Diagnostic accuracy of lung ultrasound for interstitial lung disease in patients with connective tissue diseases: a meta-analysis

G.G. Song[1], S.-C. Bae[2], Y.H. Lee[1]

[1]Division of Rheumatology, Department of Internal Medicine, Korea University College of Medicine; [2]Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea.

Abstract

Objective

The purpose of this study was to evaluate and compare the diagnostic performance of lung ultrasound (US) in respect to high-resolution computed tomography (HRCT) findings in patients with connective tissue disease-associated interstitial lung disease (CTD-ILD).

Methods

We searched the Pubmed, Embase, and Cochrane Library databases, and performed a meta-analysis on the diagnostic accuracy of lung US according to B-lines (comet tail sign) and on the correlation coefficients between lung US scores and HRCT Warrick scores in CTD-ILD patients.

Results

Five studies that included a total of 349 patients were available for this meta-analysis. The pooled sensitivity and specificity of lung US were 91.5% (95% confidence interval [CI]: 84.5–96.0) and 81.3% (95% CI: 74.6–86.9), respectively. The positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were 4.100 (2.133–7.879), 0.176 (0.006–0.363), and 34.73 (10.10–99.66), respectively. The area under the curve was 0.915 and the Q^* index was 0.848, indicating a high diagnostic accuracy. When all four studies with systemic sclerosis were considered together, the pooled sensitivity and specificity of lung US were 89.5% (95% CI 80.3–95.3) and 79.6% (69.9–87.2), respectively. A significant correlation was found between lung US B-line scores and HRCT Warrick scores in CTD-ILD (correlation coefficient: 0.783; p-value $<1 \times 10^{-9}$).

Conclusion

Our meta-analysis of published studies demonstrates that lung US has a high diagnostic accuracy, correlates well with HRCT findings, and plays an important role in the diagnosis of CTD-ILD.

Key words

connective tissue disease, interstitial lung disease, lung ultrasound, diagnostic accuracy, meta-analysis.

Gwan Gyu Song, MD, PhD
Sang-Cheol Bae, MD, PhD
Young Ho Lee, MD, PhD

Please address correspondence to:
Young Ho Lee, MD, PhD,
Division of Rheumatology,
Department of Internal Medicine, Korea University Anam Hospital,
Korea University College of Medicine,
73, Inchon-ro, Seongbuk-gu,
Seoul, 136-705, Korea.
E-mail: lyhcgh@korea.ac.kr

Received on February 16, 2015; accepted in revised form on April 8, 2015.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2016.

Funding: this study was supported in part by a grant of the Korea Healthcare technology R&D Project, Ministry for Health and Welfare, Republic of Korea (HI13C2124).

Competing interests: none declared.

Introduction

Connective tissue diseases (CTD) are a heterogeneous group of systemic disorders characterised by autoimmune response and immune-mediated organ damage that can affect multiple organ systems including systemic sclerosis, mixed connective tissue disease, and rheumatoid arthritis. Interstitial lung diseases (ILD) are a group of diffuse parenchymal lung diseases characterised by interstitial involvement resulting from inflammation and fibrosis (1). Interstitial lung disease is a common manifestation of CTD, and leads to significant morbidity and mortality (2). High-resolution computed tomography (HRCT) is a sensitive and reproducible method to assess the extent and the pattern of ILD. While HRCT represents the gold standard method for diagnosis of ILD, it potentially exposes patients to high doses of ionising radiation (3).

The role of lung ultrasound (US) in the assessment of ILD has been studied previously as both a non-invasive and non-ionising diagnostic tool (4). The use of US to assess ILD consists of detection and quantification of the B-lines fans out from the lung surface, generated by the reflection of the US beam from the thickened sub-pleural interlobar septa (4). B-lines are a reliable tool to evaluate diffuse parenchymal lung disease, and that their distribution correlates with CT signs of fibrosis (5), although mechanisms of B-lines generation are still unclear and the explanation reported is quite debated in literature. However, there are little available data on the diagnostic performance of lung US in patients with CTD-ILD (4, 6-9). The diagnostic accuracy of lung US remains unclear in the detection of ILD in patients with CTD, and thus a systematic analytical approach is needed to validate lung US for its assessment. The aim of this meta-analysis was to assess and compare the diagnostic performance of lung US in respect to HRCT, as well as evaluate the correlation between US and HRCT findings in patients with CTD-ILD.

