HHS Public Access

Author manuscript

Acad Radiol. Author manuscript; available in PMC 2016 June 01.

Published in final edited form as:

Acad Radiol. 2016 March ; 23(3): 304–314. doi:10.1016/j.acra.2015.11.013.

Novel Logistic Regression Model of Chest CT Attenuation Coefficient Distributions for the Automated Detection of Abnormal (Emphysema or ILD) versus Normal Lung

Kung-Sik Chan, PhD[1], **Feiran Jiao, MS**[1], **Marek A. Mikulski, MD, PhD**[2], **Alicia Gerke, MD**[3], **Junfeng Guo, PhD**[4], **John D Newell Jr, MD**[4], **Eric A. Hoffman, PhD**[4,5], **Brad Thompson, MD**[4], **Chang Hyun Lee, MD, PhD**[7], and **Laurence J. Fuortes, MD**[2,6]

[1]University of Iowa, Department of Statistics and Actuarial Science

[2]University of Iowa, Department of Occupational and Environmental Health

[3]University of Iowa, Department of Internal Medicine

[4]University of Iowa, Department of Radiology

[5]University of Iowa, Department of Medicine

[6]University of Iowa, Department of Epidemiology

[7]Seoul National University Hospital, Department of Radiology

Abstract

Rationale and Objectives—We evaluated the role of automated quantitative computed tomography (CT) scan interpretation algorithm in detecting Interstitial Lung Disease (ILD) and/or emphysema in a sample of elderly subjects with mild lung disease.ypothesized that the quantification and distributions of CT attenuation values on lung CT, over a subset of Hounsfield Units (HU) range [−1000 HU, 0 HU], can differentiate early or mild disease from normal lung.

Materials and Methods—We compared results of quantitative spiral rapid end-exhalation (functional residual capacity; FRC) and end-inhalation (total lung capacity; TLC) CT scan analyses in 52 subjects with radiographic evidence of mild fibrotic lung disease to 17 normal subjects. Several CT value distributions were explored, including (i) that from the peripheral lung taken at TLC (with peels at 15 or 65mm), (ii) the ratio of (i) to that from the core of lung, and (iii) the ratio of (ii) to its FRC counterpart. We developed a fused-lasso logistic regression model that can automatically identify sub-intervals of [−1000 HU, 0 HU] over which a CT value distribution provides optimal discrimination between abnormal and normal scans.

Results—The fused-lasso logistic regression model based on (ii) with 15 mm peel identified the relative frequency of CT values over [−1000, −900] and that over [−450,−200] HU as a means of discriminating abnormal versus normal, resulting in a zero out-sample false positive rate and 15%false negative rate of that was lowered to 12% by pooling information.

Reprints: Kung-Sik Chan, University of Iowa, Department of Statistics and Actuarial Science, Schaeffer Hall 241, Iowa City, IA 52242. kung-sik-chan@uiowa.edu..

Conclusions—We demonstrated the potential usefulness of this novel quantitative imaging analysis method in discriminating ILD and/or emphysema from normal lungs.

Keywords

Fused lasso; IDL; Logistic Regression; Multi-detector Volumetric CT Scanning

Interstitial Lung Disease is increasing in importance in part because of the aging population but also detection and/or incidence appear to be increasing. Data from the National Center for Health Statistics indicate that the age-adjusted mortality rate from pulmonary fibrosis has increased by 28.4% in men and 41.3% in women between 1992 and 2014. (1)

Inter- and intra- reader variability in interpretation of radiographs for pulmonary fibrosis and pneumoconiosis has been long recognized as a potential issue for screening programs, epidemiologic studies and medico-legal evaluations (2–4). Studies have found variable degrees of agreement for both parenchymal and pleural fibrosis dependent on the extent of abnormalities and the training and medical specialty of the chest-x-ray readers (5–8). The National Institute for Occupational Safety and Health (NIOSH) recommends using multiple, International Labour Organization radiographic pneumoconiosis classification system (ILO system) trained readers and median profusion scores as the preferred reconciliation protocol to increase accuracy and precision in PA film classification (9, 10). Clinical studies suggest CT imaging technology may become a gold standard for evaluation of obstructive airways disease (11), but the literature is relatively limited on CT's utility in quantifying and objectively characterizing patterns of subtle interstitial fibrosis.

High resolution CT scanning (HRCT) detects finer anatomic detail than conventional chest x-ray and its superior sensitivity in diagnosing ILD, as well as lower potential for inter-reader variability, has been established in multiple studies (12–17). This study utilizes recent advances in CT scanning, the spiral rapid CT with multi-detector volumetric CT scanning, a time and cost-effective alternative to a single-detector row CT (18, 19). Quantitative data analysis systems such as the Apollo™ (VIDA) software (20), allow for analysis of density histogram characteristics by lobe and region, similar to the ILO scheme.

Our approach leverages on the fact that scarring in the periphery of the lung results in measurable if not dramatic changes in the range, distribution and relative frequency of HU summed over sets of voxels. Similarly changes in the distribution and relative frequencies of voxels in differing ranges of HU are useful in detecting and quantifying other lung diseases such as emphysema. Our approach allows identification of ranges of CT values which lead to the best discrimination between fibrotic and normal lung. It is well known that the distribution of HU distributions and frequencies could be affected by potential confounders including gender, age and BMI. Therefore, we investigated that these confounding effects can be effectively mitigated by using the ratio function of the CT value distribution over the periphery of the lung to that over the core, including the mediastinal structures and hilum. The purpose of this study was to assess the potential of this novel CT technology in identifying and characterizing patterns of subtle interstitial changes especially in evaluation of normal aging lung and ILD.

