High-frequency Ultrasound Assessment of Systemic Sclerosis Skin Involvement: Intraobserver Repeatability and Relationship With Clinician Assessment and Dermal Collagen Content

Victoria A. Flower[1] , Shaney L. Barratt[2] , Darren J. Hart[3] , Amanda B. Mackenzie[4],
Jacqueline A. Shipley[3], Stephen G. Ward[5] , and John D. Pauling[1]

ABSTRACT. **Objective.** The modified Rodnan skin score (mRSS) remains the preferred method for skin assessment in systemic sclerosis (SSc). There are concerns regarding high interobserver variability of mRSS and negative clinical trials utilizing mRSS as the primary endpoint. High-frequency ultrasound (HFUS) allows objective assessment of cutaneous fibrosis in SSc. We investigated the relationship between HFUS with both mRSS and dermal collagen.

Methods. Skin thickness (ST), echogenicity, and novel shear wave elastography (SWE) were assessed in 53 patients with SSc and 15 healthy controls (HCs) at the finger, hand, forearm, and abdomen. The relationship between HFUS parameters with mRSS (n = 53) and dermal collagen (10 patients with SSc and 10 HCs) was investigated. Intraobserver repeatability of HFUS was calculated using intraclass correlation coefficients (ICCs).

Results. HFUS assessment of ST (hand/forearm) and SWE (finger/hand) correlated with local mRSS at some sites. Subclinical abnormalities in ST, echogenicity, and SWE were present in clinically uninvolved SSc skin. Additionally, changes in echogenicity and SWE were sometimes apparent despite objectively normal ST on HFUS. ST, SWE, and local mRSS correlated strongly with collagen quantification (r = 0.697, 0.709, 0.649, respectively). Intraobserver repeatability was high for all HFUS parameters (ICCs for ST = 0.946–0.978; echogenicity = 0.648–0.865; and SWE = 0.953–0.973).

Conclusion. Our data demonstrate excellent reproducibility and reassuring convergent validity with dermal collagen content. Detection of subclinical abnormalities is an additional benefit of HFUS. The observed correlations with collagen quantification support further investigation of HFUS as an alternative to mRSS in clinical trial settings.

Key Indexing Terms: collagen, skin, systemic sclerosis, ultrasonography

Aberrant tissue remodeling is a pathological hallmark of systemic sclerosis (SSc), resulting in fibrosis of skin and organs. SSc skin pathology is complex, comprising 3 distinct pathological phases.[1,2] An early inflammatory phase with cutaneous edema[1,3,4] often manifests as "puffy fingers." This evolves into a longer indurative phase,[1] where increased dermal collagen deposition causes thickened fibrotic skin.[5] The indurative fibrotic phase typically plateaus clinically before entering an atrophic stage, characterized by natural regression of fibrosis.[4] At this stage, lesional skin may be thinner than the skin of healthy controls (HCs), but tethering to the underlying subcutis can render the skin immobile, confounding clinical judgement of skin thickness.[1]

The modified Rodnan skin score (mRSS)[6] is a quick, noninvasive clinician assessment that correlates with histological grading

of skin fibrosis.[7] The mRSS is used in routine clinical practice and is associated with functional disability[8] and survival.[9,10] It remains the preferred method for assessment of skin involvement in clinical trials. The method is subjective and, while the intraobserver variability is acceptable,[11] there is poor interobserver agreement,[11,12] which has implications on its reliability and sensitivity to change as an outcome measure in clinical trials. A number of recent trials of promising antifibrotic interventions have failed to demonstrate statistically significant improvements in mRSS despite encouraging preclinical data and apparent improvement in composite clinical endpoints.[13,14] The mRSS also lacks the sensitivity to differentiate between clinically indistinguishable pathological phases such as inflammatory edema vs established fibrosis, which may also influence trial findings.[15]

