ORIGINAL RESEARCH

Application of the homology method for quantification of low-attenuation lung region in patients with and without COPD

Mizuho Nishio[1]
Kazuaki Nakane[2]
Yutaka Tanaka[3]

[1]Clinical PET Center, Institute of Biomedical Research and Innovation, Hyogo, Japan; [2]Department of Molecular Pathology, Osaka University Graduate School of Medicine and Health Science, Osaka, Japan; [3]Department of Radiology, Chibune General Hospital, Osaka, Japan

Background: Homology is a mathematical concept that can be used to quantify degree of contact. Recently, image processing with the homology method has been proposed. In this study, we used the homology method and computed tomography images to quantify emphysema.

Methods: This study included 112 patients who had undergone computed tomography and pulmonary function test. Low-attenuation lung regions were evaluated by the homology method, and homology-based emphysema quantification (b_0, b_1, nb_0, nb_1, and R) was performed. For comparison, the percentage of low-attenuation lung area (LAA%) was also obtained. Relationships between emphysema quantification and pulmonary function test results were evaluated by Pearson's correlation coefficients. In addition to the correlation, the patients were divided into the following three groups based on guidelines of the Global initiative for chronic Obstructive Lung Disease: Group A, nonsmokers; Group B, smokers without COPD, mild COPD, and moderate COPD; Group C, severe COPD and very severe COPD. The homology-based emphysema quantification and LAA% were compared among these groups.

Results: For forced expiratory volume in 1 second/forced vital capacity, the correlation coefficients were as follows: LAA%, −0.603; b_0, −0.460; b_1, −0.500; nb_0, −0.449; nb_1, −0.524; and R, −0.574. For forced expiratory volume in 1 second, the coefficients were as follows: LAA%, −0.461; b_0, −0.173; b_1, −0.314; nb_0, −0.191; nb_1, −0.329; and R, −0.409. Between Groups A and B, difference in nb_0 was significant (P-value = 0.00858), and those in the other types of quantification were not significant.

Conclusion: Feasibility of the homology-based emphysema quantification was validated. The homology-based emphysema quantification was useful for the assessment of emphysema severity.

Keywords: homology, emphysema, COPD, low-attenuation lung region

Introduction

According to the Global initiative for chronic Obstructive Lung Disease (GOLD) guidelines, COPD is characterized by chronic airflow limitation, which is usually progressive and associated with an inflammatory response of the lungs to noxious particles or gases.[1] The airflow limitation is usually associated with respiratory symptoms (chronic cough, dyspnea, and sputum).[1] COPD is a leading cause of morbidity and mortality worldwide and is projected to be the fourth leading cause of death globally by 2030.[2]

Technical advances in computed tomography (CT) facilitated evaluation of the structural changes caused by COPD, such as emphysema and thickening of airway walls.[3–9] Previous studies have shown that quantitative evaluation by CT is useful for

Correspondence: Mizuho Nishio
Clinical PET Center, Institute of Biomedical Research and Innovation, 2-2, Minatojimaminamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
Tel +81 78 304 5200
Fax +81 78 304 5990
Email jurader@yahoo.co.jp

assessing the severity of COPD by evaluating these structural changes. Because emphysema and airway disease are two major features of COPD and their contribution to COPD severity varies from patient to patient, the importance of CT-based phenotyping by quantitative evaluation of COPD has been increasingly recognized.[8]

For quantification of emphysema, the percentage of low-attenuation lung area (LAA%) has been frequently determined by CT, and LAA% was obtained as the percentage of the number of low-attenuation lung pixels to the total number of lung pixels,[3] where low-attenuation lung pixels were defined as pixels with CT values less than the predefined threshold (eg, −950 HU). Although LAA% has been associated with COPD severity,[9] no single type of quantification can guarantee an accurate assessment of emphysema. Because spatial distribution of low-attenuation lung regions was ignored in LAA%, it is expected that emphysema can be quantified more accurately by analyzing the spatial distribution of low-attenuation lung regions. Figure 1 shows representative images for LAA% = 30% obtained from a simulation study (in Figure 1, white pixels represent normal lung pixels, and black pixels represent low-attenuation lung pixels). Although LAA% of Figure 1A and B was the same, the spatial distribution of low-attenuation lung regions was visually different between these images. That is to say, LAA% cannot be used for evaluating the spatial distribution of low-attenuation lung regions. To overcome this problem of LAA%, it is necessary to develop other types of emphysema quantification.

Homology is a mathematical concept that can be used as a basis for the measurement of the degree of connectivity,[10,11] and the Betti number is an important index of homology. In general, tissues (structures) consist of different types of elements, and contact (boundary) occurs between the elements. Because the homology method calculates the contact degree of tissues (structures), this method has been applied in many fields.[12–16] For example, Nakane et al[12,13] proposed a new method for evaluating digital pathology images of the colon, in which the Betti numbers were extracted from the region of interest of the pathology images. Their results show that it was possible to use the Betti numbers for screening of colon cancer. This method does not require training using a library of images because the Betti numbers are derived directly from the images. Notably, the time required to calculate these numbers is very short; using an ordinary computer, calculation of the Betti numbers only requires approximately 3.0 seconds.[12,13]

In this study, we hypothesized that homology could be used for evaluating the spatial distribution of low-attenuation lung regions, and that homology-based emphysema quantification would be useful for determining the severity of emphysema. Visually, emphysema corresponds to the "holes" in the lung tissue. Originally, the concept of homology was that figures can be distinguished by their "holes". Therefore, we speculated that the homology method could be applied in order to quantify emphysema. The purpose of this study was 1) to validate the usefulness of homology-based emphysema quantification to evaluate the spatial distribution of
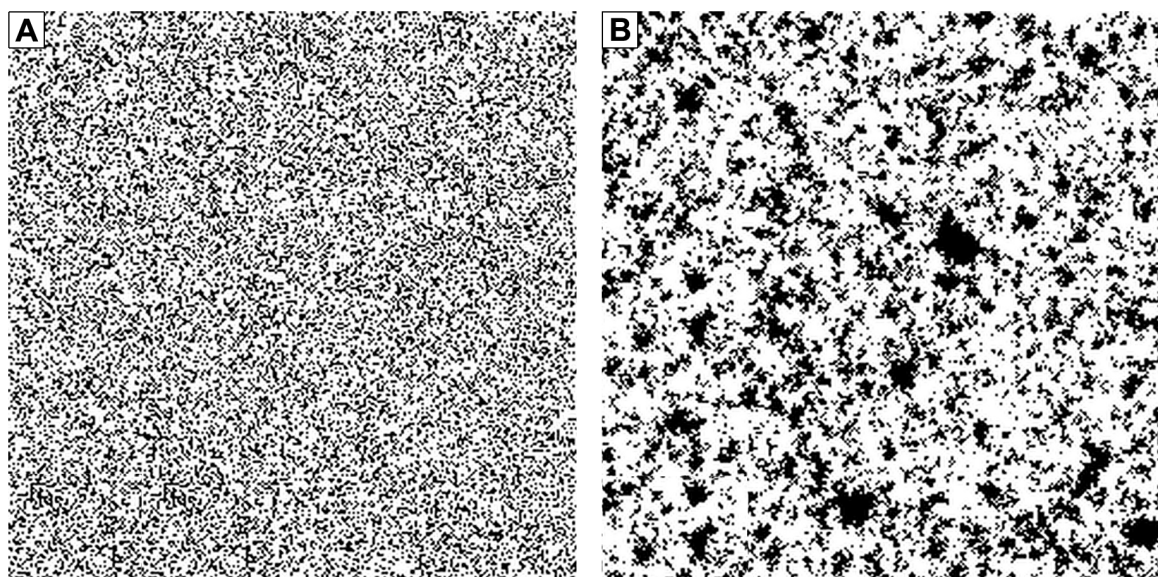

