

Computed tomography quantification of tracheal abnormalities in COPD and their influence on airflow limitation

Leticia Gallardo Estrella^{a)}

Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Nijmegen 6525 GA, The Netherlands

Esther Pompe

Department of Respiratory Medicine, University Medical Center Utrecht, Utrecht 3584 CX, The Netherlands

Jan-Martin Kuhnigk

Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen 28359, Germany

David A. Lynch

Department of Medicine, National Jewish Health, Denver, CO 80206, USA

Surya P. Bhatt

Division of Pulmonary, Allergy and Critical Care Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA

UAB Lung Health Center, University of Alabama at Birmingham, Birmingham, AL 35294, USA

UAB Lung Imaging Core, University of Alabama at Birmingham, Birmingham, AL 35294, USA

Bram van Ginneken and Eva Marjolein van Rikxoort

Department of Radiology and Nuclear Medicine, Radboud University Medical Centre, Nijmegen 6525 GA, The Netherlands

Fraunhofer MEVIS, Institute for Medical Image Computing, Bremen 28359, Germany

(Received 14 October 2016; revised 21 March 2017; accepted for publication 23 March 2017; published 22 May 2017)

Purpose: To present a method to automatically quantify tracheal morphology changes during breathing and investigate its contribution to airflow impairment when adding CT measures of emphysema, airway wall thickness, air trapping and ventilation.

Methods: Because tracheal abnormalities often occur localized, a method is presented that automatically determines the most abnormal trachea section based on automatically computed sagittal and coronal lengths. In this most abnormal section, trachea morphology is encoded using four equiangular rays from the center of the trachea and the normalized lengths of these rays are used as features in a classification scheme. Consequently, trachea measurements are used as input for classification into GOLD stages in addition to emphysema, air trapping and ventilation. A database of 200 subjects distributed across all GOLD stages is used to evaluate the classification with a k nearest neighbour algorithm. Performance is assessed in two experimental settings: (a) when only inspiratory scans are taken; (b) when both inspiratory and expiratory scans are available.

Results: Given only an inspiratory CT scan, measuring tracheal shape provides complementary information only to emphysema measurements. The best performing set in the inspiratory setting was a combination of emphysema and bronchial measurements. The best performing feature set in the inspiratory-expiratory setting includes measurements of emphysema, ventilation, air trapping, and trachea. Inspiratory and inspiratory-expiratory settings showed similar performance.

Conclusions: The fully automated system presented in this study provides information on trachea shape at inspiratory and expiratory CT. Addition of tracheal morphology features improves the ability of emphysema and air trapping CT-derived measurements to classify COPD patients into GOLD stages and may be relevant when investigating different aspects of COPD. © 2017 American Association of Physicists in Medicine [<https://doi.org/10.1002/mp.12274>]

Key words: automatic quantification, COPD, CT, lung, trachea

1. INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a chronic lung disease characterized by progressive airflow limitation that is caused by chronic bronchitis, inflammation of the airways, and emphysema, an irreversible destruction of lung tissue. The current standard for diagnosing is the pulmonary function test (PFT) and is determined by a postbronchodilator ratio of forced expiratory volume in one second and forced

vital capacity (FEV_1/FVC) of less than 0.70. Disease severity is evaluated by comparing the postbronchodilator FEV_1 value with its predicted value to assign the patients to one of the stages defined by the Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) ranging from stage 1 (mild COPD) to stage 4 (severe COPD).¹ However, COPD is a heterogeneous disease and the data obtained from PFT do not allow for differentiation between subtypes such as airway-dominant or emphysematous phenotypes.^{2,3}

Numerous studies extract quantitative measurements from thoracic computed tomography (CT) to describe different features of COPD and correlate these to clinical outcomes.^{2,3} These studies can roughly be grouped into three categories: (a) studies quantifying emphysema and/or airway disease from single inspiratory CT scans, (b) studies extracting measurements for emphysema and/or airway disease from (registered) inspiratory and expiratory CT scans, (c) studies extracting additional CT measurements that provide information about other functional aspects of the disease such as extra-pulmonary manifestations of COPD. While most studies fall into the first two groups which include the best known imaging biomarkers for COPD, this study falls into the third category, which is of special interest for analyzing possible new pathological findings associated with COPD.

Different types of tracheal abnormalities have been described in COPD subjects but their effect on airflow limitation is rarely studied. Saber-sheath trachea (SST), described as a marked narrowing of the coronal diameter associated with a sagittal elongation, is a static tracheal deformity often associated with COPD.⁴ But its relation with clinical and radiological parameters of COPD is unclear. SST is quantified in CT scans by determining the tracheal index (TI): the ratio between the coronal and sagittal diameter of the trachea. While some studies reported a significant correlation of TI with lung function and GOLD stage,^{5,6} other studies found that this correlation was not statistically significant.⁷

Dynamic tracheal abnormalities, presenting as excessive narrowing during expiration, have also been associated with COPD. Most commonly described abnormalities are tracheomalacia (TM), defined as luminal reduction caused by a softening of the tracheal cartilaginous structures, and excessive dynamic airway collapse (EDAC), characterized by an exaggerated posterior membrane movement that reduces the airway lumen without cartilage collapse.^{8,9} Other morphologies have been described such as lunate, frown, and horseshoe-shaped tracheas.¹⁰ Paired inspiratory–expiratory CT images are generally used to evaluate tracheal dynamic collapse and its relation with airflow limitation. Previous research evaluated tracheal collapsibility in COPD patients by comparing luminal areas on inspiratory and expiratory CT and no significant correlation with lung function parameters was found.^{6,11} However, the shape of the trachea in patients with COPD was different from healthy volunteers.¹² In contradiction with these findings, a more recent study found excessive airway collapse to be associated with worse respiratory quality of life in COPD patients.⁹

