COMPUTED TOMOGRAPHY

Low morphometric complexity of emphysematous lesions predicts survival in chronic obstructive pulmonary disease patients

Jeongeun Hwang[1] · Yeon-Mok Oh[2,3] · Minho Lee[4] · Seunghyun Choi[3] · Joon Beom Seo[5] · Sang Min Lee[5] · Namkug Kim[3,5]

Abstract

Objectives To investigate whether morphometric complexity in the lung can predict survival and act as a new prognostic marker in patients with chronic obstructive pulmonary disease (COPD).

Methods COPD (n = 302) patients were retrospectively reviewed. All patients underwent volumetric computed tomography and pulmonary function tests at enrollment (2005–2015). For complexity analysis, we applied power law exponent of the emphysema size distribution (D_{size}) as well as box-counting fractal dimension (D_{box3D}) analysis. Patients' survival at February 2017 was ascertained. Univariate and multivariate Cox proportional hazards analyses were performed, and prediction performances of various combinatorial models were compared.

Results Patients were 66 ± 6 years old, had 41 ± 28 pack-years' smoking history and variable GOLD stages (n = 20, 153, 108 and 21 in stages I−IV). The median follow-up time was 6.1 years (range: 0.2−11.6 years). Sixty-three patients (20.9%) died, of whom 35 died of lung-related causes. In univariate Cox analysis, lower D_{size} and D_{box3D} were significantly associated with both all-cause and lung-related mortality (both $p < 0.001$). In multivariate analysis, the backward elimination method demonstrated that D_{box3D}, along with age and the BODE index, was an independent predictor of survival ($p = 0.014$; HR, 2.08; 95% CI, 1.16–3.71). The contributions of D_{size} and D_{box3D} to the combinatorial survival model were comparable with those of the emphysema index and lung-diffusing capacity.

Conclusions Low morphometric complexity in the lung is a predictor of survival in patients with COPD.

Key Points
• *A newly suggested method for quantifying lung morphometric complexity is feasible.*
• *Morphometric complexity measured on chest CT images predicts COPD patients' survival.*
• *Complexity, diffusing capacity and emphysema index contribute similarly to the survival model.*

Keywords COPD · Emphysema · Fractals · Lung · Survival

Sang Min Lee and Namkug Kim contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5551-7) contains supplementary material, which is available to authorized users.

✉ Sang Min Lee
sangmin.lee.md@gmail.com

✉ Namkug Kim
namkugkim@gmail.com

1 Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea

2 Department of Pulmonary and Critical Care Medicine, University of Ulsan College of Medicine, Asan Medical Center, Seoul, Republic of Korea

3 Department of Convergence Medicine, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Seoul 05505, Republic of Korea

4 Biomedical Engineering Research Center, Asan Institute of Life Sciences, Asan Medical Center, Seoul, Republic of Korea

5 Department of Radiology and Research Institute of Radiology, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Seoul 05505, Republic of Korea

Abbreviations

BODE   Integrated COPD prognostic index of four factors: the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D) and exercise capacity (E), measured by the 6-min walk test

cDLCO% Percentage of diffusing capacity of the lung for carbon monoxide corrected by haemoglobin to the expected value

CI   Confidence interval

C-index Concordance index

D_{box3D} Box-counting fractal dimension of the lung parenchyma in full-3D

D_{size} Power law exponent of the size distribution of emphysema clusters

EI%   Percentage of the lung volume occupied by emphysema

HR   Hazard ratio

KOLD Korean obstructive lung disease

LAA   Low attenuation area

PLE   Power law exponent

Introduction

In the prognosis of chronic obstructive pulmonary disease (COPD), the underlying morphological abnormalities in the airways and alveoli underlie the heterogeneity of clinical outcome and mortality [1, 2]. Long-term prospective studies in various settings have succeeded in identifying significant prognostic markers [3–9]. The BODE index, a multidimensional grading system, including the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D) and exercise capacity (E), has been proposed to be a better mortality predictor than a single index [4]. Specifically, in quantitative imaging, Haruna et al [6] showed that the percentage of a low attenuation area (LAA) on chest computed tomography (CT) effectively predicts mortality. In contrast, although there is an association between CT airway measurements and COPD risk [10], the relationship with mortality remains unclear. Advanced quantitative imaging technologies have not yet been used to strengthen prediction models. We propose implementing the extent of morphological complexity in CT images of COPD patients to improve survival prediction.

Mishima et al [11] showed that the size distributions of LAAs from lung CTs of COPD patients followed the power law, implying a fractal structure. The power law exponent (PLE) is associated with pulmonary function, and can be used as an indicator of the extent of lung complexity [11–15]. The functional capacity of the lung is mainly established by its complex morphometry [14–17]. Similar previous studies have utilised fractal properties to gain a better understanding of lung structure-function relationships [11, 14, 15, 17–24]. A longitudinal study demonstrated that emphysema progression decreases heterogeneity [9], implying that PLE carries essential information about COPD pathophysiology. However, in these aforementioned studies, complexity analyses were performed using two-dimensional (2D) images; three-dimensional (3D) information was not fully utilised.

Here, we hypothesised that emphysema compromises lung morphometric complexity, diminishing lung function and consequently death. If true, a complexity measure could predict patient survival. We applied fractal dimension methods with PLE and box-counting to baseline CT data, and followed up patient survival to ascertain the predictive performance of the model.

