



Published in final edited form as:

Physiol Meas. 2014 May ; 35(5): 833–845. doi:10.1088/0967-3334/35/5/833.

Direct Assessment of Lung Function in COPD Using CT Densitometric Measures

Suicheng Gu¹, Joseph Leader¹, Bin Zheng², Qihang Chen³, Yahong Chen³, Frank Scirba⁴, Naftali Kminski⁴, David Gur¹, and Jiantao Pu^{1,5,*}

¹ Imaging Research Center, Department of Radiology, University of Pittsburgh, USA

² Department of Electrical and Computer Engineering, University of Oklahoma, USA

³ Department of Radiology, Beijing Hospital, P.R. China

⁴ Division of Pulmonary Allergy and Critical Care Medicine, Department of Medicine, University of Pittsburgh USA

⁵ Department of Bioengineering, University of Pittsburgh, USA

Abstract

Purpose—To investigate whether lung function in patients with chronic obstructive pulmonary disease (COPD) can be directly predicted using CT densitometric measures and assess the underlying prediction errors as compared with the traditional spirometry-based measures.

Materials and Methods—A total of 600 CT examinations were collected from a COPD study. In addition to the entire lung volume, the extent of emphysema depicted in each CT examination was quantified using density mask analysis (densitometry). The partial least square regression (PLS regression) was used for constructing the prediction model, where a repeated random split-sample validation was employed. For each split, we randomly selected 400 CT exams for training (regression) purpose and the remaining 200 exams for assessing performance in prediction of lung function (e.g., FEV1 and FEV1/FVC) and disease severity. The absolute and percentage errors as well as their standard deviations were computed.

Results—The averaged percentage errors in prediction of FEV1, FEV1/FVC%, TLC, RV/TLC% and DLco% predicted were 33%, 17%, 9%, 18% and 23%, respectively. When classifying the exams in terms of disease severity grades using the CT measures, 37% of the subjects were correctly classified with no error and 83% of the exams were either correctly classified or classified into immediate neighboring categories. The linear weighted kappa and quadratic weighted kappa were 0.54 (moderate agreement) and 0.72 (substantial agreement), respectively.

Conclusion—Despite the existence of certain prediction errors in quantitative assessment of lung function, the CT densitometric measures could be used to relatively reliably classify disease severity grade of COPD patients in terms of GOLD.

*Jiantao Pu, Ph.D. Department of Radiology University of Pittsburgh 3362 Fifth Avenue Pittsburgh, PA 15213 Tel: 412-641-2571 jip13@pitt.edu.

Keywords

pulmonary function; chronic obstructive pulmonary disease (COPD); computed tomography; linear prediction

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a major cause of morbidity and mortality worldwide [1-3]. As an airflow abnormality, pulmonary function testing (PFT) is routinely performed in clinical practice when diagnosing COPD by measuring how well the lungs function globally in exchanging air. Pulmonary function assessment is a mandatory procedure in many clinical situations for therapeutic decision-making (e.g., lung volume reduction for COPD patients), because pulmonary dysfunction may potentially increase postoperative morbidity and mortality [4-5]. In routine clinical practices, spirometry is widely used to efficiently and non-invasively measure lung functions. The most widely used criterion is the ratio between the forced expiratory volume in one second (FEV1) and the forced vital capacity (FVC), which reflects a patient's ability to take a deep breath. However, despite its global characteristic, the accuracy of the spirometry-based lung function test depends on patient cooperation and usually needs repetition of at least three times to ensure reproducibility. In particular, the traditional PFT test can not capture the inherent heterogeneity of the disorder and is not feasible for patients who cannot follow instructions, such as infant, preschool children, or unconscious patients [6].

In clinical practices, COPD is widely classified into two major phenotypes: emphysema and chronic bronchitis. The first one is characterized by permanent destruction of air sacs (alveoli) of the lungs and appears on CT images as regions with extremely low density. The second one is chronic bronchitis characterized by airway inflammation. The inflammation may limit the air flow and thus cause both air trapping and/or hyperinflation, ultimately leading to reduced lung density. Due to its relatively low intensity depicted on computed tomography (CT) images, the extent of emphysema or other abnormalities such as air trapping can be easily quantified using low attenuation areas (LAA). A large number of investigations [7-13] have consistently demonstrated that the density-based measure, typically the lung volume below a specific threshold (e.g., a cut-off of -950 Hounsfield Unit (HU)), correlates significantly with lung function and acts as a severity index of emphysema [14]. After finding a strong association between FEV1 the mean lung density, Moloney et al. [15] even concluded that CT densitometry could be a predictor of pulmonary function in lung cancer patients. Although a strong correlation typically implies a high accuracy in prediction, the capability of the densitometric measures in predicting lung function or disease severity as well as the associated errors remains unknown. If the lung function could be predicted accurately using CT densitometric measures, it may lead to the future development of a tool that allows non-invasive quantification of pulmonary function directly from high resolution chest CT examinations. In this tool, many other factors that are potentially associated with lung functions, such as gender, obesity, age, disease status, and even the spatial distribution patterns of low-density regions, could be combined in a plausible way using a sophisticated regression model for better prediction performance. This

would be particularly beneficial to COPD-related clinical practices. In this study, our primary objective is to specifically investigate whether lung function can be directly predicted using the CT densitometric measures and meanwhile assess the underlying errors by comparing with the traditional spirometry-based measures. A relatively large chest CT dataset (i.e., 600 exams) was used for the experiments. A description of the methods and the experimental results follows.