Ultrasound (US) was proposed as an objective method for assessing skin > 40 years ago.[16] There has been more recent research interest in US for assessing quantitative and qualitative changes in SSc skin, including skin thickness (ST),[17–24] stiffness (elastography),[21,25,26] and edema (low echogenicity).[3,18,19] The objective nature of US potentially overcomes limitations of the mRSS. Nonetheless, studies have reported poor agreement between local mRSS assessment and objective ST using high-frequency US (HFUS; > 15 MHz).[22] Most previous studies of US elastography used manual displacement of the skin (prone to procedural variation) to assess strain elastography. Newer shear wave technology measures the propagation of a "pushing" pulse generated by the US transducer through the tissue, providing consistency and reducing influence from subcutaneous tissues.[27,28,29] The interrelationship between ST, echogenicity, and elastography has not been fully explored. Further, no previous studies have directly correlated US assessment with histological analysis of dermal collagen deposition. The aim of this study was to further explore the potential of HFUS as a noninvasive tool for assessing SSc skin involvement. Specifically, we have examined the interrelationship between HFUS parameters, local mRSS, and histological dermal collagen content.

METHODS

Detailed methodology is available in Supplementary Material 1 (available with the online version of this article). Patients fulfilling the 2013 American College of Rheumatology/European League Against Rheumatism classification criteria[30] were enrolled from the SSc clinic at the Royal National Hospital for Rheumatic Diseases, Bath, UK. HCs were recruited from members of staff and relatives of participants with SSc. All participants underwent clinical and HFUS assessments.

Ethical approval. Regulatory approval was granted from the NHS Research Ethics Committee (reference 14/SW/1165). All study procedures were undertaken in line with the Declaration of Helsinki with written informed consent prior to study enrollment.

Clinical and US assessment. All study investigations were performed by the same observer (VAF). The total and local mRSS[1] at each US region of interest (ROI) were assessed by a single observer (VAF). US was performed using standardized settings on the same device (Toshiba APLIO A500) for shear wave elastography (SWE; 14 MHz), ST, and echogenicity (both 18 MHz) at 4 ROI: proximal middle finger, dorsal hand, distal dorsal forearm, and abdominal epigastrium. Variability of HFUS was assessed for each parameter, at each ROI, in all participants.

US image analysis. ST (mm) at the hand, forearm, and abdomen was measured as the distance between the external surface of the epidermis and dermo-subcutis interface. Due to challenges identifying the dermo-subcutis junction in some participants with SSc, the distance between the external surface of the epidermis down to the finger extensor tendon (clearly visible in all participants) was used to measure ST in the finger, consistent with previous studies.[31] Dermal echogenicity was recorded as the mean brightness of grayscale (scale 0–255) using Image J (https://imagej.nih.gov/ij/), such that low echogenicity suggested tissue edema and increased echogenicity was suggestive of fibrosis. Images were analyzed for ST and echogenicity in batches and without reference to mRSS scores. For SWE, the APLIO A500 built-in software calculated the mean SWE (kPa).

Skin biopsy and semiquantification of collagen density. Optional skin biopsies were obtained from the forearm at the site of HFUS assessment in accordance with a purposive sampling framework that aimed to capture a mix of early/late and limited/diffuse SSc. Formalin-fixed, paraffin-embedded tissue sections were stained for collagen with Masson trichrome (Sigma) and quantified using Image J to calculate a mean intensity (grayscale, 0–255) of blue color across the tissue section, integrated density (mean intensity × blue pixel area), and total sum of the grayscale (total sum of the intensity of each blue pixel). Histological ST was measured on the same tissue sections using Image J.

Statistical analysis. Statistical analyses were performed using SPSS (v24; IBM Corp.). Normally distributed patient demographics were compared using parametric tests including independent *t* test and chi-square test. Statistical comparison of ultrasound data between patients with SSc (whole group) and HCs utilized Mann-Whitney *U* test (presented as group median [IQR]). Comparisons made with grades of local mRSS utilized Kruskal-Wallis tests. Posthoc analysis (Dunn test) was applied for comparison between 2 individual groups when Kruskal-Wallis achieved significance.

Repeat assessments were used to calculate intraclass correlation coefficient (ICC) for intraobserver variability. SSc HFUS data for skin thickness at each site was subcategorized according to atrophic skin (< mean +2 SD of controls), thickened skin (> mean +2 SD of controls), or normal skin thickness. Multiple linear regression analysis for the 3 HFUS parameters and mRSS to predict skin collagen content was performed using backwards exclusion (for echogenicity).