MATERIALS AND METHODS

DATA

The data for this study consisted of lung CT images from 17 subjects with no radiographic and functional abnormalities and 52 subjects with ILD diagnosis confirmed by ILO review of CXR radiographs and functional testing. Table 1 summarizes demographic and clinical characteristics of the study subjects. Subjects with ILD were significantly different from normal subjects in the following demographics, at 5% level: predominantly male (92% vs. 47% in the normal group), older (77.45±8.41 years old vs. 40.35±16.66), smokers (20% never smokers vs. 100%), presence of pleural plaques (38% vs. 0%) and majority having CT evidence of emphysema (88%) and bronchiectasis (83%). Table 2 contrasts several pulmonary test functions between the two groups. Compared with the normal subjects, subjects with ILD had lower total lung capacity (TLC) and residual volume (RV), but the differences were non-significant at 5% level. The subjects with ILD also had significantly lower functional vital capacity (FVC), forced expiratory volume in first second (FEV1), diffusion capacity for the lungs of carbon monoxide (DLCO) and FEV1/FVC, all of which, except FEV1/FVC, were adjusted for age, gender and body composition (and additionally hemoglobin level for DLCO). CT imaging was performed on a Siemens SOMATOM Definition FLASH (Siemens Healthcare, Erlangen, Germany). The CT scanning protocol consisted of obtaining multi-detector CT (MDCT) images at Total Lung Capacity (TLC) and at Functional Residual Capacity (FRC) supine, and one TLC prone. All results reported below are based on data measured supine. Unique breathing instructions were required in order to obtain appropriate lung volume images. The MDCT acquisition parameters consisted of 120 kVp, 165 mAs, rotation time 0.5 seconds, and a pitch of 1. Reconstructed scan data used a 0.75mm slice thickness and 0.5mm slice spacing, and standard reconstruction algorithm (B35) for the computer analysis. A lung reconstruction algorithm (B50) was selected for the supervising radiologist to interpret the image dataset and provide a visual assessment. The images from the MDCT chest were sent via a web-based system called Medical Image File Archive and Retrieval (MIFAR), University of Iowa, using HIPPA standard de-identification procedures and secure storage of ILO data.

The CT attenuation value of a lung voxel represents the quantitative measure of the linear attenuation coefficient expressed in Hounsfield Units (HU). The relative frequency distribution of the CT values from all voxels in a lung region of interest provides a statistical description of the tissue density distribution over that lung area; below, the relative frequency distribution of the HU is simply referred to as the CT value distribution. For each lung image, we separately computed the CT value distribution over the peripheral of the lung (peel was 15 mm or 65 mm from the border of the lung) and over the core of the lung, including the lung fields, mediastinal structures and hilum. Each image was also evaluated by a panel of three chest radiologists using a scoring sheet developed for this study for interstitial and airways pathology for each lobe. The Pulmonary Analysis Software Suite (PASS) software (21, 22) was used to automatically demarcate the mediastinal and intrathoracic margins for each slice of each scan. The PASS system integrates the thoracic cavity and mediastinal borders in order to separate lung from other thoracic structure. Analyses were carried out with the original CT scan images, and compared to images

corrected for mediastinal margins to assess the hypothesis that manual modification enhances the signal in the data.

Denote by $d_{p,TLC}(\nu)$ the relative frequency of voxels in the peripheral lung image taken at TLC and having HU=ν, where ν ranges between −1000 and 0, in our application. Normal and abnormal subjects are assumed to have different patterns in their CT value distributions, with major differences occurring over a subset of the lung CT values range. Hence, it is pivotal to find a subset of the interval [−1000,0], over which some functional of $d_{p,TLC}$ furnishes a useful statistic for discriminating an abnormal scan from a normal scan. For conciseness, we denote the (random) function $d_{p,TLC}$ as V_1. However, the relevant information in $d_{p,TLC}$ is likely confounded by other factors such as age, gender, BMI, etc. One way to adjust for the confounding factors is to compute the function $V_2 = d_{p,TLC} / d_{c,TLC}$ whose value at ν HU is the ratio $d_{p,TLC}(\nu) / d_{c,TLC}(\nu)$, where $d_{c,TLC}(\nu)$ is the corresponding relative frequency of voxels in the core lung image taken at TLC. The idea is that a confounding factor may affect both the peel and core CT value distributions with an approximately identical subject-specific multiplicative factor, so taking the ratio eliminates the variation due to confounding factors. Thus, the problem becomes finding some range of [−1000,0] over which some functional of $d_{p,TLC}/d_{c,TLC}$ is a useful discriminator. However, the ratio $d_{p,TLC}(\nu)/d_{c,TLC}(\nu)$ is sensitive to measurement error when the denominator is close to 0. The problem can be mitigated by replacing the denominator by the sum of the numerator and the original denominator, i.e., by applying the instantaneous transformation $\psi(t) = t/(1 + t)$, resulting in the function $V_2'(\nu) = \psi(V_2(\nu)) = d_{p,TLC}(\nu) / \left(d_{c,TLC}(\nu) + d_{p,TLC}(\nu)\right)$. The transformation ψ effects an increasing, one-to-one mapping between the non-negative real line to the interval [0,1). Hence, $V_2'(\nu)$ is always between 0 and 1, sharply curtailing any large fluctuations in V_2. Similarly, we apply the transformation to V_1 to obtain $V_1' = \psi(V_1)$ to mitigate any large fluctuations in V_1.

The functional aspect of the lung may be partially captured by changes in the CT value distribution from FRC to TLC. This can be measured by $\frac{d_{p,TLC}}{d_{c,TLC}} / \frac{d_{p,FRC}}{d_{c,FRC}}$ denoted as V_3, and its robustified version $V_3' = \psi(V_3)$. The last function V_4 is the ratio of the relative frequency of CT values across voxels in the peripheral lung image taken at TLC versus that at FRC, denoted by $d_{p,TLC}/d_{p,FRC}$, and its robustified version $V_4' = \psi(V_4)$. These four functions and their robustified versions may potentially reveal interesting lung features from different perspectives, that can help us discriminate an abnormal scan from a normal scan.

In practice, there are a large but finite number of voxels in each lung image, so we compute the relative frequency of voxels over bins of size 10 HU. Altogether, each CT value distribution comprises the relative frequencies over 100 bins spanning from −1000 HU to 0 HU. Plots of the eight functions, namely, V_1, V_2, V_3, V_4 and their robustified counterparts V_i', i = 1,2,3,4, derived from lung images modified by radiologists and peel depth equal to 15mm, are shown in Fig. 1. Fig. 1 illustrates the presence of systematic differences between the normal and the abnormal populations, to varying degrees. For instance, the functions

from the normal lungs fluctuate tightly within narrow bands for all *V*s and *V*′s, whereas those from the abnormal subjects display greater fluctuations and more diverse functional form with greater skewness to the right (higher HU values). Moreover, plots for the robust functions V_3' and V_4', provide finer contrast between normal and abnormal lungs than V_3 and V_4.

Note that even in healthy individuals, there is loss of lung parenchyma with age. The healthy group is almost half the age of the ILD group. Due to age differences alone between the ILD and healthy groups there will be an increase in the fraction of lung tissue with a CT value between −1000 and −900 in the older group regardless of lung disease. In addition, males have larger lungs than females and will have a lower lung density at TLC. Again, this will impact the amount of lung tissue between −1000 HU and −900 HU, creating a bias for the older group, (the ILD subjects). Thus, it is pivotal to assess the efficacy of V_2', V_3' and V_4' in removing these confounding effects, which can be assessed per HU as follows. For each $\nu \in [-1000,0]$ and *V*′ equal to V_i', *i* = 1,2,3,4, we regress *V*′ (ν) on age, gender, age-gender interaction and BMI, and compute the p-value, denoted by $p_{V'}(\nu)$, for the null hypothesis that the confounders are jointly non-significant, i.e. all their coefficients are zero. By checking the location where the p-value is less than 5%, we can locate the HUs over which *V*′ is correlated with gender, age and/or BMI. Since gender, age and BMI differ significantly between the two groups of subjects, the regression must be fitted per HU, separately for the ILD group and the normal group, in order to properly assess any remaining correlations with gender, age and BMI.