Materials and methods

This study was approved by the institutional review board of the Asan Medical Center (no. 2005-0345). Written informed consent was obtained from all patients.

Patients and clinical assessments

Records of subjects from a Korean obstructive lung disease (KOLD) cohort were retrospectively reviewed. All subjects met the following criteria: a diagnosis of COPD, i.e. post-BD forced expiratory volume in 1 s (FEV1)/forced vital capacity (FVC), 0.7, and no or minimal abnormality, other than emphysema, on chest radiographs. The KOLD study design has been published previously [25]. 302 COPD patients who underwent CT scans between June 2005 and June 2015 were enrolled, and their survival status at February 2017 was ascertained. Age, sex and smoking history were recorded and body mass index, 6-min walking distance and dyspnoea scale were measured at enrolment. For all patients, all test results, measurements, spirometry testing and CT scans were acquired within 1 month of enrolment. Pulmonary function tests and CT scans were performed according to the KOLD cohort study protocol [25]. The BODE index was calculated in accordance with Celli et al [4].

CT scans

All subjects underwent multi-detector row CT. Details on the CT and automatic segmentation methods are provided in Supplementary Material 1, and in previous publications [25, 26].

Complexity analysis

For complexity analysis, two methods were applied simultaneously: the PLE of the size distribution of emphysema, D_{size}, and the box-counting fractal dimension, D_{box3D}. The

algorithm for measuring D_{size} has been described previously [26, 27] and is summarised in Supplementary Material 1.

D_{box3D} was measured using an in-house box-counting fractal analysis tool with MATLAB release 2015a (MathWorks Inc., Natick, MA, USA) that utilises the full-3D CT data. The box-counting fractal dimension was measured as first suggested by Grassberger [12] and presented by Ott [13]. Figure 1 shows a conceptual representation of this method. In brief, numbers of imaginary cubes needed to cover the lung parenchyma mask, which is automatically generated by subtracting the emphysema mask from the lung mask, are counted, using varying sizes, to yield a power-law distribution, and its exponent is considered as the fractal dimension. A higher box-counting fractal dimension would imply a greater space-filling property of the lung parenchyma in 3D space. Further details are provided in Supplementary Material 1.

To compare the performance and reliability of the full-3D method with the conventional 2D method, 2D box-counting fractal analysis was concurrently performed, and further details are provided in Supplementary Material 1.

Statistical analysis

To investigate the prediction power of D_{size} and D_{box3D} for mortality, univariate and multivariate Cox proportional hazards analyses [28] were performed. In the multivariate mortality model, multicollinear predictors with a variance inflation factor (VIF) > 3 were excluded. Conditional backward elimination was applied to remove insignificant predictors with a probability > 0.05. As there are no generally accepted optimal cut-off values for risk factors involving continuous variables for differentiating these groups, we stratified patients according to Martinez et al [8] and Haruna et al [6]. The poorest performing quintile (approximately 20% of the population) were designated the 'unhealthy' group and the rest the reference group. Data are shown as mean ± standard deviation, and p-values as well as the 95% confidence interval (95% CI) were calculated. To evaluate the performance of the prediction models, the concordance index (C-index) and area under the receiver operating characteristic curve (AUC-ROC) were analysed. AUC-ROC was determined for 6-, 9- and 12-year survival. R statistics software
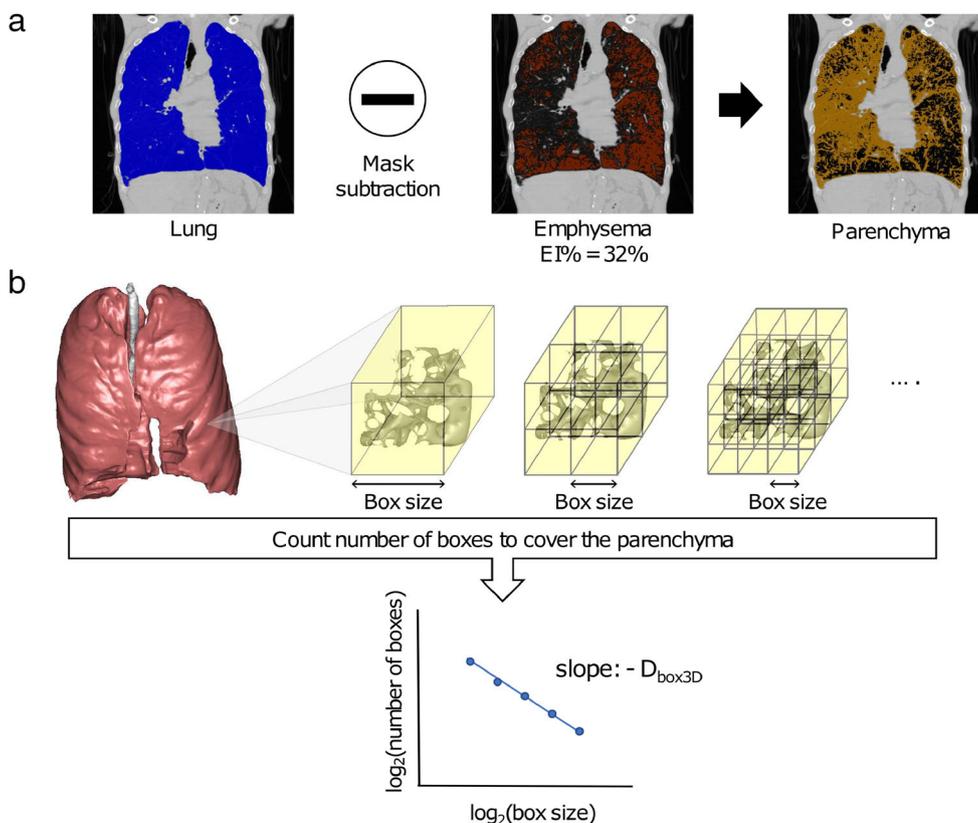