MATERIALS AND METHODS

A. Study population

A dataset consisting of 600 CT examinations was collected from a COPD study. These examinations were acquired on different subjects, whose ages were larger 40 years and had at least a 10 pack year history of tobacco use. The subjects underwent pre- and post-bronchodilator spirometry and body plethysmography, measurement of lung diffusion capacity by single breath carbon monoxide (DLCO), impulse oscillometry, a chest CT examination, demographic as well as medical history questionnaires. When performing these lung function tests, the standard methodology [16-17] and the standard reference equations [18-19] were used. The quality assurance includes the examination of test values and the evaluation of both volume-time and flow-volume curves for possible technical errors. Additional maneuvers would have been taken if any erroneous curves were detected. At least three maneuvers with repeatable results for both FVC and FEV1 were recorded. According to the global initiative for chronic obstructive lung disease (GOLD) criteria [20], these subjects (examinations) were classified into five different subgroups, namely non-COPD (223 subjects), GOLD-I (84 subjects), GOLD-II (144 subjects), GOLD-III (71 subjects), and GOLD-IV (78 subjects). The demographics of the collected cases are summarized in Table 1. All procedures were performed under a University of Pittsburgh Institutional Review Board approved protocol (#0612016) and written informed consent was obtained for each subject.

B. Acquisition of thin-section CT examinations

The CT exams were acquired on a LightSpeed VCT 64-detector scanner (GE Healthcare, Waukesha, WI) with subjects holding their breath at end inspiration without contrast at the following parameters: 64×0.625 mm detector configuration, 0.969 pitch, 120 kVp tube energy, 250 mA tube current, and 0.4 sec gantry rotation (or 100 mAs). Images were reconstructed to encompass the entire lung field in 512×512 pixel matrix using the GE “bone” kernel at 0.625 mm section thickness and 0.625 mm interval. Pixel dimensions ranged from 0.549 to 0.738 mm, depending on participant body size. The “bone” kernel was used because of its ability to analyze both the parenchyma and airways [21]. The subjects were instructed in breathing to reach TLC prior to scanning; however, no real-time measures were employed to ensure breathing compliance such spirometry-gated CT acquisition. The CT exams were reviewed to ensure compliance with the above mentioned chest CT scanning protocol, and these exams were also visually reviewed for artifacts that would contribute to poor image quality (e.g., subject motion and/or metal artifacts). Exams that did not meet these requirements were not analyzed.

C. CT Densitometric Measures

To obtain CT densitometric measures, we firstly used an available computerized scheme [22] to segment the lung parenchyma depicted in CT examinations. Thereafter, the entire lung volume and the low attenuation areas (LAA) within the lung regions were computed after the application of a threshold. In this study, we used the cut-off thresholds of both -950 HU and -910 HU. The resulting measures were denoted as LAA-910 and LAA-950, respectively. In addition to the LAA, the percentage of LAA with respect to the entire lung volume (LAA%) was computed as well and denoted separately as LAA%-910 and LAA%-950. Hence, totally, five measures were computed from the collected CT examinations: lung volume (LVOL), LAA-910, LAA-950, LAA%-910, and LAA%-950. The measure of lung volume was intended as a normalization factor to account for the variations among individuals in lung size. In this study, the capabilities of individual CT measures as well as their combinations in predicting lung function were investigated.

D. Prediction of Lung Function

When constructing the prediction model, the partial least squares regression (PLS regression) was used because of its strength in predicting the responses when the factors are many or highly collinear (correlated). The PLS regression combines merits of principal components analysis and multiple regression. In this study, the densitometric measures were used as the observable variables (X_i) and the spirometry measures were used as the predicted variables (Y_i). The objective is to identifying the underlying linear regression model (i.e., $Y=AX+B$) between X_i and Y_i . Detailed description of PLS regression can be found in [23]. In implementation, a repeated random split-sample validation was employed. For each split, we randomly selected 400 CT exams for training (regression) purpose and the remaining 200 exams for validation. Totally, the split-sample validation was performed three times. According the GOLD guideline [24], the post-bronchodilator PFT measures are used to confirm the presence of irreversible airflow limitation and distinguish asthma from COPD. Hence, the post-bronchodilator PFT measures were used in the PLS regression and included: 1) FEV1, 2) FEV₁% predicted, 3) FEV₁/FVC ratio (%), 4) TLC, 5) RV, 6) RV/TLC ratio (%) and 7) DLco% predicted. The GOLD indices were determined based on the predicted values of FEV1 % predicted and FEV1/FVC%. The MATLAB Statistics Toolbox was used to perform the PLS regression.

E. Assessment of Prediction Performance

For each split, given 200 validation (testing) exams, we computed their densitometric measures as described above and used them to predict the above mentioned six lung function measures. The prediction errors were computed by comparing with the spirometry measures using two indices. The first one is the absolute error of the predicted measures, and the second one is the percentage error of the predicted measures, which was defined as:

$$\text{Percentage Error} = \left| \frac{(\text{predictedvalue} - \text{measuredvalue})}{\text{measuredvalue}} \right| \times 100\% \quad (1)$$

When assessing the prediction errors, their standard deviations were computed. p values were computed from a Student's t -test and a value less than 0.01 was considered statistically significant.