Materials and methods

Identification of eligible studies and data extraction

We used the Pubmed, Embase, and Cochrane Library databases to identify articles published up to February 2015, in which the diagnostic accuracy of lung US according to B-lines was compared with that of HRCT findings in patients with CTD-ILD. The following keywords and subject terms were used in the search: "lung," "ultrasound," "HRCT," "sensitivity," "specificity," "pulmonary fibrosis," and "connective tissue disease." In addition, all references mentioned in the selected articles were reviewed to identify studies not indexed by the electronic databases. Studies were selected for the analysis if they included patients with CTD-ILD diagnosed on the basis of HRCT findings and sufficient data to calculate the sensitivity, specificity, or correlation coefficients, excluding studies with overlapping or insufficient data, as well as a review study. From the included studies, we extracted information on author(s), publication year, and the demographic characteristics of participants; the total number of B-lines and HRCT Warrick scores; sensitivity, specificity, and correlation coefficients; and cut-off values for B-lines. Lung US assessment was scored based on the number of B-lines, and HRCT Warrick score was obtained by adding the severity and the extension scores (10). We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria to assess the quality of each study (11).

Evaluation of statistical associations

The effect of heterogeneity was quantified by using I^2, with a range from 0% to 100%, representing the proportion of between-study variability attributable to heterogeneity rather than to chance (12). The random-effects model assumes that different studies show substantial diversity and assesses both within-study sampling error and between-study variance (13). We used a random-effects model to combine the sensitivity, specificity, positive (PLR) and negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) estimates due to heterogeneity, and analysed the summary receiver-operating characteristic (SROC) curves, area under the curve (AUC), Q^* index, and

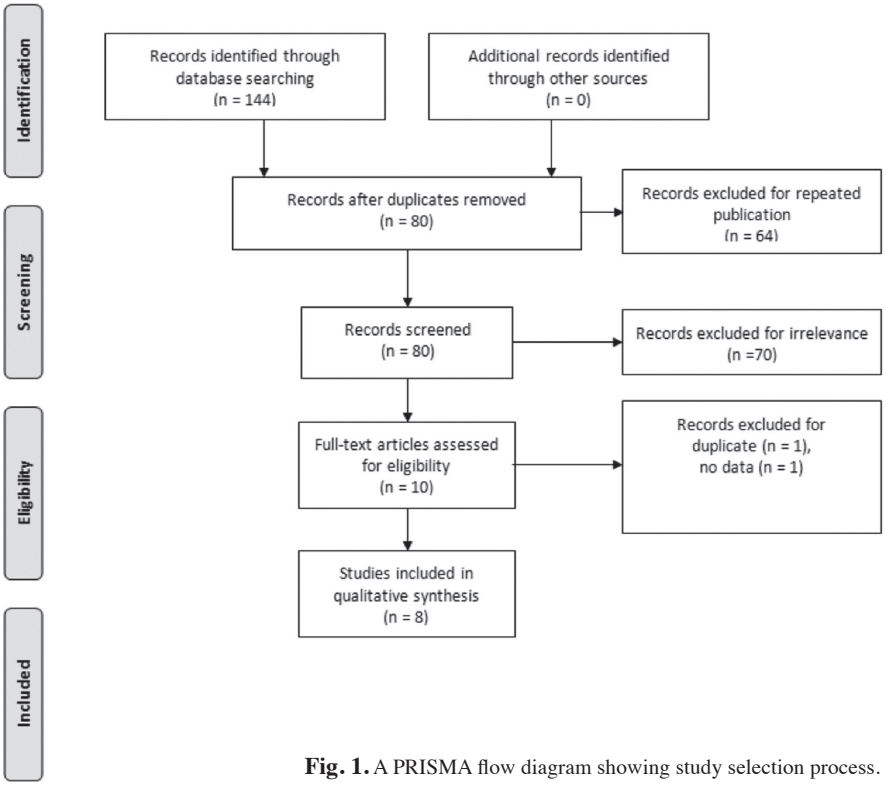

Fig. 1. A PRISMA flow diagram showing study selection process.

