

High Resolution Computed Tomography Scoring Systems for Evaluating Interstitial Lung Disease in Systemic Sclerosis Patients

Deborah Assayag¹, Sagi Kaduri², Marie Hudson³, Andrew Hirsch¹ and Murray Baron^{3*}

¹Department of Pulmonary Medicine, Jewish General Hospital, McGill University, Montreal, Canada

²Department of Radiology, Jewish General Hospital, McGill University, Montreal, Canada

³Department of Rheumatology, Jewish General Hospital, McGill University, Montreal, Canada

Abstract

Interstitial lung disease commonly develops in patients with Systemic sclerosis (SSc). High resolution computed tomography (HRCT) has become the gold standard for detection and evaluation of lung involvement in SSc. Several HRCT scoring methods have been described and used to characterize and quantify the disease. This article reviews the different scoring systems and how they have been validated clinically and applied to prognosticate patients, assess disease progression and evaluate response to treatment.

Keywords: Interstitial lung disease; Scleroderma; Computed tomography; Scoring methods

Introduction

Systemic sclerosis (SSc), also known as scleroderma, is a connective tissue disease characterized by vascular disease, immunologic abnormalities and fibrosis. The hallmark clinical feature consists of thickening of the skin. Internal organs such as the lungs, gastrointestinal tract and kidneys are also frequently affected. Pulmonary fibrosis is now a significant cause of morbidity and is the leading cause of mortality in patients with SSc [1].

High resolution computed tomography (HRCT) has become the gold standard for diagnosis of SSc related interstitial lung disease (SSc-ILD) especially for early stage disease. It is a much more sensitive diagnostic test than traditional chest radiograph, especially for detection of early or mild disease [2,3]. The most common radiologic pattern of lung disease in SSc is diffuse parenchymal lung disease characterized by prominent ground-glass opacities and fine interstitial reticular markings with lower lung predominance. The pattern is similar to that found in idiopathic non-specific interstitial pneumonia (NSIP) [4-6]. As the disease progresses, ground-glass opacifications get replaced with coarser interstitial reticulations, traction bronchiectasis and bronchiolectasis. Honeycombing also develops with time [7]. This pattern closely resembles idiopathic pulmonary fibrosis (IPF). Figure 1 illustrates the typical HRCT abnormalities that can be found in SSc-ILD.

Different systems for evaluating SSc-ILD on HRCT have been developed over the past 20 years. Several scoring methods have been used to characterize and quantify the disease, correlate with common clinical parameters, prognosticate patients, assess disease progression and evaluate response to treatment. This article reviews the different published scoring systems and their applications.

Description of Scoring Systems

Comparative scoring

Most of the scoring systems published for SSc-ILD have been developed from methods previously used to assess IPF. One of the first reading methods used to assess SSc-ILD was published by Wells et al. [8] (Table 1). They used a comparative grading system evaluating the extent of one HRCT abnormality (parenchymal disease) relative to another (reticular disease). The goal was to determine how different HRCT abnormalities correlate with lung histology, specifically inflammatory changes and fibrosis, on biopsy. This scoring method was

based on a similar protocol published by Müller et al. [9] correlating findings on HRCT of IPF with histopathology.

An even simpler scoring method was published by Morelli et al. [15]. They divided the lungs into three zones: apices to main carina, main carina to inferior pulmonary venous confluence, and pulmonary veins to diaphragms, representing upper, middle and lower zones respectively. A score of 0 was given if no abnormality was found in a zone, a score of 1 for any abnormality found in SSc-ILD except for honeycombing (ground-glass, reticular markings, bronchiectasis) and

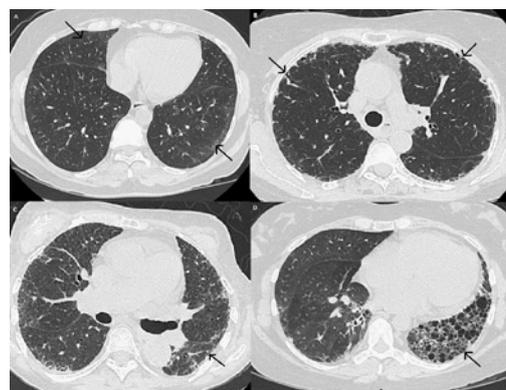


Figure 1: Typical HRCT abnormalities in SSc-ILD. A. Pure ground glass opacities can be seen peripherally, predominantly in the right middle and left lower lobes; B. Subpleural thickened reticular markings are associated with traction bronchiectasis and bronchiolectasis; C. Ground glass opacities are mixed with fine reticular septal thickening diffusely; D. Extensive macrocystic honeycombing has replaced most of the left lower lobe and is associated with significant volume loss.

