CHARACTERIZATION OF EMPHYSEMA IN LOW DOSE COMPUTED TOMOGRAPHY IMAGES USING MAXIMUM LIKELIHOOD ESTIMATION OF THE SCALE EXPONENT

Mario Ceresa, Arrate Muñoz-Barrutia and Carlos Ortiz-de-Solórzano

Cancer Imaging Laboratory, Center for Applied Medical Research (CIMA)
University of Navarra, Pamplona, Spain.

ABSTRACT

Low attenuation clusters (LACs) are associated with the presence of emphysema in low dose computer tomography (LDCT) scans. It is believed that the LACs' frequency-size distribution can be appropriately fit with power law and be well described with the scale exponent. In this work, we investigated whether the LACs' frequency-size distribution of a group of 80 LDCT scans taken from a lung cancer-screening program followed a power-law behavior. Contrary to the current thinking, only a subset of the data (60%) actually allowed a power-law fit. This suggests that the claims of power-law behavior in LAC's frequency-size distribution is a delicate issue and should be regarded with care.

Index Terms— COPD, Emphysema quantification, power-law, scaling exponent, low dose computer tomography images.

1. INTRODUCTION

Pulmonary emphysema is caused by the progressive destruction of the tissue and structures surrounding the lung alveoli. Emphysema is mainly caused by continuous exposure of the lungs to inflammatory agents (e.g., cigarette smoke). Severe emphysema causes dyspnea and shortness of breath and can eventually lead the patient to death. The most common measure of emphysema extracted from CT images is the emphysema index (EI) - the ratio between the damaged and the total lung volume - which is conceptually simple and easy to use. Moreover, its reproducibility has been confirmed for low dose computer tomography (LDCT) images, and a procedure to compensate for changes in EI due to variations in the acquisition protocol [1] has recently been proposed. However, even though several studies found correlation between EI and degree of emphysema visually diagnosed by an expert radiologist, no study to this date has found a clear correlation between both in their Low Dose counterparts. This is a very important issue, since the detection of emphysema increasingly happens as an accidental finding in CT scans originally tailored for cancer screening. In those scans, the use of low dose parameters causes the presence of image artifacts and noise that affects the EI values. An alternative to characterize the lung affectation by emphysema is based on measuring the frequency-size distribution of low attenuation clusters (LACs) Mishima *et al*. [2] were the first to study the cumulative size distribution of LACs on 2D CT slices. They conclude it follows a power law characterized by a scaling exponent inversely related with the extension of LACs. This insight has been adopted by others [3, 4, 5, 6, 7]. The common method to perform the fitting of the data to a power-law is to use a least squares regression. It takes advantage of the fact that a power-law distribution follows the linear form $\log p(x) = -\alpha \log x + c$. However, this method has a number of limitations. Namely, the regression of the data on a log-log plot is not a consistent method to estimate the scaling exponent and can result in an inaccurate guess of power-law behavior. Moreover, the typical indicator of the quality of fit (i.e., the fraction of variance accounted for by the fitting) has little capability to discriminate a power-law distribution for empirical data. Instead, we propose to use a maximum likelihood estimation (MLE), which is more reliable, accurate and allows a goodness-of-fit test to establish if a power-law is an appropriate fit for the data [8]. Using this approach on patient image data obtained from a cancer screening program, we have been able to confirm that there is some partial grounds for the idea that LACs' frequency-size distribution is actually *scale free*, although not as generalized as previously assumed. In Section 2, we describe the subjects that participate in this study, the characteristics of the CT image acquisition, the image analysis pipeline and the details of the new estimation process. In Section 3, we present the results of the analysis. In Section 4, we present our conclusions and future lines of work.