Figure 1 Representative images of the binarized images at the percentage of low-attenuation lung area =30% obtained in the simulation study.
Notes: (**A** and **B**) correspond to the results of simulations A and B, respectively. *R* of (**A** and **B**) is 0.103 and 0.380, respectively.

low-attenuation lung regions, 2) to evaluate the feasibility of the homology method for determining the severity of emphysema, and 3) to compare the homology-based emphysema quantification with LAA% and the pulmonary function test (PFT) in patients with and without COPD.

Methods

This retrospective study was approved by the institutional review boards of Institute of Biomedical Research and Innovation and Chibune General Hospital. Acquisition of informed consent was waived by the review board.

Homology

Although expert knowledge of mathematics is required to fully comprehend the concept of homology, in two-dimensional cases (eg, image analysis), use of homology is quite simple. In the two-dimensional cases, the Betti numbers consist of two numbers: b_0 (the zero-dimensional Betti number), which is the number of connected components, and b_1 (the one-dimensional Betti number), which is the number of one-dimensional or "circular" holes. In the current study, b_0 corresponded to the number of low-attenuation lung regions caused by emphysema, and b_1 corresponded to the number of normal lung regions surrounded by the low-attenuation lung regions. To calculate b_0 and b_1, we used the free software, CHomP.[12,13,17] In calculating the homology-based emphysema quantification, CT image must be converted to a binarized image in which each pixel could have two values: 0 and 1 (0 representing a low-attenuation lung pixel and 1 representing a normal lung pixel).

Simulation study

First, a simulation study was conducted in order to investigate the relationship between the homology-based emphysema quantification and LAA% and to validate whether the homology was associated with the spatial distribution of low-attenuation lung regions. A binarized image (image size 256×256) was prepared in which each pixel could have two values: 0 and 1. Before the simulation started, the values of all pixels in the binarized images were 1. Then, in one simulation (simulation A), the pixel value was randomly replaced with 0. In the other simulation (simulation B), the pixel value was randomly replaced with 0 at a probability of 20%, and the neighboring pixel value of the existing low-attenuation lung regions was replaced with 0 at a probability of 80%. In both simulations, the replacement was continued until the LAA% of the images reached the predefined value. For the predefined value of LAA%, the following values were used: 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, and 50%.

The binarized images after the replacement were evaluated by CHomP, and b_0 and b_1 were calculated. In this simulation, because LAA% was the same in simulations A and B, and the generation process of low-attenuation lung pixels was different, we speculated that the differences in the homology-based emphysema quantification reflected the spatial distribution of the normal lung regions and low-attenuation lung regions caused by the generation process. In addition, the relationship between R (= b_1/b_0)[12,13] and LAA% was examined by plotting R and LAA%.

Patients

Next, a clinical study was conducted. Patients who visited our institution because of their respiratory symptoms were examined retrospectively. If the patient underwent both CT and PFT and the interval between CT and PFT was <90 days, the patient was included in the study. Diagnosis and classification of COPD were performed according to the GOLD guideline.[1] This study included 112 consecutive patients (84 males and 28 females; age, 67.8±10.9 years). The mean smoking history of these patients was 36.8±44.7 pack-years, and the mean interval between CT and PFT was 17.7±36.7 days. Among the 112 patients, 49 patients were diagnosed as having COPD; 48 were smokers without COPD, and 15 were nonsmokers without COPD.

CT scan

Non-contrast helical CT scans were acquired from the lung apices through the lung bases using a 320-detector row scanner (Aquilion ONE; Toshiba Medical Systems, Otawara, Japan) with automated exposure control. After receiving careful instruction about breathing, the patients were scanned in the supine position during a deep inspiratory breath hold. The scan parameters were as follows: tube current, 193±67.5 mA; tube potential, 120 kV; gantry rotation time, 0.35 seconds in one patient, 0.6 seconds in two patients, and 0.5 seconds in all other patients; and noise index of automated exposure control, 10. Raw CT data were reconstructed into 5 mm thick images using a soft-tissue kernel (FC 13 or 14). The CT scanner was calibrated regularly.

PFT

The PFT was performed by an automated spirometer (HI-801 or CHESTAC-8900, Chest M.I., Inc., Tokyo, Japan). Forced expiratory volume in 1 second (FEV_1), forced vital capacity (FVC), and the ratio of FEV_1 to FVC (FEV_1/FVC) were obtained. Apart from FEV_1/FVC, these parameters were expressed as percentages of the standard predicted values.

Quantitative evaluation of emphysema

As previously described, three CT images of the upper, middle, and lower (Figure 2) lung fields were selected for each patient.[4] Then, the lungs were automatically segmented from the three CT images by region-growing and a threshold of −500 HU. The results of lung segmentation were validated visually by one board-certified radiologist (MN). The results of lung segmentation were represented as binary images, as shown in Figure 2B. After the lung segmentation, LAA% was calculated as follows:

$$LAA\% = \frac{\text{The number of low-attenuation lung pixels obtained from the three CT images}}{\text{The total number of lung pixels obtained from the three CT images}}, \quad (1)$$

where low-attenuation lung pixels were defined as pixels with CT values less than −950 HU. Approximately, LAA% corresponds to the percentage of black lung pixels in Figure 2D.