***Corresponding author:** Dr. Murray Baron, Chief of Rheumatology Department, Jewish General Hospital, Suite A 725, 3755 Cote St Catherine Rd., Montreal, Quebec, Canada, Tel: +1 514 340 8231; Fax: +1 514 340 7906; E-mail: mbaron@jgh.mcgill.ca

Received December 07, 2011; **Accepted** January 24, 2012; **Published** January 28, 2012

Citation: Assayag D, Kaduri S, Hudson M, Hirsch A, Baron M (2012) High Resolution Computed Tomography Scoring Systems for Evaluating Interstitial Lung Disease in Systemic Sclerosis Patients. *Rheumatology* S1:003. doi:10.4172/2161-1149.S1-003

Copyright: © 2012 Assayag D, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

a score of 2 if honeycombing was present. The global score was obtained by adding up these zonal scores. Comparative scoring methods assume that a higher score or higher grade represent greater severity of disease.

Semi-quantitative scoring

Semi-quantitative scoring methods have been developed to provide more precise assessment of quantity and type of ILD abnormalities. One scoring system that has been used in several studies was developed and published by Warrick et al. [16] (Table 2).

This scoring system combines severity and extent of disease. Different abnormalities corresponding to increasingly severe disease are given increasingly high scores. Extent is determined based on the total number of bronchopulmonary segments involved for each abnormality. The greater the number of segments involved, the higher the extent score. These scores are combined to obtain a global score.

Application of this scoring system, however, requires more advanced knowledge of pulmonary anatomy and proficiency in identifying bronchopulmonary segments. This may not be generally useful for clinicians.

Kazerooni et al. [25] published a scoring method to assess HRCT and to correlate with pathology in IPF. Variations of this method have since been used by several groups in patients with SS [7,26-31]. Ooi et al. [29] evaluated correlation between HRCT findings and clinical markers of disease activity. Their scoring system, which is similar to that of Kazerooni et al. [25] is detailed in Table 3.

The extent of each abnormality is estimated for each lobe. A semi-quantitative score is assigned (0-4) based on the approximate percentage of disease for each one. This scoring method makes a

distinction between pure ground-glass opacities and ground-glass mixed with reticular disease. Differentiating between pure ground-glass opacity and mixed reticular disease, and between fibrosis and honeycombing, may permit more precise assessment of the relationship between particular abnormalities and clinical parameters.

Pandey et al. [30] used the same scoring method to assess the relationship between ILD and peak systolic pulmonary arterial pressure (sPAP). They scored any ground-glass disease, reticular abnormalities and honeycombing in each of 5 lobes (scoring lingula with left upper lobe) using the same 0 to 4 semi-quantitative score. They also weighted each score for relative lobar volume using correction factors. Weighting the scores based on relative volume of each lobe may allow for more accurate estimation of global disease. A score of 3 for fibrosis in the lingula may not represent the same total amount of disease as a score of 3 in the left lower lobe, for example.

Goldin et al. [27] published the scoring method that was used to assess and follow-up ILD in the Scleroderma Lung Study population. This was also based on Kazerooni's method. Instead of scoring lobes, however, they scored 3 anatomical zones (upper, middle and lower zones) in each lung (Table 4).

The sum of these grades in all 6 zones make up the global score for each abnormality. In this case, pure ground-glass disease is scored as opposed to ground-glass mixed with reticular disease. Disease progression was assessed subsequently by comparing CT scans in a blinded fashion. Disease in each zone was qualitatively compared from one scan to another and each abnormality was scored as better, same or worse. Semi-quantitative scores were not used to evaluate progression in this case [27].

Abnormality	Grade assigned	Anatomical regions scored
Parenchymal opacification alone	1	Lobes scored independently of each others
Parenchymal opacification > reticular pattern extent	2	
Parenchymal opacification = reticular pattern extent	3	
Reticular pattern extent > parenchymal opacification	4	
Reticular pattern alone	5	
Maximum score = 5 per lobe		
Variations of this method also used by Pignone et al. [2]; Davas et al. [10]; Shahin et al. [11]; Giacomelli et al. [12]; Shah et al. [13]; Gohari Moghadam [14].		