2. MATERIAL AND METHODS

2.1. Subjects

Our study group consisted of 80 smokers enrolled in an early lung cancer detection program. To be eligible for the study, all

Correspondence: mceresa@unav.es. This work has partially been supported by the UTE Project CIMA, the Torres Quevedo program and the Spanish Science and Innovation Ministry (MICINN) under grants PSS-010000-2008-2 and DPI2009-14115-C03-03. A. Muñoz-Barrutia and C. Ortiz-de-Solórzano hold a Ramón y Cajal Fellowship of the MICINN.

Name	#	$\hat{\alpha}_{ls}$	$\hat{\alpha}_{mle}$	$\hat{p}_{mle}$	Tiffenau
(H) Healthy	29	1.19	2.35	0.34	77.93
(E) Emphysema	19	1.13	2.27(*)	0.21	75.16
(O) Obstruction	12	1.19	2.29	0.29	62.72
(EO) Emp+Obst	20	1.09	2.24(**)	0.22	59.24

Table 1. Descriptive parameters for the four groups: number, average scaling exponents with the least-square (α_{ls}) and maximum likelihood (α_{mle}) methods, goodness of fit parameter for the MLE estimation (p_{mle}) and Tiffeanu index. Note: * means that the value is significantly different from the same value of group H (p-value<0.05). ** the same but with p-value<0.01.

subjects had to be asymptomatic, at least 40 years-old, smokers of at least 10 pack-years, with no previous history of cancer. The University of Navarra institutional ethics committee approved the study protocol. One radiologist diagnosed all the patients and blindly scored the severity and extension of emphysema. Scoring was done using a scale from 0 to 4. We divided the patients into three groups VS0 (score 0, no emphysema), VS1 (score 1, low emphysema), VS2-3 (score 2 and 3, moderate and severe emphysema) according to the severity of the disease. We merged together the patients scoring 2 and 3 because of the reduced number of patients with moderate to severe affectation. In addition, forced spirometry before and after bronchodilators was performed on all the patients using a Vmax 22 spirometer (SensorMedics Corporation, Yorba Linda, CA, USA). The measurements were performed following the American Thoracic Society guidelines. The results were expressed as a percentage of the predicted reference values given by the European Respiratory Society. Patients with a post-bronchodilator FEV1/FVC lower than 70% were considered to have airway obstruction following the criteria established by the Global Initiative for Chronic Obstructive Lung Disease (GOLD). According to the values of pulmonary functional tests (PFTs) and the visual emphysema assessment of the radiologist (VS) we divided the patients in four groups: healthy (H; Tiffenau>=70, VS0), patients with emphysema (E; Tiffenau>=70, VS1, VS2-3), patients with airway obstruction (O; Tiffenau<70, VS0), and patients with both emphysema and obstruction (EO; Tiffenau<70, VS1, VS2-3).

2.2. CT acquisition

All CT exams were acquired during a single breath hold at end full inspiration after hyperventilation, using a 64-row multidetector CT scanner (Somatom Sensation 64; Siemens Healthcare, Forcheim, Germany). Low-dose CT parameters were employed (120 kVp, 40 mAs) with 2x0.6x32 mm collimation. Patients were examined in supine position and slices were obtained contiguously from the thoracic inlet to the adrenal glands without the use of intravenous contrast material. From the raw data, non-overlapping 1mm slice thickness images were obtained with a soft tissue (B40f) reconstruction algorithm. The protocol was the one used in the International Early Lung Cancer screen Program (IELCAP).

2.3. Image processing

The complete lung volume parenchyma and airway lumens were extracted from the CT image stack using the method described by Hu et al. [9]. Then the airways were subtracted from the pulmonary volume using an algorithm specifically designed for noisy LDCT scans [1]. Next, the volume of emphysematous lung was measured by applying the threshold of −960 Hounsfield units (HU), which according to literature [10] is the one that provides the best correlation between emphysema quantification and the corresponding histology.

LACs were then identified in the low attenuated lung volume as connected regions of more than one pixel (c1) within a simple 3D 6-neighbor connectivity. This particular connectivity was chosen out of the five possible values (c1, c2, c3, c4, c5) because it was the one that maximized the agreement between the automatic quantification of the scaling exponent and the visual emphysema assessment (data not shown).

