Extension of pulmonary fibrosing diseases: a comparison of quantification scoring systems.

Poster No.: E-0094
Congress: ESTI 2012
Type: Scientific Exhibit
Authors: M. Silva¹, D. Colombi¹, L. Buttarelli¹, A. Ranalli¹, V. Seletti¹, M. Mereu², C. Rossi¹, N. Sverzellati¹; ¹Parma/IT, ²Chieti/IT
Keywords: CT-High Resolution, CT, Respiratory system, Lung, Treatment effects, Observer performance, Connective tissue disorders
DOI: 10.1594/esti2012/E-0094

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
Objectives

The assessment of pulmonary fibrosis extension by Computed Tomography (CT) is a key index which has strong prognostic and therapeutic value. This method is widely used for the evaluation of drug efficacy. Several studies compared quantitative method with visual scoring CT method [1], nevertheless a standardization of visual scoring CT is lacking.

This study has two targets:

- to compare interobserver variability among the main visual scoring systems for the quantification of global extension of pulmonary interstitial disease and a new method of scoring based on HRCT coronal images
- to define which of these scores better correlate with functional deficit

Radiology 246:935-940.
Materials and Methods

Study population

39 consecutive patients affected by ILD (25 males and 14 females; mean age 68 years; range 34-91 years): 20 cases of Usual Interstitial Pneumonia (UIP), 14 cases of Non Specific Interstitial Pneumonia (NSIP), 4 cases of smoking-related ILD and 1 case of Hypersensitivity Pneumonitis (HP). Every patient underwent both HRCT and pulmonary function test (within 1 month of HRCT) for the evaluation of Diffusing Capacity for Carbon Monoxide (DL$_{co}$), used as reference for accuracy of visual scoring systems.

CT protocol

The HRCTs were performed either by two multi-detector row CT scanners: a 6-slice MDCT scanner (Somatom Emotion 6, Siemens Medical Solutions, Forchheim, Germany) or a 64-slice MDCT scanner (Sensation 64 Cardiac, Siemens Medical Solutions, Forchheim, Germany). One mm collimation HRCT images were volumetrically obtained at 1-1.5 mm intervals reconstruction, according to the departmental protocol at the time of scanning. CT images were reconstructed at 1 mm-thick sections. The scans were acquired with the patient in supine position, breath holding at full inspiration and were reconstructed by using a sharp kernel (B70f).

All images were viewed at window settings optimized for assessment of lung parenchyma (width: 1600 HU; level: -600 HU).

Visual scoring methods

The CT scans were independently reviewed by two trainee in radiology who had no experience of quantifying the ILD extent on HRCT. Before the study, the two inexperienced radiologists underwent a training session conducted by one experienced radiologist (7 years experience in quantifying ILD extent).

Visual scoring method 1

HRCTs were scored at five pre-established levels:

1. aortic arch (origin of great vessels)
2. carina
3. pulmonary venous confluence
4. between levels 3 and 5

5. 1 cm above the right hemidiaphragm.

The total extent of ILD was estimated for each lung to the nearest 5% in each of the five sections, with global extent of disease on HRCT computed as the mean of the right and left scores (Fig. 1) [2-4].

Visual scoring method 2

Lungs were divided into three zones in each side.

Upper zone: part of the lung above the level of the tracheal carina.

Middle zone: portion of the lung between the upper and lower zones.

Lower zone: part of the lung below the level of the inferior pulmonary vein.

The extent of involvement of the findings was evaluated visually and independently for each lung zone. A score was assigned on the basis of the percentage of lung parenchyma that showed evidence of an abnormality and was estimated to the nearest 10% of parenchymal involvement. Overall percentage of involvement was calculated by averaging the scores of the six lung zones (Fig. 2) [5,6].

Visual scoring method 3

Each HRCT was reformatted for MPR coronal reconstructions. Observers reviewed the coronal HRCT images at six pre-selected levels:

1. between the sternum and the anterior margin of the ascending aorta

2. center of the ascending aorta

3. tracheal bifurcation

4. between the tracheal bifurcation and the anterior wall of the thoracic vertebral bodies

5. anterior wall of the thoracic vertebral bodies

6. confluence of the thoracic vertebral body laminae.

The total extent of ILD was estimated for each lung to the nearest 5% in each of the six sections, with global extent of disease on HRCT computed as the mean of the right and left scores (Fig. 3).
Statistical analysis

We compared the levels of interobserver variation calculated for each visual scoring method by computation of the mean difference and 95% limits of agreement (LoA) using the Bland-Altman test.

To establish which methods improved the accuracy of the individual observers evaluations, correlations between DLco and either averaged or individual observers scores obtained were compared by using Spearman's rank correlation co-efficient.

P values <0.05 were taken to be statistically significant.


Fig. 1: visual scoring method based on five pre-established axial HRCT levels.

Fig. 2: upper, middle and lower lung zones established for visual scoring method based on "scrolling of HRTC images".

© Operative unit of radiological science, Parma Hospital, University of Studies of Parma - Parma/IT
**Fig. 3:** six levels established for visual scoring method based on MPR coronal reformatted images.

© Operative unit of radiological science, Parma Hospital, University of Studies of Parma - Parma/IT
Results

The 95% LoA of ILD extent between observer 1 and observer 2 for visual scoring method 1, 2, and 3 were respectively -35.2% to 20% (Fig. 4), -38.8% to 17.4% (Fig. 5), -41.3% to 24.9% (Fig. 6).

By the use of the visual scoring method 1, a fair ($r = -0.3$, $p = 0.05$) correlation between averaged observers scores and $DL_{co}$ was obtained. The lowest correlation was reported by using visual scoring method 2 ($r = -0.2$, $p = 0.2$). Despite the visual scoring method 3 showed the greatest interobserver variability, its use improved correlations between ILD extent on HRCT and $DL_{co}$ for each observer and averaged evaluations ($r = -0.4$, $p = 0.03$).
Fig. 4: 95% LoA of ILD extent between observer 1 and observer 2 for visual scoring method 1.

© Operative unit of radiological science, Parma Hospital, University of Studies of Parma - Parma/IT
Fig. 5: 95% LoA of ILD extent between observer 1 and observer 2 for visual scoring method 2.

© Operative unit of radiological science, Parma Hospital, University of Studies of Parma - Parma/IT
**Fig. 6:** 95% LoA of ILD extent between observer 1 and observer 2 for visual scoring method 3.

© Operative unit of radiological science, Parma Hospital, University of Studies of Parma - Parma/IT
Conclusions

The visual scoring method based on coronal images despite fair interobserver variability, improved correlations between ILD extent on HRCT and DL_{co}.

This study needs further assessment by experienced thoracic radiologists in a large number of patients in order to stratify the evaluation for each fibrosing lung pathology.