Fig. 1 A schematic representation of D_{box3D} measurement by a box-counting fractal dimension analysis method with a representative patient image. (**a**) Automatically segmented lung and emphysema lesion masks are the two major inputs for the analysis. A lung mask is defined by a volume with < −400 HU, with the exception of large airways, and an emphysematous lesion mask has a volume of < −950 HU in the lung. The patient shown in this figure had EI% = 32%. The lung mask subtracted from the emphysematous lesion mask represents the lung parenchyma mask. (**b**) Imaginary cubes with various sizes scan the lung space in three-dimensions and the algorithm counts the number of cubes needed to cover the lung parenchyma; a logarithmic regression is performed to estimate the D_{box3D} value. More details about measuring D_{box3D} can be found in Supplementary Material 1

version 3.4.3 (R Foundation for Statistical Computing, Vienna, Austria) and the following R packages were used: dplyr [29], survival [30] and survivalROC [31].

Results

Baseline characteristics and mortality outcome

The patients' baseline characteristics are shown in Table 1. Eleven patients were excluded for BODE measurements due to missing results for the 6-min walk test. The median follow-up time was 6.1 years (range: 0.2−11.6 years). Sixty-three patients (20.9%) died; the median survival was 5.1 years. For 35 patients, death was lung-related, and included respiratory failure (17 patients), pneumonia (11 patients), lung cancer (eight patients), sepsis (four patients) and pneumothorax (one patient), with some patients having multiple causes. The other causes of death were non-lung cancer malignancy (seven patients), cardiovascular disease (three patients), septic shock (one patient) and unknown (17 patients).

Table 1 Baseline characteristics of 302 chronic obstructive pulmonary disease patients

Characteristic	Mean ± SD (range)
Age (years)	66 ± 6 (52–88)
Sex (Male:Female)	283:19
GOLD[a] stage	I: 20 II: 153 III: 108 IV: 21
Smoking history (pack/year)	41 ± 28 (0–165)
BODE[b]	0–4: 240, 5–10: 51, missed: 11
cDLCO%[c]	72.31 ± 24.26 (22.89–147.25)
EI%[d]	20 ± 15 (0–61)
D_{size}[e]	0.4348 ± 0.1869 (0.1506–1.1601)
D_{box3D}[f]	2.9335 ± 0.0829 (2.4885–2.9999)

Data are presented as mean ± SD (range) for age, smoking history, cDLCO%, EI%, D_{size} and D_{box3D}. Counts are presented for sex and BODE. No patient data were missing, except for the BODE index in 11 patients

[a] GOLD = Global Initiative for Chronic Obstructive Lung Disease

[b] BODE = an integrated COPD prognostic index of four factors: the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D) and exercise capacity (E), measured by the 6-min walk test

[c] cDLCO% = the percentage of diffusing capacity of the lung for carbon monoxide corrected with haemoglobin to expected value

[d] EI% = the percentage of the lung volume occupied by emphysema

[e] D_{size} = the power law exponent of the size distribution of emphysema clusters;

[f] D_{box3D} = box-counting fractal dimension of the lung parenchyma in full-3D

Prognostic significance of predictors

The relationships among all-cause or lung-related mortality and candidate predictors were assessed using Cox proportional hazards analysis (Table 2). Older age (hazard ratio [HR], 4.44; 95% confidence interval [CI], 2.63–7.50), heavier smoking history (HR, 2.00; 95%CI, 1.17–3.40), higher BODE index (HR, 7.14; 95%CI, 4.27–11.95), lower percentage of diffusing capacity of the lung for carbon monoxide corrected with hemoglobin to an expected value (cDLCO%, HR, 4.46; 95%CI, 2.70–7.44), higher percentage of lung volume occupied by emphysema (EI%, HR, 3.95; 95%CI, 2.41–6.49), higher D_{size} (HR, 3.46; 95%CI, 2.09–5.70), and higher D_{box3D} (HR, 4.05; 95%CI, 2.46–6.67) were significantly associated with all-cause mortality. Sex (female sex HR, 0.73; 95%CI, 0.23–2.32) was not a significant predictor ($p = 0.059$). The same predictors also significantly predicted lung-related mortality.

