ABSTRACT
Interstitial lung disease (ILD) is one of the most serious pulmonary complications associated with connective tissue diseases (CTDs), resulting in significant morbidity and mortality. Recently, lung ultrasound examination has appeared as a useful, feasible, non-invasive, radiation-free imaging technique with high sensitivity and specificity for the diagnosis of CTD-associated ILD. In this review, we discuss the literature concerning lung ultrasound findings (B-lines and pleural irregularities) in systemic sclerosis. We also examine the possible utility and potential limitations of lung ultrasound in other CTDs including inflammatory myopathies, mixed connective tissue disease, and Sjögren’s syndrome. Specifically, we compare the diagnostic accuracy of lung ultrasound with traditional imaging techniques and pulmonary functional tests.

Introduction
Interstitial lung disease (ILD) represents one of the most severe and distressing organ involvements in connective tissue diseases (CTDs) (1). It affects about 50% of patients with systemic sclerosis (SSc) (2, 3) and 20 to 60% of patients with mixed connective tissue disease (MCTD) and inflammatory myopathies (IMs) (4, 5) and represents the second leading cause of morbidity and acute mortality (between cardiovascular and neoplastic disorders) (6). Furthermore, CTD-pulmonary involvement not only affects patient prognosis adversely, but also therapeutic choices by the rheumatologist and therefore possible exposure to infectious and immunosuppressive risks (7, 8).

From a classification and pathophysiological point of view, it is useful to distinguish the different features of CTD-pulmonary involvement by applying the Diffuse Parenchymal Lung Disease Classification proposed by the American Thoracic Society/European Respiratory Society in 2002 and updated in 2013 by Travis et al. (9). The pulmonary pictures described allow rheumatologists to distinguish clinical, radiographic (CT-scan) and prognostic features of ILD-CTD. These classification criteria are particular helpful for the distinction between predominantly fibrous UIP-like pattern (Usual Interstitial Pneumonia) more frequently diagnosed in rheumatoid arthritis (RA) and SSc and non UIP-like pattern, among these NSIP (Non Specific Interstitial Pneumonia) associated with SSc, MCTD, IMs and Sjögren’s syndrome (SS), AIP (Acute Interstitial Pneumonia) more frequent during systemic lupus erythematosus (SLE), COP (Cryptogenic Organizing Pneumonia) prevalently in dermatomyositis (DM) and polymyositis (PM) and LIP (Lymphocyte Interstitial Pneumonia) typical of SS (4, 10). In this complex scenario, ILDs represent a challenge for the rheumatologist in terms of early diagnosis (“window of opportunity”), differential diagnosis between the forms more or less responsive to immunosuppressive therapy, prognostic stratification and outcomes (10, 11). Tools available to the rheumatologist to detect CTD-ILD, in addition to clinical examination, include pulmonary functional tests (PFT) and imaging methods. However, many studies have shown that symptoms and clinical signs are often nonspecific, poor or even absent, particularly in the early phase of the disease; PFT, although sufficiently sensitive in follow-up and in therapeutic response evaluation, are not so specific and useful in the characterisation of lung injury (12). Furthermore, standard chest x-ray, widely used in clinical practice as first imaging approach, has low sensitivity in early stages. Therefore, high-resolu-
The use of ultrasound to assess lung involvement in CTDs / F. Ferro & A. Delle Sedie

B-lines are defined as discrete laser-like vertical hyperechoic reverberations that arise from the pleural line (previously described as “comet tails”), extending to the bottom of the screen without fading, and moving synchronously with lung sliding (18), erasing the normal “A line” (horizontal lines that arise from the pleural line) (19) (Fig. 1). B-lines also can be visualised in about 27% of healthy subjects, although in small numbers and primarily in the lower part of the lung (20, 21). The initial papers were not focused on patients with rheumatic diseases, but rather to establish a correlation between B-lines and the water-thickening of interlobular septa in ILD and in alveolar edema (20-23) and to provide scoring systems (22, 24)). In 2007 Gargani et al. (23) and then Jambrick et al. (25), confirmed those data on histology samples in animal models.

B-lines in systemic sclerosis

Lung US in rheumatic patients was reported initially in 2008 by Doveri et al. (26), who examined 30 SSc patients using a previous published scanning protocol (21) and the score of Picano (24). These authors used a cardiac probe (2.5-3.5 MHz, Optigo Philips), without any healthy or other diseased controls. HRCT was the comparator and a score from 0 to 3 (0: normal; 1: bibasilar fibrosis; 2: diffuse fibrosis and 3 honey combing) was used for HRCT. The number of B-lines (still called ultrasound lung comets (ULCs) at that time) was higher in Scl-70 positive (usually related to a worse prognosis) than in ACA-positive patients, and it was correlated to the HRCT score but not to the presence of pulmonary arterial hypertension. This initial study did not provide information concerning intra- or inter-reader variability, and did not use a more structured and articulated score for HRCT.

