Original Investigation

The Objective Identification and Quantification of Interstitial Lung Abnormalities in Smokers

Samuel Y. Ash, MD, Rola Harmouche, PhD, James C. Ross, PhD, Alejandro A. Diaz, MD, MPH, Gary M. Hunninghake, MD, MPH, Rachel K. Putman, MD, Jorge Onieva, MSc, Fernando J. Martinez, MD, MS, Augustine M. Choi, MD, David A. Lynch, MD, Hiroto Hatabu, MD, PhD, Ivan O. Rosas, MD, Raul San Jose Estepar, PhD, George R. Washko, MD

Abbreviations and Acronyms

CT computed tomography

COPD chronic obstructive pulmonary disease

GE General Electric

ILD interstitial lung disease

ROI region of interest

Rationale and Objectives: Previous investigation suggests that visually detected interstitial changes in the lung parenchyma of smokers are highly clinically relevant and predict outcomes, including death. Visual subjective analysis to detect these changes is time-consuming, insensitive to subtle changes, and requires training to enhance reproducibility. Objective detection of such changes could provide a method of disease identification without these limitations. The goal of this study was to develop and test a fully automated image processing tool to objectively identify radiographic features associated with interstitial abnormalities in the computed tomography scans of a large cohort of smokers.

Materials and Methods: An automated tool that uses local histogram analysis combined with distance from the pleural surface was used to detect radiographic features consistent with interstitial lung abnormalities in computed tomography scans from 2257 individuals from the Genetic Epidemiology of COPD study, a longitudinal observational study of smokers. The sensitivity and specificity of this tool was determined based on its ability to detect the visually identified presence of these abnormalities.

Results: The tool had a sensitivity of 87.8% and a specificity of 57.5% for the detection of interstitial lung abnormalities, with a c-statistic of 0.82, and was 100% sensitive and 56.7% specific for the detection of the visual subtype of interstitial abnormalities called fibrotic parenchymal abnormalities, with a c-statistic of 0.89.

Conclusions: In smokers, a fully automated image processing tool is able to identify those individuals who have interstitial lung abnormalities with moderate sensitivity and specificity.

Key Words: Pulmonary fibrosis; idiopathic; CT scanner; x-ray; pneumonia; interstitial; diffuse parenchymal lung diseases; cigarettes.

Acad Radiol 2016; ■:■■–■■

From the Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital, 75 Francis St., PBB, CA-3, Boston, MA (S.Y.A., A.A.D., G.M.H., R.K.P., I.O.R., G.R.W.); Laboratory of Mathematics in Imaging, Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts 02115 (R.H., J.C.R., J.O., R.S.J.E.); Department of Medicine, Weil Cornell Medical College, New York, New York 10065 (F.J.M., A.M.C.); Department of Radiology, National Jewish Health, Denver, Colorado 80206 (D.A.L.); Department of Radiology, Brigham and Women's Hospital, Boston, Massachusetts 02115 (H.H.). Received July 23, 2016; revised August 9, 2016; accepted August 19, 2016. Samuel Y. Ash and Rola Harmouche contributed equally to this work. Raul San Jose Estepar and George R. Washko contributed equally to this work. Funding/Sponsor: This work was supported by NIH grants: 5-T32-HL007633-30 (Ash, Putman), R01-HL107246 (Washko, Estepar, Harmouche, Onieva Onieva), R01-HL116933 (Washko, Estepar, Ross, Harmouche, Onieva), R01-HL111024 (Hunninghake), P01-HL114501 (Choi, Rosas, Washko), and R01-HL089856 (Washko, Lynch, Estepar, Ross). **Address correspondence to:** S.Y.A. e-mail: syash@partners.org

http://dx.doi.org/10.1016/j.acra.2016.08.023

INTRODUCTION

Interstitial lung diseases (ILD) such as idiopathic pulmonary fibrosis have long been described based on clinical, radiographic, and pathologic findings. However, there is a growing recognition that a broader definition is required to identify the early stages of ILD (1). This has become increasingly important with the introduction of new therapies that slow, but do not reverse, the progression of idiopathic pulmonary fibrosis, which is a progressive and frequently fatal disease (2,3). Interstitial lung abnormalities are radiographic precursor lesions to ILD, which themselves are associated with reduced total lung capacity, decreased exercise capacity, and increased mortality (4–8). In populations at risk for ILD, such as smokers, the rate of interstitial lung abnormalities seen visually on computed tomography (CT) scans can be as high as 9.7%, but visual analysis alone may be insensitive to very

early pathology, and grading systems built using visual analysis may be inaccurate (8–10).

Certain types of objective analysis of CT images, including densitometric and textural-based approaches, have been shown to be sensitive for detecting ILD in patients at high risk, and have been found to correlate with pulmonary function and mortality in those with known ILD (11–15). However, the role of objective CT as a screening tool for the detection of interstitial lung abnormalities in a large cohort of smokers is unknown.

