2002) pp. 103–113.

[23] M. Montaudon, P. Desbarats, P. Berger, G. de Dietrich, R. Marthan, and F. Laurent, "Assessment of bronchial wall thickness and lumen diameter in human adults using multidetector computed tomography: comparison with theoretical models." Journal of anatomy **211**, 579–588 (2007).

[24] P. Lo, B. V. Ginneken, J. Reinhardt, T. Yavarna, P. A. de Jong, B. Irving, C. Fetita, M. Ortner, R. Pinho, J. Sijbers, M. Feuerstein, A. Fabijanska, C. Bauer, R. Beichel, C. S. Mendoza, R. Wiemker, J. Lee, A. P. Reeves, S. Born, O. Weinheimer, E. M. van Rikxoort, J. Tschirren, K. Mori, B. Odry, D. P. Naidich, I. Hartmann, E. A. Hoffman, M. Prokop, J. H. Pedersen, and M. de Bruijne, "Extraction of airways from CT (EXACT'09)," (TMI) IEEE Transactions on Medical Imaging **31**, 2093–2107 (2012).

[25] K. Mori, S. Ema, T. Kitasaka, Y. Mekada, I. Ide, H. Murase, Y. Suenaga, H. Takabatake, M. Mori, and H. Natori, "Automated nomenclature of bronchial branches extracted from CT images and its application to biopsy path planning in virtual bronchoscopy," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, 8th International Conference, Proceedings, Part II* (2005) pp. 854–861.

[26] A. Feragen, J. Petersen, M. Owen, P. Lo, L. Thomsen, M. Wille, A. Dirksen, and M. de Bruijne, "Geodesic atlas-based labeling of anatomical trees: Application and evaluation on airways extracted from CT," Medical Imaging, IEEE Transactions on **34**, 1212–1226 (2015).

[27] G. McGuinness, D. Naidich, B. Leitman, and D. Mccauley, "Bronchiectasis: CT Evaluation," American Journal of Roentgenology **160**, 253–259 (1993).

[28] F. R. Long, R. S. Williams, and R. G. Castile, "Structural airway abnormalities in infants and young children with cystic fibrosis," Journal of Pediatrics **144**, 154–161 (2004).

[29] N. Kapur, J. P. Masel, D. Watson, I. B. Masters, and A. B. Chang, "Bronchoarterial ratio on high-resolution CT scan of the chest in children without pulmonary pathology: Need to redefine bronchial dilatation," Chest **139**, 1445–1450 (2011).

[30] E. U. Mumcuoglu, F. R. Long, R. G. Castile, and M. N. Gurcan, "Image analysis for cystic fibrosis: Computer-assisted airway wall and vessel measurements from low-dose, limited scan lung CT images," Journal of Digital Imaging **26**, 82–96 (2013).

[31] B. Odry, A. Kiraly, C. Novak, D. Naidich, and J. Lerallut, "Automated airway evaluation system for multi-slice computed tomography using airway lumen diameter, airway wall thickness and broncho-arterial ratio," in *SPIE, Medical Imaging: Physiology, Function, and Structure from Medical Images*, Vol. 6143 (2006).

[32] B. L. Odry, A. P. Kiraly, C. L. Novak, D. P. Naidich, and J.-F. Lerallut, "An evaluation of automated broncho-arterial ratios for reliable assessment of bronchiectasis," in *Proceedings of SPIE*, Vol. 6915 (2008) p. 69152M.

[33] C. Fetita, P.-Y. Brillet, C. Brightling, and P. A. Grenier, "Grading remodeling severity in asthma based on airway wall thickening index and bronchoarterial ratio measured with MSCT," in *SPIE, Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, Vol. 9415 (2015).

[34] P. de Jong, Y. Nakano, W. Hop, F. Long, H. Coxson, P. Paré, and H. Tiddens, "Changes in Airway Dimensions on Computed Tomography Scans of Children with Cystic Fibrosis," American Journal of Respiratory and Critical Care Medicine **172**, 218–224 (2005).

[35] P. Lo, J. Sporring, H. Ashraf, J. J. H. Pedersen, and M. de Bruijne, "Vessel-guided airway tree segmentation: A voxel classification approach." Medical Image Analysis **14**, 527–538 (2010).

[36] R. Duda, P. Hart, and D. Stork, *Pattern classification*, Pattern Classification and Scene Analysis: Pattern Classification (Wiley, 2001).

[37] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," Pattern Recognition Letters **15**, 1119–1125 (1994).

[38] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," Pattern Analysis and Machine Intelligence, IEEE Transactions on **26**, 1124–1137 (2004).

[39] Https://bitbucket.org/opfront/opfront.

[40] P. H. Quanjer, T. J. Cole, G. L. Hall, and B. H. Culver, "Mutiethnic reference values for spirometry for thee 3-95 years age range: the global lung function 2012 equations," Eur Respir J **40**, 1324–1343 (2013).

[41] J. D. Newell, *Contemporary Medical Imaging: CT of the Airways*, edited by P. M. B. Lynch and D. A. (Humana Press, Totowa, NJ, 2008) pp. 213–235.