2.4. Estimation of the scaling exponent

A continuous power-law distribution is described by a probability density

$$p(x) = Cx^{-\alpha}, \text{ for } x \geq x_{\text{min}} \tag{1}$$

where α is the scaling exponent, $C = (\alpha - 1)x_{\text{min}}^{\alpha-1}$is the normalization constant. x_{min} is the first value of x for which a probable power law behavior is observed.

The probability density function $p(x)$ was estimated from the histogram of the data[1].

The estimator of α using MLE is:

$$\hat{\alpha}_{mle} = 1 + n\left(\sum_{i=1}^{n} \ln \frac{x_i}{x_{\text{min}}}\right)^{-1} \tag{2}$$

where $x_i,\ i = 1, \cdots, n$, are the observed values of x such that $x_i \geq x_{\text{min}}$.

An initial, rough, estimation of $\hat{x}_{\text{min}}$ was computed using the Bayesian information criterion (BIC). A more precise, final estimate was then computed as the value of x_{min} that minimizes the Kolmogorov-Smirnov (KS) test, which measures

[1]To fit the data to the model we used the code made freely available by the authors of [8] in http://tuvalu.santafe.edu/∼aaronc/powerlaws/

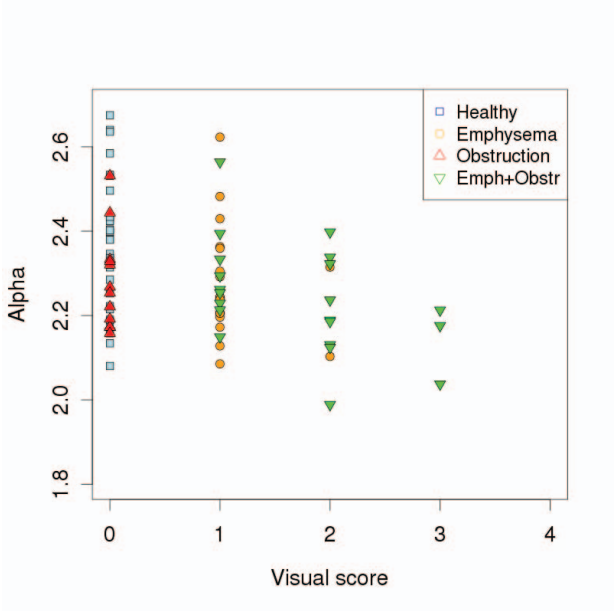


Fig. 1. : α_{mle} estimates versus emphysema visual score

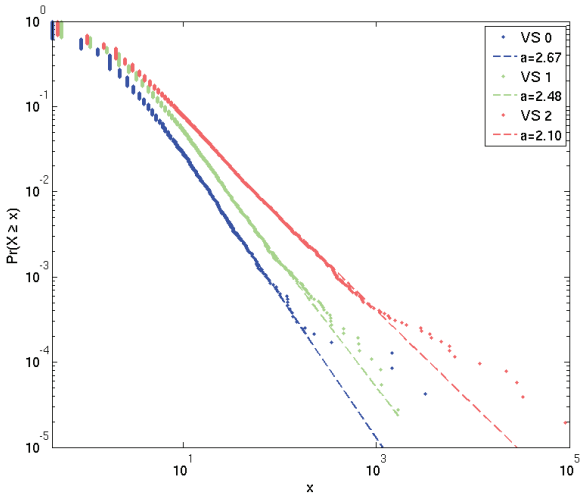


Fig. 2. Cumulative LACs size distribution fits for three patients with increasing emphysema severity (VS0: α_{mle}=2.67, p_{val}=0.155; VS1: α_{mle}=2.48, p_{val}=0.437; VS3: α_{mle}=2.10, p_{val}=0.226).

the maximum distance between the cumulative density functions (CDF) of the data and the fitted model. Additional details on the fitting procedure can be found in [8].

$$D = \max_{x \geq x_{\text{min}}} |S(x) - P(x)|, \tag{3}$$

where $S(x)$ is the CDF of the data and $P(x)$ is the CDF for the power-law model that best fits the data, both in the region $x \geq x_{\text{min}}$.