The development of automatic quantification methods for tracheal morphology allows for the assessment of static and dynamic shape changes. However, all studies so far fail to evaluate differences in tracheal shape next to collapsibility. Together with potential differentiation of COPD patients, detection of tracheal deformities in COPD patients is important since they commonly reflect the pathology of existing tracheal disorders. For example, a frown-like tracheal shape is a result of an excessive flaccidity of the posterior membranous wall, while a lunate configuration reflects a lengthening

of the membranous wall and a widening of the cartilaginous rings in the coronal plane. Patients showing these abnormalities, in which a flaccid membranous wall is the main contributor to the excessive airway collapse, could benefit from specific treatments such as tracheoplasty, that can improve their quality of life.¹⁰ However, patients in which the excessive collapse is produced by the anterior or lateral airway wall would not benefit from this treatment.¹⁰

This paper presents a method to automatically quantify trachea shape and dynamic shape changes from inspiration to expiration and investigates the contribution of tracheal morphology features to airflow limitation. To our knowledge, this is the first study that presents a fully automated method to quantify trachea shape changes during breathing. In addition, to evaluate the added value of trachea measurements when assessing COPD severity, we implement a set of previously proposed quantitative features for COPD quantification in two settings: a) when only inspiratory scans are accessible; b) when both inspiratory and expiratory scans are available.

2. QUANTIFICATION OF COPD FROM CHEST CT SCANS

Quantitative CT methods have been developed to describe different features of COPD, such as emphysema, airway morphology, and air trapping, either extracted from individual inspiratory CT scans or (registered) inspiratory and expiratory CT combined.^{2,3,13,14}

Emphysema is observed as low attenuation areas in the lungs, therefore, it is usually quantified from inspiratory chest CT scans using density-based measurements, such as the emphysema score (ES): percentage of voxels below a certain Hounsfield Unit (HU) threshold. ES has been proven to correlate well with PFT and pathology.^{2,3} Furthermore, the presence of emphysema in the lower lobes has been shown to have a larger influence on airflow limitation than emphysema in the upper lobes.¹³

The large and medium airways (lumen diameter > 2 mm) can reliably be visualized on inspiratory CT scans. Its morphology is generally expressed in different measurements obtained from airway segmentations such as wall area, wall thickness, and lumen area, which have been shown to correlate well with PFT.² However, small airways (< 2 mm diameter) cannot be visualized due to the limited resolution of current CT technology. As a result, no direct measurements of small airway disease can be obtained from inspiratory CT scans. Expiratory CT scans are used to indirectly quantify small airways disease by detection of air trapping: abnormal retention of air in the lungs during expiration. Air trapping shows in an expiratory CT scan as regions of abnormally decreased attenuation after exhalation, and is quantified using density-based measurements, such as the percentage of voxels below a HU threshold.² Recently developed registration techniques allow joint analysis of inspiratory and expiratory CT scans. Murphy *et al.*¹³ used registered scans to compute a set of ventilation measurements based on parenchymal densities. They showed that these measurements

have a better correlation with FEV₁ values than measurements taken from the individual scans. Furthermore, they performed a lobar analysis of different features revealing that measurements from lower lobes have more effect in the global pulmonary function. Galbán *et al.*¹⁴ used parametric response mapping (PRM), a registered voxel-wise technique that allows to discern the phenotype contributions of small airway disease and emphysema using information from the registered scans. This way, they overcome the inability to distinguish between emphysema and small airways disease inherent in the air trapping measurement computed in the expiratory scan alone.

3. MATERIALS

A database was constructed consisting of 200 subjects with various stages of COPD and control smokers without COPD taken from the COPDGene study.¹⁵ Subjects in the database were equally distributed over the different GOLD stages with 40 subjects in every GOLD stage. Every subject underwent inspiratory and expiratory CT, and PFT. One hundred and three (51.5%) subjects were male, and the mean age was 64 yr (range 45–86). Reconstruction settings are specified in Appendix A.

4. METHODS

This section describes the automatic analysis for trachea shape, both in the setting of single CT scan analysis and joint registered inspiratory and expiratory CT scans analysis. To determine the added value of trachea shape measurements to the prediction of GOLD stage, quantification methods for emphysema and airway abnormalities were added, and an experiment was set up to classify subjects into their GOLD stage based on different combinations of features.

4.A. Preprocessing

Prerequisites for the quantitative analysis are segmentation of the lungs and lobes and image registration between inspiratory and expiratory CT scans.

4.A.1. Segmentation

All anatomical structures of interest were automatically segmented using previously proposed methods integrated into CIRRUS Lung Quantification (Diagnostic Image Analysis Group, Radboud UMC, Nijmegen, The Netherlands; Fraunhofer MEVIS, Bremen, Germany). First, lung segmentation was performed using a method based in region growing and morphological smoothing.¹⁶ Then, the airway tree was segmented using a method based on adaptive thresholding.¹⁷ To allow for regional analysis in the lungs, a lobe segmentation was performed using a multi-atlas based approach.¹⁸ All lung and lobe segmentations were visually checked and edited where needed.

4.A.2. Image registration

Registration of inspiratory and expiratory CT scans is used to analyze changes in trachea morphology as well as to extract quantitative measurements of emphysema, air trapping, and ventilation. Registration was performed using the method described by Rùhaak *et al.*,¹⁹ deforming the expiratory scan to match the corresponding inspiratory scan. First, an affine registration was applied to the lung masks, aligning the lung boundaries, followed by nonrigid registration. Second, a gradient-based distance measure that focuses on image edges instead of intensities is used, since the different levels of inspiration lead to considerable intensity changes. To speed-up the processing, the algorithm uses a multilevel strategy from coarse to fine deformation resolution in the optimization process.