In addition, we studied the capability of the CT densitometric measures in classifying the examinations in terms of GOLD-based diseases severity. Considering that GOLD II category is relatively wide (50-80% predicted FEV1 and covers a large category of severity), we subdivided the GOLD II category by half into two sub-categories, namely GOLD II-A (FEV1% predicted below 65%) and GOLD II-B (FEV1% predicted above 65%), and presented the analysis results. Totally, the subjects were classified into six categories, namely, NONE, GOLD I, GOLD II-A, GOLD II-B, GOLD III and GOLD IV. The contingency table and the weighted kappa statistics [25] were computed to assess the agreement between the true GOLD category and the estimated GOLD category.

RESULTS

Table 2 summarized the mean absolute errors (MAE) when predicting lung functions using different CT densitometric measures, while the percentage errors in lung function prediction were shown in Table 3-Table 8. The averaged errors in prediction FEV1 and TLC for the involved CT measures ranged from 0.51 - 0.79 liter, and from 0.57 - 0.99 liter, respectively. When predicting FEV1/FVC%, the errors associated with these CT measures ranged from 8.24 to 14.23; while for RV/TLC%, the prediction errors ranged from 7.57 to 10.27. The best averaged percentage errors in predicting FEV1, FEV1/FVC%, TLC, RV/TLC% and DLco% predicted were 33%, 17%, 9%, 18% and 23%, respectively. When the same CT densitometric measures were used to classify the exams into different disease severity categories in terms of GOLD, the distribution of the classification was summarized in Table 9 and Figure 2. The linear weighted kappa and quadratic weighed kappa were 0.54 (Moderate agreement) and 0.72 (Substantial agreement), respectively. On average, 37% of the subjects were correctly classified with no error and 83% of the exams were either correctly classified or classified into immediate neighboring categories.

DISCUSSION

In the past, the underlying association between the extent of emphysema based on densitometric measures and pulmonary function measures had been extensively investigated [26-28]. As an index of the extent of emphysema, CT densitometric measures, such as %LAA, correlates significantly with RV/TLC, FEV₁, FEV₁% predicted, FEV₁/FVC ratio, and DLco% predicted, and have been verified as an index of lung function decline [29] and as a risk factor of lung cancer [30]. Given the close relationship between the extent of emphysema and pulmonary function, it may be interesting to know whether the quantitative CT densitometry is capable of evaluating pulmonary functions and the associated errors. Unfortunately, no investigation has been performed in this regard to date. Under the help of a relative large diverse cohort of chest CT examinations from an available COPD study, we investigated the capability of different densitometric measures in predicting lung function by assessing the prediction errors. It is notable that lung densities could be reduced by disorders (e.g., air trapping) or other factors (e.g., obesity). Prediction of lung functions on the basis of

lung densitometric measures might reflect the impact of these disorders or other unknown factors on lung function without explicitly incorporating them into a regression model. In other words, the actual impact of the factors that may cause the change of lung density and lung functions are implied in the lung density measures. Also, our primary objective is to specifically investigate whether the lung function measures and/or disease severity could be reliably and directly predicted from CT densitometric measures, but not to build and test a sophisticated prediction model that may need to incorporate patient demographic information, such as age, weight, and height, as well as disorders that potentially affect lung functions. This simple and specific strategy could aid in a clear understanding of the predictability (or the role) of the CT densitometric measures in lung function and disease severity.

First, our experiments showed that different CT densitometric measures as well as their combinations varied somewhat in their capabilities in prediction lung function. For example, the combination, namely “LVOL+LAA910+LAA950”, has a significant stronger capability in predicting FEV1 ($p<0.01$). The combination, namely “LAA910%+LAA950%”, has a stronger capability in predicting FEV1% and FEV1/FVC% as compared with measures, such as LVOL and LAA910 ($p<0.01$). The lung volume measure (LVOL) has the poorest capability in prediction lung functions ($p<0.01$) except TLC ($p>0.01$). Here, we have to mention that it is very difficult to identify an optimal cut-off threshold for each single CT examinations when performing densitometric analysis. Previously, Müller et al. [40] suggested a threshold of -910 HU, while Gevenois et al. [41] recommended a threshold of -950 HU. Recently, Madani et al. [42] mentioned that a cut-off of -960 HU or -970 HU might lead to more accurate assessment. Hence, we tested separately the densitometric measures obtained at both -910 HU and -950 HU as observable variables and found that they had no significant difference in prediction of PFT measures ($p>0.1$).

Second, when studying the prediction capability of the densitometric measures for subpopulations with different disease severity (Table 3-Table 8), we found that the percentage prediction errors for FEV1, FEV1% Predicted, and FEV1/FVC% were consistently smaller for the subpopulations with less severe diseases. For example, the best percentage prediction errors for FEV1/FVC% were less than 13% for the subpopulations rated as non-COPD, GOLD-I, and GOLD-II, but the percentage errors increased to 25% and 55% for the GOLD-III and GOLD-IV subpopulations. In contrast, for lung capacity measures (i.e., TLC, RV), the combination “LVOL+LAA910+LAA950” achieved the best performance in prediction and there was no obvious difference across the categories in terms of disease severity. For TLC, the percentage prediction error ranged from 8% to 10% with an average of 9%; for RV, the percentage prediction error ranged from 19% to 23% with an average of 21%. If we defined the percentage error less than 15% as being clinically meaningful, lung function measures such as FEV1/FVC% and FEV1_Predicted for less severe disease could be predicted in a relatively reliable way using “LAA910%+LAA950%”. Similarly, lung capacity measures such as TLC could be predicted more reliably using “LVOL+LAA910+LAA950.” This is straightforward because CT-based lung volume is somewhat equivalent to plethysmographic TLC. In contrast, predicting TLC using LAA% does not make sense. Hence, when assessing lung functions and lung capabilities, different CT measures should be used to achieve an optimal prediction performance.