Table 1: Comparative scoring method, Wells et al. [8].

Severity score		Extent score	
Abnormality	Grading	Bronchopulmonary segments – for each abnormality, score by number of segments involved	Grading
Ground-glass opacities	1	1 to 3 segments involved	1 2 3
Irregular pleural margin	2	4 to 9 segments involved	
Septal or subpleural lines	3	>9 segments involved	
Honeycombing	4	N.B. extent of disease measured for each of the abnormalities	
Subpleural cyst	5		
Maximal severity score	15	Maximal extent score	15
Also used by Diot E et al. [17]; Orlandi I et al. [18]; Afeltra et al. [19]; Camiciottoli et al. [20]; Yiannopoulos et al. [21]; Bellia et al. [22]; Savarino et al. [23]; Daoussis D et al. [24].			

Table 2: Semi-quantitative scoring method: Warrick et al. [16].

Abnormality	Grading for each abnormality		Anatomical regions scored
	Percentage disease extent	Score	
Ground-glass opacity alone	0	0	Lobes are scored independently Lingula is considered a separate lobe 6 total lobes
Mixed ground glass and reticular disease	1-25%	1	
Reticular fibrosis alone	26-50%	2	
Honeycombing	51-75%	3	
	>75%	4	
Global score: summation of scores for each abnormality, in all lobes			
Also used by Choi et al. [26]; Mok et al. [28]; Pandey et al. [30]; Tiev et al. [31].			

Table 3: Semi-quantitative scoring : Ooi et al. [29].

Quantitative scoring

Wells et al. [32] proposed a quantitative scoring method, using estimation of extent of disease as a percentage of the anatomical region assessed, which consisted of 5 separate levels. This method had been previously developed in patients with IPF [33]. Details are outlined in Table 5.

Scoring Methods Applied

Figure 2 is an example of a typical HRCT scan of a patient with SSc. This scan can be used as an example to illustrate how these scoring methods can be applied and how they differ.

Whole pulmonary lobes, bronchopulmonary segments or anatomical zones are scored in the comparative scoring system and the semi-quantitative methods described above. All the HRCT slices would be required to adequately assess disease extent in each lobe or zone, however for convenience, four slices will be used here and scored independently. Using Wells' comparative score, this HRCT would get graded as 4, 5, 3 and 4 for images A, B, C and D respectively, due to the predominance of reticular pattern over ground glass opacities. In contrast, using Warrick's score, this scan would receive a severity score of 10 out of 15 because of the presence of ground glass opacities (score of 1), irregular pleural margins (score of 2), septal lines (score of 3) and honeycombing (score of 4). It would be impossible to score extent of disease given that bronchopulmonary segments cannot be identified with only four cuts. When combining the images and applying the

method used by the Scleroderma Lung Study group, a grade of 1 (1-25% involvement) is obtained for ground glass opacity, 2 (26-50%) for fibrosis and 1 for honeycombing (although technically, right and left lungs should be scored separately). Finally, when applying Well's quantitative method, global disease extent is estimated at 25%, with a grade of coarseness of reticulation of 2 because of the microcystic honeycombing.

Validity

The construct validity of a scoring system can be defined as its ability to predict an accepted measure of an underlying construct such as disease severity. The most common outcomes used to assess the validity of the different scoring methods included pulmonary function tests, findings on bronchoalveolar lavage, histopathology and survival. Details of correlation between HRCT scores for each method and clinical parameters can be found in the supplementary table (appendix 1).

Studies using comparative scoring methods found correlation between higher HRCT severity grades and low DLCO, low FVC and low TLC [2,11,14]. When compared to biopsy specimen, mixed ground-glass and reticular thickening (HRCT grade 3) corresponded to more inflammatory changes on histopathology. In contrast, a predominance of reticular disease (HRCT grade 4) corresponded to predominant fibrosis histopathologically [8]. CT discriminated correctly between inflammation and fibrosis in 80% of biopsy specimens. This scoring method suggested a relationship between particular CT abnormalities

Abnormality	Grading per abnormality		Anatomical regions scored
	Percentage disease extent	Score	
Pure ground-glass opacity	0	0	Zone 1: Apex to aortic arch Zone 2: Aortic arch to inferior pulmonary veins Zone 3: Inferior pulmonary veins to diaphragms Right and left lung scored separately
Fibrosis (including thickened reticular markings, bronchiectasis and bronchiolectasis)	1-25%	1	
	26-50%	2	
	51-75%	3	
Honeycombing	>75%	4	

Table 4: Semi-quantitative scoring; Scleroderma Lung Study [27].