3. RESULTS

The average values of the estimated scaling exponents calculated with the least-square (α_{ls}) and maximum likelihood (α_{mle}) methods are given in Table 1. The additional value p_{mle} is the goodness of fit parameter for the MLE estimation. The differences for the α_{mle} estimate between the control group (H) and the E or EO groups are significative (p_{val} 0.03 and 0.002, respectively). Not statistically significant differences between groups were found for the α_{ls} estimate.

Figure 1 shows the result of the fit of $\hat{\alpha}_{mle}$ for the four groups (H, E, O, EO) and the agreement with the visual classification made by the radiologist. The large spread of the scale exponent within each emphysema class is mainly due to the differences in the internal pulmonary structure of patients. Noise and artifacts could also contribute to the inter-patient variability. The estimated scale exponents are very similar for the VS0 and VS1 groups, but decrease significantly for the VS2-3 group (p_{val}=0.0006). A similar tendency was described in [2], being supported by the fact that emphysema starts with small clusters which eventually coalesce as the disease progresses. So, clusters of higher sizes are formed which could result in lower estimates for the scale exponent. Figure 2 shows the cumulative size distribution of LACs and the fit for three representative cases with increasing severity of emphysema (VS0: α=2.67, p_{val}=0.155; VS1: α=2.48, p_{val}=0.437; VS3: α=2.10, p_{val}=0.226). Figure 3 shows the corresponding 3D lung reconstructions with the LACs in red. Since we used a MLE estimation, we can compute the goodness-of-fit to test whether the fitted data actually supports power-law behavior. In 40% of the cases the fit was low (p-val = 0.1). As expected, least squares regression was unable to see this misfit as the fraction of variance accounted for the fit were in the range 0.88 ± 0.03. When we visually reviewed the cases, the fit was good in images where emphysema was present as medium sized clusters but deviated from the power law behavior in the ones that displayed large emphysema clusters, located at the tail of the distribution. When analyzing the 60% of the cases with cluster distributions that did fit a power law, we could not find statistically significant differences between the three visually classified emphysema groups.

4. CONCLUSIONS

In this paper, we apply a theoretically sound method to demonstrate that a power-law distribution is a good fit for some but not all LACs frequency-size distributions in LDCT images. This is against the traditional belief that the power-law fit is good in all cases, and it could not be detected when using the classical least square regression method. Further re-