To calculate the homology-based emphysema quantification, preprocessing of the three CT images was necessary. First, small pulmonary vessels in the results of lung segmentation were deleted by morphological closing to emphasize the low-attenuation lung pixels caused by emphysema. Representative images of lung segmentation before and after the deletion of small pulmonary vessels are shown in Figure 2B and C, respectively. Then, the results of lung segmentation and CT images were analyzed pixel by pixel. If the pixel belonged to a lung pixel and the CT value of the pixel was <−950 HU, the value of the results of lung segmentation was replaced with 0. After this step, 1 in
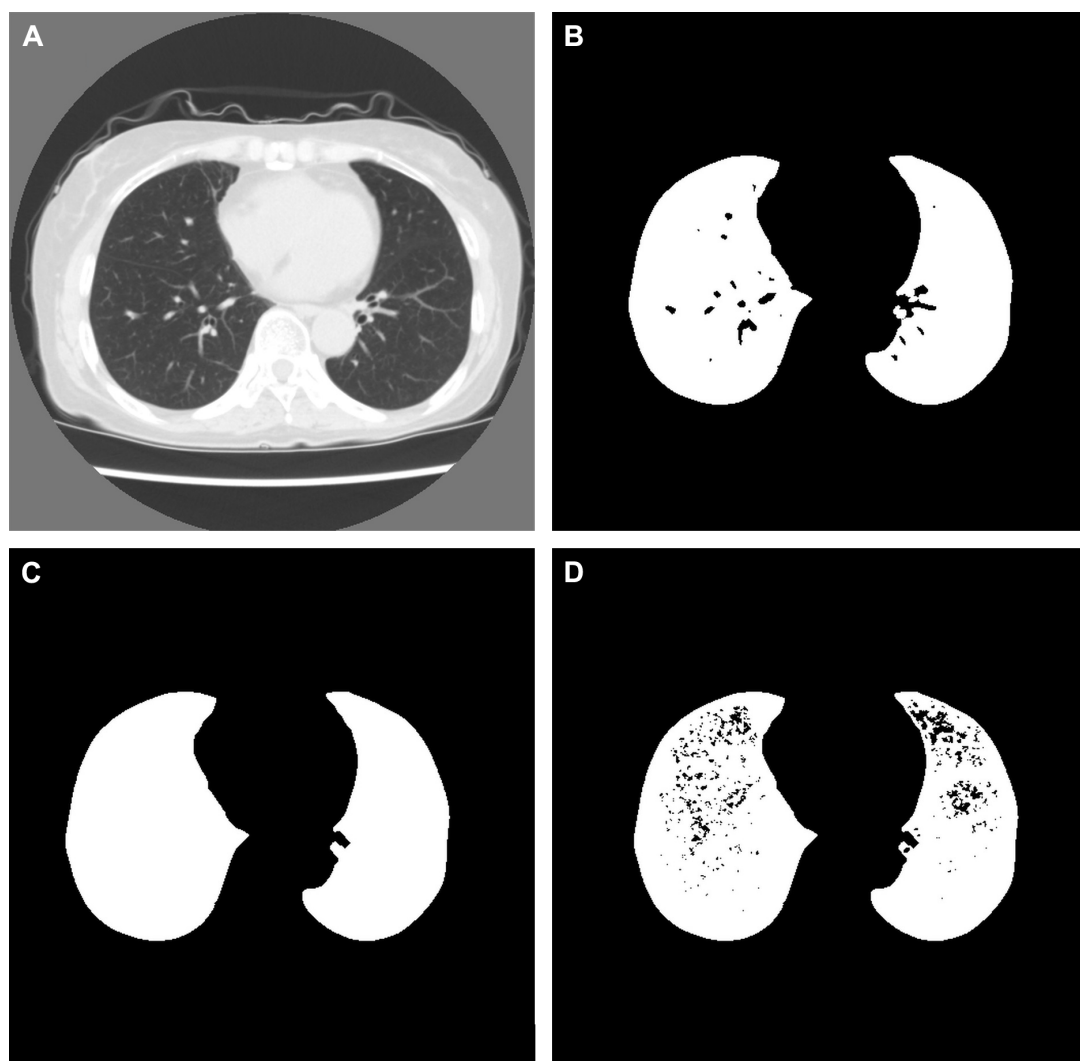

Figure 2 Representative images of computed tomography (CT) image and results of preprocessing.
Notes: (**A**) CT image of the lower lung field. (**B**–**D**) Binarized images after lung segmentation, after deletion of small pulmonary vessels, and after all preprocessing, respectively. (**D**) Used for evaluation of the homology-based emphysema quantification.

the binarized image indicated a normal lung pixel, and 0 indicated a non-lung pixel or low-attenuation lung pixel. A representative image after the whole preprocessing step is shown in Figure 2D.

After preprocessing, the binarized images were evaluated with CHomP to calculate the homology-based emphysema quantification. For each patient, b_0 and b_1 for the three binarized images were summed, and the summed value was used for the following statistical analyses. $R (= b_1/b_0)$ was also calculated. In addition, b_0 and b_1 were normalized to the total number of lung pixels obtained from the three CT images and referred to as nb_0 and nb_1, respectively.

Statistical analysis

To test whether the quantification reflected the severity of COPD, Pearson's correlation coefficients were calculated between the results of CT quantification and PFT. Correlation between LAA% and the homology-based emphysema quantification was also evaluated.

Next, the 112 patients were divided into three groups: Group A, nonsmokers; Group B, smokers without COPD, mild COPD, and moderate COPD; Group C, severe COPD and very severe COPD. Here, COPD severity was determined based on the GOLD guidelines.[1] The homology-based emphysema quantification and LAA% were compared among these three groups. The difference in the homology-based emphysema quantification and LAA% was tested by a Student's t-test, and P-values <0.0167 determined by Bonferroni correction were considered to indicate statistical significance.

Finally, linear models were used to investigate the relationship between the PFT results and emphysema quantification. One linear model was built to predict FEV_1, and another was built to predict FEV_1/FVC. In each model, the homology-based emphysema quantification (nb_0, nb_1, and R), patient sex, patient age, and smoking history were included as predictive variables, and uninformative variables were excluded from the models by stepwise regression on the basis of Akaike information criterion (AIC) values.[18] For comparison, linear models with LAA%, patient sex, patient age, and smoking history were also evaluated by stepwise regression and AIC. The model with a lower AIC was better for predicting PFT results. Because the correlation coefficients between R and LAA% were high, LAA% and the homology-based emphysema quantification were not included in the same linear model. The coefficients of the predictive variables were evaluated according to their P-values. P-values <0.05 were considered to indicate statistical significance. All analyses were performed by

R-3.1.1 (available at http://www.r-project.org/). The MASS package was used for stepwise regression.

Results

The results of homology-based emphysema quantification in the simulation study are shown in Table 1. Figure 1A and B show representative images for LAA% = 30% obtained from the simulations A and B, respectively. Figure S1 shows all images used in the simulations. Table 1 shows that b_0, b_1, and R were different between simulations A and B even when LAA% was the same. Because the generation process of low-attenuation lung pixels was different between simulations A and B, this result suggests that the differences in the spatial distribution of normal lung regions and low-attenuation lung regions were associated with those of the homology-based emphysema quantification. In Figure 3, R and LAA% of Table 1 are plotted on semi-log graphs. According to Figure 3, the relationships between R and LAA% were exponential and represented as:

$$R = 0.000307 \times e^{0.193 \times LAA\%} \tag{2}$$

in Figure 3A and

$$R = 0.0496 \times e^{0.0682 \times LAA\%} \tag{3}$$

in Figure 3B.