Figures S1 and S2 plot the p-value against $\nu \in [-1000,0]$ (in fact, over each of the one hundred 10-HU bins), separately for normal subjects and IDL subjects, for V_1', V_2', V_3' and V_4'; superimposed in each figure is a horizontal line at the nominal 5% value. Because V_1' does not adjust for any confounding factors, it is expected that it may be correlated with gender, age and/or BMI over an interval at least extending from −1000 HU to −900 HU, while if V_2', V_3' and V_4' are successful in removing these confounding effects, then they will not be associated with these confounders. Fig. S1 shows that for the normal subjects, V_1' is significantly correlated with gender, age and/or BMI over the very narrow interval from −880 HU to −870 HU, but for the ILD subjects, V_1' is significantly correlated with gender, age and/or BMI over a wider interval from −940 HU to −680 HU, which is consistent with the aforementioned age and gender effects over the range between −1000 HU and −900 HU. On the contrary, Figures S1 and S2 show that V_2' and V_3' are uncorrelated with gender, age and/or BMI, within each of the two groups of subjects, across the entire range from −1000 HU to 0 HU, except for V_2' over the interval from −490 HU to −480 HU for the IDL group, and V_3' over the interval from −820 HU to −800 HU for the normal group. Interestingly, V_4' is significantly correlated with gender, age and/or BMI over the interval from −190 HU to −150 HU for the normal group, and also over the interval from −520 HU to −460 HU for the IDL group. Thus, V_2' and V_3' have successfully eliminated the confounding effects due to gender, age and BMI, whereas V_4' has done so with partial success.

Logistic Regression Model

We developed a logistic regression model for discriminating between normal and abnormal subjects. Let V stand for one of the functions (V_i, i=1,2,3,4 or V_i', i=1,2,3,4). The idea is to use some functional of V, in the form of $\beta_0+\int_{-1000}^{0}\beta(\nu)V(\nu)d\nu$, to be a linear predictor of the probability of abnormality, on the logistic scale, where β_0 is the intercept term and the coefficient function $\beta(\nu)$, $-1000\leq\nu\leq 0$ is assumed to be a sparse, piecewise constant function. Sparsity of a function means that it is mostly equal to a zero function, based on the belief that the CT value distribution of a normal scan mainly differs from that of an abnormal scan over a small subset of [−1000 HU, 0 HU]. As we only have a discretized version of V, in the form of a histogram with bin size of 10 HU, the predictor becomes $\beta_0+\sum_{j=1}^{M}\beta_j V(\nu_j)$. These considerations lead to the following logistic regression model:

$$\text{logit}(p)=\beta_0+\Sigma_{j=1}^{M}\beta_j V(\nu_j) \quad (1)$$

where $\text{logit}(p)=\log\{p/(1-p)\}$ is the logistic transformation, p is the probability that the subject is abnormal, i.e. having lung disease, M is the total number of bins, here 100, ν_j the center of the j^{th} bin and $\beta_j=\beta(\nu_j)$ the coefficient. The sparse, piece-wise constancy assumption on β entails that the β_j's are piece-wise constant and sparse.

The fused lasso estimator—Denote by $\boldsymbol{\beta}=(\beta_0,\beta_i, i=1,\ldots,M)^{\intercal}$ the coefficient vector. Assuming that the data comprise two independent random samples from the normal and abnormal populations, the maximum likelihood (ML) estimator of $\boldsymbol{\beta}$ is obtained by maximizing the following negative log likelihood function

$$l(\boldsymbol{\beta})=\Sigma_{i=1}^{n}w_i\{y_i\,log(p_i)+(1-y_i)\,log(1-p_i)\} \quad (2)$$

where w_i are known weights, y_i is 1 if the i^{th} subject is abnormal and 0 otherwise, p is given by Eqn. (1) with V there replaced by that of the i^{th} subject, and n=17+52=69, the total sample size. The weights are generally set to be identically 1, but to mitigate the unbalanced group sizes (17 normal scans vs. 52 abnormal scan), we set the weights to be proportionally 52 for each normal scan and 17 for each abnormal scan. The ML estimator is, however, generally neither sparse nor piece-wise constant. So, a different approach of estimation is desirable.

The sparsity and piecewise constancy properties, however, can be prompted by restricting the parameter space to the set of parameters deemed to be sparse and piece-wise constant. The non-sparsity of $\boldsymbol{\beta}$ can be measured by its L^1-norm $\sum_{i=1}^{M}|\beta_i|$, i.e. the sparser $\boldsymbol{\beta}$ is the smaller its L^1 norm is. Piecewise constancy is equivalent to stipulating that consecutive β_i's (excluding β_0) are generally identical, which amounts to requiring the sparsity of $(\beta_2-\beta_1,\beta_3-\beta_2,\ldots,\beta_M-\beta_{M-1})^{\intercal}$, the first difference of $\boldsymbol{\beta}$, with β_0 excluded. The sparsity and piecewise constancy conditions may then be enforced by maximum likelihood estimation over a constrained parameter space defined by bounding $\sum_{i=1}^{M}|\beta_i|$ and $\sum_{i=2}^{M}|\beta_i-\beta_{i-1}|$ by some suitable upper bounds. The preceding strategy can, however, be more conveniently

formulated in terms of maximizing the penalized log likelihood function with the fused lasso penalty (23):

$$l(\boldsymbol{\beta}) - \lambda_1 \Sigma_{j=1}^{M} |\beta_j| - \lambda_2 \Sigma_{j=2}^{M} |\beta_j - \beta_{j-1}| \quad (3)$$

where λ_1 and λ_2 are two non-negative tuning parameters to be determined by cross-validation; the penalized likelihood estimator is then obtained by maximizing (3). The two tuning parameters effectively determine the degree of sparsity and piece-wise constancy in the function estimate $\hat{\beta}$. For instance, when both tuning parameters are zero, the estimation becomes unconstrained maximum likelihood estimation resulting in a generally non-sparse estimator that is not piece-wise constant. On the other hand, for very large $\lambda_1(\lambda_2)$, the β estimates will be mostly zero (approach a constant function with few jumps). Thus, the choice of the two tuning parameters is pivotal.