Abnormality	Grading	Anatomical regions scored
Global disease extent	Estimated to the nearest 5%	Five levels are assessed : 1) Origin of the great vessels 2) Main carina 3) Pulmonary venous confluence 4) Halfway between levels 3 and 5 5) Immediately above right hemidiaphragm
Extent of reticulation	As proportion of total disease extent	
Percent of ground-glass	As proportion of total disease extent	
Coarseness of reticulation :		
- Normal	0	
- Fine intralobular fibrosis	1	
- Microcystic honeycombing	2	
- Macrocystic honeycombing	3	
Total disease extent corresponds to the mean of the percentage of disease at each level		
Also used by Desai et al. [6]; Hoyles et al. [34]; Goh et al. [35].		

Table 5: Quantitative and semi-quantitative scoring: Wells et al. [32].

Authors	Scoring method	Readers	HRCT feature assessed	Inter-observer agreement (kappa)
Desai et al. [6]	Wells 1997	Chest radiologists	Overall grade	K = 0.74
			Coarseness of fibrosis	K = 0.88
Goh et al. [35]	Wells 1997	Clinicians	Disease extent	K = 0.64
		Trainees	Disease extent	K = 0.41
Camiciottoli et al. [20]	Warrick 1991	Radiologists	Disease extent	K = 0.69
Goldin et al. [27]	SLS	Chest radiologists	Presence or absence of ground-glass	K = 0.72
			Presence or absence of fibrosis	K = 0.61
			Presence or absence of honeycombing	K = 0.39
Goldin et al. [39]	SLS	Chest radiologists	Change over time in ground-glass	K = 0.36
			Change over time in fibrosis	K = 0.51
			Change over time in honeycombing	K = 0.16

Table 6: Inter-observer agreement.

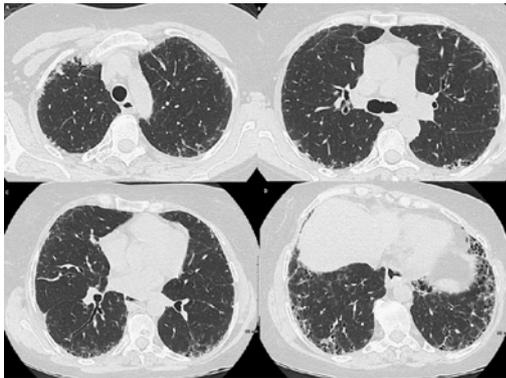


Figure 2: HRCT scan of a patient with SSc-ILD. This figure represents four cuts of an HRCT for a patient with SSc: A. Level of the aortic arch; B. Main carina; C. Pulmonary venous confluence; D. Origin of the diaphragm. The disease is more severe in the lower lung zones, with more pronounced reticular fibrosis and peripheral honeycombing.

and histopathological disease. Wells et al. [36] also showed a correlation between extent of disease (in both IPF and SSc) as scored by CT, and neutrophil levels within each lobe.

Using semi-quantitative scoring systems, global score on HRCT has been shown consistently to be significantly and inversely correlated with TLC, DLCO and FVC% predicted [17,19-20,22,29]. A global score of 7 using the score described by Warrick et al. [16] was predictive of PFT abnormalities with a positive predictive value of 0.82, a sensitivity of 0.6 and specificity of 0.83 [17]. Global scores have also been correlated with elevated pulmonary arterial pressures, poor exercise tolerance and gastroesophageal reflux [23,29-30]. When assessing each abnormality separately, fibrosis scores were again correlated with low FVC, DLCO and TLC. One study reported correlation between ground-glass disease and pulmonary function test abnormalities [37]. In contrast, in the Scleroderma Lung Study, pure ground-glass disease did not correlate well with pulmonary function tests, but was associated with evidence of alveolitis on bronchoalveolar lavage, albeit weakly [27]. These results suggest that pure ground-glass on CT scan may be reversible and may represent inflammation, but that mixed ground-glass corresponds to microscopic fibrosis, as it is equally correlated with PFT abnormalities as reticular disease.

