Pleural irregularity, a new ultrasound sign for the study of interstitial lung disease in systemic sclerosis and antisynthetase syndrome

I. Pinal-Fernandez1,3, E. Pallisa-Nuñez2, A. Selva-O’Callaghan1, E. Castella-Fierro2, C.P. Simeon-Aznar1, V. Fonollosa-Pla1, M. Vilardell-Tarres1

1Autoimmune Systemic Diseases Unit (Department of Internal Medicine); 2Department of Radiology, Vall d’Hebron University Hospital, Universitat Autonoma de Barcelona, Barcelona, Spain; 3National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland, USA.

Iago Pinal-Fernandez, MD, PhD
Esther Pallisa-Nuñez, MD
Albert Selva-O’Callaghan, MD, PhD
Eva Castella-Fierro, MD
Carmen Pilar Simeon-Aznar, MD, PhD
Vincent Fonollosa-Pla, MD, PhD
Miquel Vilardell-Tarres, MD, PhD

Please address correspondence and reprint requests:
Dr Iago Pinal-Fernandez,
Muscle Disease Unit,
Laboratory of Muscle Stem Cells and Gene Expression,
National Institute of Arthritis and Musculoskeletal and Skin Diseases,
National Institutes of Health,
50 South Drive, Room 1141, Building 50,
MSC 8024, Bethesda, MD 20892, USA.
E-mail: iago.pinalfernandez@nih.gov

Received on April 25, 2015; accepted in revised form on July 10, 2015.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2015.

Key words: pleural irregularity, interstitial lung disease, systemic sclerosis, inflammatory myopathy, antisynthetase syndrome, ultrasound

ABSTRACT

Objective. To evaluate a new ultrasound sign, pleural irregularity (PI), for the study of interstitial lung disease (ILD), enrolled from 2011 to 2014. Chest high-resolution computed tomography (HRCT), pulmonary function tests (FVC and DLCO) and chest sonography were performed in each patient. Ultrasound PI and B-lines were quantified using a 72-sonographic point score and HRCT lung abnormalities were quantified using PFTs and categorised through Goh’s algorithm. PI was correlated with HRCT and pulmonary function test parameters and its diagnostic performance to detect and classify the extent of ILD was evaluated and compared with B-lines.

Results. Thirty-seven patients were studied. PI correlated with the Warrick score both in SSc (r=0.6, p=0.01) and ASS patients (r=0.6, p=0.005), showing a higher performance to detect ILD than using B-lines (p=0.01). In SSc patients PI also correlated with Wells score (r=0.7, p<0.001) and with DLCO (r=0.5, p=0.05), showing a high diagnostic value for detecting ILD (AUC=0.85, 95% CI 0.64–1) and classifying it into limited or extensive (AUC=0.81, 95% CI 0.57–1). A modification of the Goh algorithm including PI was developed as a screening tool to avoid the use of HRCT in SSc patients without ultrasound evidence of extensive ILD.

Conclusion. PI is useful for evaluation of ILD in SSc and ASS patients, and can be incorporated into a diagnostic algorithm in SSc patients to reducing the need for exposure to ionising radiation.

Introduction

Systemic sclerosis (SSc) and myositis, especially antisynthetase syndrome (ASS), are systemic autoimmune diseases associated with a high prevalence of interstitial lung disease (ILD) (1). In these disorders, ILD can cause considerable morbidity if it is not treated adequately and close monitoring is highly recommended both before and after ILD is diagnosed. Pulmonary function tests (PFTs) and high-resolution computed tomography (HRCT) are the main tools for functional and structural follow-up of the lung in these patients, but the ionising radiation of HRCT is a matter of concern, as it is a recognised cause of cancer (2). Chest ultrasound (US) with B-lines study has emerged over the last years as an alternative method to evaluate ILD in patients with connective tissue diseases, particularly SSc (3), but also in ASS (4) or rheumatoid arthritis (5).

During the preliminary phase of our previous study about the utility of B-lines in AAS patients with ILD (4) and in accordance with reported observations of other authors (6-9), a marked irregularity of the pleural contour was noted on US examination of these patients. The objective of the present study was to determine the potential of pleural irregularity (PI) to detect ILD, correlate this parameter with HRCT and PFT findings, and assess its performance compared to B-lines detection in SSc and ASS patients.

Material and methods

Patients

A convenience sample of patients from the SSc and ASS cohorts of Vall d’Hebron Hospital, in active follow-up and with known varying degrees of ILD, was recruited between 2011 and 2014. All SSc patients met the LeRoy...
the other exams or the patients’ clinical characteristics, performed all the tests. The principal investigator collected the data as each test was completed. The hospital ethics committee approved the study protocol (PR [AG] 186/2011) and patients gave informed consent to undergo the tests carried out.