correlation coefficient, as the random-effects model is a conservative method. The AUC presents an overall summary of test performance and displays the trade-off between sensitivity and specificity (13). In addition, the Q^* index is another useful global estimate of test accuracy for comparing SROC curves. Statistical manipulations for this meta-analysis were performed by using Meta-DiSc, v. 1.4 (Hospital Universitario Ramon y Cajal, Madrid, Spain) (14) and a comprehensive meta-analysis computer programme (Biosta, Englewood, NJ).

Results

Studies included in the meta-analysis

We identified 144 studies by electronic and manual searching, and 10 were selected for full-text review based on title and abstract. Two of these were excluded because they had no data or included duplicate data with the study by Tardella *et al*. (9). Thus, eight articles met the inclusion criteria, with a total of 349 patients (159 patients with CTD-ILD and 190 patients with CTD) (4, 6-9, 15-17) (Fig. 1). Six of these studies showed the diagnostic accuracy of lung US for diagnosis of CTD-ILD (6-8, 15-17), and four of these studies provided correlation coefficients between lung US B-line scores and HRCT Warrick scores in CTD-ILD patients (4, 6, 7, 9). Table I shows the characteristic features of the studies' participants, as well as the studies' reported quality assessments of diagnostic accuracy.

Diagnostic accuracy of lung ultrasound in CTD-ILD

When the three studies examining the diagnostic accuracy of lung US were considered together, the sensitivity estimates of lung US ranged from 73.58% to 100%, and the specificity estimates ranged from 56% to 100% (Table I). The pooled sensitivity and specificity of lung US were 91.5% (95% confidence interval [CI]: 84.5–96.0) and 81.3% (95% CI: 74.6–86.9), respectively (Table II). In summary, the PLR, NLR, and DOR of lung US were 4.100 (2.133–7.879), 0.176 (0.006–0.363), and 34.73 (10.10–99.66), respectively (Table II). Figure 1 shows the performance of lung US in the form of SROC curves. The AUC of lung US was 0.915 and the Q^* index was 0.848, indicating a high diagnostic accuracy (Table II; Fig. 2). Spearman rank correlation test showed no presence of threshold effects in the lung US meta-analysis (Spearman correlation coefficient: -0.500, p=0.667). Data that were limited to studies of SSc were similar to those from all 6 studies. When all four SSc studies were considered together, the pooled sensitivity and specificity of lung US were 89.5% (95% CI 80.3–95.3) and 79.6% (69.9–87.2), respectively (Table II).

Meta-analysis of correlation coefficients between lung US score and HRCT score in CTD-ILD

Meta-analysis of correlation coefficients from four studies revealed a significant correlation between lung US B-line scores and HRCT Warrick scores in CTD-ILD (correlation coefficient: 0.783; p-value $<1 \times 10^{-9}$) (Table III; Fig. 3). Some heterogeneity was observed, but there was no evidence of publication bias (Table III).

Discussion

High-resolution computed tomography is the gold standard method for the diagnosis, activity, and therapy of ILD (3). However, HRCT uses high doses of ionising radiation, thus increasing the risk of radiation exposure (18). In contrast, US represents a non-ionising, non-invasive, inexpensive, and safe method of assessment, and has been used as an important diagnostic tool in rheumatologic diseases because of its merits (19). Several studies have shown that lung US using B-line assessment may be a reliable additional method for the evaluation of ILD in patients with CTD (4, 6-9); however, its diagnostic value needs to be validated either by studies with large numbers of subjects or by meta-analysis.

In this meta-analysis, we combined evidence of the diagnostic accuracy of lung US for ILD in patients with CTD. This meta-analysis of five studies with a total of 249 patients showed that lung US has a high diagnostic accuracy. The pooled sensitivity and specificity of lung US were 91.5% and 81.3%, respectively, and the AUC and the Q^* index were 0.915 and 0.848, respectively, indicating highly accurate diagnostic

Table I. Characteristics of individual studies included in meta-analysis.

