Evaluation of quantitative methods for the interstitial lung disease extent assessment in MDCT

Poster No.: P-0011
Congress: ESTI 2014
Type: Scientific Poster
Authors: N. Papapanagiotou, A. Kazantzi, P. Korfiatis, S. Skiadopoulos, L. Costaridou, D. Daoussis, C. Kalogeropoulou; Patras/GR
Keywords: Lung, Thorax, CAD, CT-Quantitative, Computer Applications-Detection, diagnosis, Physics, Tissue characterisation
DOI: 10.1594/esti2014/P-0011

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.
As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method ist strictly prohibited.
You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.
Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.
www.myESR.org
Purpose

Interstitial lung diseases are a heterogeneous group of disorders and MDCT is the modality of choice for determining their extent and for predicting the clinical outcomes, as the scoring of fibrosis correlates well with mortality rate. We evaluate four different methods for the assessment of interstitial lung disease extent.
Materials and Methods

Thirty patients with diagnosis of collagen vascular disease, meeting the criteria of the American College of Rheumatology, were referred to our institution from 2007-2009 for chest CT scan. The case sample analyzed consisted of 30 patients (22 females and 8 males) with mean age of 56.4 years and mean disease duration 7.2 years. Fifteen patients were diagnosed with scleroderma, 10 patients with rheumatoid arthritis, 2 with systemic lupus erythematosus, 1 with Sjogren syndrome and 2 mixed connective tissue disease. Informed consent was obtained from all subjects participating in this study. All patients were scanned with a 16-row multidetector CT scanner (GE Lightspeed 16, General Electric Medical Systems, Milwaukee, Wisconsin, USA) at 120 kVp, rotation time of 0.5s, automatic modulation of mA, collimation thickness of 16x0.625 mm and slice thickness of 1.25 mm, using a protocol obtaining volumetric 3D data at full inspiration, in supine position. Each scan volume comprised of approximately 200-250 slices per patient. The mean volume CT dose index and the mean dose-length product were 11.5 mGy and 270.9 mGy·cm, respectively. Assuming 0.017 mSv/mGy·cm for a standard chest CT examination, the effective radiation dose for the volumetric chest CT protocol used was 4.6 mSv, complying with European Working Group for Guidelines on Quality Criteria in CT.

STEP 1: SEMI-QUANTITATIVE VISUAL SCORING: Semi-quantitative scoring of total, ground glass and reticular pattern extents was visually performed by the radiologist (RAD1) on 150 axial slices, resulting from the 30 patients’ CT scans. Radiologist estimated percentages of disease extent with an 1% step for the first 10% of disease extent and a 5% step consequently. The disease extent was assessed at five representative levels, proposed by Desai et al: origin of great vessels (i), carina (ii), pulmonary venous confluence (iii), between levels (iii) and (v) and 1cm above the right hemidiaphragm (Fig.1). For each level, RAD1 provided three quantitative values for total, reticular and ground glass disease extent respectively. The time required for each estimation was approximately 3 minutes. The monitor and viewing conditions were identical to those provided for the pixel-based reference and sample data were evaluated in random order.

STEP 2: PIXEL - BASED VISUAL QUANTIFICATION OF ILD: Disease extent was assessed by RAD1 at the same five representative lung anatomic levels for each patient, that have been used for semi-quantitative scoring purposes up to now in a blinded manner, after having received a short training. The time required for disease extent estimation on each axial slice was approximately 20 minutes.
For the free-hand delineation of the affected lung parenchyma, a digital matrix (Wacom Intuos 3 Tokyo, Japan) with an active surface of 305×305 mm with 5.080 dpi and precision ±0.25 mm was utilized. Images were displayed on a high-resolution (1536x2048) gray scale diagnostic LCD monitor (Barco, Coronis3MP, Belgium).

The GUI supports editing by enabling manual delineation of ILD image segments, displayed as color overlays on the original 3D CT data. Editing is performed on an axial slice basis in the upper middle window of the GUI, shown (Fig.2), while the upper left window displays the current instance of editing.

A free-hand drawn segment on the axial plane is automatically and consistently propagated to the other two planes (i.e. sagittal and coronal), depicted in the left lower and middle lower windows of the GUI, respectively.

After initial delineation, radiologist in-trainee reviewed segments and performed corrections. Following segment editing, total, ground glass and reticular disease extents were provided by the GUI as percentages of the lung parenchyma area, according to eq. [1]-[3] (Table1):

**STEP 3: DISEASE EXTENT ASSESSMENT WITH HISTOGRAM THRESHOLDING:**
The first step of this process refers to vessel segmentation. The user selects values manually by scrolling a bar, according to visual criteria, at a random slice of the CT dataset and the system automatically implements thresholding to entire 3D data (Fig.3).