Considerable multicollinearity (VIF > 3) between cDLCO%, EI%, D_{size}, and D_{box3D} were found, and thus, only one should be included in multivariate analysis. In univariate analysis, cDLCO% performed best in predicting all-cause mortality (HR, 4.46), followed by D_{box3D} (HR, 4.05), whereas D_{box3D} performed best in predicting lung-related mortality (HR, 6.16), followed by EI% (HR, 5.71). Because the superiority among cDLCO%, EI%, D_{size} and D_{box3D} as a predictor was difficult to distinguish, four models including cDLCO%, EI%, D_{size} or D_{box3D}, with variables age, sex, smoking history and BODE index, were tested using conditional backward elimination. In the model including D_{box3D}, the variables age, smoking history, BODE index and D_{box3D} were identified as significant predictors for all-cause mortality (C-index, 0.790) and lung-related mortality (C-index, 0.785) (Table 3). Other models including cDLCO%, EI% or D_{size} instead of D_{box3D} showed qualitatively similar results (Supplementary Tables S1−3). The analysis revealed that the D_{box3D} value was a significant predictor of mortality of COPD patients, independent of age and the BODE index.

Predicting survival by fractal dimension

The prediction power of D_{box3D} on COPD patient mortality was assessed using Kaplan-Meier survival curves of the D_{box3D} subgroups (Fig. 2). The low-D_{box3D} subgroup is the lowest quintile (n = 60) and the high-D_{box3D} subgroup comprised all the remaining patients (n = 242) (Table 2). For both all-cause death and lung-related death, the survival curves for the high- and low-D_{box3D} subgroups differed significantly ($p < 0.001$).

Contribution of complexity measures to combinatorial prediction models

Table 4 shows combinatorial prediction models and their prediction performances. A multivariate Cox proportional hazard model including age, smoking history and BODE index was

Table 2 Univariate Cox proportional hazards regression analysis showing prognostic significance of demographic parameters, smoking history, BODE index, cDLCO%, EI%, D_{size} and D_{box3D}

Variable	No.	HR[a] (95% CI[b]) All-cause death (n = 63)	p-value	HR (95% CI) Respiratory (n = 35)	p-value
Age at baseline, years					
74–88	58	4.44 (2.63–7.50)	<0.001	4.56 (2.28–9.10)	<0.001
52–73	244	Reference		Reference	
Sex					
Female	13	0.73 (0.23–2.32)	0.592	0.89 (0.21–3.71)	0.873
Male	289	Reference		Reference	
Smoking history, pack-years					
55–165	60	2.00 (1.17–340)	0.011	2.20 (1.10–4.44)	0.027
0–54	242	Reference		Reference	
BODE Index[c]					
5–10	51	7.14 (4.27–11.95)	<0.001	8.38 (4.21–16.68)	<0.001
0–4	240	Reference		Reference	
cDLCO%[d]					
23–51	59	4.46 (2.70–7.44)	<0.001	5.69 (2.90–11.15)	<0.001
51–147	243	Reference		Reference	
EI%[e]					
34–61	60	3.95 (2.41–6.49)	<0.001	5.71 (2.92–11.16)	<0.001
0–34	241	Reference		Reference	
D_{size}[f]					
0.1506–0.2682	60	3.46 (2.09–5.70)	<0.001	4.79 (2.46–9.30)	<0.001
0.2685–1.1601	242	Reference		Reference	
D_{box3D}[g]					
2.4885–2.8859	60	4.05 (2.46–6.67)	<0.001	6.16 (3.15–12.06)	<0.001
2.8867–2.9999	242	Reference		Reference	

For each variable, patients were stratified into an unhealthy group or a reference group. The poorest performing quintile was defined as the unhealthy group, and the remainder as the reference group

[a] HR = Hazard ratio

[b] CI = Confidence interval

[c] BODE = an integrated COPD prognostic index of four factors: the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D) and exercise capacity (E), measured by the 6-min walk test

[d] cDLCO% = the percentage of diffusing capacity of the lung for carbon monoxide corrected with haemoglobin to expected value

[e] EI% = the percentage of the lung volume occupied by emphysema

[f] D_{size} = the power law exponent of the size distribution of emphysema clusters

[g] D_{box3D} = box-counting fractal dimension of the lung parenchyma in full-3D

used as a basis to which cDLCO%, EI%, D_{size} and D_{box3D} were added, individually. D_{size} and D_{box3D} were each added to the base model plus cDLCO% and base model plus EI%. The C-index and AUC-ROC for 6-, 9- and 12-year survival of the resulting nine combinatorial models were evaluated. The base model was concordant (C-index, 0.777) and highly predictive of 6-, 9- and 12-year survival (AUC-ROC, 0.828, 0.837 and 0.824, respectively). Adding D_{size} or D_{box3D} to the base model enhanced its performance subtly. Adding cDLCO% or EI% instead of D_{size} or D_{box3D} similarly enhanced prediction performance. Combining the base model,

cDLCO% or EI%, and D_{size} or D_{box3D} minimally affected prediction performance. These effects were too subtle to indicate superiority among the predictors. Thus, the contribution of D_{size} or D_{box3D} is comparable to that of EI% or cDLCO%.