Extent of disease assessed using the quantitative scoring method by Wells et al. [32] also correlated inversely with DLCO, FVC and TLC. This scoring method was applied to a cohort of 215 subjects with SSc to help predict prognosis [35]. Global extent of disease greater than 20% was associated with significantly increased mortality (hazard ratio (HR) of death 2.48) and worse progression-free survival. When combined with FVC of less than 70% predicted, HR was 3.46, and this was highly statistically significant.

In terms of assessing disease progression and response to treatment, the Scleroderma Lung Study group found that patients with more severe fibrosis at baseline (or higher fibrosis scores) had a greater mean of decline in FVC compared to groups with no or moderate fibrosis [38]. The pulmonary fibrosis score progressed in a greater proportion of patients on placebo compared to those receiving cyclophosphamide, but there was no difference in ground-glass or honeycombing [39]. Severity of fibrosis, but not ground glass, at baseline correlated with better response to treatment with cyclophosphamide [40]. The results of this study, again, put into doubt the interpretation that ground glass changes represent those pathologic abnormalities which perhaps

lead to fibrosis and which may be reversible. Launay et al. [7] using a scoring system similar to the one used in the Scleroderma Lung Study showed that over time, patients had higher mean extent scores and higher proportion of fibrosis, which was associated with lower DLCO and oxygen saturation on follow-up, but not with FVC decline. Kim et al. [41] used the quantitative scoring method described by Wells and found that change in HRCT disease was weakly associated with decrease in DLCO but not with FVC. In contrast, the scoring system used by Warrick et al. [16] when applied to a small cohort (n = 13) of patients receiving cyclophosphamide, did not show progression of HRCT scores despite worsening PFTs [21]. This may have been due to the small sample studied.

Reliability

There is significant variability in the different scoring systems in the literature, although many variations of similar methods have been described. Several systems have been found to have moderate to good inter-rater reliability (Table 6) [6,20,27,35,39]. Intra-rater test-retest reliability has also been shown in one scoring system to be good [20]. Assessing change in disease severity over time however has been shown to be much less reliable.

Discussion

Several quite different scoring systems for HRCT abnormalities in SSc have been used. We have classified these scoring methods as comparative, semi-quantitative or quantitative. Comparative methods assess one abnormality (ground-glass) relative to another (fibrosis) to determine disease severity. Semi-quantitative methods are characterized by estimating disease extent and assigning a grade, with higher grades corresponding to a higher percentage of disease, whereas the actual percentage of lung involvement is used in quantitative methods.

Most of the systems assessed several abnormalities but generally looked at ground glass, reticular interstitial thickening, bronchiectasis and bronchiolectasis and honeycombing. The major differences between systems relate to: 1) the anatomical regions assessed: either entire lung zones [27], whole lobes [29-31] or a defined number of HRCT slices [32], and 2) the use of a severity score roughly related to a broad range of percentage involvement of a segment [27,29] versus a severity score directly related to the actual percentage involvement (to the nearest 5%) [32]. An advantage of scoring lobes instead of zones is the ability to correlate lobar scores with findings on histopathology or BAL. This can be a useful tool in the study of the pathogenesis of disease. Assessing disease extent using a limited number of thin-section CT slices [32] may result in decreased accuracy compared with scoring entire lobes or zones. However, Kazerooni et al. [25] showed similar accuracy when using a 3-slice method compared to a full CT in patients with IPF. This may be due to relative gradual change of disease severity over the lung fields. Whether this applies to SSc-ILD remains to be determined.

Most of the systems have demonstrated adequate reliability and validity so deciding which system to use may be based on other factors. Many have been shown to have good inter-rater reliability, although chest radiologists do better than clinicians and trainees and the agreement between observers for change over time is less than that for the score at one time. The implications of these findings are that scoring requires some expertise and that using HRCT scores for longitudinal multi-center trials may be limited by considerable error. Simpler scoring systems that have little inter-observer variability may be easier to use in a clinical setting, especially by physicians who are not

radiologists. Assigning a percentage of disease is easier than identifying disease extent by counting broncho-pulmonary segments, which requires more advanced knowledge of pulmonary anatomy. Similarly, scoring lobes can be more complex for the clinician than scoring upper, middle and lower lung zones. However, simpler methods with less variability will be less sensitive to small changes over time and will have less discriminant validity.

