A Comparative Study of HRCT Image Metrics and PFT Values for Characterization of ILD and COPD

Gang Song, MS, Eduardo Mortani Barbosa Jr., MD, Nicholas Tustison, DSc, Warren B. Gefter, MD, Maryl Kreider, MD, MSCE, James C. Gee, PhD, Drew A. Torigian, MD, MA

Rationale and Objectives: The aim of this study was to compare the performance of various image-based metrics computed from thoracic high-resolution computed tomography (HRCT) with data from pulmonary function testing (PFT) in characterizing interstitial lung disease (ILD) and chronic obstructive pulmonary disease (COPD).

Materials and Methods: Fourteen patients with ILD and 11 with COPD had undergone both PFT and HRCT within 3 days. For each patient, 93 image-based metrics were computed, and their relationships with the 21 clinically used PFT parameters were analyzed using a minimal-redundancy-maximal-relevance statistical framework. The first 20 features were selected among the total of 114 mixed image metrics and PFT values in the characterization of ILD and COPD.

Results: Among the best-performing 20 features, 14 were image metrics, derived from attenuation histograms and texture descriptions. The highest relevance value computed from PFT parameters was 0.47, and the highest from image metrics was 0.52, given the theoretical bound as [0, 0.69]. The ILD or COPD classifier using the first four features achieved a 1.92% error rate.

Conclusions: Some image metrics are not only as good discriminators as PFT for the characterization of ILD and COPD but are also not redundant when PFT values are provided. Image metrics of attenuation histogram statistics and texture descriptions may be valuable for further investigation in computer-assisted diagnosis.

Key Words: Interstitial lung disease; ILD; chronic obstructive pulmonary disease; COPD; computed tomography; statistical computing; computer-assisted image analysis.

Chronic lung disease constitutes a major worldwide public health care problem and is the fourth leading cause of morbidity and mortality in the United States (1). On the basis of clinical, imaging, and pathologic characteristics, most types of chronic lung disease can be grouped into two basic categories: interstitial lung disease (ILD) and chronic obstructive pulmonary disease (COPD). ILD is a heterogeneous group of diseases in which the hallmark is chronic, progressive, predominantly interstitial inflammation with varying degrees of fibrosis of the lung parenchyma, often leading to reduced lung volume, decreased lung compliance, and restrictive physiology. COPD is characterized by chronic airflow limitation due to airway inflammation and lung parenchymal destruction that is not fully reversible and is usually progressive.

The diagnosis, differentiation, and classification of the severity of ILD and COPD rely on clinical assessment, thoracic imaging (using computed tomography and chest radiography), and pulmonary function testing (PFT). PFT is a noninvasive method of assessing the integrated mechanical function of the lung, chest wall, and respiratory muscles. It is the current reference standard for pulmonary functional assessment. Using PFT, the heterogeneous group of ILDs typically exhibit a restrictive physiology pattern, whereas COPD typically manifests an obstructive physiology pattern.

PFT strictly permits a global assessment of lung physiology. In contrast, image analysis using high-resolution computed tomography (HRCT) is a powerful tool with the potential for regional as well as global quantification of pulmonary diseases. Although generally effective, radiologic interpretation of HRCT is time consuming, requires considerable expertise, is largely qualitative, and is prone to interobserver diagnostic variability. As a result, various computed tomographic image metrics have been proposed with the goal of more consistent and quantitative assessment. Early investigations into computed tomographic lung analysis used relatively simple metrics, such as the mean attenuation value or first-order statistical measurements obtainable from the attenuation histogram (2–4). More sophisticated metrics relying on

Acad Radiol 2012; 19:857–864

From the Penn Image Computing and Science Laboratory (G.S., J.C.G.), Department of Radiology (E.M.B., W.B.G., D.A.T.), and the Department of Medicine (M.K.), University of Pennsylvania School of Medicine, 3600 Market Street, Suite 370, Philadelphia, PA 19104; and the Department of Radiology, University of Virginia, Charlottesville, Virginia (N.T.). Received November 5, 2011; accepted March 2, 2012. Drs Gee and Torigian contributed equally to this work. **Address correspondence to:** G.S. e-mail: songgang@seas.upenn.edu

doi:10.1016/j.acra.2012.03.007

texture descriptions of the parenchyma have shown promise in recent studies (5).