A. Characteristics of individual studies

Authors	Country	Disease	Numbers		Probe frequency (MHz)	Cut-off value (B-line)	Lung ultrasound				Study quality
			ILD (+)	ILD (-)			Sensitivity	Specificity	Correlation coefficient	Sample size	
Moazedi, 2015 (8)	Austria	SSc	9	5	3,5	NA	81.8	100	NA	14	10
Moazedi, 2014 (15)	Austria	RA	17	47	3–5	NA	97.1	97.1	NA	64	10
Cogliati, 2014 (6)	Italy	RA	13	26	2–5	10	92	56	0.806	39	11
Mohammadi, 2014 (7)	USA	SSc	17	53	7–10	5	73.58	88.23	0.695	70	7
Barskova, 2013 (16)	Italy	SSc	36	22	2.5-3.5	5	100	59	NA	58	11
Delle, 2010 (17)	Italy	SSc	14	13	2.5-3.5	5	85	70	NA	27	7
Gargani, 2009 (4)	Italy	SSc	17	16	2.5–3.5	10	NA	NA	0.72	33	11
Tardella, 2012 (9)	Italy	CTD*	36	8	2–7	10	NA	NA	0.875	44	10

ILD: interstitial lung disease; RA: rheumatoid arthritis; SSc: systemic sclerosis; and CTD: connective tissue disease; * 26 SSc; NA: not avaialble.

B. Demographic characteristics of study populations

Authors	Mean age (range), years±standard deviation	Number of LUS operators	Reproducibility between operators	Number of intercostal spaces or regions studied
Moazedi, 2015 (8)	54(28-74)	2	NA	18
Moazedi, 2014 (15)	59 ± 12	2	Kappa = 0.92	18
Cogliati, 2014 (6)	64.87 ± 9.9	2	Kappa =0.78	Each intercostal space
Mohammadi, 2014 (7)	50.29 ± 9.7	≥ 2	Kappa = 0.838	10
Barskova, 2013 (16)	51 ± 15	2	Inter-observer variability 7.4%	8
Delle, 2010 (17)	53 ± 10.5	2	Intra-class correlation = 0.681	8
Gargani, 2009 (4)	53 ± 14	2	Inter-observer variability 7.4%	8
Tardella, 2012 (9)	57.0 ± 12.96	2	Kappa = 0.46-0.980	14

LUS: lung ultrasound; NA: not available.

C. Study quality using QUADAS tool.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Sum
Moazedi, 2015 (8)	Y	Y	Y	U	Y	Y	Y	Y	N	Y	U	Y	Y	U	10
Moazedi, 2014 (15)	Y	Y	Y	U	Y	Y	Y	Y	N	Y	U	Y	Y	U	10
Cogliati, 2014 (6)	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Y	U	11
Mohammadi, 2014 (7)	Y	U	U	U	Y	Y	Y	N	U	Y	U	Y	Y	U	7
Barskova, 2013 (16)	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Y	U	11
Delle, 2010 (17)	Y	U	Y	U	Y	Y	Y	N	N	Y	U	Y	U	U	7
Gargani, 2009 (4)	Y	Y	Y	U	Y	Y	Y	Y	Y	Y	U	Y	Y	U	11
Tardella, 2012 (9)	Y	Y	Y	U	Y	Y	Y	Y	N	Y	U	Y	Y	U	10

QUADAS: Quality Assessment of Diagnostic Accuracy Studies, Y: Yes, N: No, U: Umclear, Item 1. Was the spectrum of patients representative of the patients who will receive the test in practice? 2. Were selection criteria clearly described? 3. Is the reference standard likely to correctly classify the target condition? 4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? 5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis? 6. Did patients receive the same reference standard regardless of the index test result? 7. Was the reference standard independent of the index test (*i.e.* the index test did not form part of the reference standard)? 8. Was the execution of the index test described in sufficient detail to permit replication of the test? 9. Was the execution of the reference standard described in sufficient detail to permit its replication? 10. Were the index test results interpreted without knowledge of the results of the reference standard? 11. Were the reference standard results interpreted without knowledge of the results of the index test? 12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? 13. Were uninterpretable/ intermediate test results reported? 14. Were withdrawals from the study explained?