4.B. Trachea morphology

4.B.1. Inspiratory CT scan

Saber-sheath trachea (SST) is characterized by an elongation of the sagittal diameter associated with a reduction of the coronal diameter and is commonly measured by computing the tracheal index (TI), defined as:

$$TI = \frac{\text{coronal length inspiration}}{\text{sagittal length inspiration}} \quad (1)$$

Since the saber-sheath shape usually is very localized, we first detect the most abnormal cross-section in the trachea and perform the shape analysis there. The procedure to extract trachea features is described below:

- As a first step, a rough trachea and main bronchi segmentation is automatically computed using a method based on region growing.¹⁶ The center of gravity of this structure serves as the initial seed point for the trachea segmentation.
- The segmentation uses a wavefront propagation method in which the first front is provided by the initial seed point. The new wavefront is composed by all unprocessed voxels that meet the following voxel criteria: its density value is lower than a threshold t , or the HU value in a $3 \times 3 \times 3$ neighborhood around the voxel is $< t$. The threshold t was set to -750 HU.
- The wavefront is allowed to propagate until a split is detected and the carina is marked as the point where the wavefront bifurcates.
- Once the carina is detected, the region of interest in the trachea is limited by the section located 2 cm above the detected carina (to avoid the bifurcation of the main bronchi), and the uppermost apex of the lungs.
- Centerlines are extracted for the region of interest using a skeletonisation algorithm²⁰ and cross-sectional image planes perpendicular to the local tracheal direction are analyzed. These cross-sectional planes are created using multiplanar reformatting with trilinear interpolation.

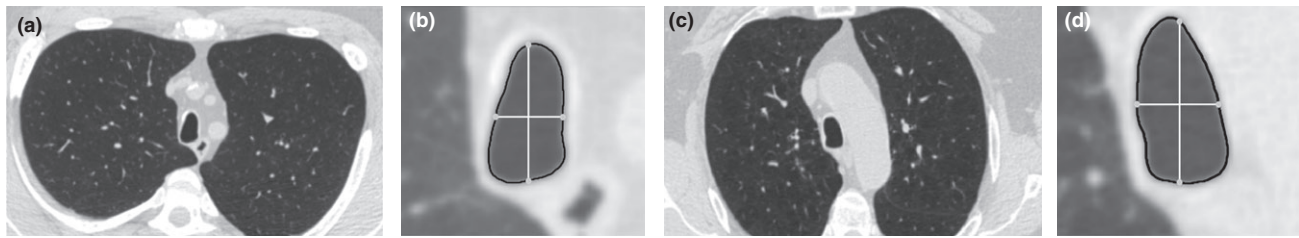


FIG. 1. Illustration of saber-sheath trachea across stages. (a) Axial inspiratory scan image showing a saber-sheath trachea from a GOLD 1 patient (TI = 0.49). (b) Detail view of the tracheal cross-section in (a). (c) Axial inspiratory scan image showing a saber-sheath trachea from a GOLD 4 patient (TI = 0.47). (d) Detail view of the tracheal cross-section in (c). In the detail views, the trachea segmentation is outlined in black. The intersection of the four equiangular rays with the trachea segmentation is highlighted in with grey circles. The white rays represent the distances used for the computation of the TI.

- In each plane, four equiangular rays (in the sagittal and coronal direction) are cast from the center of gravity outwards, defining the sagittal and coronal length. TI is calculated according to the formula (1) for every cross-sectional plane and the one with the minimum TI is selected as being the most abnormal location in the trachea.
- Tracheal morphology is encoded for the selected cross-section as the length of the four rays, normalized by the length of the longest ray to disregard differences in tracheal lumen size. An illustration is shown in Fig. 1.

4.B.2. Inspiratory and expiratory CT

When inspiratory and expiratory scans are available, changes in trachea shape during expiration can be quantified. As for static tracheal deformities, dynamic tracheal narrowing during expiration is often localized. Therefore, we again first determine the cross-sectional plane in which maximal tracheal narrowing occurs and consequently quantify the change in airway morphology in that plane. Since the different types of expiratory tracheal disorders are classified by looking at the sagittal and coronal narrowing that occurs during breathing,⁸ we developed an index called T to detect the region of the trachea where the maximum narrowing occurs. T is defined as:

$$T = \frac{2 \cdot T_{\text{sagittal}} \cdot T_{\text{coronal}}}{T_{\text{sagittal}} + T_{\text{coronal}}} \quad (2)$$

where

$$T_{\text{sagittal}} = \frac{\text{sagittal length exp}}{\text{sagittal length insp}} \quad (3)$$

$$T_{\text{coronal}} = \frac{\text{coronal length exp}}{\text{coronal length insp}} \quad (4)$$

T_{sagittal} accounts for changes in the sagittal axis, whereas T_{coronal} represent the changes in the coronal axis between inspiration and expiration. Thus, if there is a considerable, either sagittal or coronal, narrowing between inspiration and expiration, T will tend to zero. Therefore, the section with the lowest T is selected for analysis. The usage of T is more

focused in detecting abnormal tracheal narrowing occurring in one of the axes, rather than quantifying tracheal area reduction. The procedure to extract trachea shape features is as follows:

- For every inspiratory section, the anatomically equivalent section in the expiratory scan is found by propagating the center of gravity of the trachea. Then, cross-sectional tracheal-paired sections in inspiratory and expiratory scans are analyzed as described in section 4.B.1.
- T is computed for every pair of sections, and the pair with the minimum T is selected for the feature extraction.
- The tracheal shape is encoded by the length of the eight rays, normalized by dividing by the length of the longest ray in the inspiratory scan. This way, we implicitly include information about the area change between inspiration and expiration. The distances used to compute T are contained in the eight rays. An illustration is shown in Fig. 2.

4.C. Quantitative measurements of COPD

This section describes commonly used quantitative measures of COPD from chest CT scans, which will be included in the final GOLD stage classification to determine the additional value of the trachea morphology analysis. To investigate the influence of the different lobar regions, all the measurements computed in the lung parenchyma (i.e., emphysema, air trapping, and ventilation) are calculated per lobe.

4.C.1. Emphysema quantification

Emphysema was quantified using emphysema score (ES): the percentage of voxels in the lung below -950 HU in the inspiratory scan.