Patient characteristics, PFT results, and emphysema quantification results are summarized in Table 2. These values in nonsmokers, smokers without COPD, mild COPD, moderate COPD, severe COPD, and very severe COPD are summarized in Tables S1 and S2 of the Supplementary materials.

Table 1 Results of the homology-based emphysema quantification for different values of LAA% in the simulation study

LAA%	Simulation A			Simulation B		
	b_0	b_1	R	b_0	b_1	R
10	4,325	7	0.00162	1,109	125	0.113
15	5,049	31	0.00614	1,513	203	0.134
20	5,116	94	0.0184	1,657	305	0.184
25	4,670	183	0.0392	1,700	443	0.261
30	3,867	400	0.103	1,669	634	0.380
35	2,941	798	0.271	1,603	782	0.488
40	1,958	1,287	0.657	1,433	1,091	0.761
45	1,192	2,087	1.751	1,244	1,306	1.050
50	675	3,180	4.711	969	1,638	1.690

Notes: b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; R, b_1/b_0.
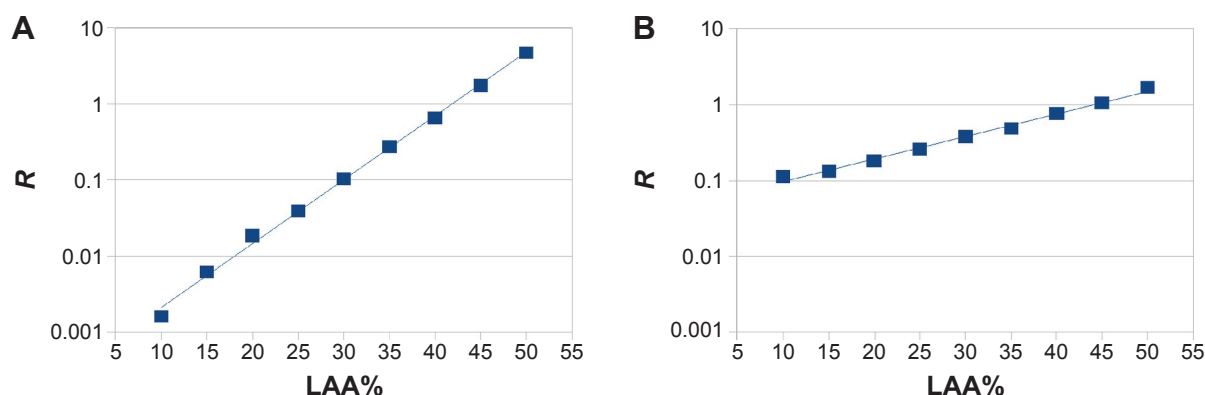**Abbreviation:** LAA%, percentage of low-attenuation lung area.

Figure 3 R and the percentage of low-attenuation lung area (LAA%) in the simulation study.
Notes: (**A** and **B**) correspond to the results of simulations A and B, respectively. The relationships between R and LAA% are represented as $R = 0.000307 \times e^{0.193 \times \text{LAA}\%}$ in (**A**) and $R = 0.0496 \times e^{0.0682 \times \text{LAA}\%}$ in (**B**). The determinants of the coefficients were 0.997 and 0.993 in (**A** and **B**), respectively.

Figures S2–S4 in the Supplementary materials show representative CT images and the binarized images after preprocessing.

Table 3 shows the Pearson's correlation coefficients among the homology-based emphysema quantification, PFT results, and LAA%. Table 3 shows that b_1, nb_1, and R had moderate correlation with FEV_1. In contrast, the correlation coefficients of b_0 and nb_0 were weak. For FEV_1/FVC, each of the homology-based emphysema quantifications had moderate correlation. The Pearson's correlation coefficients between R and FEV_1 and between R and FEV_1/FVC were −0.409 and −0.574, respectively. The Pearson's correlation coefficients between LAA% and FEV_1 and between LAA% and FEV_1/FVC were −0.461 and −0.603, respectively.

Table 2 Summary of patient characteristics, PFT results, and emphysema quantification

Variables	Mean	SD
Age (years)	67.8	11.0
Sex (M:F)	84:28	
Smoking history (pack year)	44.7	36.8
FVC (%)	89.6	23.8
FEV_1/FVC (%)	68.1	16.8
FEV_1 (%)	73.0	26.2
VC (%)	95.4	23.8
LAA%	10.7	14.1
b_0	902	490
b_1	114	159
nb_0	0.00580	0.00258
nb_1	0.000665	0.000870
R	0.880	0.115

Notes: The PFT results are expressed as percentages of the standard predicted values except for those of FEV_1/FVC. b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0 normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
Abbreviations: PFT, pulmonary function test; FEV_1, forced expiratory volume in 1 second; FEV_1/FVC, ratio of forced expiratory volume in 1 second to forced vital capacity; FVC, forced vital capacity; VC, vital capacity; LAA%, percentage of low-attenuation lung area; SD, standard deviation.

The homology-based emphysema quantification had relatively strong correlation with LAA%. Especially, the coefficient between R and LAA% was more than 0.9.

The emphysema quantification among Groups A, B, and C are summarized in Table S1. And, the results of Student's t-tests among Groups A, B, and C are summarized in Table 4. As shown in Table 4, between Groups A and B, the difference in nb_0 was statistically significant, and those in the other types of quantification were not statistically significant. On the contrary, between Groups B and C, the difference in nb_0 was not statistically significant.