**Chest sonography**

Chest US was performed with a MyLabTwice system (Esaote, Genoa), using a 5-MHz, 5-cm linear probe (LA 527). US examinations were performed with patients in supine position to record the anterior and anterolateral sonographic points (up to 28), and in sitting position for the posterior and posterolateral ones (up to 44). The 72 sonographic points corresponded to the anatomical distribution defined by Gargani (13). We excluded sonographic points with scarce lung parenchyma or absence of pleural interface, where it was not possible to find ILD-associated signs. A B-line was defined as a vertical hyperechoic artifact perpendicular to the pleural line extending to the edge of the sonographic window. Pleural irregularity was defined as the loss of the normal hyperechoic linear pleural contour (Fig. 1).

**High-resolution computed tomography**

High-resolution computed tomography examinations of the chest were performed on a spiral CT scanner (Philips Brilliance 64) with 64 detector rows and 0.75 s rotation time. All patients underwent a preliminary A-P scout view. Subsequently, 20 to 25 slices were acquired in apnea at the end of inspiration from the lung apex to base, either in supine or prone decubitus position. The acquisition parameters were as follows: sequential mode, 1-mm collimation and 10-mm interval, 150 mA average tube current (depending on the patient’s build), and 120 kV tube voltage. A bone plus reconstruction with lung window was used. No intravenous contrast material was administered. The duration of CT acquisition was 20 to 30 s, matrix was 512-512 and absorbed dose was under 7 mSv in all cases. Pulmonary involvement was quantified with use of the Warrick (14) and Wells (15) scores. The Goh algorithm (16) was applied to differentiate between extensive and limited ILD.

**Pulmonary function tests**

Pulmonary function tests, including the forced vital capacity (FVC) and diffusing capacity of the lung for carbon monoxide (DLCO), were performed using a MasterLab Pro lung function measurement system (Jaeger GmbH, Wuerzburg, Germany) and adjusted according to the reference values proposed by Roca (17) for the Mediterranean population.

**Autoantibody detection**

Antisynthetase antibodies (anti-Jo-1, -PL-7, -PL-12, -EJ and -OJ) were identified by ELISA or line immunoassay (Myositis Profile Euroline®, Euroimmun, Lübeck, Germany) (18) and confirmed by RNA or protein immunoprecipitation assay with radiolabeled HeLa cells (19). Anti Scl-70 and anti-centromere antibodies were detected using ELISA.

**Continuous variables were expressed as the mean and standard deviation (SD). A sonographic point was considered positive for PI if it showed an irregular pleural contour. To properly describe the results of US examination and to adjust them to the patient’s lung size (thus avoiding underscoring of patients with small lungs due to ILD), the number of sonographic points with PI were divided by the total number of sonographic points studied in each patient to yield percentages, which were used as the PI ultrasound score. Likewise, B-lines were quantified as previously described (4). Correlations between the HRCT, PFT findings and PI score were calculated using Pearson’s correlation coefficient (r). The Wilcoxon rank-sum test was used to assess differences between US scores in patients with and without ILD. The PI score to diagnose ILD was evaluated using receiver operating characteristic (ROC) curves. The area under the curve (AUC) with the 95% confidence interval (CI) was calculated, considering the following reference:
AUC 0.5 to 0.6, a bad test; 0.6 to 0.75, a regular test; 0.75 to 0.9, a good test; 0.9 to 0.97, a very good test, and 0.97 to 1 an excellent test. The chi-square test was used to compare the AUCs of the different scores, and the point that maximised Youden’s index was selected as the optimal cut-off to detect ILD.

Subanalyses of SSc and ASS patients separately were also performed, examining the same parameters in each group as in the complete sample. In addition, the US scores for SSc patients were correlated with the Wells’ score, and the diagnostic value of US scores to detect ILD and classify the condition into limited or extensive according to the Goh algorithm was assessed using ROC curves.

Based on the results of previous reports on the subject (20, 21), it was calculated that the study had a statistical power of 99.9% to detect a correlation of 0.7 in the total sample of 37 patients assuming an alpha error of 5%. For the SSc (16 patients) and ASS (21 patients) sample sizes, the statistical power was 88% and 96%, respectively.

Statistical analyses were performed with Stata v.12. P-values less than 0.05 were considered statistically significant.