In all numerical work reported below, the tuning parameters are determined by 5-fold cross validation as follows. The data are randomly split into five blocks of approximately equal size, with one block serving for validating the model fit to the other four blocks of data. For each pair of (λ_1,λ_2), we obtain $\hat{\beta}(\lambda_1, \lambda_2)$, the penalized estimator of $\boldsymbol{\beta}$ by maximizing (3) using data from 4 blocks, and then evaluate the (validatory) log likelihood at $\hat{\beta}(\lambda_1, \lambda_2)$ using the remaining block of data; we then compute the cross-validatory log-likelihood of the tuning parameter vector (λ_1,λ_2) by averaging the validatory log-likelihood over the 5 blocks of data. The tuning parameters $\hat{\lambda}_1$, $\hat{\lambda}_2$ are chosen by maximizing the cross-validatory log-likelihood. Finally, $\hat{\beta}\left(\hat{\lambda}_1, \hat{\lambda}_2\right)$ is the fused lasso estimator. Optimization of Eqn. (3) was carried out by the Matlab package SLEP (Sparse Learning with Efficient Projections; (24, 25))

Misclassification Rates

We fit the model with *V* as one of the eight functions, one by one, and compared them based on their misclassification rates. A subject is classified as abnormal if the odds of having an abnormal lung is not less than 1, i.e. $\hat{p}/(1-\hat{p})$, is ≥ 1 and otherwise classified as being normal. The performance of a classification scheme, as effected by model (1) with a particular choice for *V*, is characterized by its in-sample and out-sample misclassification rates. In-sample misclassification rates concerns the error rates when classification is done based on the model fit using all data. In contrast, out-sample misclassification rates attempt to measure the error rates of a proposed classification scheme when it is applied to the general population. This is done by computing the error rates with the classification of each subject done based on the model fit using all data except the data from the subject to be classified. The out-sample misclassification rates are generally more indicative of the accuracy of a proposed classification method in real practice. Two commonly used misclassification rates are false positive rate (FPR) and false negative rate (FNR). FPR measures the proportion of normal subjects who are misclassified as abnormal. The true positive rate is obtained by subtracting FNR from 1. FNR refers to the proportion of abnormal subjects incorrectly classified as normal. Similarly, the true negative rate is one minus FNR. The total false rate (TFR) is the ratio of the number of false positives and

negatives to the total number of subjects; it measures the overall error rate of the method. The in-sample and/or out-sample misclassification rates can be used as a basis for choosing which of the 8 functions (V_i,i=1,2,3,4 or V_i',i=1,2,3,4) leads to the best discriminator between normal and abnormal subjects.

Information Pooling

Instead of picking which of the 8 functions to use, an alternative, perhaps better approach is combining the information contained in a set of functions. A simple way to combine information occurs in the case that the (random) functions are stochastically independent given the disease status of the subject. In particular, V_i,i = 1,2,3,4 may be assumed to satisfy the preceding assumption, (and so may V_i') because these functions measured from a normal (abnormal) subject are likely independently distributed about their normal (abnormal) patterns. But across a mixed population of normal and abnormal subjects, these random functions may become dependent through the latent disease status. Denote by A the event that a (random) subject is abnormal and N the event that the subject is normal. Let $P(A|V_i)$ be the probability that a subject is abnormal, computed based on the logistic regression model (1) using V_i. Similarly, $P(N|V_i)$ is the corresponding probability that a subject is normal. It is shown in the Appendix that under the further assumption that a subject has even prior odds of being abnormal, then the posterior odds of abnormality given all the information in V_i,$i \in S$, where S is a subset of {1,2,3,4}, is given by the following formula:

$$\frac{P(A|V_i, i \in S)}{P(N|V_i, i \in S)} = \Pi_{i \in S} \frac{P(A|V_i)}{P(N|V_i)}. \quad (4)$$

In other words, given all in the information contained in V_i,$i \in S$, the posterior odds that a subject is abnormal is proportional to the product of the corresponding posterior odds given the information in individual functions to be pooled.

Data used in this analysis were collected from a study on the implications of and association between radiographic evidence of interstitial lung disease and spirometry in an elderly population that was approved by our institutional review board. Written informed consent was obtained from all subjects.

RESULTS

We fit the fused lasso logistic regressions using the SLEP package for each of the eight functions derived from lung images acquired under several conditions, namely, whether image reconstruction was modified, (anatomic boundaries edited or defined by hand), or not, (anatomic borders read automatically). (In the models, yes modification =True (T) or no modification = False (F)), the peel depth was defined as either 15mm or 65mm, and the breathing phase defined (TLC or FRC). Hence, eight logistic regression models were fit, per each combination of modification and peel depth. Fig. 2 plots the fused lasso coefficient estimate as a function of HU, over the range from −1000 to 0, for the case of modified lung images with 15mm peel depth. Supplementary Fig. S3 shows the estimated coefficients for 15mm peel depth and Modification = F. Notice that there are little differences in the

function fits whether or not the lung image reconstruction was modified, at 15mm peel depth. Figs. S4 and S5 show the corresponding function estimates, for the case of 65mm peel depth.

Misclassification rates were calculated for assessing the classification performance for each of the eight functions, and under different means of processing the lung images. Results (not reported), however, show that pooling information across the robustified variables, i.e. $V_i^{'}$,*i*=1,2,3,4, generally yield the lowest misclassification rates. Tables 3--6 show results using each robustified variables. Tables 7--9 show some results using three different combinations of the robustified functions.

In terms of out-sample misclassification rates, the best discriminatory performance was achieved by pooling information in $V_2^{'}$, $V_3^{'}$, $V_4^{'}$ with the second best achieved by pooling information in $V_2^{'}$, $V_3^{'}$, that were derived from lung images with 15mm peel depth; the same performance was achieved whether or not image modification took place. This suggests the excellence of the automatic image reconstruction procedure, as no further improvement results from manual modification of anatomic margins by radiologists. However, as $V_2^{'}$, $V_3^{'}$ are largely free of any confounding effects due to age, gender and/or BMI while $V_4^{'}$ is only partially successful in removing these confounding effects, it is perhaps clinically more relevant to use information in $V_2^{'}$, $V_3^{'}$ for classifying the subjects. Moreover, the out-sample error rates using $V_2^{'}$, $V_3^{'}$ are about 0% FPR and 12% FNR, indicating that the proposed method has about 88% power of detecting a patient with lung disease at about 0% type I error rate.