Table II. Summary results of meta-analysis.

Method Population	Study No.	Numbers		Sensitivity (95% CI)	Heterogeneity		Specificity (95% CI)	Heterogeneity		PLR (95% CI)	NLR (95% CI)	DOR (95% CI)	AUC (SE)	Q* (SE)
		CTD-ILD	CTD		I^2	*p*-value		I^2	*p*-value					
Lung US Overall	6	106	166	0.915 (0.845-0.960)	66.6	0.010	0.813 (0.746-0.869)	84.1	<0.001	4.100 (2.133-7.879)	0.176 (0.006-0.363)	34.73 (10.10-99.66)	0.915 (0.026)	0.848 (0.030)
Lung US SSc	4	76	93	0.895 (0.803-0.953)	74.1	0.009	0.796 (0.699-0.872)	72.5	0.012	3.523 (1.965-6.315)	0.223 (0.117-0.425)	25.86 (9.491-70.48)	0.902 (0.033)	0.834 (0.036)
Lung US RA	2	30	73	0.967 (0.828-0.999)	41.8	0.190	0.836 (0.730-0.912)	95.0	<0.001	7.726 (0.254-235.0)	0.080 (0.017-0.386)	111.4 (1.853-6696)	NA	NA

CTD: connective tissue disease; ILD: interstitial lung disease; SSc: systemic sclerosis; RA: rheumatoid arthritis; US: ultrasound; PLR: positive likelihood ratio; NLR: negative likelihood ratio; DOR: diagnostic OR; AUC: area under the curve; CI: confidence interval; SE: standard error; NA: not available; and US: ultrasound. The number 1 means 100% in sensitivity and specificity.

Table III. Meta-analysis of correlation coefficient between US B-line score and HRCT score.

Method	No. of studies	No. of subjects	Test of association			Test of heterogeneity			Publication bias *p*-value
			Correlation coefficient	95% CI	*p*-value	Model	I^2	*p*-value	
Lung US *vs*. HRCT	4	186	0.783	0.676–0.858	$<1 \times 10^{-8}$	R	57.1	0.072	0.589

US: ultrasound; HRCT: high-resolution computed tomography; CI: confidence interval; and R: random-effects model.