The initial study was soon followed by a report from Gargani et al. (27), using the same scanning protocol and US machine but modifying the Picano cutoff (24) to consider pathologic a total number of B-lines higher than 10, and improve the HRCT by using the Warrick score. Examining 33 SSc patients, this report confirmed the correlation between B-lines and the HRCT score. A weak but significant correlation was also found with total lung capacity, vital capacity and diffusion capacity of the lung for carbon monoxide (DLCO). The intra and inter-observer variability between two blinded and independent sonographers was evaluated and confirmed the data previously published by Jambrick (21).

Another important step in LUS was performed when Delle Sedie et al. (28) demonstrated that B-lines also were detectable using a linear probe. In this report, a linear (6 MHz, Toshiba Powervision 6000) and a cardiac transducer (2.5-3.5 MHz, Optigo Philips), previously used in other studies, were compared to the “gold standard” provided by HRCT using the Warrick score. The scanning protocol was modified from the previously described method, not performing mid axillary line scanning.

Fig. 1. Grey-scale US of the lung with the probe on the anterior axillary line showing US lung comet, or B-lines, (arrow heads). White arrows indicate pleural line (51).
to avoiding duplicate counting of B-lines, because of the different width of the linear probe footprint (6 cm). The results in 25 SSc patients indicated a significant intra-class correlation (ICC) between both probes (ICC=0.681), with better (but not statistically significant) results for the anterior than posterior chest region. Moderate to good intra-class correlation was shown between a cardiac or linear probe and HRCT (ICC=0.547 and 0.600, respectively). Specificity and sensitivity of US with respect to the HRCT were 70% and 85% for the cardiac probe, and 60% and 85% for the linear one. Pathologic cut-offs for B-lines were calculated for each probe used (resulting in 5 and 11, for cardiac and linear probe, respectively).

A less time-consuming methodology than the comprehensive scanning protocol was reported by Gutierrez et al. (29) in 36 patients with different connective tissue diseases (mainly SSc). The patients were examined using both US (using a MyLab 70 Esaote machine with a 2–7 MHz convex transducer) and HRCT (adopting Warrick score). A comprehensive US assessment was performed according to the Jambrick approach (21) and a simple post-hoc analysis of 50 sites in the standard US comprehensive assessment identified 14 sites for the simplified US assessment. The study confirmed the positive correlation between the number of B-lines and the Warrick score in both methods; the inter-operators agreement was very high (K-value>0.8), even if minimally lower than in the comprehensive assessment (K-value=0.9), and the mean time spent to perform the assessment passed from 23.3 minutes (for the comprehensive one) to 8.3 minutes (for the simplified one).

Since 2011, a few other reports have confirmed the association between B-lines and HRCT (applying different scanning protocols) (30-32) and providing new information on the correlation between B-lines and vascular damage (33).

In a pilot study on 25 SSc patients, Moazed-Fuerst et al. (30) compared LUS and HRCT. Using a LOQ 7 (GE) US machine and a convex 3.5-MHz transducer, but also a linear probe for the pleural changes, they scanned 18 regions of the thoracic wall. They found that B-lines were more common in SSc patients than in normal controls and a significant difference was seen between patients with and without radiographic ILD. The results about the pleural line are reported in the next section of the paper.

Soon after, Barskova et al. (31) demonstrated that LUS is sensitive to detect ILD even in patients with very early SSc. In 58 consecutive patients who were examined using both HRCT and LUS, 32 met criteria for very early SSc (VEDOSS). ILD was found in the 88% of the patients and in 41% of the VEDOSS group using HRCT. B-lines were significantly higher in patients with ILD than in those with no ILD on HRCT; the concordance rate was 83%. Sensitivity and negative predictive value for B-lines were 100% in both SSc and VEDOSS (all discordant cases between HRCT and LUS were false positive for the second imaging technique).

In 2014 Mohammadi et al. (32) confirmed a positive correlation between HRCT (Warrick score) and LUS in 70 SSc patients. Using a Medison Accuvix V20 (Medison, South Korea) US machine equipped with a 7–10 MHz broad band linear multi-frequency transducer, they scanned only 10 intercostal spaces, selected according to the higher prevalence of ILD in SSc and to accessibility by LUS. The sensitivity, specificity, positive and negative predictive value of LUS were 73.58%, 88.23%, 95.12% and 51.72%, respectively, with an excellent intra-observer K-value of 0.838. The same author published the same results in a smaller number of RA patients the year before (34).