Comparing the 2D versus 3D box-counting methods

Figure 2 shows the difference between the 2D and 3D box-counting methods in prediction performance. Compared with the 2D box-counting (grey dots) for a small sampling number, such as three or five, the 3D box-counting method (blue line)

Table 3 Significant predictors in multivariate Cox proportional hazards regression analysis: age at baseline, smoking history, BODE index, and D_{box3D}. The concordance indices of the multivariate prediction model with all four significant predictors are shown

Variable	No.	HR[a] (95% CI[b])			HR (95% CI)		
		All-cause death (n = 63)	p-value	C-index[c]	Respiratory (n = 35)	p-value	C-index
Age at baseline, years				0.790			0.785
74–88	58	2.71 (1.56–4.72)	<0.001		2.49 (1.20–5.16)	0.014	
52–73	244	Reference			Reference		
Smoking history, pack-years							
55–165	60	1.93 (1.11–3.34)	0.019		2.16 (1.05–4.43)	0.036	
0–54	242	Reference			Reference		
BODE Index[d]							
5–10	51	4.27 (2.31–7.89)	<0.001		4.11 (1.82–9.30)	0.001	
0–4	240	Reference			Reference		
D_{box3D}[e]							
2.4885–2.8859	60	2.08 (1.16–3.71)	0.014		3.26 (1.49–7.13)	0.003	
2.8867–2.9999	242	Reference			Reference		

[a] HR = Hazard ratio

[b] CI = Confidence Interval

[c] C-index = Concordance index

[d] BODE = an integrated COPD prognostic index of four factors: the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D), and exercise capacity (E), measured by the 6-min walk test;

[e] D_{box3D} = box-counting fractal dimension of the lung parenchyma in full-3D

outperformed the 2D method in both predictive significance and reliability. Forty or more 2D slices should be sampled to achieve a similar performance as the 3D method.

Discussion

We here defined methods to quantify lung parenchymal morphometric complexity, and propose two complexity measures,

D_{size} and D_{box3D}; both are novel and significant predictors of survival in COPD patients.

In this study, we applied box-counting fractal analysis to CT scans of COPD patients, used the full-3D data, which yielded better prediction performance than 2D slice sampling, and revealed that morphometric complexity is closely associated with survival.

Lower morphometric complexity has been observed in patients with more severe COPD [11, 26]. We found that age, smoking history, BODE index, cDLCO%, EI%, D_{size}, and
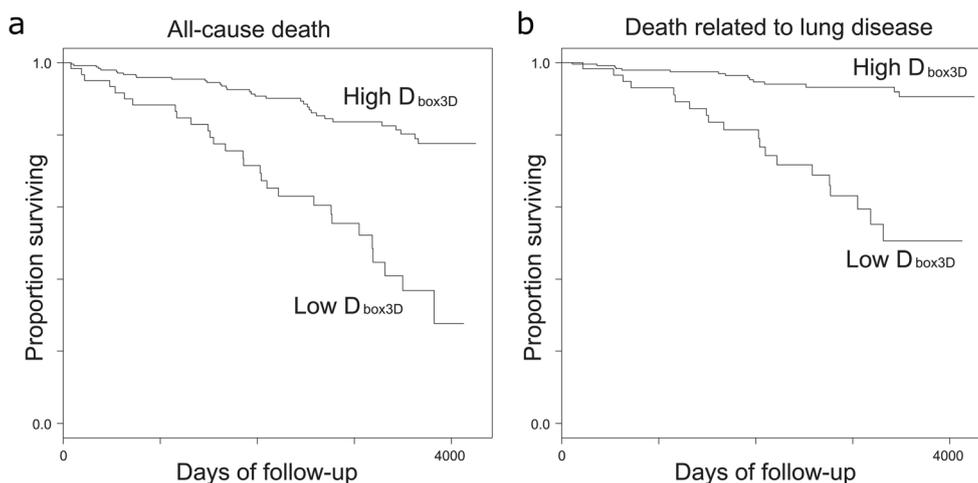

Fig. 2 Kaplan-Meier survival curves showing the outcomes of chronic obstructive pulmonary disease patients stratified by D_{box3D}, for all-cause death (**a**) and death related to lung (**b**). The patients were stratified into two groups, the lowest D_{box3D} quintile (n = 60) and the remaining patients (n = 242) with relatively higher D_{box3D}. The patients with high D_{box3D} showed a better prognosis than those with a low D_{box3D}, for both all-cause death and death related to lung. D_{box3D}: the Box-counting fractal dimension of the lung parenchyma shown in full in three dimensions

Table 4 Prediction performances of combinations of predictors. Concordance index (C-index) and area under the receiver operating characteristics (AUC-ROC) curves for various combinations of predictors are shown. Considering a prediction model comprising age, smoking history, and BODE index a base model, addition of D_{size} or D_{box3D} enhanced the model performance subtly. D_{size} or D_{box3D} benefitted the prediction model to a similar extent as cDLCO% or EI%

Predictors	C-index[a]	AUC-ROC[b]: 6-year survival	AUC-ROC: 9-year survival	AUC-ROC: 12-year survival
Base model: Age, smoking history, BODE[c] index	0.777	0.828	0.837	0.824
Base model, cDLCO%[d]	0.787	0.829	0.866	0.846
Base model, EI%[e]	0.795	0.841	0.851	0.844
Base model, D_{size}[f]	0.793	0.847	0.850	0.856
Base model, D_{box3D}[g]	0.790	0.843	0.846	0.845
Base model, cDLCO%, D_{size}	0.792	0.849	0.864	0.867
Base model, cDLCO%, D_{box3D}	0.789	0.837	0.859	0.858
Base model, EI%, D_{size}	0.799	0.851	0.856	0.860
Base model, EI%, D_{box3D}	0.793	0.845	0.852	0.847