Table 5 shows the results of the linear models. As shown in Table 5, the model predicting FEV_1 was:

$$FEV_1 = 97.1 - 288 \times R - 2{,}327 \times nb_0 + 31{,}646 \times nb_1 - 0.142 \times (\text{smoking history}). \quad (4)$$

Stepwise regression excluded patient age and patient sex from the model. Although the model included nb_0, the coefficients of this variable were not statistically significant. In this model, the adjusted R squared and AIC were 0.239 and 1,025.84, respectively. In the LAA% model predicting FEV_1, LAA% and smoking history were selected by stepwise regression, and adjusted R squared and AIC were 0.228 and 1,025.52, respectively. The model predicting FEV_1/FVC was:

$$\frac{FEV_1}{FVC} = 93.2 - 201 \times R - 3{,}148 \times nb_0 + 21{,}877 \times nb_1 - 0.0826 \times (\text{smoking history}). \quad (5)$$

Stepwise regression excluded patient age and patient sex from the model. In this model, the adjusted R squared and AIC were 0.440 and 892.00, respectively. In the LAA%

Table 3 Pearson's correlation coefficients for homology-based emphysema quantification, PFT results, and LAA%

Variables	FEV_1		FEV_1/FVC		LAA%	
	Coefficient	*P*-value	Coefficient	*P*-value	Coefficient	*P*-value
b_0	−0.173	0.0677	−0.460	$3.34×10^{-7}$	0.597	$<1.0×10^{-10}$
b_1	−0.314	0.000726	−0.500	$1.97×10^{-8}$	0.868	$<1.0×10^{-10}$
nb_0	−0.191	0.04338	−0.449	$6.96×10^{-7}$	0.53	$1.8×10^{-9}$
nb_1	−0.329	0.000393	−0.524	$3.11×10^{-9}$	0.88	$<1.0×10^{-10}$
R	−0.409	$7.70×10^{-6}$	−0.574	$<1.0×10^{-10}$	0.932	$<1.0×10^{-10}$

Notes: b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0, normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
Abbreviations: PFT, pulmonary function test; FEV_1, forced expiratory volume in 1 second; FEV_1/FVC, ratio of forced expiratory volume in 1 second to forced vital capacity; LAA%, percentage of low-attenuation lung area.

model predicting FEV_1/FVC, LAA% and smoking history were selected by stepwise regression, and the adjusted R squared and AIC were 0.372 and 902.93, respectively. According to the AIC values, for predicting FEV_1 and FEV_1/FVC, the models with homology-based emphysema quantification were comparable with or better than the models with LAA%.

Discussion

To our knowledge, this was the first study to perform homology-based emphysema quantification. The results of the simulation study clarified the usefulness of the homology-based emphysema quantification for assessing the spatial distribution of low-attenuation lung regions and validated that the relationship between R and LAA% was exponential. The clinical study demonstrated that the correlation between the homology-based emphysema quantification and PFT was comparable with that between LAA% and PFT. In addition, Table 4 shows that nb_0 was more useful for evaluating early changes in COPD than was LAA%. The results of linear models predicting PFT showed that the models with

Table 4 Statistical tests of emphysema quantification among the three groups

	Group A vs Group B	Group B vs Group C	Group A vs Group C
LAA%	0.274	0.000265*	0.000102*
b_0	0.0237	0.0100*	0.000168*
b_1	0.139	0.000950*	0.000146*
nb_0	0.00858*	0.0579	0.000343*
nb_1	0.0777	0.000683*	0.0000578*
R	0.162	0.000792*	0.000182*

Notes: The numbers in the table are *P*-values. *Indicates a *P*-value <0.0167. Group A, nonsmokers; Group B, smokers without COPD, mild COPD, moderate COPD; Group C, smokers with severe COPD and very severe COPD; b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0 normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
Abbreviation: LAA%, percentage of low-attenuation lung area.

homology-based emphysema quantification were comparable with or better than the models with LAA%. According to these results, the homology-based emphysema quantification obtained by the dedicated software was useful for assessing emphysema.

Because the disease distribution of emphysema was ignored in LAA%, it is expected that emphysema can be quantified more accurately by analyzing the size and spatial distribution of low-attenuation lung regions. For example, Mishima et al showed that D, obtained by analyzing the size distribution of low-attenuation lung regions, is a parameter for detection of the terminal airspace enlargement.[4] In line with this trend, we hypothesized that the spatial distribution of emphysema could be evaluated by the homology method. The two parameters of homology, b_0 and b_1, correspond to the numbers of connected components and "circular" holes, respectively.[10,11] The connected components corresponded to the low-attenuation lung regions in the present study. Therefore, b_0 is large in patients with emphysema. Concerning b_1, the "circular" holes corresponded to the normal lung regions surrounded by the low-attenuation lung regions in the present study. Therefore, on the CT images of the patients without emphysema, b_1 was low. In contrast, in the patients with emphysema, because the normal lung regions were divided into several regions by the low-attenuation lung regions, b_1 was increased. Considering these properties of the two parameters, we speculated that b_0 and b_1 are associated with the spatial distribution of normal lung regions and low-attenuation lung regions. The results of the simulation study also suggest that it was possible to evaluate the difference in the spatial distribution of normal lung regions and low-attenuation lung regions by the homology method. In addition, the results of Table 3 show that the homology-based emphysema quantification correlated with FEV_1 and FEV_1/FVC. These results validated the usefulness of homology for emphysema quantification.

Table 5 Results of the linear model predicting PFT results using the homology-based emphysema quantification

Model	Predictive variable	Coefficient	SE of coefficient	P-value	AIC	Adjusted R squared
For FEV_1					1,025.84	0.239
	R	−288	74.3	0.000182		
	nb_0	−2,327	1,337	0.0846		
	nb_1	31,646	11,231	0.00576		
	Smoking history	−0.142	0.0599	0.0200		
For FEV_1/FVC					892.00	0.440
	R	−201	40.9	3.11×10^{-6}		
	nb_0	−3,148	735	4.07×10^{-5}		
	nb_1	21,877	6,179	0.000593		
	Smoking history	−0.0826	0.0330	0.0137		

Notes: b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0 normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
Abbreviations: AIC, Akaike information criterion; PFT, pulmonary function test; FEV_1, forced expiratory volume in 1 second; FEV_1/FVC, ratio of forced expiratory volume in 1 second to forced vital capacity; SE, standard error.

According to Table 4, the difference in nb_0 values was statistically significant between Groups A and B, and that in LAA% was not significant. This implies that nb_0 was more useful for evaluating early changes in COPD than was LAA%. LAA% was easily affected by noise, especially when LAA% was low. Consequently, it was difficult to differentiate between Groups A and B using LAA%. We speculate that, by evaluating the spatial distribution of low-attenuation lung regions, nb_0 can provide more accurate information about early changes in COPD than can LAA%. In contrast, nb_0 was less sensitive to the progression of COPD than was LAA%. According to the simulation study (Table 1), b_0 at low LAA% was higher than b_0 at high LAA%. Because large low-attenuation lung regions were formed in high LAA%, b_0 was low in high LAA%. In short, in severe emphysema, b_0 or nb_0 was not suitable for assessing emphysema. Considering these results, we suggest that although nb_0 is less useful in severe emphysema than was LAA%, nb_0 can be used for evaluating early changes in COPD.