The graphical user interface allows the constant observation of the image histogram and the changes applied by the user. After vessel extraction, a new 3D dataset is created.

Consequently, the user applies a new histogram thresholding to the new dataset in order to discriminate normal from abnormal parenchyma (Fig.4). This process is implemented in the same fashion as vessel extraction at random axial CT slices and the system automatically applies the new histogram values to all 3D data. The GUI allows also in this case the animation of the changes implemented using different color overlays for normal and abnormal lung. Once user has decided about the minimum and maximum intensity values of thresholding and the system has applied them to all data provided, a third step takes place. In the last phase of the procedure, user needs to decide about the pattern of lesion, i.e. whether the abnormal lung parenchyma represents ground glass or reticular pattern of disease (Fig.5). The third and last application of thresholding technique has the same features with the previous steps. After this final estimation, the system calculates the percentages of total, ground glass and reticular extent corresponding to the three times applied threshold values.

**STEP 4: DISEASE EXTEND ASSESSMENT BY A PROTOTYPE CAD QUANTIFICATION TOOL:** A CAD prototype tool based on a recently proposed ILD algorithm by Korfiatis et al was utilized. The algorithm was developed in the Department of Medical Physics of the School of Medicine of Patras University and is based on voxel classification of lung parenchyma volume, into normal and existing ILD voxel patterns,
employing a k-nearest neighbor classifier and 3D co-occurrence texture features, taking advantage of lung field and vessel tree segmentation.

Although the CAD tool produces 3D image output, the current evaluation was performed on 2D axial slice-basis, in order to be comparable to pixel-based reference of disease extent assessment provided by the radiologist. Specifically, the CAD tool is employed to quantify total, ground glass and reticular disease extent on 150 axial slices, originating from the same dataset of 30 patients. An example of CAD output for mixed pattern disease is shown in Fig.6.

To ensure an acceptable reference for our study, in the absence of a "gold standard", two radiologists in consensus with 10 and 20 years of experience in chest CT (A.K. and C.K, respectively), aware of patient history, provided in consensus the independent visual pixel-based reference standard (Rcons) for the CT data of the 30 patient scans (150 slices). To achieve consensus, the two radiologists discussed and agreed on pattern type and extent per slice prior to delineating segments. After initial delineation, both radiologists reviewed segments and performed corrections.

Methods have been compared pairwise by means of Bland-Altman analysis, utilized in order to assess by inspection the degree of agreement for varying disease extent. In this analysis, the differences were plotted against average values for each pair of disease extent estimation. Additionally, the Intraclass Correlation Coefficient index has been calculated for all pairs compared. The degree of agreement was characterized as almost perfect (ICC = 0.81-1.00), substantial (ICC = 0.61-0.81), moderate (ICC = 0.41-0.61), weak (ICC = 0.21-0.41) or very weak to negligible (ICC = 0.00-0.21).

Inter-observer variation between RAD1 and RADcons was also studied for visual scoring assessment (a) and visual pixel-based scoring evaluation (b) of ILD disease extent. In order to investigate the degree of agreement between radiologists, Bland-Altman and reliability analysis was also implemented.

Statistical analysis was performed using the IBM SPSS Statistics software package (SPSS Release 20.0, SPSS Inc., Chicago, IL, USA).
Fig. 1: Anatomical levels of disease extent assessment (Desai et al. 2004)

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Fig. 2: Free-hand segmentation process. First step: The user outlines the abnormal parenchyma (left). Second step: User decides about the type of the pattern - reticular (blue overlay)

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

Table 1: Equations for the calculation of ground glass, reticular and total disease extent using pixel-based scoring method

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
**Fig. 3:** First step: Vessel extraction. Green overlays correspond to the extracted vessel tree. Note that system has included areas of disease (reticular) in the extracted data.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

**Fig. 4:** Second step: Segmentation of abnormal parenchyma. Red overlays correspond to areas of total disease.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Fig. 5: Third step: Classification of abnormal parenchyma to ground glass and reticular patterns (purple overlay corresponds to reticular and green to ground glass pattern).

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

Fig. 6: Mixed pattern -Ground glass and reticular. The left picture corresponds to original data and the second to pixel based disease extent estimation by the CAD tool in axial plane. Areas of affected lung parenchyma are shown with green overlays for ground glass and blue overlays for reticular.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Results

In Fig.7,8 the performances of all applied methods for the disease extent assessment are visually presented for two different cases.

The comparison of the two methods based on visual observation, semi-quantitative visual scoring (SEMIQ.RAD1) and visual pixel based scoring (RAD1) did not manifest a high degree of agreement (Fig.9).

The inter-observer agreement seems to be higher when radiologists quantify the interstitial disease extent using pixel based methods than visual scoring (Fig10,11,12).