4.C.2. Airway morphology

Airway morphology is assessed in inspiratory scans by measuring lumen and wall dimensions of the airways. The

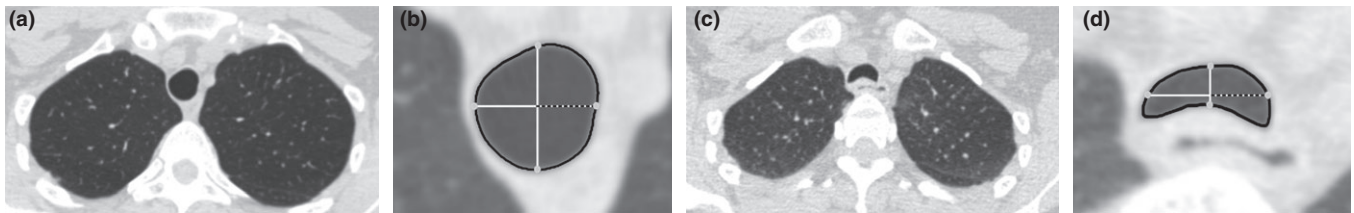


FIG. 2. Illustration of maximum narrowing detection between inspiration and expiration. The figure shows the inspiration-expiration pair of sections with the minimum T . (a) Axial inspiratory scan image showing the inspiration section with the lowest T . (b) Detail view of the tracheal cross-section in (a). (c) Axial expiratory scan corresponding to the inspiratory image in (a). (d) Detail view of the tracheal cross-section in (c). In the detail views, the trachea segmentation is outlined in black. The four equiangular rays cast (at 0, 90, 180 and 270 degrees) from the center to the boundary of the trachea are shown in white, and the intersections of the rays with the trachea segmentation are highlighted with grey circles. These rays also represent the distances used for the computation of the T . The dashed rays depict the distance quantified by the tracheal features at 0 degrees.

airway morphology was summarized as the square root of the wall area of a hypothetical airway of 10 mm internal perimeter (Pi10).² The airway tree was first skeletonized²⁰ and for each centerline voxel, cross-sectional image planes perpendicular to the local bronchial direction were analyzed. In each plane, a set of rays pointing outwards was cast, and the inner and outer wall borders were detected using an intensity-integration-based method.²¹ For those sections where the bronchial wall segmentation was automatically determined to be successful, the square root of the wall area versus the lumen perimeter was plotted and a regression line was calculated from which Pi10 was derived.

4.C.3. Air trapping quantification

Air trapping was quantified with an air trapping score (AS) as the relative lung area below -856 HU in the expiratory scan.

4.C.4. Ventilation defects

By employing image registration between the inspiratory and expiratory CT scans, ventilation measurements can directly be obtained from the scans. Similarly to the study of Murphy et al.,¹³ ventilation was depicted as the voxelwise ratio of inspiratory and expiratory intensity values. It was then summarized as the median of these ratios (medRatio).

4.C.5. Parametric response map

To compute the parametric response map (PRM), inspiratory and expiratory scans were registered and the thresholds for emphysema and air trapping quantification were applied as on individual scans. Every voxel was classified as emphysema if its intensity in inspiration was lower than -950 HU and its intensity in expiration was lower than -856 HU; as air trapping if its intensity in inspiration was higher than -950 HU and in expiration was lower than -856 HU; and as healthy parenchyma otherwise. This way we extracted the percentage of voxels classified as emphysema (prmES) and air trapping (prmAS) per lobe.

4.D. GOLD stage classification and feature selection

All the features computed for the experiments are listed in Table I. Except for trachea and airway features, all the measurements are computed in a per lobe basis. A kNN classifier was used to classify subjects into GOLD stages based on different feature sets combining the features in Table I. All classification experiments were performed in a leave-one-subject-out fashion, with k empirically set to 10, based on pilot experiments using an external dataset described in detail in Appendix B. Feature sets were manually constructed to test the contribution of tracheal morphology and other quantitative measurements, when using only an inspiratory CT scan and when using both inspiratory and expiratory CT. In addition, a sequential forward floating selection (SFFS)²² was

TABLE I. Quantitative measurements computed from the CT scans. Number of features are indicated between parenthesis.

Name	Description
idistTI α^a	Length of the ray cast from the center of the trachea to the border with an angle α in inspiration (4)
ES	Percentage of voxels inside each lobe with CT values below -950 HU in inspiration (5)
Pi10	Square root of the wall area of a hypothetical airway of 10mm internal perimeter (1)
idist α^b	Length of the ray cast from the center of the trachea to the border with an angle α in inspiration (4)
edist α^b	Length of the ray cast from the center of the trachea to the border with an angle α in expiration (4)
AS	Percentage of voxels in each lobe with CT values below -856 HU in expiration (5)
medRatio	Median of the ratio of the CT values at expiration to the CT values at inspiration per lobe (5)
prmES	Percentage of voxels in each lobe with CT values below -950 HU in inspiration and below -856 HU in expiration (5)
prmAS	Percentage of voxels in each lobe with CT values above -950 HU in inspiration and below -856 HU in expiration (5)

^aThis is used in inspiration only, and the section selected is the one with the minimum tracheal index (TI).

^bThis is used when inspiratory and expiratory scans are available, and the section selected is the one with the minimum T .

performed on the complete feature set to find those features that contain the most important information for the classifier. A detailed description of the SFFS procedure can be found in Appendix C.

It is important to note that PFTs have a considerable within-patient variability, which is increased in the case of patients with airflow obstruction.²³ This variability may lead to a subject being categorized into a higher or lower GOLD stage, depending on the factors that may affect them on the day that the test is done. Therefore, to evaluate the performance of the classifier, we computed, not only the averaged accuracy and macro-averages of precision, recall and F-score,²⁴ but also the percentage of subjects assigned to either the correct class or a neighbouring one (% within one class). Since we have ordinal classes, we cannot assume that any misclassification is equally bad. For this reason, instead of the misclassification error rate, we computed the mean absolute deviation (MAD) defined as:

$$\sum_{i=1}^N \frac{1}{N} |(classification\ label)_i - (true\ label)_i| \quad (5)$$

where N is the total number of samples. MAD is a variant of the error rate where the error weight is proportional to the distance from the real label.