The following year Moazed-Fuerst et al. (35) examined a small mixed population of patients with diagnoses of RA, SSc and SLE. LUS. ILD on HRCT was present in 64% of the 14 SSc patients. Once again, LUS pathologic (assessing B-lines, subpleural nodes and irregularities of the pleura) were significantly more frequent in the ILD group than in the non-ILD group.

More recently, Gigante et al. (33), examined 39 SSc pts according to clinical examination, HRCT, pulmonary function tests and LUS for B-lines, and showed a positive correlation between B-lines and the HRCT score (r=0.81, p<0.0001) and a negative correlation between B-lines and DLCO (r=-0.63, p<0.0001). The authors also demonstrated that the number of B-lines were increasing along with the progression of the capillaroscopic damage, and a statistically significant difference in the number of B-lines was found between patients with and without digital ulcers.

B-lines in other connective tissue diseases

Some authors applied LUS to other diseases. Tardella et al. (35) investigated a group of 34 patients with a diagnosis of CTD (not specified further). Each patient underwent clinical examination, pulmonary function test (PFT), chest HRCT, and lung US (including 50 intercostal spaces). In each intercostal space, the number of B-lines was recorded, summed, and graded according to a semiquantitative scoring (grade 0 = normal (<10 B-lines); grade 1 = mild (11 to 20 B-lines); grade 2 = moderate (21 to 50 B-lines); and grade 3 = marked (> 50 B-lines)). As already demonstrated in SSc, a significant linear correlation was found between US and HRCT scores (p<0.001, correlation coefficient ρ=0.875), as well as a negative correlation between B-lines and DLCO (p=0.014). Finally, reliability between two sonographers was excellent.

A few years later, Pinal Fernandez et al. (36) explored the use of LUS in 21 patients with the anti-synthetase syndrome, using both HRCT (Warrick score) and LUS (B-lines) at several intercostal spaces in the anterior and posterior chest wall. B-lines were present in 44.1% of the patients and a positive association between the number of bronchopulmonary segments showing ground glass findings and the percentage of sonographic points with B-lines (Rho=0.502; p=0.02) was demonstrated.

Buda et al. (37) studied 52 patients with diagnosed diffuse ILD with LUS and HRCT, reporting the presence of a new artefact that was called Am-line.
This new finding is a reverberation artifact presenting a fusion of both A-line and B-line, which probably arises from the multiple reflection of US waves between two border surfaces (the pleural line and the wall of emphysematous bulla or subpleural cyst). Am lines were significantly correlated to honeycomb (\(p=0.002\)) and subpleural cysts (\(p=0.0014\)) in HRCT. The sensitivity and specificity for Am line coexisting with a blurred pleural line with respect to HRCT were 0.9 and 0.75, respectively.

Vasco et al. (38) explored the accuracy of LUS to diagnose ILD in 13 patients with Sjögren’s syndrome (SS), all presenting with abnormal pulmonary function tests or respiratory symptoms. The intra-rater reliability for LUS (B-lines) was perfect (\(k=1\)); sensitivity and specificity of LUS (with respect to the HRCT) were 1 and 0.89, respectively. The excellent correlation between LUS and HRCT was confirmed in SS patients affected by ILD. Overall, the existing literature has highlighted several advantages of using B-lines, not only in SSc but also in other CTDs. However, the available data have also shown some potential pitfalls that should be overcome in order to use this technique in a clinical setting, as discussed below.

**Limitations of LUS**

Lung US assessment in rheumatology is a new application and the validation process remains incomplete at this time. An OMERACT Ultrasound Working Group has been recently created to validate and standardize the US assessment of the lung in rheumatic diseases, reflecting growing interest in the field. In the first submitted paper (39), LUS appears to pass the OMERACT filter in term of face and content validity as well as for feasibility (40). There appears a need for definitions for elementary lesions, even if a definition of B-line was provided by a panel of experts in point-of-care lung ultrasound (18). Furthermore, there remain many unanswered questions about scanning methodology, including the number of areas or intercostal spaces to be examined, time needed for every single space to be assess, optimal probe to use (convex, cardiac, linear or more than one for a better evaluation of all of the pathological findings), scoring system (i.e. count all of the pathological findings, grade the positivity of the findings, count the positive spaces), cut-off for ILD (different authors used different cut-offs, not always based on a statistic analysis) and to discriminate which of the actual possibility to assess ILD by US could be best (i.e. B-lines or pleural irregularity). Data concerning the minimal significant change would appear needed before using LUS as an outcome for clinical trials on ILD treatment. All of those points were considered to be missing by the OMERACT group, which indicated that data for a full confirmation of criterion validity, reliability and sensitivity to change over time remain incomplete.