Semi-quantitative scoring systems, using grades that correspond to ranges in percentage, automatically causes approximation of the regional extent of disease. Therefore there is inherent imprecision to these methods. Six percent involvement of fibrosis would be graded the same way as 24% involvement of a region, but these are clearly different and may have different clinical implications. Semi-quantitative systems, compared to quantitative scores using actual percentage of disease may also be less sensitive in detecting small changes. This may impair the ability of these methods to assess disease progression over time. Variations that may be clinically or physiologically significant, but still remain within the same scoring range, will be unaccounted for. This is even more problematic for a global score that groups different abnormalities. Progression of one abnormality, for example fibrosis or honeycombing, but with simultaneous regression of another, for example ground-glass opacity, may be reflected as a lower global score but in fact may represent worsening overall lung physiology. If we knew that ILD in SSc clearly progressed through a series of stereotypical changes, then comparative systems may provide a different method of assessing progressive disease. For example, a progression from ground glass to honeycombing to subpleural cysts may equate with disease progression and increased severity [8].

Scoring systems quantifying different abnormalities separately allow for measuring contribution of each abnormality, which cannot be done by scoring global percentage of disease only. This is important in view of recent evidence that ground glass opacities do not necessarily correspond to reversible disease [13]. Although pure ground glass may represent inflammatory changes, it may also represent very fine reticular fibrosis beyond the resolution of current HRCT technique. Determining the correlation of each abnormality with clinical parameters may provide more insight into the relationship between specific CT findings and underlying disease pathophysiology and disease progression. The relationship of each abnormality with prognosis, histology, or clinical parameters can then be assessed empirically in any study, without any a priori assumption about the meaning of the abnormality.

Conclusion

A good scoring system should be reliable, valid, sensitive to change and have prognostic implications. Scoring systems for research purposes should adhere closely to these principals and thus may require a high level of complexity. In a clinical setting, the issue of practicality, ease of use and applicability by most physicians who are following patients suffering from SSc is important.

Of the scoring systems discussed, the method described by the investigators of the Scleroderma Lung Study most closely approximates the requirements for a research scoring system. One of the limitations of this method might be a lack of sensitivity given the semi-quantitative nature of the method, especially when assessing for small changes over time.

In the clinical setting, the quantitative method described by Wells and colleagues may be more useful given that it is easier to apply and correlates with prognosis. Further studies however are required to

evaluate how sensitive it is to detect change over time, and therefore how it can be used in assessing response to therapy.

Judging by the number of published CT scoring methods, there is little consensus among research groups regarding the optimal way to assess ILD in SSc. A method that would be uniformly adopted for research would allow for pooling of studies and could dramatically improve the strength of the evidence available. A scoring method, perhaps not the same as the best research method, that can be easily applied in clinical settings, could be integrated as part of a routine evaluation of patients with SSc.