5. EXPERIMENTS

We designed two experimental settings to simulate situations with only inspiratory scans available, and with paired inspiratory and expiratory scans available. In this section, we describe the different feature sets selected for every setting. In Appendices D–F we also investigated the relation between TI , T , and tracheal morphology with GOLD stages.

5.A. Inspiratory scans

In this situation only inspiratory scans are available, therefore, the only measurements we can compute were ES, Pi10, and idistTI α . To assess the added value of the trachea measurements next to emphysema and bronchial measures, we manually selected several feature sets as listed in Table II.

5.B. Inspiratory and expiratory scans

The full set of features is computed in this setting: ES, Pi10, AS, idist α , edist α , medRatio, prmES, and prmAS. Here, we evaluate the added value of trachea features compared to every feature set. We also compare the performance of non registered emphysema and air trapping features (ES

TABLE II. Quantitative measurements automatically computed from the CT scans. Initials between brackets indicate the lobe where the measurements were computed: left lower lobe (LLL), left upper lobe (LUL), right lower lobe (RLL), right middle lobe (RML), right upper lobe (RUL). Emph = emphysema, air trap = air trapping, SFFS = sequential forward floating selection.

Feature set	Features								
	idistTI	idist	edist	ES	Pi10	AS	prmES	prmAS	MedRatio
<i>Insp</i>									
Trachea	X								
Emph				X					
Emph, trachea	X			X					
Bronchi					X				
Bronchi, trachea	X				X				
Emph, bronchi				X	X				
Full	X			X	X				
<i>Inspiration-expiration</i>									
Trachea		X	X						
Emph, air trap				X		X			
Emph, air trap, trachea		X	X	X		X			
Emph, air trap, bronchi				X	X	X			
Emph, air trap, bronchi, trachea		X	X	X	X	X			
PRM							X	X	
PRM, trachea		X	X				X	X	
PRM, bronchi					X		X	X	
PRM, bronchi, trachea		X	X		X		X	X	
Ventilation									X
Ventilation, trachea		X	X						X
Ventilation, bronchi					X				X
Ventilation, bronchi, trachea		X	X		X				X
Features selected by SFFS	prmES[LLL], ES[RUL], ES[RLL], AS[RLL], medRatio[LLL], Pi10, ES[LLL], medRatio[RLL], prmAS[LLL], edist0								

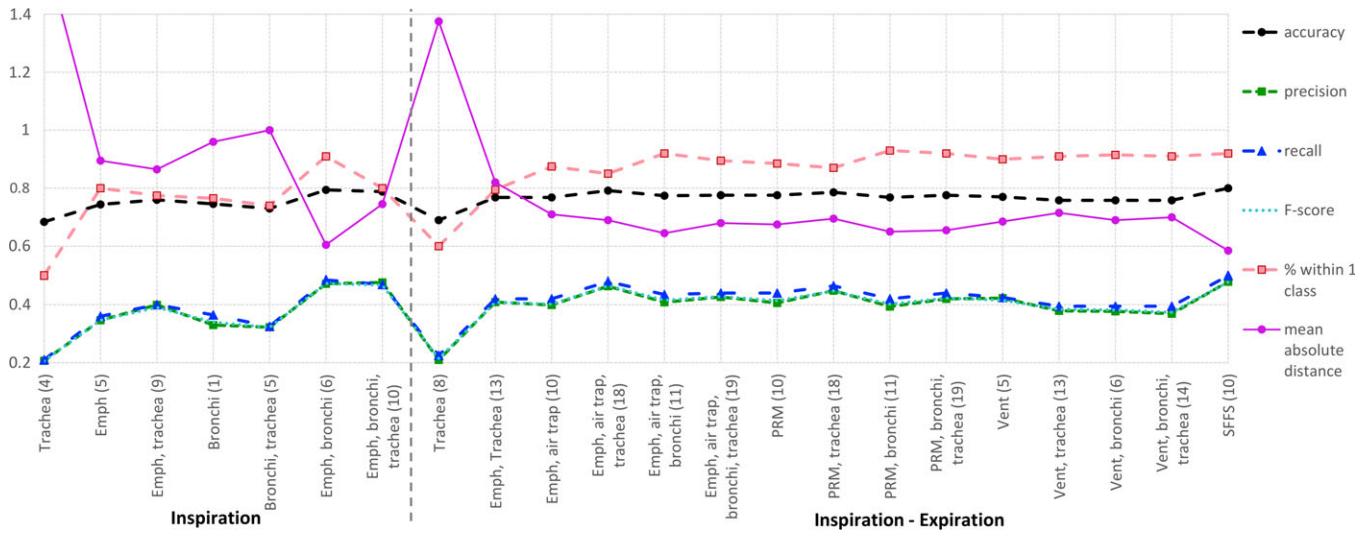


FIG. 3. Results of the classification into GOLD stages. The number of features for every setting is stated between parenthesis. Emph = emphysema; air trap = air trapping, SFSS = sequential forward floating selection. Dashed lines depict measures that indicate better performance with higher values. Solid lines depict measures that indicate better performance with lower values. [Color figure can be viewed at wileyonlinelibrary.com]

and AS) vs the registered ones (prmES and prmAS). Furthermore, we test the feature set obtained applying SFSS to the entire feature set. The feature sets evaluated are specified in Table II.

6. RESULTS

Table II shows the features selected by the SFSS algorithm as the most useful for the performance of the kNN classifier. The performance metrics of the classifier are presented in Fig. 3. The number of features used is also reported for every feature set tested.