**Pleural irregularities (PI)**

As previously noted, most of clinical studies from 2008 to date have focused attention on the evaluation of the number and distribution of B-lines as a marker of “interstitial syndrome” in order to validate LUS as a useful tool for screening in the preclinical stages of ILD-CTDs and as a sensitive and specific imaging technique compared to HRCT and PFTs. More recently, some authors evaluated the diagnostic power of a different US findings, the alteration of the pleural lines. Wohlgemant et al. (41) and Reißig and Kroegel (42) were the first authors to describe thickening, irregularities and pleural fragmentation in patients with ILD (Fig. 2-3) (42). After many years of neglect, in 2012 Moazedi-Fuerst et al. (30) analysed the pulmonary US profile, evaluating both PI and B-lines in 25 HRCT-diagnosed ILD patients compared with 40 healthy subjects. The results not only confirmed the diagnostic sensitivity of both variables to higher than the standard chest x-ray, but also the higher negative predictive power of PI score compared to B-line score. No healthy subject had alterations of the pleural profile of more than 2.8 mm thickened in more than two lung segments explored, although 3/40 healthy subjects showed B-lines in the absence of ILD. These data were confirmed in a second study published in 2015, analysing patients affected by RA, SSc and SLE. Pathological US patterns were significantly different in ILD group than in non-ILD group (comet tail artifacts/B-pattern: 100% vs. 12%, \(p<0.001\); subpleural nodules: 55% vs. 17%, \(p=0.006\); thickenings of the pleural line: 95% vs. 12.5%, \(p<0.001\)). An irregular pleural line >3 mm was documented in 100% of SSs and SLE patients with ILD vs. 86% of ILD patients suffering from RA (\(p=ns\)) (35).

Subsequently, Pinal Fernandez et al. recruited 37 patients, 16 with anti-synthetase syndrome and 21 with SSc, quantifying and describing the distribution of pleural irregularities and B-lines, compared with clinical, spirometric and DLCO indices, scoring sys-
tems (Warrick and Wells) for the LUC diagnosis and quantification of ILD severity using HRCT (43). The study demonstrated that the pleural profile was an equally reliable finding in the diagnosis and stratification of the severity in terms of extent of ILD, showing a higher diagnostic performance of PI compared to B-lines. Alterations of the pleural profile, expressed as thickening, interruption, and fragmentation appear to be more sensitive markers in the diagnosis of ILD. Furthermore, US-pleural irregularity appears to be useful in discriminating B-lines false positivity due to age (44) or pulmonary congestion (24) in patients not affected by ILD.

**New applications and future perspectives**

In recent years the use of LUS in clinical practice has expanded the instrumental equipment of the rheumatologist in CTDs-ILD diagnosis and follow-up. Most studies concerning SSc and B-lines have confirmed LUS as a reliable method in the early diagnosis, including the pre-clinical phase of the disease and, in stratification of severity and follow-up care. The last 3 years resulted in a growing interest in lung ultrasound in SS, PM and DM, in which interstitial lung involvement appears to be less frequent and relatively less severe than in SSc, although sometimes underestimated (45). From 2015 to date, preliminary results have been presented at Italian, European and International Rheumatology Congresses on the potential role of PI in SS (46, 47). US assessment of PI has proved to be a very sensitive and specific tool in ILD-SS vs. HRCT, being able to recognise early PFT alterations (particularly in the initial phase of ILD) and to assess the follow-up after therapy. Moreover, PI results were strongly associated with the Warrick scores in course of PM and DM, confirming diagnostic power also in those autoimmune inflammatory myopathies.

Lung Ultrasound Surface Wave Elastography (LUSWE) represents a new echographic application that has recently appeared. This method, firstly proposed for pulmonary assessment by Zhang et al. in 2016 on experimental models, represents a non-invasive and non-ionising technique with capacity to measure the elasticity of superficial lung tissue (48). The data obtained and published in 2017 and in February 2018 indicated significant differences in speed of propagation of US wave through the lung surface between patients with ILD-SSc compared to healthy volunteers (49, 50). These preliminary data are encouraging and LUSWE may represent a useful tool for a quantitative assessment of CTDS-ILD to reduce ionising radiation exposure and assess post-therapy modifications.

**Conclusions**

Lung US represents a promising application that remains under development. The last few years showed an increasing interest on the field that provide newer approaches and attempts to simplify the methodology of examination of the lung. We hope that this interest will soon lead to a fully standardisation for the scanning technique as well as for the scoring system. This should lead to increased application of US to evaluate ILD in rheumatic patients.

**References**