Third, when studying the capability of the CT measures in classifying the disease severity grades (Figure 1), we found that the CT densitometric measures demonstrated a relatively high reliability. For example, the CT measure “LAA910%+LAA950%” could classify 37% of the involved subjects correctly without any error in disease severity grade. When the classification error was no more than one grade, around 83% of the subjects could be classified accurately. In particular, the linear weighted kappa and quadratic weighed kappa were 0.54 (moderate agreement) and 0.72(substantial agreement), respectively. Therefore, despite the relative large percentage error in prediction, the CT densitometric measure could be used in practice to assess the disease severity by classifying CT examinations into corresponding GOLD categories. In general, the percentage errors for more severe GOLD stages were smaller for predicting plethysmographic values than spirometric values.

We observed that there existed somewhat errors in predicting lung functions when only densitometric measures were used. For example, some examinations rated as GOLD-II or GOLD-III were incorrectly predicted as non-COPD, and some examination rated as non-COPD were incorrectly predicted as GOLD-II. This demonstrated that lung function could be affected by a wide range of factors (e.g., age/height [31], sex [32], pulmonary blood flow [33-35], arterial stiffness [36-37], rib muscles [38-39], and airway morphology [43]). In particular, it is notable that the existence of image noise or artifacts may also contribute to this incorrect prediction; however, it may not be an easy task to accurately define image noises/artefacts and filter out them. Nevertheless, it is beyond any doubt that combining all these factors together may significantly enhance regression model for accurately assessing pulmonary function and/or GOLD stages. However, we did not perform the investigation in this way because our primary goal of this study is to understand whether the traditional pulmonary function test measures could be predicted directly from CT-based density measures, but not to constructive a sophisticated prediction model that may need to incorporate a number of factors. Given this simplified strategy, we only used a linear model instead of using any other nonlinear regression model, which is a nonlinear combination of the model parameters and depends on multiple variables. Hence, given our specific issues with relatively simple variables, we believe that the linear model is sufficient for the mentioned purpose. In fact, a direct assessment of lung function using CT densitometric could be potentially used to identify the cases that are not significantly affected by the LAA, thereby aiding in a further investigation of the factors that affect lung functions. In the future, we will investigate how to combine additional factors (e.g., the distribution patterns of low-density regions) using other nonlinear regression models (e.g., nonlinear support vector machine (SVM) [44]) to achieve a better prediction performance.

Finally, we are aware that there are other limitations with this study. For example, we did not take the CT acquisition protocols into account when computing the densitometric measures. Actually, the studies from other investigators [14, 39] demonstrated that different reconstruction kernels had limited impact on densitometric measures. Boedeker et al. [39] concluded in their study that the “bone” kernel might shift the obtained density mask volume by 2.4% as compared to the “standard” kernel, which could be ignored in practice. In addition, the diversity of the involved CT exams may not be sufficient and large enough in terms of factors affecting lung function, such as age, obesity, and disease severity. The unequal/biased distributions may affect the performance of the prediction. For an accurate

prediction of the lung functions, a more diverse database is desirable. Despite these limitations, we investigation showed that the CT densitometric measures play a critical role in predicting pulmonary functions from a different perspective and they should be considered in any prediction model.

CONCLUSION

In this study, we investigated the capability of lung CT densitometric measures in predicting lung function. The traditional spirometry-based measures were used as “ground truth”. Totally, 600 CT examinations collected from a COPD study at our medical center were used. The PLS regression was used for prediction modeling and a repeated random split-sample validation was employed for performance assessment. Our experiments showed that lung function measures, such as FEV1/FVC%, could be more accurately predicted when the diseases were less severe (<GOLD-III), and lung capability measures, such as TLC, could be reliably predicted across all disease severities. In particular, the CT measure combinations, such as “LVOL+LAA910+LAA910” and “LAA910%+LAA950%”, have better prediction capabilities in assessing lung function as compared to other single or combined measures. Although there exist certain errors in directly predicting lung function, our experimental results demonstrate that the densitometric measures could be used to classify the severity grade of COPD patients in terms of GOLD.

Acknowledgments

This work is supported in part by grants from the National Heart, Lung, and Blood Institute of the National Institutes of Health to the University of Pittsburgh under Grants and RO1 HL096613, P50 CA090440, P50 HL084948, and the Bonnie J. Addario Lung Cancer Foundation.