Table 5 shows that the homology-based emphysema quantification was useful for predicting airflow limitations measured by PFT. In the models with homology-based emphysema quantification of Table 5, R was the strongest variable among the predictors. Because R was highly correlated with LAA%, airflow limitation was mainly associated with the percentage of emphysema. In contrast, because nb_0 and nb_1 were also used in the linear models, the homology-based emphysema quantification provided additional value for assessment of airflow limitation. According to a previous study, combined use of LAA% and D led to better evaluation of COPD.[19] The results of the previous study and the current study suggest that COPD severity could be quantified more

accurately by analyzing the size and spatial distribution of low-attenuation lung regions.

Although the results of the simulation study show that the relationship between R and LAA% was exponential, this relationship was retained only when the generation process of low-attenuation lung pixels was fixed. In the clinical cases, the generation process of low-attenuation lung pixels, which is associated with the mechanism of emphysema progression, was not apparent. Conversely, we speculated that by analyzing the spatial distribution of normal lung regions and low-attenuation lung regions with b_0, b_1, and R, the generation process of low-attenuation lung pixels could be evaluated. For example, according to the simulation study, R was lower in Figure 1A than in Figure 1B, whereas LAA% was the same. Considering this example, combined use of LAA% and R can help to evaluate whether the spatial distribution of low-attenuation lung regions is random or not. Previous studies have shown that, by evaluating the spatial distribution of emphysema, homogeneous distribution of emphysema tended to result in worse pulmonary function in COPD patients.[20,21] In these studies, LAA% values were used to assess the spatial distribution of emphysema. Although our results are preliminary, LAA% and the homology-based emphysema quantification may provide more detailed information about the spatial distribution of emphysema than can LAA% only.

There were several limitations in this study. First, this was a retrospective study, and the number of patients was relatively small. Second, two-dimensional image analyses were performed. Recently, CT quantification based on thin-slice CT images has been used frequently. In the present study, two-dimensional image analyses were used to validate

the feasibility and usefulness of the homology-based emphysema quantification. In the future, we will extend our method to the three-dimensional cases. Finally, the homology-based emphysema quantification was mainly related to the spatial distribution of low-attenuation lung regions, and the size distribution was ignored in this method. To evaluate emphysema severity accurately, we think that the size and spatial distribution of low-attenuation lung regions should be evaluated for accurate quantification of emphysema.

Conclusion

In conclusion, feasibility of the homology-based emphysema quantification was validated. By evaluating the spatial distribution of low-attenuation lung regions, the homology-based emphysema quantification could provide more accurate information about early changes in COPD than can LAA%. The homology-based emphysema quantification was useful for the assessment of emphysema severity.

Acknowledgment

Disclosure

The authors report no conflicts of interest in this work.

References

1. Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187(4): 347–365.
2. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.
3. Müller NL, Staples CA, Miller RR, Abboud RT. "Density mask". An objective method to quantitate emphysema using computed tomography. *Chest*. 1988;94(4):782–787.
4. Mishima M, Hirai T, Itoh H, et al. Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. *Proc Natl Acad Sci U S A*. 1999;96(16):8829–8834.
5. Nakano Y, Muro S, Sakai H, et al. Computed tomographic measurements of airway dimensions and emphysema in smokers. Correlation with lung function. *Am J Respir Crit Care Med*. 2000;162(3 Pt 1): 1102–1108.
6. Lee YK, Oh YM, Lee JH, et al. Quantitative assessment of emphysema, air trapping, and airway thickening on computed tomography. *Lung*. 2008;186(3):157–165.
7. Mets OM, Murphy K, Zanen P, et al. The relationship between lung function impairment and quantitative computed tomography in chronic obstructive pulmonary disease. *Eur Radiol*. 2012;22(1):120–128.
8. 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(11):1711–1715.
9. Tanabe N, Muro S, Hirai T, et al. Impact of exacerbations on emphysema progression in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2011;183(12):1653–1659.
10. Herzog J, Hibi T. *Monomial Ideals*. London: Springer-Verlag; 2010.
11. Gross JL, Tucker TW. *Topological Graph Theory*. Wiley Interscience; 1987. Paperback edition by Dover Publications; 2001.
12. Nakane K, Takiyama A, Mori S, Matsuura N. Homology-based method for detecting regions of interest in colonic digital images. *Diagn Pathol*. 2015;10:36.
13. Nakane K, Tsuchihashi Y, Matsuura N. A simple mathematical model utilizing topological invariants for automatic detection of tumor areas in digital tissue images. *Diagn Pathol*. 2013;8(Suppl 1):S27.
14. Ishida M, Kida K, Mizobe K, et al. Relation between the Betti Number of Fatigue Fracture Surfaces and Stress Intensity Factors of Low Carbon Steel (JIS, S45C). *Advanced Materials Research*. 2015;1102:59–63.
15. Nakane K, Santos EC, Honda T, et al. Homology analysis of structures of high carbon bearing steel: effect of repeated quenching on prior austenite grain size. *Materials Research Innovations*. 2014;18(Suppl 1): 33–37.
16. Adcock A, Rubin D, Carlsson G. Classification of hepatic lesions using the matching metric. *Computer Vision and Image Understanding*. 2014; 121:36–42.
17. Mischaikow K. CHomP, Computational Homology Project. Available from: http://chomp.rutgers.edu/. Accessed July 14, 2016.
18. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. 1974;19(6):716–723.
19. Nishio M, Matsumoto S, Koyama H, et al. Emphysema quantification by combining percentage and size distribution of low-attenuation lung regions. *AJR Am J Roentgenol*. 2014;202(5):W453–W458.
20. Ju J, Li R, Gu S, et al. Impact of emphysema heterogeneity on pulmonary function. *PLoS One*. 2014;9(11):e113320.
21. Tanabe N, Muro S, Tanaka S, et al. Emphysema distribution and annual changes in pulmonary function in male patients with chronic obstructive pulmonary disease. *Respir Res*. 2012;13:31.
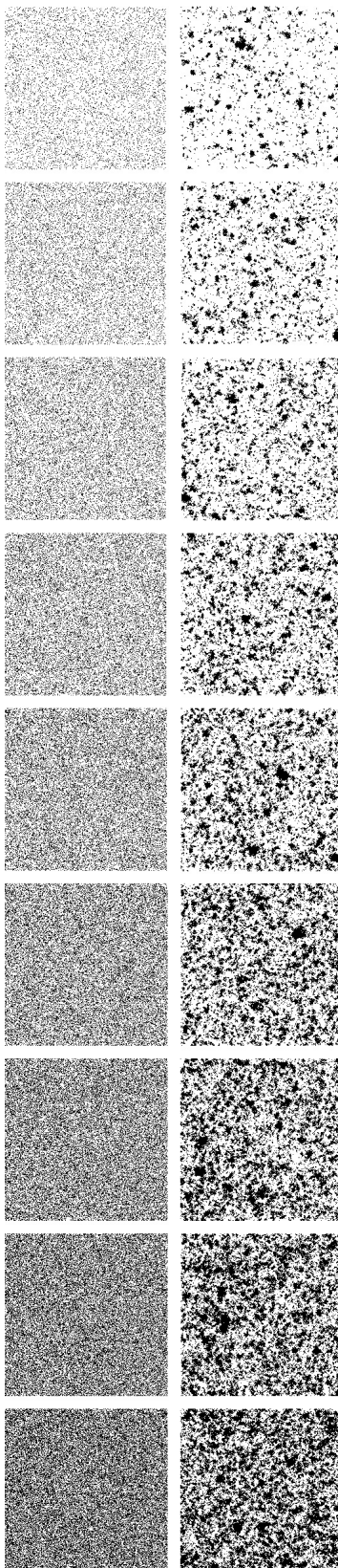
Supplementary materials