We have developed an objective and fully automated method that uses the local histogram pattern of lung density proposed by Mendoza et al., but which in addition also incorporates the distance to the pleural surface to identify and quantify the volume of radiographic features consistent with interstitial lung abnormalities on CT imaging of 9501 smokers (16). We hypothesized that the objective measurement of these interstitial radiographic features would accurately identify those individuals who have visually identified interstitial lung abnormalities, and may also identify those individuals who have the visually defined subtype of interstitial abnormalities called fibrotic parenchymal abnormalities.

MATERIALS AND METHODS

Study Design and Image Acquisition

The Genetic Epidemiology of COPD (COPDGene) is a longitudinal investigation focused on the epidemiologic and genetic risk factors for the development of chronic obstructive pulmonary disease (COPD) in smokers. A total of 10,300 smokers were enrolled, and each underwent a protocolized study visit that included an extensive interview, volumetric high-resolution CT scanning of the chest, and spirometric assessments of lung function. Smokers between the ages of 45 and 80 with at least a 10 pack-year history of tobacco smoke exposure were eligible to enroll. Subjects were excluded if they had active lung diseases other than asthma or COPD. These exclusion criteria included the presence of ILD as determined by the visual review of CT scans by the COPDGene imaging core (17).

Volumetric CT scans of the chest were performed at both maximal inflation and relaxed exhalation. Images were acquired with the following CT protocol: for General Electric (GE) Lightspeed 16, GE VCT-64, Siemens Sensation-16 and -64, and Philips 40- and 60-slice scanners with 120 kVp, 200 mAs, and 0.5-second rotation time. Images were reconstructed using a standard algorithm at 0.625 mm slice thickness and 0.625 mm intervals for GE scanners; using a B31f algorithm at 0.625 (Sensation-16) or 0.75 mm slice thickness and 0.5 mm intervals for Siemens scanners; and using a B algorithm at 0.9 mm slice thickness and 0.45 mm intervals for Philips scanners.

All subjects who participated in the COPDGene study provided written informed consent, and the study was approved by the institutional review boards at all of the participating centers.

Objective CT Analysis

The objective detection and quantification of the volume of radiographic interstitial feature subtypes was performed using an approach similar to that which we designed for evaluating subtypes of emphysema, and which combines the attenuation properties of the local tissue and distance from the pleural surface (18). First, segmentation of the lung parenchyma from the chest wall and surrounding structures was performed on the inspiratory scans using a fully automated approach that uses particles, thin plate splines, and maximum a posteriori estimation. This method has been described previously and was implemented through the Chest Imaging Platform (http://acil.med.harvard.edu/chest-imaging-platform) (19).

To train the objective detection and quantification tool, two experts placed 33,865 fiducials in 138 CT scans on radiologic features unique to each disease type. The training CT scans were randomly selected by subject identification number. The radiologic features included normal parenchyma, interstitial feature subtypes (reticular, honeycombing, centrilobular nodule, linear scar, nodular, subpleural line, ground-glass), and emphysema subtypes (centrilobular and paraseptal). (Note that panlobular emphysema was not identified in the training cases likely because patients with alpha 1 antitrypsin disease were not represented in the cohort.) This was done to build a library of points to be used as tissue classifiers.

Regions of interest (ROI) consisting of 30 by 30 in-plane voxels centered about these training points were constructed to use as training data. For each ROI, a local histogram of densities within that region was built using kernal density estimation. This information was combined with the distance of the ROI from the pleural surface to create a feature vector (Fig 1). Kernel density estimation is a nonparametric method used to estimate a probability distribution over the

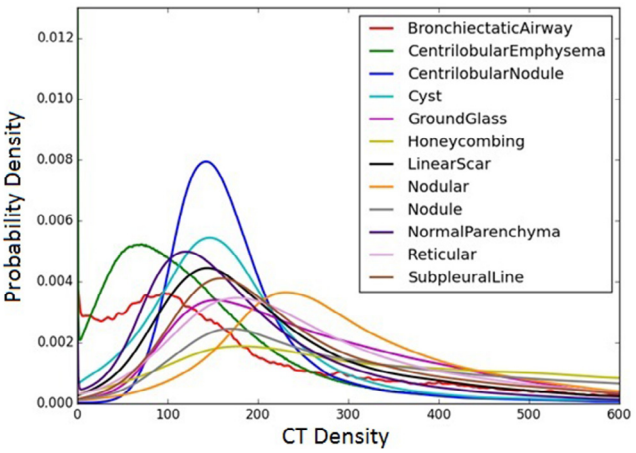

Figure 1. Averaged kernel density estimates over the training samples for each radiographic type. Each kernel density estimate approximates the probability density function over computed tomography density values in the training patch.