6.A. Inspiratory scans

Trachea features increased the precision and recall, and diminished the MAD and % within one class of the classifier when added to emphysema measurements. The best classification performance was achieved when the combination of emphysema and bronchial features was used.

6.B. Inspiratory and expiratory scans

Similarly to the inspiration setting, adding trachea features to emphysema measurements yielded better results for the GOLD stage classification in terms of recall, precision, and MAD metrics. However, the combination of emphysema and bronchi measurements still showed a superior performance. Adding trachea measurements to feature sets combining emphysema and air trapping information raised precision and recall metrics while diminishing MAD and % within one class. The emphysema and air trapping set, and PRM set, which represent different ways of measuring the same components of the disease, showed similar performance. Ventilation measurements, which try to quantify the air flow in each lobe, also performed similarly to

TABLE III. Confusion matrices obtained for the classification experiment with the emphysema and air trapping set (above), and the emphysema, air trapping and trachea set (below).

		Classification				
GOLD		0	1	2	3	4
0		27	10	3	0	0
1		21	9	7	3	0
2		12	8	10	8	2
3		1	3	12	10	14
4		0	0	1	11	28

		Classification				
GOLD		0	1	2	3	4
0		31	5	4	0	0
1		13	18	8	1	0
2		8	10	14	5	3
3		2	6	9	9	14
4		1	0	5	10	24

Bold numbers indicate correct classifications.

the emphysema and air trapping set and the PRM set. Finally, the best results were obtained for the features selected by the SFSS algorithm, which consisted of a combination of tracheal and bronchial measurements, together with features computed mostly in the lower lobes. Further investigation of our results showed that adding trachea features to measurements of emphysema and air trapping, or ventilation, improves classification of lower stages while worsening the classification of the higher stages. This is illustrated by Tables III and IV that show the confusion matrices for the emphysema and air trapping set and for the ventilation set.

TABLE IV. Confusion matrices obtained for the classification experiment with the ventilation set (above), and ventilation and trachea set (below).

GOLD	Classification				
	0	1	2	3	4
0	21	17	2	0	0
1	24	8	8	0	0
2	8	10	13	5	4
3	1	2	9	13	15
4	0	1	2	7	30

GOLD	Classification				
	0	1	2	3	4
0	28	9	3	0	0
1	20	11	8	1	0
2	2	15	11	9	3
3	1	2	10	9	18
4	1	1	4	14	20

Bold numbers indicate correct classifications.

7. DISCUSSION

7.A. Trachea analysis

Tracheal abnormalities have been described to be present in COPD patients^{4–6}, but their influence on lung function has been barely studied. This paper presents an automatic method to quantify trachea morphology both in a static and a dynamic setting, with which we investigate the influence on GOLD stage classification. The most common inspiratory tracheal shape associated with COPD is saber-sheath trachea (SST) which is defined as an exaggerated decrease of the coronal diameter associated with a sagittal widening.^{4,6} We detected the most saber-sheath-like tracheal section by selecting the cross-section with the minimum TI in the inspiratory setting. We observed that adding trachea shape features to emphysema quantification improved the assessment of COPD severity, whereas adding them to bronchial features did not. This suggests that tracheal abnormalities might comprise a bronchial component. Thus, when tracheal features are added to emphysema measurements, they complement these features by providing an indirect bronchial component. However, when emphysema and bronchial features are combined, trachea shape quantification does not have any added value. These results suggest that SST may be a consequence of the presence of chronic bronchitis. It is believed that this shape is a result of remodelling and fixation of the tracheal cartilage induced by chronic coughing and inflammation.^{4,25} These are very typical symptoms in COPD patients suffering from chronic bronchitis, which shows in CT scans as thickened airway walls. We also analyzed the relation of TI with GOLD stages (data shown in Appendix D), and we found that TI values for GOLD 1–4 were significantly smaller than for GOLD 0. This is in line with previous studies where TI was correlated to lung function and GOLD stage.^{5,6} However, these findings need further investigation, since the presence of SST

(defined as $TI < 0.5$) in our cohort is not very high. Eleven patients (5.5%) presented saber-sheath trachea, distributed along all the GOLD stages as: two patients in GOLD 1, two patients in GOLD 2, four patients in GOLD 3, and three patient in GOLD 4. Figure 1 illustrates SST across stages by comparing a GOLD 1 subject with a GOLD 4 subject, both with similar tracheal shape feature values.

The tracheal features in the inspiration–expiration setting aim to analyze dynamic alterations in tracheal morphology during breathing, namely tracheomalacia (TM) or excessive dynamic airway collapse (EDAC). The former refers to tracheal narrowing due to a softened cartilage, while the latter alludes to exaggerated airway collapse due to a laxity of the posterior membrane.^{8,26} Our results show that, similarly to the inspiration setting, adding trachea features to bronchitis quantification in any feature set, the performance of the classifier is barely altered. This may suggest that, as in SST, tracheal abnormalities during breathing could be associated with the presence of chronic bronchitis. Some authors²⁶ consider the excessive collapse occurring in EDAC to be a result of alterations in the velocity of airflow due to narrowing of peripheral airways that is often produced by mucus or wall thickening. If this causes a significant resistance to airflow, the transmural pressure will increase, facilitating airways collapse. Additionally, the degeneration of the tracheal cartilage observed in TM is believed to be a result of the chronic irritation and coughing that occurs in COPD patients. The repeated mechanical stress from coughing might cause weakening and damage to the cartilage in the trachea.²⁷

The relationship between alterations in tracheal shape during breathing and lung function parameters has been previously investigated. In the study of Boiselle *et al.*,¹¹ it was shown that excessive tracheal narrowing does not correlate with physiologic impairment, even though it is observed in a subset of patients with COPD. However, another study with a large database of COPD subjects showed that excessive collapse was associated with greater airflow obstruction.⁹ We analyzed the relation of T with GOLD stages in our dataset (Appendix E) and we only found differences between GOLD 0 and GOLD 4 patients. However, as illustrated in Appendix F, we found that tracheal morphology in GOLD 1–4 patients was different from GOLD 0 subjects, which is in agreement with previous studies.¹² Additionally, our results seems to indicate that trachea shape changes during breathing hold valuable information about COPD severity, especially in early stages. It is observed that trachea shape features improve classification of lower stages, but worsen the classification of the higher stages when they are combined with features that quantify emphysema and air trapping, or ventilation. Furthermore, edist0 which accounts for the coronal narrowing that may occur in EDAC or TM in expiration (as illustrated in Fig. 2), was selected by the SFFS algorithm as a relevant feature for the inspiration–expiration setting.