Additionally, the two computerized systems that have been implemented in our study, CAD and histogram thresholding, have shown substantial agreement (Fig.13).

The visual pixel-based - CAD pair has demonstrated the higher degree of agreement among all comparisons (Fig.14).

None of the pairwise comparisons exhibited a high degree of agreement concerning ground glass extent estimation.
Fig. 7: Case 1: Reticular pattern. Visual presentation of performances of methods applied for ILD quantification.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

Fig. 8: Case 2: Mixed pattern. Visual presentation of performances of methods applied for ILD quantification.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Fig. 9: Bland-Altman plots corresponding to the differences in total, reticular and ground glass disease extent assessment between semi-quantitative visual scoring and pixel-based visual quantification by radiologist-in-trainee are depicted. Almost perfect agreement of the two methods concerning total disease assessment (ICC: 0.817 - CI: 0.652-0.893) and a substantial agreement for reticular pattern disease extent (ICC: 0.731 - CI: 0.647-0.798) is indicated. For the assessment of ground glass extent, ICC equals 0.200 (CI: 0.038-0.352), indicating a weak agreement between the two methods.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

Fig. 10: Figure 10 presents Bland-Altman plots corresponding to the differences in ground glass disease extent assessment between RAD1 and RADcons using semi-quantitative (left) and pixel-based scoring method (right). A value of ICC 0.320 (CI: 0.171-0.455) indicates a weak agreement between radiologists concerning the
assessment of ground glass extent for the visual scoring method, whereas a value of ICC 0.596 (CI: 0.379-0.732) indicates a moderate agreement between the radiologists when they assess the disease extent pixelwise.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

**Fig. 11:** Bland - Altman plots corresponding to the differences in reticular disease extent assessment between RAD1 and RADcons using semi-quantitative (left) and pixel-based scoring method (right). ICC values show substantial agreement between the radiologists concerning reticular pattern (ICC 0.724- CI 0.574-0.816) for visual scoring and almost perfect agreement for the pixel-based method (ICC: 0.816- CI: 0.597-0.901)

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR

**Fig. 12:** Bland - Altman plots corresponding to the differences in reticular disease extent assessment between RAD1 and RADcons using semi-quantitative (left) and pixel-based scoring method (right). ICC values show substantial agreement between the radiologists concerning reticular pattern (ICC 0.724- CI 0.574-0.816) for visual scoring and almost perfect agreement for the pixel-based method (ICC: 0.816- CI: 0.597-0.901)

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Fig. 13: Bland-Altman plots corresponding to the differences in total, reticular and ground glass disease extent assessment between CAD system and Histogram thresholding technique are depicted. Almost perfect agreement of the two methods concerning total disease (ICC: 0.833 - CI: 0.775-0.877) and substantial agreement for reticular pattern (ICC: 0.771 - CI: 0.697-0.828) are indicated by the ICC value. For the assessment of ground glass extent, ICC is 0.212 (CI: 0.059-0.357) indicating a weak agreement between the two methods.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Fig. 14: Bland-Altman plots corresponding to the differences in total, reticular and ground glass disease extent assessment between CAD system and visual pixel-based quantification by radiologist-in-trainee are depicted. Almost perfect agreement of the two methods concerning both total disease (ICC 0.861 - CI 0.813-0.897) and reticular pattern (ICC 0.827 - CI 0.707-0.891) is indicated by the ICC value. For the assessment of ground glass extent, ICC is 0.296 (CI: 0.134-0.440) showing a weak agreement between the two methods.

© Department of Radiology, Medical School University of Patras, Patras University Hospital - Patras/GR
Conclusions

In this study, four different methods for the evaluation of interstitial lung disease extent were applied to the same clinical data set and their performances were investigated in pairwise comparison in terms of agreement. During this study, a visual hand-segmentation technique (pixel-based) was assumed to be among the most accurate and precise one, and therefore was regarded as the "ground truth" method and used to "anchor" the rest of the techniques. The two visual scoring techniques applied by the radiologist-in-trainee, i.e. semiquantitative visual scoring and visual pixel-wise method did not manifest high degree of agreement, indicating that each of these techniques capture different aspects of information. Consequently, visual semi-quantitative scoring is not proposed for ILD extent quantification.

The evaluated chest CT CAD system is a reliable and reproducible disease extent quantification tool that could be used in extent estimation. These algorithms do not by any means replace radiologists in medical practice and especially in evaluating lung disease. The final decision is made by the radiologist, not the computer. Ultimately, the goal of CAD is to reduce search and interpretation errors, and reduce variation between and within observers.

Our study is characterized by various limitations; the radiologist that performed the evaluation of the disease extent is in training and therefore the experience level may have affected the results, although he has received training in evaluating lung diseases and has been familiarized with the technical equipment used. Limitations include also technical issues that refer to each method applied.
References