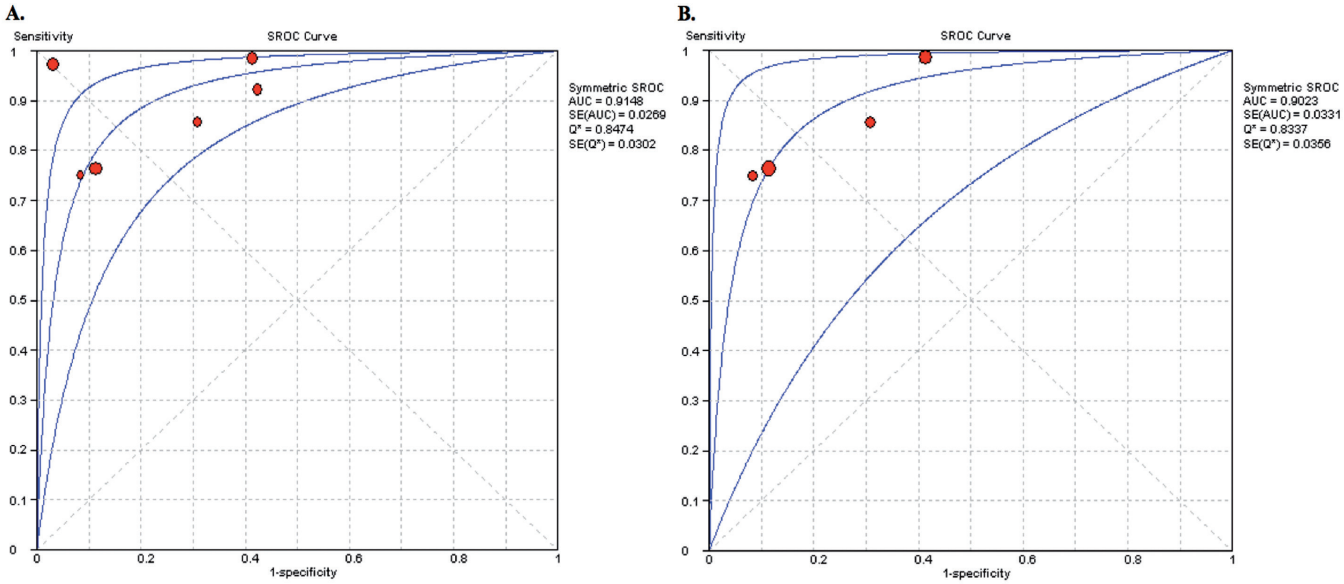

Fig. 2. Summary receiver-operating characteristic curves for lung ultrasound for the diagnosis of interstitial lung disease in connective tissue diseases (**A**) and systemic sclerosis (**B**). Solid circles represent individual studies included in this meta-analysis. The curve shown is a regression line that summarises the overall diagnostic accuracy. SE (AUC): standard error of the area under the curve; Q^*: an index defined by the point on the SROC curve where the sensitivity and specificity are equal; and SE (Q^*): Q^* index standard error.

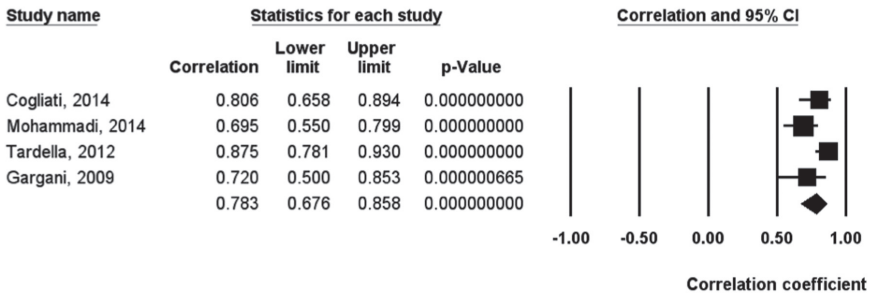

Fig. 3. Meta-analysis of the correlation coefficient between ultrasound B-line score and high-resolution computed tomography Warrick score.

performances. In addition, the results of lung US have shown promising correlation with HRCT findings, the gold standard method. These meta-analysis data suggest lung US can be used as part of the diagnosis of CTD-ILD.

To the best of our knowledge, our meta-analysis represents the first study providing combined evidence for the diagnostic performance of lung US using B-lines for detection of ILD in CTD. However, lung US assesses only the lung surface, while HRCT is able to assess the entire lung parenchyma. Thus, lung US may be useful as an adjunct method in monitoring patients with CTD-IDL during both initial treatment and follow-up, because it has no radiation exposure risk.

This study has several shortcomings that should be considered. First, the number of studies included was small and heterogeneous in demographic characteristics, such as mean age of the patients, the number of LUS operators, the reproducibility between operators, and the number of intercostal spaces. The machines and probes used are really different as well as the scoring system used. This could affect the final result because of the definition of B-lines used and the heterogeneity. Second, between-study heterogeneity was encountered in this meta-analysis. This between-study heterogeneity may have affected our results, which may be compounded by the limited information provided on clinical status and disease severity in the populations involved. This limited data did not allow further analysis, such as subgroup analysis or meta-regression. Third, we included most papers using the Warrick score. However, the Warrick score is not the only CT score employed to semiquantify ILD (*e.g.* Wells or Kazerooni scores). Fourth, to summarise the studies in me-

ta-analysis is important, but the measure points of B-lines were different (from 18 to 8 points). It remains unclear whether more measure points are more accurate or the measurement of small number points such as 8 is enough, although a significant positive linear correlation was found between B-lines and Warrick scores (4). Nevertheless, this meta-analysis also has its strengths. The number of the patients from individual studies ranged from 33 to 70; our pooled analysis included a total of 250 patients. Compared to individual studies, our study was able to provide more accurate data on the diagnostic tests by increasing the statistical power and resolution through pooling the results of independent analyses.