The above analysis was done using image data from both (left and right) lungs. We have repeated the analysis using only one lung, but analysis with data from both lungs generally yielded lower error rates than using data from left (right) lung alone.

DISCUSSION

The fused-lasso function estimates, $\hat{\beta}_i$, *i* = 1, 2, …, 100, shown in Fig. 2 (or Fig. S3) provide some clues on the major difference between the normal and abnormal scans. For example the function estimate of the logistic regression model using $V_2^{'}$ equals zero, except that it is approximately equal to 0.6 and 1.3 between −1000 HU to −900 HU, and −400 HU to −200 HU, respectively, which, letting $V(\nu_1,\nu_2)$ be the sum of $V(\nu_j)$ with $\nu_1 \le \nu \le \nu_2$, can be written as:

$$\text{logit}\,(p) \approx \text{constant} + 0.6 \times V_2^{'}\,(-1000, -900) + 1.3 \times V_2^{'}\,(-400, -200)\,.$$

In other words, the log odds of abnormality is, up to an additive constant, equal to the sum of 0.6 times the adjusted density of lung tissues between −1000 HU and −900 HU plus 1.3 times the adjusted density of lung tissues between −400 HU to −200 HU. Hence, the odds for abnormality increases with the fraction of lung tissues with CT values between −1000 HU and −900 HU, (emphysema range), and that between −400 HU and −200 HU, (fibrosis

range). Thus, abnormal scans appear to have higher percent of hyperaerated lung tissue (26) that may be due to emphysema, and higher percent of poorly aerated lung tissue, e.g. lung scarring. The information contained in V_3' may be inferred similarly:

$$\text{logit}\,(p) \approx \text{constant} - 1.8 \times V_3'\,(-1000, -950) + 0.4 \times \left\{ V_3'\,(-950, -900) - 2 \times V_3'\,(-900, -800) + V_3'\,(-700, -600) \right\},$$

where the term enclosed in curly brackets is equal to $\left\{ V_3'\ (-700, -600) - V_3'\ (-900, -800) \right\} - \left\{ V_3'\ (-900, -800) - V_3'\ (-950, -900) \right\}$ which can be loosely interpreted as curvature, i.e., the second derivative of the V_3' over the interval between −950 HU to −600 HU. Thus, the odds of abnormality then increases with reduced hyperaerated lung tissue from FRC to TLC, and it also increases with the second derivative of the V_3' over the interval between −950 HU to −600 HU. Indeed, Fig. 1 shows that for normal scans, V_3' is generally a concave function between −950 HU to −600 HU, whereas it may become a convex function for abnormal scans. The physiological basis for the concavity in the normal population versus the convexity in some of the subjects with abnormal scans is an interesting future research problem. Fig. 2 displays that the function estimate for the logistic regression model using V_4' implies the following model:

$$\text{logit}\,(p) \approx \text{constant} - 0.4 \times V_4'\,(-1000, -900) + 1.9 \times V_4'\,(-800, -700) - V_4'\,(-500, -450)\,.$$

Thus, the odds of abnormality increases with decreased proportion of hyperaerated lung tissue from FRC to TLC, increased proportion of lung tissue between −800 and −700 HU from FRC to TLC, and decreased proportion of lung tissue between −500 HU and −450 HU from FRC to TLC. Fig. 1 indicates that for normal subjects, V_4' takes a high value between −1000 HU and −800 HU and then drops sharply to a lower level between −800 HU and 0 HU. On the other hand, for abnormal scans, the drop in V_4' is generally much more gradual. Hence, the estimated logistic regression model attempts to use the rate of change in the drop of V_4' for discriminating between normal and abnormal scans. A caveat is in order: recall that we have earlier demonstrated that V_4' contains signal from age, gender and/or BMI over the interval between −520 HU and −460 HU so that V_4' derives some of its discriminatory power from these confounders. Studies of populations of elderly subjects with both normal and abnormal physiology and radiography will help refine these diagnostic models. Note that the function estimates for V_i, i = 1,2,3,4 are non-zero over relatively narrow ranges, as compared to their robustified counterparts. This indicates that robustification enhances the signal in the data.

This method is relatively simple, eliminates the need for between scanner calibration and could be applied to automated screening of interstitial disease and perhaps more importantly in providing an objective measure of degree of interstitial disease which could be used to assess disease progression and or response to treatment (27).

We had an a priori assumption that the percent of voxels in the lung periphery in the HU range of −1000 to −850 should serve as an indicator of emphysema and similarly that the percent of voxels in the lung periphery in HU range of −450 to −200 might serve as an indicator of fibrosis. This prior assumption is essentially confirmed by the proposed method that infers the relevant HU intervals for disease detection, based on the data alone and not using any prior information. In contrast, earlier works focused on finding biomarkers of interstitial disease in terms of summary statistics of the entire CT value distributions. In particular, others have evaluated mean lung attenuation, skewness, and kurtosis as indicators of fibrosis with alteration of skew to the left decreased kurtosis or sharpness of the histogram peak associated with fibrosis (28), revealing strong correlations of these summary statistics with several physiologic and quality-of-life variables. Another approach developed by Yilmaz et al (29) uses the CT values to compute lobar summary statistics (averages, variance and inter- and intra-lobar coefficient of variation) of the tissue volume, air volume and fractional tissue volume. Authors demonstrated that these lobar summary statistics were strongly correlated with several pulmonary test functions. None of these earlier works attempted to develop formal tools for discriminating between abnormal and normal lungs, or pooling information across different phases of inspiration.

In an attempt to account for differences in BMI, penetration and possible technique differences the range and distribution of HU in the periphery of the lung was adjusted by the range of HU in the mediastinum providing an internal control for water and or soft tissue density. Our approach is simpler than that of Yilmaz et al who calibrated the CT attenuation coefficients with subject-specific references, namely, those of intra-thoracic air and air-free tissues. Moreover, our method makes use of the entire CT value distribution in classifying a lung, and quantifies the relevant range of HUs over which the functional statistics V_2' and V_3' can be compared with the normal profile for detecting abnormality in the lung. An important feature of V_2' and V_3' is that based on our limited data, they are free of any confounding effects due to gender, age and BMI, which, if confirmed in a future large-scale study, makes them potentially useful tools for lung disease diagnosis.