References

1. Steen VD, Medsger TA (2007) Changes in causes of death in systemic sclerosis, 1972-2002. *Ann Rheum Dis* 66: 940-944.
2. Pignone A, Matucci-Cerinic M, Lombardi A, Fedi R, Fargnoli R, et al. (1992) High resolution computed tomography in systemic sclerosis. Real diagnostic utilities in the assessment of pulmonary involvement and comparison with other modalities of lung investigation. *Clin Rheumatol* 11: 465-472.
3. Sergiacomi G, De Nardo D, Capria A, Manenti G, Fabiano S, et al. (2004) Non-invasive diagnostic and functional evaluation of cardiac and pulmonary involvement in systemic sclerosis. *In Vivo* 18: 229-235.
4. Chan TY, Hansell DM, Rubens MB, du Bois RM, Wells AU (1997) Cryptogenic fibrosing alveolitis and the fibrosing alveolitis of systemic sclerosis: morphological differences on computed tomographic scans. *Thorax* 52: 265-270.
5. Wells AU, Steen V, Valentini G (2009) Pulmonary complications: one of the most challenging complications of systemic sclerosis. *Rheumatology (Oxford)* 48: iii40-iii44.
6. Desai SR, Veeraraghavan S, Hansell DM, Nikolakopoulou A, Goh NS, et al. (2004) CT features of lung disease in patients with systemic sclerosis: comparison with idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia. *Radiology* 232: 560-567.
7. Launay D, Remy-Jardin M, Michon-Pasturel U, Mastora I, Hachulla E, et al. (2006) High resolution computed tomography in fibrosing alveolitis associated with systemic sclerosis. *J Rheumatol* 33: 1789-1801.
8. Wells AU, Hansell DM, Corrin B, Harrison NK, Goldstraw P, et al. (1992) High resolution computed tomography as a predictor of lung histology in systemic sclerosis. *Thorax* 47: 738-742.
9. Müller NL, Staples CA, Miller RR, Vedral S, Thurlbeck WM, et al. (1987) Disease activity in idiopathic pulmonary fibrosis: CT and pathologic correlation. *Radiology* 165: 731-734.
10. Davas EM, Peppas C, Maragou M, Alvanou E, Hondros D, et al. (1999) Intravenous cyclophosphamide pulse therapy for the treatment of lung disease associated with scleroderma. *Clin Rheumatol* 18: 455-461.
11. Shahin AA, Sabri YY, Mostafa HA, Sabry EY, Hamid MA, et al. (2001) Pulmonary function tests, high-resolution computerized tomography, alpha-1-antitrypsin measurement, and early detection of pulmonary involvement in patients with systemic sclerosis. *Rheumatol Int* 20: 95-100.
12. Giacomelli R, Valentini G, Salsano F, Cipriani P, Sambo P, et al. (2002) Cyclophosphamide pulse regimen in the treatment of alveolitis in systemic sclerosis. *J Rheumatol* 29: 731-736.
13. Shah RM, Jimenez S, Wechsler R (2007) Significance of ground-glass opacity on HRCT in long-term follow-up of patients with systemic sclerosis. *J Thorac Imaging* 22: 120-124.
14. Gohari Moghadam K, Gharibdoost F, Parastandehchehr G, Salehian P (2011) Assessments of pulmonary involvement in patients with systemic sclerosis. *Arch Iran Med* 14: 22-26.
15. Morelli S, Ferri C, Poletti E, Bellini C, Gualdi GF, et al. (1995) Plasma endothelin-1 levels, pulmonary hypertension, and lung fibrosis in patients with systemic sclerosis. *Am J Med* 99: 255-260.
16. Warrick JH, Bhalla M, Schabel SI, Silver RM (1991) High resolution computed tomography in early scleroderma lung disease. *J Rheumatol* 18: 1520-1528.
17. Diot E, Boissinot E, Asquier E, Guilmet JL, Lemarié E, et al. (1998) Relationship between abnormalities on high-resolution CT and pulmonary function in systemic sclerosis. *Chest* 114: 1623-1629.