[a] C-index = Concordance index

[b] AUC-ROC = Area under the receiver operating characteristic curve

[c] BODE = an integrated COPD prognostic index of four factors: the body mass index (B), the degree of airflow obstruction (O) and dyspnoea (D) and exercise capacity (E), measured by the 6-min walk test

[d] cDLCO% = the percentage of diffusing capacity of the lung for carbon monoxide corrected with haemoglobin to expected value

[el] EI% = the percentage of the lung volume occupied by emphysema

[f] D_{size} = the power law exponent of the size distribution of emphysema clusters

[g] D_{box3D} = box-counting fractal dimension of the lung parenchyma in full-3D

D_{box3D} are significant predictors of mortality, which was in agreement with previous studies [3–8], although the different study designs prohibited direct comparisons.

D_{size} and its correlations with COPD severity and progression, and its robustness to the breath-hold level have previously been studied [9, 11, 26]; however, its significance in mortality prediction has not yet been reported. Although D_{size} is a significant predictor, its performance is similar to or slightly inferior to those of EI%, cDLCO% and the BODE index.

The prognostic significance of D_{box3D} for both all-cause and lung-related mortality is equivalent to those of EI% and cDLCO%, and comparable to that of the BODE index. D_{size} accounts only for the size distribution of the emphysema lesions, whereas D_{box3D} also reflects the lesions' spatial distribution pattern.

The predictive significance of the BODE index and cDLCO% was similar in this and previous studies [4, 6–8]. The BODE index [4, 8] and/or its individual elements, FEV1 [3, 4, 6–8], body mass index [4, 6–8], dyspnoea scale [4, 7] and 6-min walking distance [4, 7, 8] were identified as significant survival predictors in our and previous studies. Similar to our results, DLCO% [8] and DLCO/VA [6] have previously been proposed as mortality predictors.

A positive association with smoking history was significant in our study, in contrast to previous studies [3–8]. We deduce that smoking history only marginally influences mortality in the KOLD cohort, in agreement with other studies. A relatively shorter smoking history (41 pack years ± 28; Table 1) in the KOLD cohort than subjects in previous studies may underlie the discrepancy.

EI% was a highly significant predictor of mortality in our and previous studies [5–7], but not in the study of Martinez et al [8], who found that the percentage of emphysema is not significant, but rather that the difference in the percentage of emphysema between the upper and lower lung was a predictive factor [8]. Their findings were in contradiction to those of Dawkins et al [5] and they proposed that it may be due to differences in the patient selection criteria [8]. Upper or lower lung predominance in terms of EI% and/or D_{box3D} should be investigated in future to elucidate this matter.

cDLCO%, EI%, D_{size} and D_{box3D} had similar prediction performances, and were also considerably collinear. To avoid multi-collinearity in multivariate analysis, we chose the variable D_{box3D} as it had the best HR in predicting lung-related death and the second best in predicting all-cause death. Although the HR of cDLCO% was slightly higher than that of D_{box3D} in all-cause death (4.46 vs. 4.05), it ranked only third in predicting lung-related death.

The conditional backward elimination method revealed D_{box3D} to be a significant variable, along with age, smoking history and BODE index; their predictive performances were similar. D_{box3D} had smaller HR values than those of the BODE index (Tables 2 and 3), but was not inferior.

D_{box3D} and all other predictors performed better for lung-related death than all-cause death (Tables 2 and 3 and Fig. 2). A D_{box3D} value in the lowest quintile, indicating relatively low

lung morphometric complexity, was a significant risk factor for mortality. The low-D_{box3D} quintile 10-year survival for lung-related diseases halved, while nearly 90% of the patients with high-D_{box3D} values survived.

D_{size} or D_{box3D} contributed to a combinatorial survival prediction model comprising age, smoking history and the BODE index, similar to EI% or cDLCO%, where the BODE index contributed the most.

We also investigated the benefits of utilising full-3D data over 2D data. The extent of agreement among the repeated 2D measurements for sampling 3, 5, 10, 20, 30 and 40 slices was almost saturated to 1 between 10–20 slices (Fig. 3, Supplementary Material 2). Even three sampling slices yielded a very high inter-class correlation coefficient (> 0.998) over 20 repetitions. D_{box2D} values from the sampled slices were highly concordant between repeated measurements (Supplementary Material 2). However, considering the prediction performance, at least 40 slices should be sampled to obtain stable HR values similar in quality to those of the 3D method (Fig. 3). Overall, we recommend the full-3D method rather than applying the 2D box-counting method to an arbitrary number of sampled slices, to improve mortality prediction.