These models confirmed the hypothesis that discrimination between normal and abnormal films could be improved by correcting for the ratio of increased attenuation in the peel to the core, (comparing peripheral to lung fields with mediastinal structures). Although the most rigorous model included both the end inhalation, (TLC), and end exhalation, (FRC), images, the model using only TLC, V_2' above, resulted in an out-sample FPR of 0% and a 13% FNR suggesting the possible utility of this method as both a screening and clinical tool. (Equivalently, the true positive rate (TPR) is 100% and true negative rate (TNR) is 87 %.) We have demonstrated the proposed quantitative imaging analysis method has good empirical power for discriminating ILD or emphysema from normal lungs, but it is pertinent to conduct further validation of the proposed method with more extensive data. It is an interesting problem to assess the potential of adapting the proposed method to monitor progression of ILD or emphysema, which will be useful for disease management and assessment. Ultimately technology may improve the current methods of diagnosing and assessing the extent of interstitial radiographic abnormalities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported, in part, by National Institutes of Health grant NIHR01HL112986, National Institute for Environmental Health Sciences through the University of Iowa Environmental Health Sciences Research Center, NIEHS/NIH P30 ES05605, NIH U01 HL114494, NIH RO1 HL089897, and by a pilot project research training grant from the Heartland Center for Occupational Health and Safety at the University of Iowa. The Heartland Center supported by Training Grant No. T42OH008491 from the Centers for Disease Control and Prevention/ National Institute for Occupational Safety and Health. E.A.H. is a founder and shareholder of VIDA Diagnostics, Inc., a company commercializing lung image analysis software developed, in part, at the University of Iowa. K.S.C., and J.D.N. are paid consultants for VIDA Diagnostics, Inc. J.D.N. has stock options with VIDA. J.G. owns VIDA stock. Siemens Healthcare has provided in-kind support for hardware and software residing at the University of Iowa and used in this project.

Appendix

Proof of Eqn. (4)

We shall prove a slightly more general result that indicates how to modify Eqn. (4) when the prior odds of abnormality need not be 1. Let D be the disease status of the subject, which takes the value 1 if the subject is abnormal and 0 otherwise. In other words, the event A (N) occurs if and only if D=1 (0). Let S be a subset of {1,2,3,4} and denote by $|S|$ the number of elements in S. Let $p(v_i, i \in S; d)$ be the joint probability density function of $V_i, i \in S$ and D, evaluated at $v_i \in R^M, i \in S$ and $d \in \{0,1\}$. (Technically, the V's and D are assumed to admit a probability density function that is the Radon-Nikodym derivative w.r.t. the product measure of the Lebesgue measure on $R^{M \times |S|}$ and the counting measure on $\{0,1\}$.) Let $p(d)=p(D=d)$ be the prior probability that $D=d$. The assumption that $V_i, i \in S$ are conditionally independent given D implies that $p(v_i, i \in S|d)=\Pi_{i \in S} p(v_i|d)$. We aim to compute $p(d|v_i, i \in S)$, the posterior probability that $D=d$ given the observations. To do so, consider the product

$$\begin{aligned} & p\left(d|\nu_i, i \in S\right) \times p\left(\nu_i, i \in S\right) \\ = \; & p\left(\nu_i, i \in S; d\right) \\ = \; & p\left(d\right) p\left(\nu_i, i \in S|d\right) \\ = \; & p\left(d\right) \prod_{i \in S} p\left(\nu_i|d\right) \\ = \; & \left\{p\left(d\right)\right\}^{-|S|+1} \prod_{i \in S} p\left(\nu_i; d\right) \\ = \; & \left\{p\left(d\right)\right\}^{-|S|+1} \prod_{i \in S} p\left(d|\nu_i\right) p\left(\nu_i\right) \end{aligned}$$

Hence, the posterior odds of abnormality is

$$\begin{aligned} & \frac{p(D=1)|\nu_i, i \in S)}{p(D=0)|\nu_i, i \in S)} \\ = \; & \frac{\{p(D=0)\}^{|S|-1}}{\{p(D=1)\}^{|S|-1}} \prod_{i \in S} \frac{p(D=1|\nu_i)}{p(D=0|\nu_i)}, \end{aligned}$$

which becomes $\Pi_{i\in S}\frac{p(D=1|\nu_i)}{p(D=0|\nu_i)}$, the product of the posterior odds when the prior odds $\frac{p(D=1)}{p(D=0)}=1$.

REFERENCES

1. Olson AL, Swigris JJ, Lezotte DC, et al. Mortality from pulmonary fibrosis increased in the United States from 1992 to 2003. American journal of respiratory and critical care medicine. 2007; 176(3): 277–84. [PubMed: 17478620]
2. Fletcher C, Oldham P. Problem of Consistent Radiological Diagnosis in Coalminers' Pneumoconiosis: An Experimental Study. British journal of industrial medicine. 1949; 6(3):168. [PubMed: 18146274]
3. Impivaara O, Zitting AJ, Kuusela T, et al. Observer variation in classifying chest radiographs for small lung opacities and pleural abnormalities in a population sample. American journal of industrial medicine. 1998; 34(3):261–5. [PubMed: 9698995]
4. Mulloy KB, Coultas DB, Samet JM. Use of chest radiographs in epidemiological investigations of pneumoconioses. British journal of industrial medicine. 1993; 50(3):273–5. [PubMed: 8457495]
5. Mikulski MA, Hartley PG, Sprince NL, et al. Risk and significance of chest radiograph and pulmonary function abnormalities in an elderly cohort of former nuclear weapons workers. J Occup Environ Med. 2011; 53(9):1046–53. [PubMed: 21866051]
6. Muir D, Bernholz CD, Morgan W, et al. Classification of chest radiographs for pneumoconiosis: a comparison of two methods of reading. British journal of industrial medicine. 1992; 49(12):869. [PubMed: 1472446]
7. Musch D, Higgins I, Landis J. Some factors influencing interobserver variation in classifying simple pneumoconiosis. British journal of industrial medicine. 1985; 42(5):346–9. [PubMed: 3986146]
8. Yerushalmy J, Garland L, Harkness J, et al. An evaluation of the role of serial chest roentgenograms in estimating the progress of disease in patients with pulmonary tuberculosis. American review of tuberculosis. 1951; 64(3):225–48. [PubMed: 14868842]
9. International Labour Organization (ILO). International Labour Office. Guidelines for the use of the ILO international classification of radiographs of pneumoconioses. ILO; Geneva: 2002. Occupational Safety and Health Series No. 22, Rev 20002002
10. National Institute of Occupational Safety and Health (NIOSH). NIOSH, Centers for Disease Control and Prevention (CDC). Atlanta, GA: 2010. Issues in classification of chest radiographs.
11. Busacker A, Newell JD, Keefe T, et al. A multivariate analysis of risk factors for the air-trapping asthmatic phenotype as measured by quantitative CT analysis. CHEST Journal. 2009; 135(1):48–56.
12. Al Jarad N, Poulakis N, Pearson M, et al. Assessment of asbestos-induced pleural disease by computed tomography—correlation with chest radiograph and lung function. Respiratory medicine. 1991; 85(3):203–8. [PubMed: 1882109]
13. Bessis L, Callard P, Gotheil C, et al. High-resolution CT of parenchymal lung disease: precise correlation with histologic findings. Radiographics. 1992; 12(1):45–58. [PubMed: 1734481]
14. Friedman AC, Fiel SB, Fisher MS, et al. Asbestos-related pleural disease and asbestosis: a comparison of CT and chest radiography. American Journal of Roentgenology. 1988; 150(2):269–75. [PubMed: 3257311]
15. Harkin TJ, McGuinness G, Goldring R, et al. Differentiation of the ILO boundary chest roentgenograph (0/1 to 1/0) in asbestosis by high-resolution computed tomography scan, alveolitis, and respiratory impairment. Journal of occupational and environmental medicine. 1996; 38(1):46–52. [PubMed: 8871331]
16. Mathieson J, Mayo J, Staples C, Müller N. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. Radiology. 1989; 171(1):111–6. [PubMed: 2928513]