REFERENCES

1. <http://www.nhlbi.nih.gov>
2. http://www.nhlbi.nih.gov/health/public/lung/other/copd_fact.pdf
3. <http://www.copdinamerica.org/background.html>
4. Parimon T, Madtes DK, Au DH, Clark JG, Chien JW. Pretransplant lung function, respiratory failure, and mortality after stem cell transplantation. *American Journal of Respiratory and Critical Care Medicine*. 2005; 172(3):384–390. [PubMed: 15894602]
5. Crema E, Benelli AG, Silva AV, Martins AJ, Pastore R, Kujavao GH, Silva AA, Santana JR. Assessment of pulmonary function in patients before and after laparoscopic and open esophagogastric surgery. *Surgical Endoscopy*. 2005; 19(1):133–136. [PubMed: 15549632]
6. Sly PD, Lombardi E. Measurement of lung function in preschool children using the interrupter technique. *Thorax*. 2003; 58:742–744. [PubMed: 12947127]
7. Dourmes G, Laurent F. Airway Remodelling in Asthma and COPD: Findings, Similarities, and Differences Using Quantitative CT. *Pulm Med*. 2012; 2012:670414. [PubMed: 22448324]
8. Camiciottoli G, Bartolucci M, Maluccio NM, Moroni C, Mascalchi M, Giuntini C, Pistolesi M. Spirometrically gated high-resolution CT findings in COPD: lung attenuation vs lung function and dyspnea severity. *Chest*. Mar; 2006 129(3):558–64. [PubMed: 16537852]
9. Heussel CP, Herth FJ, Kappes J, Hantusch R, Hartlieb S, Weinheimer O, Kauczor HU, Eberhardt R. Fully automatic quantitative assessment of emphysema in computed tomography: comparison with pulmonary function testing and normal values. *Eur Radiol*. Oct; 2009 19(10):2391–402. [PubMed: 19458953]

10. Washko GR, Criner GJ, Mohsenifar Z, Sciruba FC, Sharafkhaneh A, Make BJ, Hoffman EA, Reilly JJ. Computed tomographic-based quantification of emphysema and correlation to pulmonary function and mechanics. *COPD*. Jun; 2008 5(3):177–86. [PubMed: 18568842]
11. Mets OM, Murphy K, Zanen P, Gietema HA, Lammers JW, van Ginneken B, Prokop M, de Jong PA. The relationship between lung function impairment and quantitative computed tomography in chronic obstructive pulmonary disease. *Eur Radiol*. Jan; 2012 22(1):120–8. [PubMed: 21837396]
12. Madani A, Van Muylem A, Gevenois PA. Pulmonary emphysema: effect of lung volume on objective quantification at thin-section CT. *Radiology*. Oct; 2010 257(1):260–8. [PubMed: 20663967]
13. Mohamed Hoessein FA, van Rikxoort E, van Ginneken B, de Jong PA, Prokop M, Lammers JW, Zanen P. CT-quantified emphysema distribution is associated with lung function decline. *Eur Respir J*. Oct; 2012 40(4):844–50. [PubMed: 22323577]
14. Wang Z, Gu S, Leader JK, Kundu S, Tedrow JR, Sciruba FC, Gur D, Siegfried JM, Pu J. Optimal threshold in CT quantification of emphysema. *Eur Radiol*. 2013; 23(4):975–84. [PubMed: 23111815]
15. Moloney F, McWilliams S, Crush L, Laughlin PD, Kennedy M, Henry M, O' Connor O, Maher MM. CT Densitometry as a Predictor of Pulmonary Function in Lung Cancer Patients. *Open Respir Med J*. 2012; 6:139–44. [PubMed: 23264834]
16. Standardization of spirometry, 1994 update. American thoracic society. *American journal of respiratory and critical care medicine*. 1995; 152:1107–1136. [PubMed: 7663792]
17. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CPM, Gustafsson P, et al. Standardisation of spirometry. *Eur Respir J*. 2005; 26:319–338. [PubMed: 16055882]
18. Lung function testing: Selection of reference values and interpretative strategies. American thoracic society. *The American review of respiratory disease*. 1991; 144:1202–1218. [PubMed: 1952453]
19. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. Population. *American journal of respiratory and critical care medicine*. 1999; 159:179–187. [PubMed: 9872837]
20. <http://www.goldcopd.org/other-resources-gold-spirometry-guide.html>
21. Pauls S, Gulkin D, Feuerlein S, Muche R, Krüger S, Schmidt SA, Dharaiya E, Brambs HJ, Hetzel M. Assessment of COPD severity by computed tomography: correlation with lung functional testing. *Clin Imaging*. May-Jun;2010 34(3):172–8. [PubMed: 20416480]
22. Pu J, Roos J, Yi CA, Napel S, Rubin GD, Paik DS. Adaptive border marching algorithm: automatic lung segmentation on chest CT images. *Comput Med Imaging Graph*. 2008; 32(6):452–62. [PubMed: 18515044]
23. Wold, Herman. Partial least squares. In: Kotz, Samuel; Johnson, Norman L., editors. *Encyclopedia of statistical sciences* 6. Wiley; New York: 1985. p. 581-591.
24. Pauwels RA, Buist AS, Calverley PM, et al. Global strategy for the diagnosis, management and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease [GOLD] Workshop summary. *Am J Resp Crit Care Med*. 2001; 163:1256–76. [PubMed: 11316667]
25. Sim J, Wright C C. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*. 2005; 85(3):257–268. [PubMed: 15733050]
26. Marsh S, Aldington S, Williams MV, Nowitz MR, Kingzett-Taylor A, Weatherall M, Shirtcliffe PM, McNaughton AA, Pritchard A, Beasley R. Utility of lung density measurements in the diagnosis of emphysema. *Respir Med*. Jul; 2007 101(7):1512–20. [PubMed: 17307350]
27. Kim WJ, Silverman EK, Hoffman E, Criner GJ, Mosenifar Z, Sciruba FC, Make BJ, Carey V, Estépar RS, Diaz A, Reilly JJ, Martinez FJ, Washko GR. NETT Research Group. CT metrics of airway disease and emphysema in severe COPD. *Chest*. Aug; 2009 136(2):396–404. [PubMed: 19411295]
28. Dournes G, Montaudon M, Berger P, Laurent F. In Vivo Computed Tomography as a Research Tool to Investigate Asthma and COPD: Where Do We Stand? *J Allergy*. 2012; 2012:972479.