We speculate that morphological complexity successfully predicts patients' survival because it represents the integrity of the lung parenchyma. Patients with healthier, more intact lung parenchyma have a higher probability of surviving longer than patients with a less complex lung. Considering the 'lung at risk' concept [32, 33], a lung parenchymal area that is apparently normal but under mechanical stress leads to disease progression; thus, a loss of morphometric complexity may be a prelude to progressive emphysema. By definition, fractal geometry allows irregular structures to be quantitatively characterised in geometric terms, even if they are difficult to describe by classical geometry [14, 16]. The box-counting fractal dimension method best represents the so-called

'space-filling property', which is similar to, but not equivalent to, EI%. More than simply estimating the volumetric ratio of the LAA with respect to the whole lung, the fractal dimension measures the complexity, the space-filling property, and the integrity of the lung parenchyma. Thus, the concept of 'a good lung', designed under fractal rules, is a highly space-filling and complex lung, as formulated by Weibel [14, 15, 17].

We also speculate that the predictive performance of D_{box3D} would have been superior to EI% or cDLCO% in a larger cohort and with a longer follow-up period. Although both EI% [6] and cDLCO% [8] are proven predictors of mortality in COPD patients, they have variable reproducibility [26, 34, 35]. EI% is susceptible to intra- and/or inter-subject variations in the breath-hold-level in CT imaging, whereas fractal measures are more robust [26]. The DLCO test also has variable reproducibility [35], requires additional time and effort to obtain, and cannot be performed properly by some patients [34]. D_{box3D} may be used in conjunction with cDLCO% in such circumstances.

The limitations of the current study are as follows: (1) The CT data were analysed only at baseline and not at follow-up. In future studies, follow-up CT data of our patients and follow-up lung clinical indices, such as the BODE index and cDLCO%, may reveal the nature and temporal evolution of these indices. (2) The causes of deaths were not completely followed up. Among 63 deaths, the cause of death in 17 patients remained unknown. (3) Accounting for the full-3D CT data acquired in quantification is a virtue of D_{box3D}; however, the full-greyscale data were excluded from the current analysis. The high collinearity between D_{box3D} and EI% may be due to the same -950 Hounsfield units-threshold binary masks that these parameters share. In future studies, a 4D fractal analysis could be performed with the inclusion of an additional greyscale dimension, which currently would require excessive computational time. (4) The study population was only Korean. External validation that includes populations
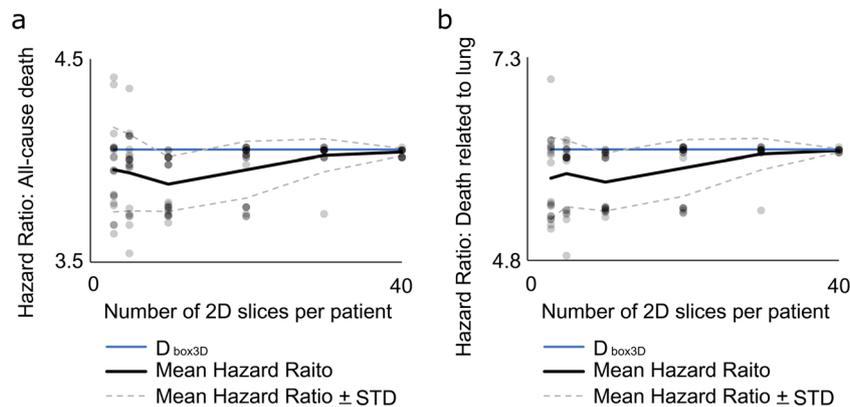

Fig. 3 The prediction significance of a two-dimensional slice selection schema for all-cause death (a) and death related to lung (b). When a small number of slices were selected, such as three or five, some random trials give higher hazard ratios than the three-dimensional (3D) measurement, but the average performance is lower than that of the 3D measurement.

Moreover, the predictive significance varies from trial to trial, which implies that the measurement is unreliable. Over 40 slices should be measured to achieve a similar performance to that of 3D box-counting fractal analysis

of multinational, multi-ethnic COPD patients may be needed. Nevertheless, we did demonstrate that morphological complexity is a significant predictor of mortality in COPD patients, and is unique and non-inferior to conventional prognostic indices.

In conclusion, we have demonstrated two methods for measuring the morphometric complexity in the lung on chest CT images, and determined that low morphometric complexity in the lung is a risk factor for mortality in patients with COPD, which has not been reported previously. Further investigations of lung complexity may facilitate our understanding of the nature of COPD and enhance the ability to predict mortality in clinical practice.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Namkug Kim.

Conflict of interest The authors of this article declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Nayoung Kim kindly provided statistical advice for this manuscript.

Informed consent Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval Institutional Review Board approval was obtained.

Study subjects or cohorts overlap The study subjects are from Korean Obstructive Lung Disease (KOLD) cohort. There are 54 articles by KOLD study group, and more than 100 articles stating the exact phrase "Korean Obstructive Lung Disease". We suppose that many of those would have a substantial extent of subject overlaps with our current study. However, the key point of our current study was to suggest new imaging biomarkers, and those biomarkers, D_{box} and D_{size}, have never been measured for any of the subjects before, making our current finding novel.