**Results**

Over the 36-month study period, 37 patients were included (mean age 51.5 years, SD 14.2 years; 24 women): 21 had ASS (2 without ILD) and 16 SSc (6 without ILD). Chest US and HRCT were performed less than 6 days apart, on average (SD 19.2). Among the ASS patients, 8 were performed less than 6 days apart, (2 without ILD) and 16 SSc patients (2 without ILD) and 16 SSc patients were diagnosed with ILD with a sensitivity of 79% and a specificity of 100%.

**Correlation of PI with HRCT and PFT results**

Analysing SSc and ASS patients together, the PI score showed a significant correlation with the Warrick score (r=0.59; p<0.001), DLCO (r=0.36; p=0.04), and number of bronchopulmonary segments with ILD involvement (r=0.48; p=0.003), ground glass appearance (r=0.49; p=0.002), septal or subpleural lines (r=0.42; p=0.009), subpleural cysts (r=0.34; p=0.04), and irregular pleural margins (r=0.46; p=0.004) on HRCT (Table I).

**Performance of pleural irregularity as a diagnostic tool to detect ILD**

Patients with ILD showed a significantly higher PI score than those without it (35.3% vs. 6%; p<0.001). The AUC of the PI score for the diagnosis of ILD (AUC=0.93, 95% CI 0.85–1) was similar to that of the Warrick score (AUC=0.93, 95% CI 0.83–1) and significantly higher (p=0.01) than that of the B-line score (AUC=0.63, 95% CI 0.4–0.86) (Fig. 2a).

The PI score that maximised the Youden index for the diagnosis of ILD was 24%. With this cut-off, the sensitivity for the diagnosis of ILD was 79% and the specificity 100%. This means that an individual patient with 24% or more segments showing PI can be diagnosed with ILD with a sensitivity of 79% and a specificity of 100%.

**Subanalysis in patients with systemic sclerosis**

In the group of SSc patients, the PI score was higher in those with ILD than in those without (33.3% vs. 4.1%; p=0.03). Similarly, the PI score was higher in patients with extensive disease than in those with limited disease, evaluated with the Goh score (34.2% vs. 7.1%; p=0.04).

The PI score correlated with the Warrick (r=0.63; p=0.01), and Wells score (r=0.74; p=0.001), as well as with the DLCO (r=0.53; p=0.05), and the number of bronchopulmonary segments with ILD involvement (r=0.65; p=0.009), segments with a ground glass appearance (r=0.69; p=0.003), irregular pleural margins (r=0.64; p=0.008), and septal or subpleural lines (r=0.62; p=0.01), assessed by HRCT (Table I).

The area under the ROC curve to detect ILD for the PI score was 0.85 (95% CI 0.64–1), was higher, but not significantly, than the B-lines score (AUC=0.65, 95% CI 0.32–0.98). The PI score that maximised Youden’s index for the diagnosis of ILD was 16% (sensitivity 80%, specificity 100%). Thus, a given SSc patient with 16% or more segments showing PI can be diagnosed with ILD with a sensitivity of 80% and a specificity of 100% (Fig. 2b). To differentiate between limited and extensive disease, the US scores for SSc patients were correlated with the Wells’ score, and the diagnostic value of US scores to detect ILD and classify the condition into limited or extensive according to the Goh algorithm was assessed using ROC curves.
extensive lung disease according to the Goh algorithm, the PI score showed an AUC of 0.81 (95% CI 0.57–1), also superior, but not significantly, to B-line (AUC=0.57, 95% CI 0.25–0.89). The PI score that maximised Youden’s index to differentiate between limited and extensive lung disease was 28% (sensitivity 67%, specificity 100%). A specific SSc patient with 28% or more segments showing PI can be diagnosed with an extensive lung disease with a sensitivity of 67% and a specificity of 100% (Fig. 2c).

Subanalysis in patients with antisynthetase syndrome
We were unable to properly evaluate US for diagnosing ILD in ASS patients because of the high prevalence of ILD in the sample (19 out of 21). The PI score was higher (but only approaching significance; p=0.06) in patients with ILD than in those without (36.3% vs. 11.9%). The PI score, however, correlated with the Warrick score (r=0.59; p=0.005) and the number of bronchopulmonary segments with honeycombing (r=0.64; p=0.002) (Table I).

Discussion
In this study, we present pleural irregularity as a new ultrasound sign for use in the evaluation of ILD due to SSc and ASS. Our results show that this parameter can reliably differentiate between patients with and without ILD, showing a high specificity and that its performance for this purpose seems to be higher than using B-lines. Moreover, it is effective for determining limited and extensive lung disease in SSc patients. The best cut-off points for each of these purposes were calculated.