29. Wilson DO, Weissfeld JL, Balkan A, Schragin JG, Fuhrman CR, Fisher SN, Wilson J, Leader JK, Siegfried JM, Shapiro SD, Scirba FC. Association of radiographic emphysema and airflow obstruction with lung cancer. *Am J Respir Crit Care Med*. Oct 1; 2008 178(7):738–44. [PubMed: 18565949]
30. Kiefer EM, Hankinson JL, Barr RG. Similar relation of age and height to lung function among Whites, African Americans, and Hispanics. *Am J Epidemiol*. 2011; 173(4):376–87. [PubMed: 21242304]
31. Ezeugwu VE, Olaogun M, Mbada CE, Adedoyin R. Comparative Lung Function Performance of Stroke Survivors and Age-matched and Sex-matched Controls. *Physiother Res Int*. Jan 29.2013 doi: 10.1002.
32. Tiefenbrun J, Kim SI, Shoemaker WC. The relation of the distribution of pulmonary blood flow to lung function during hemorrhagic shock. *Surg Gynecol Obstet*. 1974; 138(4):557–61. [PubMed: 4816143]
33. Stevens AD, Lumbers ER. Effects of reduced uterine blood flow on fetal cardiovascular, renal, and lung function. *Am J Physiol*. 1990; 259(5 Pt 2):R1004–11. [PubMed: 2240260]
34. Dijk PH, Heikamp A, Bambang Oetomo S. Surfactant nebulisation: lung function, surfactant distribution and pulmonary blood flow distribution in lung lavaged rabbits. *Intensive Care Med*. 1997; 23(10):1070–6. [PubMed: 9407243]
35. Ayer JG, Belousova EG, Harmer JA, Toelle B, Celermajer DS, Marks GB. Lung function is associated with arterial stiffness in children. *PLoS One*. 2011; 6(10):e26303. [PubMed: 22046271]
36. Weiler Z, Zeldin Y, Magen E, Zamir D, Kidon MI. Pulmonary function correlates with arterial stiffness in asthmatic patients. *Respir Med*. 2010; 104(2):197–203. [PubMed: 19892539]
37. Gayan-Ramirez G, Decramer M. Mechanisms of striated muscle dysfunction during acute exacerbations of COPD. *J Appl Physiol*. 2013; 114(9):1291–9. [PubMed: 23372146]
38. LoMauro A, Pochintesta S, Romei M, D'Angelo MG, Pedotti A, Turconi AC, Aliverti A. Rib cage deformities alter respiratory muscle action and chest wall function in patients with severe osteogenesis imperfecta. *PLoS One*. 2012; 7(4):e35965. [PubMed: 22558284]
39. Boedeker KL, McNitt-Gray MF, Rogers SR, Truong DA, Brown MS, Gjertson DW, Goldin JG. Emphysema: effect of reconstruction algorithm on CT imaging measures. *Radiology*. Jul; 2004 232(1):295–301. [PubMed: 15220511]
40. Müller NL, Staples CA, Miller RR, Abboud RT. “Density mask”. An objective method to quantitate emphysema using computed tomography. *Chest*. 1988; 94(4):782–7. [PubMed: 3168574]
41. Gevenois PA, de Maertelaer V, De Vuyst P, Zanen J, Yernault JC. Comparison of computed density and macroscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med*. 1995; 152(2):653–7. [PubMed: 7633722]
42. Madani A, De Maertelaer V, Zanen J, Gevenois PA. Pulmonary Emphysema: Radiation Dose and Section Thickness at Multidetector CT Quantification—Comparison with Macroscopic and Microscopic Morphometry. *Radiology*. 2007; 243(1):250–7. [PubMed: 17392257]
43. Pu J, Leader JK, Meng X, Whiting B, Wilson D, Scirba FC, Reilly JJ, Bigbee WL, Siegfried J, Gur D. Three-dimensional airway tree architecture and pulmonary function. *Acad Radiol*. 2012; 19(11):1395–401. [PubMed: 22884402]
44. Hsu CW, Lin CJ. A comparison of methods for multiclass support vector machines. *IEEE Trans Neural Netw*. 2002; 13(2):415–25. [PubMed: 18244442]