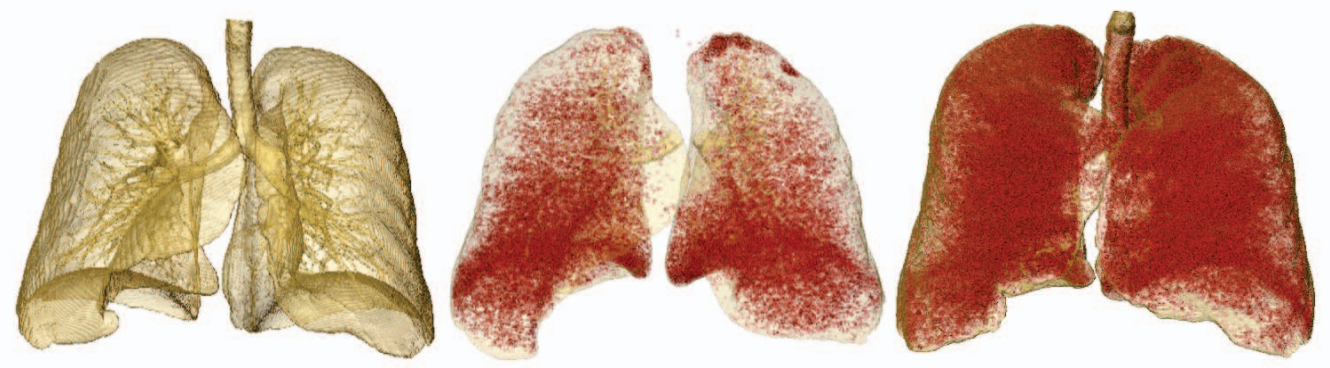


Fig. 3. 3D lung reconstructions with the law attenuation clusters in red for the three cases which fits are shown in Figure 2. From left to right, the affectation for emphysema increases.

search will help us determine whether the larger clusters that cause the distribution to escape from the power law behavior are noise contamination due to the LDCT technique or reflect some internal mechanism of the initiation and progression of emphysema, which in turn might have relevant implications in the diagnosis and prognosis of the disease.

5. REFERENCES

[1] M. Ceresa, G. Bastarrika, J.P. de Torres, L.M. Montuenga, J.J. Zulueta, C. Ortiz de Solorzano, and A. Muñoz-Barrutia, "Robust, standardized quantification of pulmonary emphysema in low dose CT exams," *Academic Radiology*, vol. 18, no. 11, pp. 1382–1390, Nov. 2011, PMID: 21852160.

[2] M. Mishima, T. Hirai, H. Itoh, Y. Nakano, H. Sakai, S. Muro, K. Nishimura, Y. Oku, K. Chin, M. Ohi, T. Nakamura, J. H. Bates, A. M. Alencar, and B. Suki, "Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease," *P Natl Acad Sci USA*, vol. 96, no. 16, pp. 8829–8834, 1999.

[3] A. Madani, A. van Muylem, V. de Maertelaer, J. Zanen, and P. A. Gevenois, "Pulmonary emphysema: Size distribution of emphysematous spaces on multidetector CT images - Comparison with macroscopic and microscopic morphometry," *Radiology*, vol. 248, no. 3, pp. 1036–1041, 2008.

[4] E. A. Hoffman, R. Jiang, H. Baumhauer, M. A. Brooks, J. J. Carr, R. Detrano, J. Reinhardt, J. Rodriguez, K. Stukovsky, and N. D. Wong, "Reproducibility and validity of lung density measures from cardiac CT Scans. The Multi-Ethnic study of atherosclerosis (MESA) lung study1," *Acad Radiol*, vol. 16, no. 6, pp. 689–699, 2009.

[5] B. M. Keller, A. P. Reeves, C. I. Henschke, R. G. Barr, and D. F. Yankelevitz, "Variation of quantitative emphysema measurements from CT scans," in *Proc. of SPIE*, San Diego, CA, USA, 2008, pp. 69152I–69152I–8.

[6] S. K. Alford, E. J. R. van Beek, G. McLennan, and E. A. Hoffman, "Heterogeneity of pulmonary perfusion as a mechanistic image-based phenotype in emphysema susceptible smokers," *P Natl Acad Sci USA*, vol. 107, no. 16, pp. 7485–7490, 2010.

[7] F. Mitsunobu, "Complexity of terminal airspace geometry assessed by computed tomography in asthma," *American Journal of Respiratory and Critical Care Medicine*, vol. 167, no. 3, pp. 411–417, Feb. 2003.

[8] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-Law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661, 2009.

[9] S. Hu, E. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantification of volumetric X-ray CT images," *IEEE T Med Imaging*, vol. 20, pp. 490–498, 2001.

[10] A. Madani, J. Zanen, V. de Maertelaer, and P. A. Gevenois, "Pulmonary emphysema: Objective quantification at multi-detector row CT - Comparison with macroscopy and microscopy morphometry," *Radiology*, vol. 238, no. 3, pp. 1036–1093, 2006.