Sperandeo (6), Moazedi-Fuerst (7, 9) and the international recommendations for point-of-care lung ultrasound (8) have recognised pleural line irregularity as a chest ultrasound sign of ILD. However, the international recommendations do not mention how to quantify pleural irregularity and the two first authors focused on measuring the pleural line thickening (defined as >3mm or >2.8mm) rather than assessing pleural irregularity itself. Some patients in our study showed clearly irregular pleural lines without pleural thickening (Fig. 1b, irregular pleural line, but maximum thickness of 2mm) and others, significant blurring of the pleural contour (Fig. 1c), which made objective calculations of pleural thickness very difficult. Thus, we believe that direct quantification of pleural irregularity is preferable to the use of pleural thickness as a surrogate marker.
Pleural irregularity in SSc and ASS ILD / I. Pinal-Fernandez et al.

Fig. 3. Therapeutic decision tree, a modification of Goh’s algorithm, for patients with systemic sclerosis, including the pleural irregularities score as a screening tool to diagnose and classify the extent of interstitial lung disease.

marker. To our knowledge, this is the first study to directly assess the irregularity of the pleural line as a US sign for ILD evaluation.

While the genesis of the B-lines has been well described (22), the origin of pleural irregularity in ILD remains obscure. Serositis is not a common feature in SSc or in ASS. Therefore, it would not be reasonable to attribute blurring of the pleural to direct inflammation of the pleura. A more plausible theory is that increased echogenicity in the peripheral lung due to subpleural ILD would somewhat equal the echogenicity of the pleura, thereby making it difficult to discriminate between the pleural line and the abnormally hyper-echogenic lung periphery.

Notwithstanding the good correlation of PI with HRCT scores in both ASS and SSc patients, the number of HRCT parameters that correlated with PI score was higher in SSc than in ASS, even though the sample size was larger in the latter group. This may suggest that the utility of PI may be lower in ASS than in SSc. Non-specific interstitial pneumonia in myositis has been reported to be less peripheral than idiopathic non-specific interstitial pneumonia (23). If these findings and our theory about the origin of PI are true, US study of the lung periphery in patients with ASS would be less informative than in other, more peripheral, causes of ILD.

In our experience, PI shows higher reliability to detect ILD than B-lines. Known the heterogeneity of B-line quantification methods proposed in the literature, we used also the consensus system of the international recommendations for point-of-care lung ultrasound (8) to quantify B-lines and we did not find improvement of its performance compared with PI (data not shown). Moreover, PI and B-line quantification was performed at the same time using the same technical parameters in each patient, consequently both techniques were directly comparable and subject to the same potential bias.

The extent of ILD measured by HRCT in patients with SSc has been associated with mortality. In clinical practice, Goh’s algorithm (16) is used to identify patients that would benefit the most from immunosuppressive treatment. Because of the good capability of PI to detect ILD and to discriminate between limited and extensive ILD in patients with SSc, we propose a modification of Goh’s algorithm including PI to classify the extent of ILD in SSc patients (Fig. 3). This algorithm would help to reduce the need for HRCT with its ionising radiation, especially in patients with mild ILD, a population with longer survival and consequently, the most prone to be affected by the potential long-term harms of ionising radiation.

Our study has some limitations. First, given the rarity of ASS and SSc, the number of patients included was relatively small. Nonetheless, the statistical power was optimal to detect the correlations reported in the literature for chest US findings in SSc. Second, the high prevalence of ILD in the ASS sample prevented assessment of the value of PI for ILD diagnosis in this patient group. Third, convenience sampling is suitable to determine the discriminative capacity and best cut-offs of a diagnostic tool (in this case PI), but it does not allow estimation of the predictive values. Finally, detection of any ultrasound sign, including PI, depends to a certain extent on the expertise of the operator. Determining the inter- and intra-observer variability of PI findings was not the aim of the study, but preliminary testing was carried out. The same author that performed the chest ultrasound (ECF) evaluated 276 random static images of US points from the study patients twice. Good intra-observer variability between the two readings was observed (κ=0.71).

In conclusion, we present PI as a new ultrasound sign to incorporate in the study of ILD in SSc and ASS patients. PI showed a correlation with several HRCT and PFT parameters and performed well as a diagnostic tool for detecting and determining the extent of ILD in these patients. We propose an algorithm to categorise the extent of ILD in SSc that can be used to reduce exposure to ionising radiation in these patients.

References
4. PINAL FERNANDEZ I, PALLISA NUNEZ E,
Pleural irregularity in SSc and ASS ILD / I. Pinal-Fernandez et al.