PFT values are used clinically to diagnose ILD and COPD and to stratify disease severity. With the increasing amount of proposed image metrics, research inquiry concerns the effectiveness of these metrics compared to the reference standard PFT values. Previous research (6–13) has been conducted to investigate the correlations between various quantitative image metrics and different PFT values.

Instead of performing a correlation study, we addressed the question primarily from the point of view of feature selection. We put more emphasis on what image metrics and what PFT values can characterize ILD and COPD in a quantitative framework. Image metrics and PFT values are viewed as candidates for selecting which best characterize the corresponding diagnosis (ie, ILD or COPD). We are interested in whether and what features from image metrics provide additional information for differential diagnosis when PFT values are provided. This also differs from the classification work of Xu et al (5) in the sense that we do not tend to train any classifiers directly. The selected features can be used as inputs for any available classifier. We use support vector machines (SVMs) (14) as an example to test the efficiency of the feature selection results. One advantage of SVMs is that the training is not affected by the order of input features, which is not considered in some previous work (13).

A minimal-redundancy-maximal-relevance (mRMR) information framework was introduced by Peng et al (15) for such a feature selection task. The ideal selected features satisfy two constraints: maximal relevance and minimal redundancy. The *relevance* of both image and PFT features concerns the ability of such features in matching an existing classification (in our case, from clinical and radiologic diagnosis). It is usually computed in terms of mutual information, correlation, or statistical tests. However, to get a compact subset of features to classify different types of disease, it is not enough to consider only the features with highest relevance. The selected features need to be as independent from one another as possible. This is known as the criterion of minimal *redundancy*, which makes the features more compact for certain diseases, in comparison to previous studies (6–13) focusing on the relevance (more specifically, correlation) between all the available image features and PFT features.

In this report, we provide a systematic relevance and redundancy analysis comparing 31 various statistical image metrics and 21 PFT values obtained in patients with diagnosed ILD and COPD. We first describe the analysis framework and then provide the results of our analysis in comparing the characterization performance of both image and PFT features. This is followed by a discussion of the results and the possible clinical applicability of our approach.

MATERIALS AND METHODS

Fourteen patients with ILD and 11 patients with COPD were retrospectively identified. The demographics of these selected patients are illustrated in Tables 1 and 2. Every patient underwent both thoracic HRCT image acquisition and PFT within 3 days of each other. HRCT was performed for both inspiration and expiration using a 64–detector row computed tomographic scanner (Siemens Medical Solutions, Malvern, PA) with reconstruction of contiguous 1-mm axial images with a 1-mm interval and a B41f kernel. Inspiratory and expiratory image data sets were then analyzed through computational software developed in our laboratory that is capable of generating several hundred distinct metrics encompassing various aspects of lung physiology (eg, which include pulmonary volumetric and gross tissue indices, attenuation histogram statistics, deformation indices, co-occurrence [16] and run-length [17] matrix texture indices, and attenuation mask indices) gleaned from the relevant literature. For this study, we focused only on a subset of these metrics.

Our whole image-processing pipeline (Fig 1) is fully automatic. First, a segmentation algorithm (18) is applied to

TABLE 1. Statistics of Patients with ILD (n = 14)

Variable	Value
Gender	
Female	8
Male	6
Age (y)	
Mean	56.1
Standard deviation	12.0
Median	56
Degree of PFT restriction	
Normal	3
Mild	3
Moderate	4
Severe	2
Moderately severe	1
Very severe	1

ILD, interstitial lung disease; PFT, pulmonary function testing.

TABLE 2. Statistics of Patients with COPD (n = 11)

Variable	Value
Gender	
Female	6
Male	5
Age (y)	
Mean	53.1
Standard deviation	7.7
Median	54
Degree of PFT restriction	
Normal	0
Mild	3
Moderate	1
Severe	3
Very severe	4

COPD, chronic obstructive pulmonary disease; PFT, pulmonary function testing.