Conclusion

The present meta-analysis of lung US for the detection of ILD in 349 patients with CTD demonstrates that lung US has a high diagnostic accuracy, correlates well with HRCT findings, and is a valuable, additional option in the diagnosis of CTD-ILD. Given the small number of included studies, however, further studies with large population samples are needed to definitively determine the diagnostic value of lung US for CTD-ILD.

References

1. TRAVIS WD, KING JR T, BATEMAN E *et al*.: American Thoracic Society/European Respiratory Society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2002; 162: 277-304.
2. SHAHIN AA: Pulmonary involvement in systemic sclerosis. *Treat Respir Med* 2006; 5: 429-36.
3. DEVARAJ A, WELLS AU, HANSELL DM: Computed tomographic imaging in connective tissue diseases. *Semin Respir Crit Care Med* 2007; 389-97.
4. GARGANI L, DOVERI M, D'ERRICO L *et al*.: Ultrasound lung comets in systemic sclerosis: a chest sonography hallmark of pulmonary interstitial fibrosis. *Rheumatology* 2009; 48: 1382-7.
5. VOLPICELLI G, ELBARBARY M, BLAIVAS M *et al*.: International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med* 2012; 38: 577-91.
6. COGLIATI C, ANTIVALLE M, TORZILLO D *et al*.: Standard and pocket-size lung ultrasound devices can detect interstitial lung disease in rheumatoid arthritis patients. *Rheumatology* (Oxford) 2014; 53: 1497-503.
7. MOHAMMADI A, OSHNOEI S, GHASEMI-RAD M: Comparison of a new, modified lung ultrasonography technique with high-resolution CT in the diagnosis of the alveolo-interstitial syndrome of systemic scleroderma. *Med Ultrason* 2014; 16: 27-31.
8. MOAZEDI-FUERST F, KIELHAUSER S, BRICKMANN K *et al*.: Sonographic assessment of interstitial lung disease in patients with rheumatoid arthritis, systemic sclerosis and systemic lupus erythematosus. *Clin Exp Rheumatol* 2015; 33 (Suppl 91): S87-91.
9. TARDELLA M, GUTIERREZ M, SALAFFI F *et al*.: Ultrasound in the assessment of pulmonary fibrosis in connective tissue disorders: correlation with high-resolution computed tomography. *J Rheumatol* 2012; 39: 1641-7.
10. WARRICK JH, BHALLA M, SCHABEL SI, SILVER RM: High resolution computed tomography in early scleroderma lung disease. *J Rheumatol* 1991; 18: 1520-8.
11. WHITING P, RUTJES AW, REITSMA JB, BOSSUYT PM, KLEIJNEN J: The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; 3: 25.
12. HIGGINS JP, THOMPSON SG: Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; 21: 1539-58.
13. DER SIMONIAN R, LAIRD N: Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177-88.
14. ZAMORA J, ABRAIRA V, MURIEL A, KHAN K, COOMARASAMY A: Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006; 6: 31.
15. MOAZEDI-FUERST F, KIELHAUSER S, SCHEIDL S *et al*.: Ultrasound screening for interstitial lung disease in rheumatoid arthritis. *Clin Exp Rheumatol* 2014; 32: 199-203.
16. BARSKOVA T, GARGANI L, GUIDUCCI S *et al*.: Lung ultrasound for the screening of interstitial lung disease in very early systemic sclerosis. *Ann Rheum Dis* 2013; 72: 390-5.
17. DELLE SEDIE A, DOVERI M, FRASSI F *et al*.: Ultrasound lung comets in systemic sclerosis: a useful tool to detect lung interstitial fibrosis. *Clin Exp Rheumatol* 2009; 28 (Suppl. 62): S54-S.
18. SARMA A, HEILBRUN ME, CONNER KE, STEVENS SM, WOLLER SC, ELLIOTT CG: Radiation and chest CT scan examinations: what do we know? *Chest* 2012; 142: 750-60.
19. SOLDATI G: Sonographic findings in pulmonary diseases. *Radiol Med* 2006; 111: 507-15.