18. Orlandi I, Camiciottoli G, Diciotti S, Bartolucci M, Cavigli E, et al. (2006) Thin-section and low-dose volumetric computed tomographic densitometry of the lung in systemic sclerosis. *J Comput Assist Tomogr* 30: 823-827.
19. Afeltra A, Zennaro D, Garzia P, Gigante A, Vadacca M, et al. (2006) Prevalence of interstitial lung involvement in patients with connective tissue diseases assessed with high-resolution computed tomography. *Scand J Rheumatol* 35: 388-394.
20. Camiciottoli G, Orlandi I, Bartolucci M, Meoni E, Nacci F, et al. (2007) Lung CT densitometry in systemic sclerosis: correlation with lung function, exercise testing, and quality of life. *Chest* 131: 672-681.
21. Yiannopoulos G, Pastrovas V, Antonopoulos I, Katsiberis G, Kalliolias G, et al. (2007) Combination of intravenous pulses of cyclophosphamide and methylprednisolone in patients with systemic sclerosis and interstitial lung disease. *Rheumatol Int* 27: 357-361.
22. Bellia M, Cannizzaro F, Scicilione N, Riili M, Triolo G, et al. (2009) HRCT and scleroderma: semiquantitative evaluation of lung damage and functional abnormalities. *Radiol Med* 114: 190-203.
23. Savarino E, Bazzica M, Zentilin P, Pohl D, Parodi A, et al. (2009) Gastroesophageal reflux and pulmonary fibrosis in scleroderma: a study using pH-impedance monitoring. *Am J Respir Crit Care Med* 179: 408-413.
24. Daoussis D, Liossis SN, Tsamandas AC, Kalogeropoulou C, Kazantzi A, et al. (2010) Experience with rituximab in scleroderma: results from a 1-year, proof-of-principle study. *Rheumatology (Oxford)* 49: 271-280.
25. Kazerooni EA, Martinez FJ, Flint A, Jamadar DA, Gross BH, et al. (1997) Thin-section CT obtained at 10-mm increments versus limited three-level thin-section CT for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *AJR Am J Roentgenol* 169: 977-983.
26. Choi HJ, Shin YK, Lee HJ, Kee JY, Shin DW, et al. (2008) The clinical significance of serum N-terminal pro-brain natriuretic peptide in systemic sclerosis patients. *Clin Rheumatol* 27: 437-442.
27. Goldin JG, Lynch DA, Strollo DC, Suh RD, Schraufnagel DE, et al. (2008) High-resolution CT scan findings in patients with symptomatic scleroderma-related interstitial lung disease. *Chest* 134: 358-367.
28. Mok MY, Fung PC, Ooi C, Tse HF, Wong Y, et al. (2008) Serum nitric oxide metabolites and disease activity in patients with systemic sclerosis. *Clin Rheumatol* 27: 315-322.
29. Ooi GC, Mok MY, Tsang KW, Wong Y, Khong PL, et al. (2003) Interstitial lung disease in systemic sclerosis. *Acta Radiol* 44: 258-264.
30. Pandey AK, Wilcox P, Mayo JR, Sin D, Moss R, et al. (2010) Predictors of pulmonary hypertension on high-resolution computed tomography of the chest in systemic sclerosis: a retrospective analysis. *Can Assoc Radiol J* 61: 291-296.
31. Tiev KP, Cabane J, Aubourg F, Kettaneh A, Ziani M, et al. (2007) Severity of scleroderma lung disease is related to alveolar concentration of nitric oxide. *Eur Respir J* 30: 26-30.
32. Wells AU, Hansell DM, Rubens MB, King AD, Cramer D, et al. (1997) Fibrosing alveolitis in systemic sclerosis: indices of lung function in relation to extent of disease on computed tomography. *Arthritis Rheum* 40: 1229-1236.
33. Collins CD, Wells AU, Hansell DM, Morgan RA, MacSweeney JE, et al. (1994) Observer variation in pattern type and extent of disease in fibrosing alveolitis on thin section computed tomography and chest radiography. *Clin Radiol* 49: 236-240.
34. Hoyles RK, Ellis RW, Wellsbury J, Lees B, Newlands P, et al. (2006) A Multicenter, Prospective, Randomized, Double-Blind, Placebo-Controlled Trial of Corticosteroids and Intravenous Cyclophosphamide Followed by Oral Azathioprine for the Treatment of Pulmonary Fibrosis in Scleroderma. *Arthritis & Rheum* 54: 3962-3970.
35. Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ, et al. (2008) Interstitial lung disease in systemic sclerosis: a simple staging system. *Am J Respir Crit Care Med* 177: 1248-1254.
36. Wells AU, Hansell DM, Haslam PL, Rubens MB, Cailles J, et al. (1998) Bronchoalveolar lavage cellularity: lone cryptogenic fibrosing alveolitis compared with the fibrosing alveolitis of systemic sclerosis. *Am J Respir Crit Care Med* 157: 1474-1482.
37. De Santis M, Bosello S, La Torre G, Capuano A, Toluoso B, et al. (2005) Functional, radiological and biological markers of alveolitis and infections of the lower respiratory tract in patients with systemic sclerosis. *Respir Res* 6: 96.
38. Khanna D, Tseng CH, Farmani N, Steen V, Furst DE, et al. (2011) Clinical course of lung physiology in patients with scleroderma and interstitial lung disease: analysis of the Scleroderma Lung Study Placebo Group. *Arthritis Rheum* 63: 3078-3085.
39. Goldin J, Elashoff R, Kim HJ, Yan X, Lynch D, et al. (2009) Treatment of scleroderma-interstitial lung disease with cyclophosphamide is associated with less progressive fibrosis on serial thoracic high-resolution CT scan than placebo: findings from the scleroderma lung study. *Chest* 136: 1333-1340.
40. Roth MD, Tseng CH, Clements PJ, Furst DE, Tashkin DP, et al. (2011) Predicting treatment outcomes and responder subsets in scleroderma-related interstitial lung disease. *Arthritis Rheum* 63: 2797-2808.
41. Kim EA, Johkoh T, Lee KS, Ichikado K, Koh EM, et al. (2001) Interstitial pneumonia in progressive systemic sclerosis: serial high-resolution CT findings with functional correlation. *J Comput Assist Tomogr* 25: 757-763.