Methodology
• retrospective
• diagnostic or prognostic study
• multicentre study

References

1. Labaki WW, Martinez CH, Martinez FJ et al (2017) The role of chest computed tomography in the evaluation and management of the patient with COPD. Am J Respir Crit Care Med 196:1372–1379
2. Vogelmeier CF, Criner GJ, Martinez FJ et al (2017) Global strategy for the diagnosis, management, and prevention of chronic

obstructive lung disease 2017 report: GOLD executive summary (vol 53, pg 128, 2017). Arch Bronconeumol 53:411–412
3. Anthonisen NR, Wright EC, Hodgkin JE, Hopewell PC, Levin DC, Stevens PM (1986) Prognosis in chronic obstructive pulmonary-disease. Am Rev Respir Dis 133:14–20
4. Celli BR, Cote CG, Marin JM et al (2004) The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. N Engl J Med 350:1005–1012
5. Dawkins PA, Dowson LJ, Guest PJ, Stockley RA (2003) Predictors of mortality in alpha(1)-antitrypsin deficiency. Thorax 58:1020–1026
6. Haruna A, Muro S, Nakano Y et al (2010) CT scan findings of emphysema predict mortality in COPD. Chest 138:635–640
7. Kessler R, Faller M, Fourgaut G, Mennecier B, Weitzenblum E (1999) Predictive factors of hospitalization for acute exacerbation in a series of 64 patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 159:158–164
8. Martinez FJ, Foster G, Curtis JL et al (2006) Predictors of mortality in patients with emphysema and severe airflow obstruction. Am J Respir Crit Care Med 173:1326–1334
9. Tanabe N, Muro S, Sato S et al (2012) Longitudinal study of spatially heterogeneous emphysema progression in current smokers with chronic obstructive pulmonary disease. PLoS One 7:e44993
10. Kirby M, Tanabe N, Tan WC et al (2018) Total airway count on computed tomography and the risk of chronic obstructive pulmonary disease progression findings from a population-based study. Am J Respir Crit Care Med 197:56–65
11. Mishima M, Hirai T, Itoh H et al (1999) 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 96:8829–8834
12. Grassberger P (1983) On the fractal dimension of the henon attractor. Phys Lett A 97:224–226
13. Ott E (1993) Chaos in dynamical systems. Cambridge University Press, Cambridge
14. Weibel ER (1991) Fractal geometry - a design principle for living organisms. Am J Physiol 261:L361–L369
15. Weibel ER (2009) What makes a good lung? The morphometric basis of lung function. Swiss Med Wkly 139:375–386
16. Mandelbrot B (1983) The fractal geometry of nature. Freeman, New York
17. Weibel ER (2013) It takes more than cells to make a good lung. Am J Respir Crit Care Med 187:342–346
18. Vuidel G PFaCT Fractal analysis software. research team "Mobilities, city and transport" of the research centre ThéMA., France. Available via http://www.fractalyse.org/. Accessed 31 July 2017
19. Gilliard N, Pappert D, Spragg RG (1995) Fractal analysis of surfactant deposition in rabbit lungs. J Appl Physiol 78:862–866
20. Glenny R, Robertson HT (1991) Spatial correlation - a corollary of fractal pulmonary perfusion. FASEB J 5:A404–A404
21. Glenny RW, Robertson HT (1990) Fractal properties of pulmonary blood-flow - characterization of spatial heterogeneity. J Appl Physiol 69:532–545
22. Glenny RW, Robertson HT (1991) Fractal modeling of pulmonary blood-flow heterogeneity. J Appl Physiol 70:1024–1030
23. Glenny RW, Robertson HT, Yamashiro S, Bassingthwaighte JB (1991) Applications of fractal analysis to physiology. J Appl Physiol 70:2351–2367
24. Horsfield K (1990) Diameters, generations, and orders of branches in the bronchial tree. J Appl Physiol 68:1089–1097
25. Park TS, Lee JS, Seo JB et al (2014) Study design and outcomes of korean obstructive lung disease (Kold) cohort study. Tuberc Respir Dis 76:169–174
26. Hwang J, Lee M, Lee SM et al (2016) A size-based emphysema severity index: robust to the breath-hold-level variations and correlated with clinical parameters. Int J Chron Obstruct Pulmon Dis 11:1835–1841

27. Lee M KN, Lee SM, Seo JB, Oh SY (2015) Size-based emphysema cluster analysis on low attenuation area in 3D volumetric CT: comparison with pulmonary functional testProc SPIE 9417, Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging, Orlando, FL, USA, p 91472V

28. Cox DR (1972) Regression models and life-tables. J Royal Stat Soc B 34:187–220

29. Wickham H, Francois R, Henry L, Muller K (2017) dplyr: a grammar of data manipulation. R package version 0.7.3. Available via https://CRAN.R-project.org/package=dplyr

30. Therneau TM (2015) A package for survival analysis in S. Available via https://CRAN.R-project.org/package=survival

31. Saha-Chaudhuri PJ, Hapb P (2013) survivalROC: Time-dependent ROC curve estimation from censored survival data. Available via https://CRAN.R-project.org/package=survivalROC

32. Bhatt SP, Bodduluri S, Hoffman EA et al (2017) Computed tomography measure of lung at risk and lung function decline in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 196:569–576

33. Parr DG (2017) Quantifying the lung at risk in chronic obstructive pulmonary disease does emphysema beget emphysema? Am J Respir Crit Care Med 196:535–536

34. Ley B, Flicker BM, Hartman TE et al (2014) Idiopathic pulmonary fibrosis: CT and risk of death. Radiology 273:570–579

35. MacIntyre N, Crapo RO, Viegi G et al (2005) Standardisation of the single-breath determination of carbon monoxide uptake in the lung. Eur Respir J 26:720–735