Trachea morphology changes during breathing have been described and visually scored in COPD,^{8,10} but this is the first paper to quantitatively assess them, allowing a larger database and evaluation. The method presented here can facilitate

future studies that may provide valuable insight regarding the diagnostic criteria, aetiology, and management of excessive tracheal narrowing. Since tracheal shape is related to pathology, tracheal morphology analysis can be of great importance when planning airway interventions in patients suffering from abnormal tracheal narrowing. For example, patients showing a frown-like expiratory shape are ideal candidates for tracheoplasty, whereas patients suffering from excessive collapse of lateral walls do not benefit from this procedure, nor the use of bronchodilators.^{10,27}

7.B. GOLD stage classification

In this paper, we presented a system that quantifies the different components that affect lung function in COPD patients and use these measurements to predict airflow limitation. The standard for diagnosis and staging of COPD is performed by means of PFT,¹ but this test only provides information about how severe the airflow limitation is, not about the underlying causes. The presented method, not only is able to predict COPD GOLD stage with a reasonable accuracy but also provides quantitative information about the different components that are limiting the airflow in COPD patients. To analyze the feasibility of using automatic CT quantitative measurements to assess COPD severity, a kNN classifier was tested to assign GOLD stages to a set of subjects. The results of the classification are summarized in Fig. 3. Given that the reference standard for COPD staging is a test completely independent of the quantified measurements obtained from a CT scan, and taking into account that this is a five class problem, we can consider the results to be quite promising.

In the case of only inspiratory scans being available, emphysema features, which are fairly easy to obtain, yielded a reasonable performance. Yet Pi10, being only one feature, had a similar performance to emphysema quantification for COPD classification. The best performance is achieved when combining bronchi and emphysema measurements. These results are in agreement with previous literature. Mohamed Hoessein *et al.*³ compared emphysema and airway wall thickening with lung function parameters and demonstrated that while emphysema was best correlated with FEV_1/FVC (which accounts for the diagnosis of COPD), $FEV_1\%$ predicted (which accounts for COPD severity) was most influenced by airway wall thickness. Mets *et al.*² implemented different models to diagnose COPD using emphysema, airway wall thickness and air trapping measurements computed in paired inspiratory and expiratory scans. Even though the best performance was achieved when using the three components, the diagnostic ability of the model was still acceptable using only inspiratory measurements of emphysema and airway wall thickness. Comparing the two experimental settings, it is observed that using only inspiration features we can achieve a quite good performance, very similar to the best results in the inspiration–expiration setting. However, although inspiratory measurements may be acceptable for GOLD stage classification, inspiratory–expiratory measurements can better characterize the underlying causes of the disease.

The best performing feature set is composed by Pi10, one tracheal feature and a combination of measurements taken in the lower lobes. This may indicate that disease present in the lower lobes has more influence on airflow limitation. This is consistent with the results reported by Murphy *et al.*,¹³ in which the best classification setting only used features measured in the lower lobes.

In the study of Murphy *et al.*¹³ they evaluated the ability of parenchyma-based features to classify COPD patients into different GOLD stages. In this work, we added quantification of large and small airways morphology to account for the chronic bronchitis component that is one of the main phenotypes in COPD. The addition of these measurements improved the classification results compared to the use of only parenchyma-based features.

7.C. Limitations

This study has some limitations. First, we used CT scans obtained at full inspiration and end expiration instead of dynamic expiratory CT imaging, which is more sensitive at detecting tracheal collapse.¹² Dynamic CT scans were not available in our dataset, since static end expiratory imaging is commonly used in COPD patients as it is the preferred method to assess air trapping. However, our method could be also used in dynamic CT to assess changes during coughing or dynamic breathing. Second, our dataset is not very large, so our results should be further investigated. Third, spirometric gating was not used during CT imaging, therefore we cannot guarantee that full expiration was achieved at the time of scanning. Fourth, even though it is observed that adding tracheal measures to some feature sets improves the performance of the classifier, this finding should be taken with caution, since it is not a substantial improvement.

8. CONCLUSIONS

We have presented a fully automatic system that quantifies tracheal morphological changes during breathing and investigated its contribution to lung function impairment in COPD patients in combination with other components of the disease. We used these measurements to classify patients into GOLD stages using a kNN classifier. The system has been evaluated in a database consisting of well distributed patients across the five stages. Different measurements for the assessment of emphysema, air trapping, ventilation defects, and airway remodeling were used as an input for the classifier. We compared the performance of the classifier in two settings: (a) only inspiratory scans are accessible, (b) both inspiratory and expiratory scans are taken. Both settings showed similar performance.

To our knowledge, this is the first paper to assess the association of tracheal shape with COPD severity using an automatic quantification algorithm. Tracheal shape features provide relevant information in the COPD severity classification task, and their usage may be relevant for therapy planning or further research to better understand underlying causes of abnormal trachea morphology.

CONFLICTS OF INTEREST

D. A. L. is a consultant for Parexel Inc, Veracyte Inc, Boehringer Ingelheim Inc, Genentech/Roche Inc, and Siemens Inc. S. P. B. is supported by NIH under Grant No. K23HL133438. B. v. G. receives research funding and royalties from MeVis Medical Solutions, Delft Imaging Systems and Toshiba. He is co-founder and shareholder of Thirona BV. E.M. v. R. is co-founder and shareholder of Thirona BV.