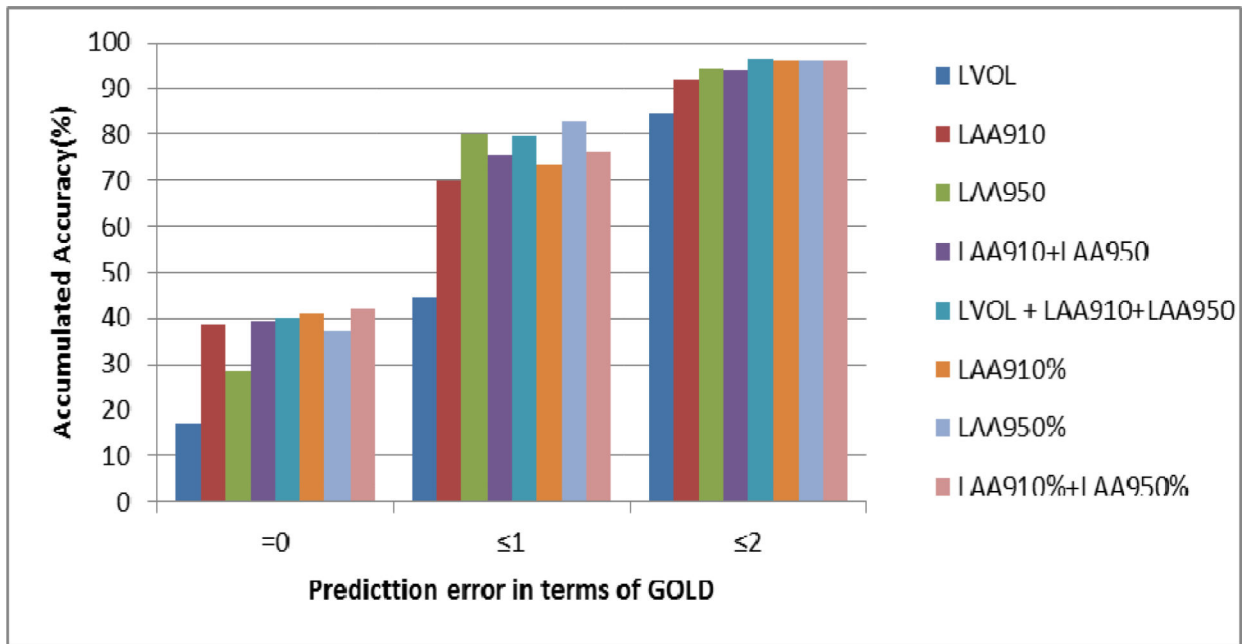


Figure 1. Percentage distribution of the cases in predicting disease severity in terms of GOLD (Unit: %).

Table 1Subject Demographics ($n = 600$)

Parameter	Mean (\pm std) or count (%)
Sex male	323 (53.9%)
Age	63.8 (\pm 5.4)
Pack years	57.8 (\pm 33.0)
Height(cm)	169.3 (\pm 9.3)
Weight(kg)	80.0 (\pm 16.1)
BMI	27.7 (\pm 4.5)
TLC(litre)	6.27 (\pm 1.29)
RV/TLC %	45.0 (\pm 13.2)
RV (litre)	2.85 (\pm 1.17)
FEV ₁ % predicted	73.1 (\pm 29.3)
FEV ₁	2.10(\pm 0.96)
FEV ₁ /FVC %	59.7 (\pm 18.3)
DLco % predicted	65.8(\pm 23.4)
Five-category classification	
NONE-COPD	223 (37.2%)
GOLD I	84 (14.0%)
GOLD II A	81 (13.5%)
GOLD II B	63(10.5%)
GOLD III	71 (11.8%)
GOLD IV	78 (13.0%)

Abbreviations: TLC – total lung capacity, RV – residual volume, FVC – functional vital capacity, FEV₁ – forced expiratory volume in one second. GOLD II-A: FEV₁ % predicted < 65%, GOLD II-B: FEV₁ % predicted \geq 65%.

Table 2Mean Errors(\pm std) in lung function prediction using different CT densitometric measures

CT Densitometry	Pulmonary Function Measures						
	FEV1(Liter)	FEV1%PRED	FEV1/FVC%	TLC(Liter)	RV(Liter)	RV/TLC %	DLco %Pred
LVOL	0.79 \pm 0.54	23.69 \pm 15.71	14.23 \pm 9.44	0.58 \pm 0.51	0.74 \pm 0.68	10.27 \pm 8.19	19.4 \pm 13.0
LAA910	0.67 \pm 0.49	17.29 \pm 13.33	9.27 \pm 7.30	0.87 \pm 0.62	0.61 \pm 0.55	8.50 \pm 7.09	15.2 \pm 10.9
LAA950	0.63 \pm 0.45	16.64 \pm 12.41	9.47 \pm 7.09	0.95 \pm 0.68	0.61 \pm 0.56	7.92 \pm 6.80	13.6 \pm 10.1
LAA910+LAA 950	0.63 \pm 0.45	16.57 \pm 12.41	9.08 \pm 6.97	0.85 \pm 0.60	0.60 \pm 0.54	7.93 \pm 6.81	13.5 \pm 10.1
LVOL+LAA910+LAA950	0.51 \pm 0.39	15.94 \pm 12.01	8.90 \pm 6.66	0.57 \pm 0.50	0.57 \pm 0.55	7.81 \pm 6.43	12.6 \pm 9.6
LAA910%	0.60 \pm 0.45	15.88 \pm 12.29	8.35 \pm 6.61	0.97 \pm 0.69	0.65 \pm 0.56	7.99 \pm 6.62	13.7 \pm 10.1
LAA950%	0.60 \pm 0.42	15.71 \pm 11.87	8.92 \pm 6.88	0.99 \pm 0.71	0.64 \pm 0.57	7.59 \pm 6.49	12.6 \pm 9.4
LAA910%+LAA950%	0.59 \pm 0.43	15.37 \pm 11.79	8.24 \pm 6.49	0.97 \pm 0.68	0.64 \pm 0.56	7.57 \pm 6.48	12.7 \pm 9.5