17. Ochsmann E, Carl T, Brand P, et al. Inter-reader variability in chest radiography and HRCT for the early detection of asbestos-related lung and pleural abnormalities in a cohort of 636 asbestos-exposed subjects. International archives of occupational and environmental health. 2010; 83(1): 39–46. [PubMed: 19618201]
18. Flohr TG, Schaller S, Stierstorfer K, et al. Multi–Detector Row CT Systems and Image-Reconstruction Techniques 1. Radiology. 2005; 235(3):756–73. [PubMed: 15833981]
19. Kalra MK, Maher MM, Toth TL, et al. Techniques and Applications of Automatic Tube Current Modulation for CT 1. Radiology. 2004; 233(3):649–57. [PubMed: 15498896]
20. VIDA diagnostics. Accessed at www.vidadiagnostics.com/solutions/APOLLO.htm
21. Guo J, Fuld MK, Alford SK, et al. Pulmonary Analysis Software Suite 9.0: Integrating quantitative measures of function with structural analyses. First International Workshop on Pulmonary Image Analysis. 2008:283–92.
22. Hu S, Hoffman EA, Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. Medical Imaging, IEEE Transactions on. 2001; 20(6):490–8.
23. Tibshirani R, Saunders M, Rosset S, et al. Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2005; 67(1):91–108.
24. Guide MUs. The mathworks. Inc. Natick, MA: 1998. p. 3335
25. Liu, J.; Ji, S.; Ye, J. SLEP: Sparse learning with efficient projections. Arizona State University; 2009.
26. Gattinoni L, Caironi P, Pelosi P, Goodman LR. What has computed tomography taught us about the acute respiratory distress syndrome? American journal of respiratory and critical care medicine. 2001; 164(9):1701–11. [PubMed: 11719313]
27. Lynch DA, Godwin JD, Safrin S, et al. High-resolution computed tomography in idiopathic pulmonary fibrosis: diagnosis and prognosis. American journal of respiratory and critical care medicine. 2005; 172(4):488–93. [PubMed: 15894598]
28. Hartley PG, Galvin JR, Hunninghake GW, et al. High-resolution CT-derived measures of lung density are valid indexes of interstitial lung disease. Journal of Applied Physiology. 1994; 76(1): 271–7. [PubMed: 8175517]
29. Yilmaz C, Watharkar SS, de Leon AD, et al. Quantification of regional interstitial lung disease from CT-derived fractional tissue volume: a Lung Tissue Research Consortium Study. Academic radiology. 2011; 18(8):1014–23. [PubMed: 21596593]

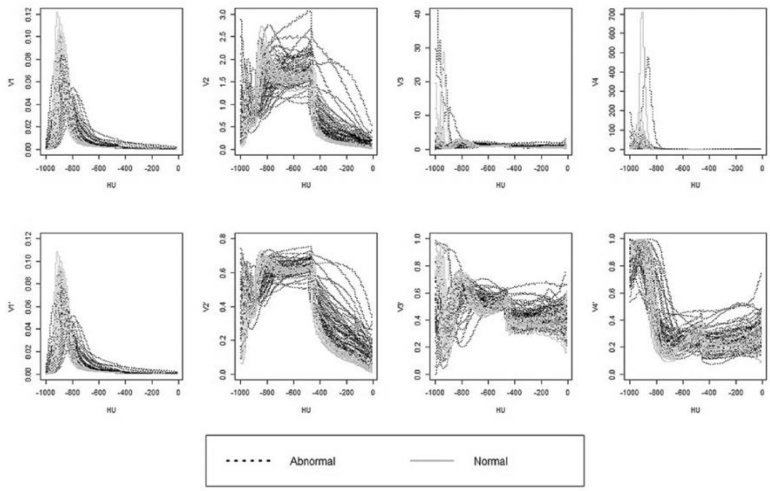

Fig. 1.

Plots of the eight function summaries of the CT attenuation value distribution of lung parenchyma. Diagrams in the upper panel, from left to right, plot the realizations of V_i,i=1,2,3,4, with those of the normal lungs drawn as gray solid lines and the abnormal counterparts as dashed lines. Diagrams in the lower panel are those of V_i',i=1,2,3,4. These plots show that V_s and V' s of the normal lungs fluctuated within narrow bands, while those of the abnormal lungs had greater fluctuations with higher skewness to the right. The preceding contrast is generally amplified in the plots of the V' s.

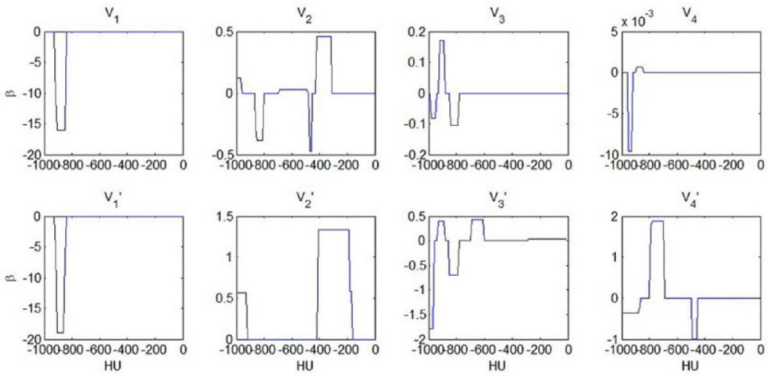

Fig. 2.