^{a)}Author to whom correspondence should be addressed. Electronic mail: leti.g.estrella@gmail.com.

REFERENCES

- Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management and prevention of COPD. <http://www.goldcopd.org>. 2014.
- Mets OM, Schmidt M, Buckens CF, et al. Diagnosis of chronic obstructive pulmonary disease in lung cancer screening computed tomography scans: independent contribution of emphysema, air trapping and bronchial wall thickening. *Respir Res*. 2013;14:59.
- Mohamed Hoessein FAA, de Jong PA, Lammers J-WJ, et al. Contribution of CT quantified emphysema, air trapping and airway wall thickness on pulmonary function in male smokers with and without COPD. *COPD*. 2014;11:503–509.
- Greene R. “saber-sheath” trachea: relation to chronic obstructive pulmonary disease. *Am J Roentgenol*. 1978;130:441–445.
- Eom JS, Lee G, Lee HY, et al. The relationships between tracheal index and lung volume parameters in mild-to-moderate COPD. *Eur J Radiol*. 2013;82:e867–e872.
- Lee HJ, Seo JB, Chae EJ, et al. Tracheal morphology and collapse in COPD: correlation with CT indices and pulmonary function test. *Eur J Radiol*. 2011;80:e531–e535.
- Muro S, Nakano Y, Sakai H, et al. Distorted trachea in patients with chronic obstructive pulmonary disease. *Respiration*. 2000;67:638–644.
- Leong P, Bardin PG, Lau KK. What's in a name? Expiratory tracheal narrowing in adults explained. *Clin Radiol*. 2013;68:1268–1275.
- Bhatt SP, Terry NLJ, Nath H, et al. Association between expiratory central airway collapse and respiratory outcomes among smokers. *JAMA*. 2016;315:498–505.
- Boiselle PM, Ernst A. Tracheal morphology in patients with tracheomalacia: prevalence of inspiratory lunate and expiratory “frown” shapes. *J Thorac Imaging*. 2006;21:190–196.
- Boiselle PM, Michaud G, Roberts DH, et al. Dynamic expiratory tracheal collapse in COPD: correlation with clinical and physiologic parameters. *Chest*. 2012;142:1539–1544.
- Boiselle PM, O'Donnell CR, Bankier AA, et al. Tracheal collapsibility in healthy volunteers during forced expiration: assessment with multidetector CT. *Radiology* 2009;252:255–262.
- Murphy K, Pluim JPW, van Rikxoort EM, et al. Toward automatic regional analysis of pulmonary function using inspiration and expiration thoracic CT. *Med Phys*. 2012;39:1650–1662.
- Galbán CJ, Han MK, Boes JL, et al. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. *Nat Med*. 2012;18:1711–1715.
- Regan EA, Hokanson JE, Murphy JR, et al. Genetic epidemiology of COPD (COPDGene) study design. *COPD*. 2010;7:32–43.
- van Rikxoort EM, de Hoop B, Viergever MA, Prokop M, van Ginneken B. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Med Phys*. 2009; 36: 2934–2947.
- van Ginneken B, Baggerman W, van Rikxoort EM. Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. In: Metaxas D, Axel L, Fichtinger G, Szekely G, eds. *Med. Image Comput. Comput. Assist. Interv., Lect. Notes Comput. Sci.* Vol. 5241. Berlin: Springer Berlin Heidelberg; 2008:219–226.
- van Rikxoort EM, Prokop M, de Hoop B, Viergever M, Pluim J, van Ginneken B. Automatic segmentation of pulmonary lobes robust against incomplete fissures. *IEEE Trans Med Imaging*. 2010;29:1286–1296.
- Rühaak J, Heldmann S, Kipshagen T, Fischer B. Highly accurate fast lung CT registration. In: Ourselin S, Haynor DR, eds. *Proc. of the SPIE, Medical Imaging*, Vol. 8669. Lake Buena Vista (Orlando Area), FL: SPIE; 2013. doi:10.1117/12.2006035.
- Selle D, Preim B, Schenk A, Peitgen H-O. Analysis of vasculature for liver surgical planning. *IEEE Trans Med Imaging*. 2002;21:1344–1357.
- Schmidt M, Kuhnigk J-M, Krass S, Owsijewitsch M, de Hoop B, Peitgen H-O. Reproducibility of airway wall thickness measurements. In: Karssemeijer N, Summers RM, eds. *Proc. of the SPIE, Medical Imaging*, Vol. 7624. San Diego, CA: SPIE; 2010:7624IO.
- Pudil P, Novovicova J, Kittler J. Floating search methods in feature selection. *Pattern Recognit Lett*. 1994;15:1119–1125.
- Rozas CJ, Goldman AL. Daily spirometric variability: normal subjects and subjects with chronic bronchitis with and without air ow obstruction. *Arch Intern Med*. 1982;142:1287–1291.
- Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. *Inf Process Manag*. 2009;45:427–437.
- Hayes D Jr, Ballard HO. Saber-sheath trachea in a patient with bronchiolitis obliterans syndrome after lung transplantation. *Chron Respir Dis*. 2009;6:49–52.
- Park JG, Edell ES. Dynamic airway collapse: distinct from tracheomalacia. *J Bronchol*. 2005;12:143–146.
- Sverzellati N, Rastelli A, Chetta A, et al. Airway malacia in chronic obstructive pulmonary disease: prevalence, morphology and relationship with emphysema, bronchiectasis and bronchial wall thickening. *Eur Radiol*. 2009;19:1669–1678.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Appendix A. Reconstruction settings.

Appendix B. Classification performance for different values of k in the classifier.

Appendix C. Sequential forward floating selection (SFFS).

Appendix D. Tracheal index (TI) and GOLD stages.

Appendix E. T and GOLD stages.

Appendix F. Tracheal morphology and GOLD stages.