Figure S1 All binarized images used in the simulation study.
Notes: Left: the binarized images obtained from simulation A. Right: the binarized images obtained from simulation B. LAA% of the two images in the same row was identical. For each row, LAA% increases by 5%.
Abbreviation: LAA%, percentage of low-attenuation lung area.

Table S1 Results of emphysema quantification in the three groups

Variables	Group A		Group B		Group C	
	Mean	SD	Mean	SD	Mean	SD
LAA%	4.54	9.83	7.66	9.66	26.26	19.13
b_0	617.8	366.8	884.3	500.4	1,169.4	404.5
b_1	42.73	96.71	87.51	133.51	259.00	198.01
nb_0	0.00429	0.00168	0.00579	0.00266	0.00694	0.00231
nb_1	0.000239	0.000517	0.000529	0.000742	0.001461	0.001036
R	0.0358	0.0697	0.0650	0.0786	0.2086	0.1648

Notes: Group A, nonsmoker; Group B, smokers without COPD, mild COPD, and moderate COPD; Group C, severe COPD and very severe COPD; b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0 normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
Abbreviations: LAA%, percentage of low-attenuation lung area; SD, standard deviation.

Table S2 Summary of the emphysema quantification in all the patients, nonsmokers, smokers without COPD, and COPD patients

Variables	All patients		Nonsmokers		Smokers		COPD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LAA%	10.73	14.08	4.54	9.83	4.69	6.89	18.53	16.54
b_0	902.1	490.2	617.8	366.8	780.0	514.2	1,108.7	417.0
b_1	113.7	159.5	42.7	96.7	66.4	139.0	181.7	169.2
nb_0	0.00580	0.00258	0.00429	0.00168	0.00522	0.00278	0.00684	0.00220
nb_1	0.000665	0.000870	0.000239	0.000517	0.000401	0.000769	0.001054	0.000900
R	0.0880	0.1147	0.0358	0.0697	0.0456	0.0726	0.1455	0.1335

Variables	Mild COPD		Moderate COPD		Severe COPD		Very severe COPD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
LAA%	12.91	7.51	12.72	12.09	20.66	16.77	35.35	20.26
b_0	926.3	393.6	1,079.5	436.2	1,096.5	447.0	1,287.9	314.7
b_1	94.0	77.5	127.3	121.8	207.2	171.5	343.1	220.3
nb_0	0.00615	0.00237	0.00683	0.00218	0.00670	0.00250	0.00733	0.00206
nb_1	0.000614	0.000490	0.000765	0.000671	0.001234	0.000992	0.001830	0.001064
R	0.0941	0.0551	0.0987	0.0817	0.1587	0.1269	0.2896	0.1946