Table 3

Percentage errors (%) in predicting FEV1 when using different CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	25±14	22±12	25±23	95±47	255±93	63±88
LAA910	22±15	25±15	35±30	63±52	147±94	47±59
LAA950	20±13	22±13	35±31	69±49	128±84	44±53
LAA910+LAA 950	20±13	21±12	35±30	73±48	131±83	44±54
LVOL+LAA910+LAA950	17±11	17±13	29±23	53±37	89±64	33±39
LAA910%	22±16	23±15	34±29	49±45	102±68	39±44
LAA950%	20±13	21±12	35±30	66±46	102±69	40±44
LAA910%+LAA950%	20±14	21±13	34±30	62±45	97±67	39±43

Table 4

Percentage errors (%) in predicting FEV1% predicted when using CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	22±11	22±11	16±15	83±29	232±74	55±78
LAA910	15±11	17±13	24±18	51±38	124±74	36±48
LAA950	13±9	12±10	23±18	64±36	119±72	35±47
LAA910+LAA 950	13±9	13±10	23±18	61±36	117±71	35±47
LVOL+LAA910+LAA950	14±10	13±10	23±18	55±36	102±69	32±43
LAA910%	14±11	17±12	24±18	44±36	97±63	31±40
LAA950%	13±9	12±10	23±16	63±39	100±70	32±43
LAA910%+LAA950%	13±10	13±10	23±17	56±37	93±66	31±40

Table 5

Percentage errors (%) in predicting TLC when using different CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	10±8	9±8	10±10	8±6	9±8	10±8
LAA910	15±11	15±8	16±15	14±9	13±13	15±12
LAA950	17±13	16±9	17±17	15±10	14±14	16±14
LAA910+LAA 950	15±11	15±9	15±14	12±9	12±12	14±11
LVOL+LAA910+LAA950	10±7	9±7	10±11	8±7	9±8	9±8
LAA910%	16±13	16±9	17±16	15±10	16±15	16±14
LAA950%	17±14	16±9	18±17	16±10	16±15	17±14
LAA910%+LAA950%	16±13	16±9	18±16	15±10	15±15	16±13

Table 6

Percentage errors (%) in predicting RV when using different CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	29±24	30±28	18±17	20±12	34±11	26±21
LAA910	23±22	26±27	24±19	20±14	20±15	23±20
LAA950	23±22	23±22	23±20	24±18	20±15	23±20
LAA910+LAA 950	22±21	24±25	23±19	22±16	19±15	22±20
LVOL+LAA910+LAA950	20±19	23±25	21±19	23±16	19±14	21±19
LAA910%	25±23	27±27	26±21	21±15	22±17	25±22
LAA950%	24±22	24±23	24±20	25±19	22±17	24±21
LAA910%+LAA950%	23±22	25±25	25±21	24±17	21±17	24±21

Table 7

Percentage errors (%) in predicting RV/TLC% when using different CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	26±21	28±23	15±14	19±10	32±10	24±18
LAA910	19±18	25±25	19±17	17±12	21±9	20±18
LAA950	17±16	21±21	18±18	19±15	19±10	18±17
LAA910+LAA 950	17±16	21±21	18±18	20±15	19±10	18±17
LVOL+LAA910+LAA950	18±16	20±20	19±17	18±14	17±10	18±16
LAA910%	18±17	24±24	19±17	15±12	17±9	19±17
LAA950%	16±16	21±21	18±17	19±14	16±10	18±16
LAA910%+LAA950%	16±15	21±21	18±17	19±14	16±10	18±16

Table 8

Percentage errors (%) in predicting DLco%Pred when using different CT densitometric measures

CT Densitometry	Disease Severity					
	NO COPD	GOLD I	GOLD II	GOLD III	GOLD IV	ALL
LVOL	20±11	22±17	27±25	78±70	121±75	41±52
LAA910	18±13	22±16	26±23	46±47	65±48	29±32
LAA950	15±11	18±15	25±22	39±35	55±42	26±27
LAA910+LAA 950	15±11	18±15	25±22	40±36	57±43	26±27
LVOL+LAA910+LAA950	15±11	18±15	23±21	35±31	46±36	23±24
LAA910%	18±13	20±15	24±22	37±34	45±35	25±24
LAA950%	15±11	17±13	23±21	38±35	42±35	23±24
LAA910%+LAA950%	15±11	18±13	23±21	37±33	42±35	23±23

Table 9

Distribution of the cases in predicting disease severity in terms of GOLD classification distribution in terms of GOLD. The linear weighted kappa is 0.54 (moderate agreement) and the quadratic weighed kappa is 0.72 (substantial agreement).

	None	GOLD I Est	GOLD II-A Est	GOLD II-B Est	GOLD III Est	GOLD IV Est	SUM
NONE	48	14	11	1	0	0	74
GOLD I	6	10	10	2	0	0	28
GOLD II-A	9	6	7	4	0	1	27
GOLD II-B	2	4	7	6	2	0	21
GOLD III	2	0	5	6	8	3	24
GOLD IV	0	1	1	8	11	5	26
SUM	67	35	41	27	21	9	200