Plot of the coefficient estimates $\hat{\beta}$ as a function of HU; 15mm peel depth and modification = T. The leftmost figure in the upper panel shows the fussed-lasso estimate of β_j as a function of v_j in model (1) with V there being V_1 computed with data derived from lung images modified by radiologists and peel depth equal to 15mm. Note that the function estimate is essentially a piecewise constant function that is non-zero over a small subset of [−1000 HU, 0 HU], as a result of the fussed-lasso penalty enforced in maximizing the objection function (3). Other figures in the upper panels, from left to right, are similar figures with V in (1) replaced by V_i,i=2,3,4 Figures in the lower panel are counterparts of V_i',i=1,2,3,4.

Table 1

Demographic and clinical characteristics of study populations.

	ILD	Normal	P-value
N	52	17	
Sex			0.0002
Male	48 (92%)	8 (47%)	
Female	4 (8%)	9 (53%)	
Age, year	77.35±8.41	40.35±16.66	<0.0001
Race			0.16
African American	0 (0%)	1 (6%)	
Caucasian	50 (96%)	16 (94%)	
Hispanic	2 (4%)	0 (0%)	
BMI	28.81±5.25	25.65±3.69	0.025
Smoking history			0.00006
Never	20 (38%)	17 (100%)	
Former	25 (48%)	0 (0%)	
Current	7 (14%)	0 (0%)	
Pack-years	42.39±34.47	0±0	0.003
Plaques			<0.0001
No	32 (62%)	17 (100%)	
Yes	20 (38%)	0 (0%)	
Emphysema			<0.0001
No	6 (12%)	17(100%)	
Yes	46 (88%)	0 (0%)	
Bronchiectasis			<0.0001
No	9 (17%)	17(100%)	
Yes	43 (83%)	0 (0%)	

Table 2

Pulmonary function in patients with ILD compared with normal subjects.

	ILD	Normal	P-value
N	52	17	
TLC	96.41±22.20	103.41±8.52	0.21
RV	96.08±31.26	100.18±17.44	0.61
FVC	82.71±24.11	109.76±19.62	<0.0001
FEV1	78.40±24.40	112.82±20.66	<0.0001
DL CO	62.25±19.91	126.71±15.81	<0.0001
FEV1/FVC	68.88±14.14	82.94±6.43	0.0002

Table 3

Misclassification rates based on classification using V_1'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	3/17≈0.18	15/52≈0.29	18/69≈0.26	4/17≈0.24	15/52≈0.29	19/69≈0.28
15mm, False	3/17≈0.18	15/52≈0.29	18/69≈0.26	4/17≈0.24	15/52≈0.29	19/69≈0.28
65mm, True	3/17≈0.18	15/52≈0.29	18/69≈0.26	4/17≈0.24	16/52≈0.31	20/69≈0.29
65mm, False	4/17≈0.24	15/52≈0.29	19/69≈0.28	4/17≈0.24	16/52≈0.31	20/69≈0.29

Table 4

Misclassification rates based on classification using V_2'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	0/17=0	7/52≈0.13	7/69≈0.10	0/17=0	8/52≈0.15	8/69≈0.12
15mm, False	0/17=0	6/52≈0.12	6/69≈0.09	0/17=0	7/52≈0.13	7/69≈0.10
65mm, True	2/17≈0.12	9/52≈0.17	11/69≈0.16	2/17≈0.12	11/52≈0.21	13/69≈0.19
65mm, False	2/17≈0.12	9/52≈0.17	11/69≈0.16	2/17≈0.12	14/52≈0.27	16/69≈0.23

Table 5

Misclassification rates based on classification using V_3'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	2/17≈0.12	6/52≈0.12	8/69≈0.12	3/17≈0.18	7/52≈0.13	10/69≈0.14
15mm, False	2/17≈0.12	6/52≈0.12	8/69≈0.12	3/17≈0.18	7/52≈0.13	10/69≈0.14
65mm, True	1/17≈0.06	8/52≈0.15	9/69≈0.13	1/17≈0.06	9/52≈0.17	10/69≈0.14
65mm, False	1/17≈0.06	8/52≈0.15	9/69≈0.13	1/17≈0.06	8/52≈0.15	9/69≈0.13

Table 6

Misclassification rates based on classification using V_4'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	0/17=0	9/52≈0.17	9/69≈0.13	1/17≈0.06	10/52≈0.19	11/69≈0.16
15mm, False	0/17=0	9/52≈0.17	9/69≈0.13	1/17≈0.06	10/52≈0.19	11/69≈0.16
65mm, True	0/17=0	8/52≈0.15	8/69≈0.12	2/17≈0.12	8/52≈0.15	10/69≈0.14
65mm, False	0/17=0	8/52≈0.15	8/69≈0.12	2/17≈0.12	8/52≈0.15	10/69≈0.14

Table 7

Misclassification rates based on classification using V_1', V_2', V_3' and V_4'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	0/17=0	3/52≈0.06	3/69≈0.04	0/17=0	4/52≈0.08	4/69≈0.06
15mm, False	0/17=0	3/52≈0.06	3/69≈0.04	0/17=0	4/52≈0.08	4/69≈0.06
65mm, True	0/17=0	5/52≈0.10	5/69≈0.07	1/17≈0.06	6/52≈0.12	7/69≈0.10
65mm, False	0/17=0	5/52≈0.10	5/69≈0.07	1/17≈0.06	6/52≈0.12	7/69≈0.10

Table 8

Misclassification rates based on classification using V_2', V_3' and V_4'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	0/17=0	3/52≈0.06	3/69≈0.04	0/17=0	3/52≈0.06	3/69≈0.04
15mm, False	0/17=0	3/52≈0.06	3/69≈0.04	0/17=0	3/52≈0.06	3/69≈0.04
65mm, True	0/17=0	4/52≈0.08	4/69≈0.06	1/17≈0.06	5/52≈0.10	6/69≈0.09
65mm, False	0/17=0	4/52≈0.08	4/69≈0.06	1/17≈0.06	5/52≈0.10	6/69≈0.09

Table 9

Misclassification rates based on classification using V_2', and V_3'.

Depth, Modification	In-Sample			Out-Sample		
	FPR	FNR	TFR	FPR	FNR	TFR
15mm, True	0/17=0	4/52≈0.08	4/69≈0.06	0/17=0	6/52≈0.12	6/69≈0.09
15mm, False	0/17=0	4/52≈0.08	4/69≈0.06	0/17=0	6/52≈0.12	6/69≈0.09
65mm, True	2/17≈0.12	8/52≈0.15	10/69≈0.14	2/17≈0.12	9/52≈0.17	11/69≈0.16
65mm, False	2/17≈0.12	8/52≈0.15	10/69≈0.14	2/17≈0.12	9/52≈0.17	11/69≈0.16