Notes: b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; nb_0, b_0 normalized to the total number of lung pixels; nb_1, b_1 normalized to the total number of lung pixels; R, b_1/b_0.
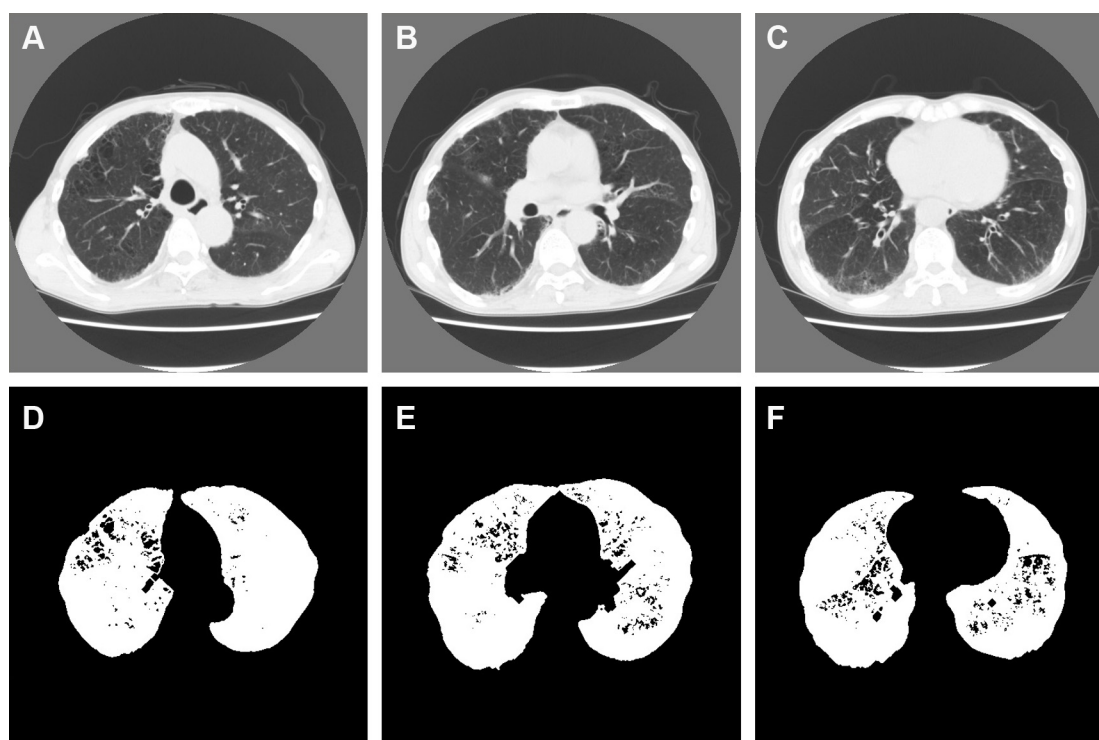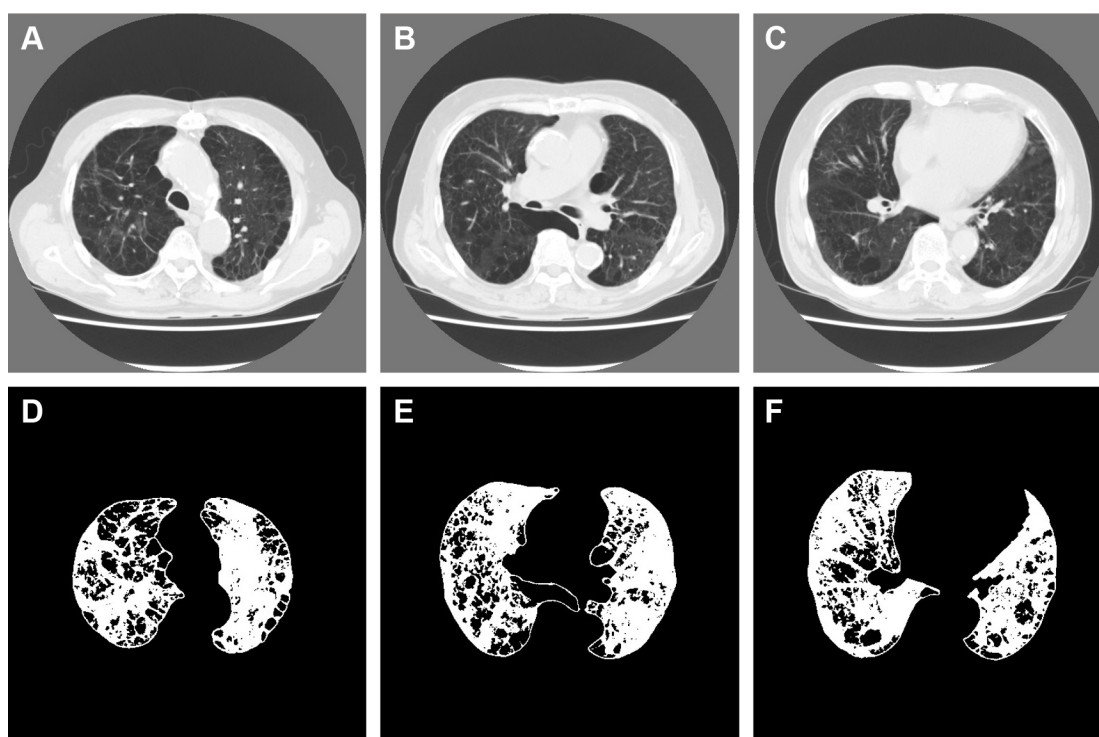**Abbreviations:** LAA%, percentage of low-attenuation lung area; SD, standard deviation.
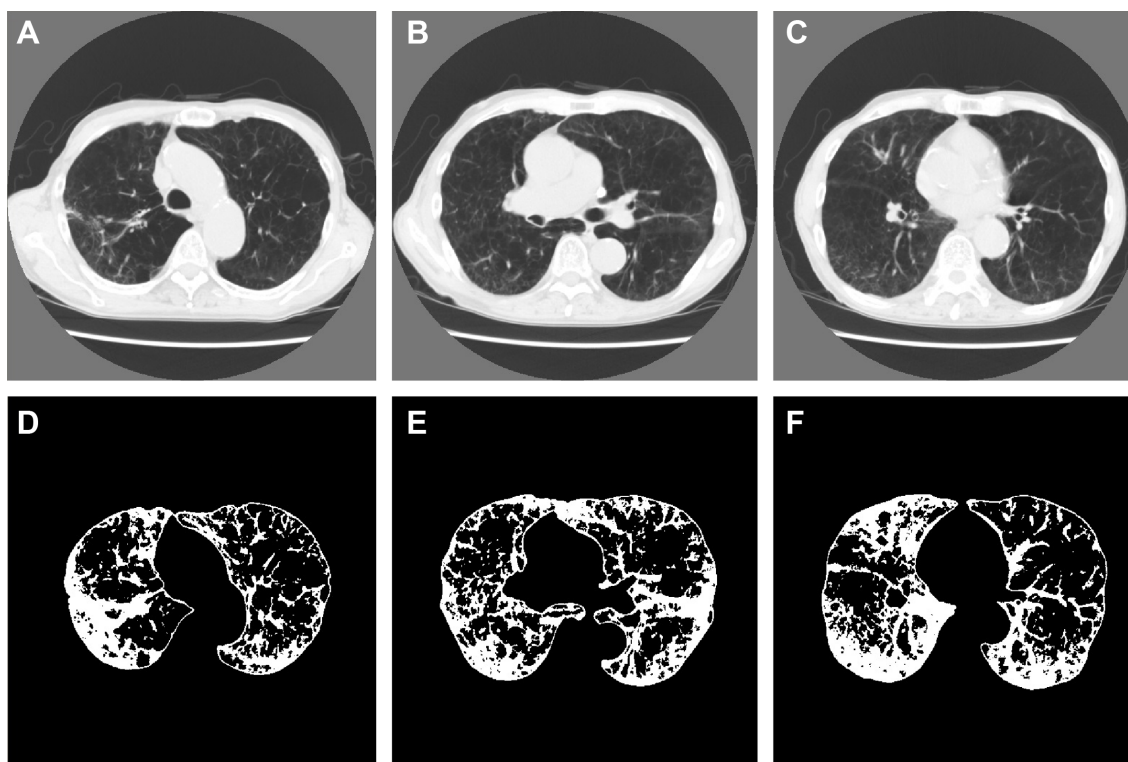
Figure S2 Representative computed tomography (CT) images and results of preprocessing of the upper, middle, and lower lung fields in a 64-year-old man with COPD.
Notes: Forced expiratory volume in 1 second (FEV_1)/forced vital capacity; FEV_1; b_0, b_1, R; and the percentage of low-attenuation lung area were 67.4%; 61.4%; 944, 39, 0.0413; and 5.13%, respectively (b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; R, b_1/b_0). The CT images and the binarized images are shown in (**A–C**) and (**D–F**), respectively.

Figure S3 Representative computed tomography (CT) images and results of preprocessing of the upper, middle, and lower lung fields in a 77-year-old man with COPD.
Notes: Forced expiratory volume in 1 second (FEV_1)/forced vital capacity; FEV_1; b_0, b_1, R; and the percentage of low-attenuation lung area were 42.8%; 50.2%; 1,420, 321, 0.226; and 26.5%, respectively (b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; R, b_1/b_0). The CT images and the binarized images are shown in (**A–C**) and (**D–F**), respectively.

Figure S4 Representative computed tomography (CT) images and results of preprocessing of the upper, middle, and lower lung fields in a 75-year-old man with COPD.
Notes: Forced expiratory volume in 1 second (FEV_1)/forced vital capacity; FEV_1; b_0, b_1, R; and the percentage of low-attenuation lung area were 28.1%; 27.5%; 977, 443, 0.453; and 56.6%, respectively (b_0, the zero-dimensional Betti number; b_1, the one-dimensional Betti number; R, b_1/b_0). The CT images and the binarized images are shown in (**A**–**C**) and (**D**–**F**), respectively.