RESULTS

Participants. Fifty-three patients with SSc were enrolled (45 with limited cutaneous SSc [lcSSc] and 8 with diffuse cutaneous SSc [dcSSc]) alongside 15 HCs. A summary of the demographics and group median HFUS findings of the participants with SSc and HCs is presented in Table 1. Mean SSc disease duration was 11.7 years, with 26% (14 SSc, including 10 lcSSc) being within 5 years of disease onset or lcSSc to dcSSc transition. Patients with SSc were significantly older and were more likely to be administering vasodilator therapies compared to HCs. Across the whole cohort, skin thickness and SWE were generally higher at each ROI in patients with SSc compared to HCs (Table 1). Dermal echogenicity was lower at the finger but higher at the forearm in patients with SSc compared to HCs. HFUS data values according to mRSS are illustrated in Supplementary Table 1 (available with the online version of this article).

Relationship between HFUS assessment of ST and clinical assessment using mRSS. There was a linear relationship between HFUS median ST and the local mRSS at the hand (*P* = 0.034) with significantly higher skin thickness in patients with mRSS = 1 and mRSS = 2 compared with HCs (*P* = 0.043 and *P* = 0.015, respectively; Figure 1). Similar trends were evident at the finger (*P* = 0.137, not significant [NS]) and forearm (*P* = 0.012;

Table 1. Demographic and ultrasound data for SSc and healthy control groups.

	Healthy Controls, n = 15	Patients with SSc, n = 53	*P*
Age, yrs, mean (SD)	49.6 (8.9)	62.2 (11.2)	< 0.01
Female, n (% of subgroup)	12 (80.0)	47 (88.7)	0.767
Ethnicity, n (% of subgroup)			
White	15 (100)	51 (96.2)	–
Asian	0 (0)	2 (3.8)	0.445
Smoking status, n (% of subgroup)			
Current smoker	1 (6.7)	6 (11.3)	–
Ex-smoker	3 (20)	11 (20.8)	–
Nonsmoker	11 (73.3)	36 (67.9)	0.861
Disease duration, yrs, mean (SD)	–	11.7 (11.6)	–
Subgroup, n (% of SSc)			
LcSSc	–	45 (84.9)	–
DcSSc	–	8 (15.1)	–
Autoantibody, n			
Anticentromere	–	30	–
Anti-Scl70	–	9	–
Anti-RNAP III	–	2	–
Current total mRSS median (IQR)			
LcSSc	–	2 (2–5)	–
DcSSc	–	15 (2–28)	–
Drug exposure (%)			
Current vasodilator use	13.3	66.0	< 0.01
Previous DMARD ever	–	43.4	–
Previous cyclophosphamide	–	18.9	–
Previous rituximab	–	1.9	–
Current DMARD	–	32.1	–
HFUS skin assessment			
Proximal phalanx			
Skin thickness, mm, median (IQR)	2.9 (2.6–3.4)	3.4 (2.8–4.05)	0.047
Echogenicity, grayscale 0–255, mean (IQR)	67 (55–81)	48 (40.5–63.5)	< 0.001
SWE, kPa, mean (IQR)	33.3 (21.8–40.6)	39.8 (32.4–53.3)	0.02
Dorsal hand			
Skin thickness, mm, median (IQR)	1.4 (1.2–1.5)	1.6 (1.3–1.8)	0.052
Echogenicity, grayscale 0–255, mean	55 (37–74)	46 (34.5–58)	0.112
SWE, kPa	27 (15.9–40.4)	38.9 (31.8–50.1)	0.003
Distal forearm			
Skin thickness, mm, median (IQR)	1.5 (1.3–1.7)	1.4 (1.2–1.6)	0.183
Echogenicity, grayscale 0–255, mean	64 (57–71)	73 (63–84)	0.025
SWE, kPa, mean	27.9 (21.1–41.1)	35.4 (28.6–43.2)	0.117
Abdomen			
Skin thickness, mm, median (IQR)	2.0 (1.9–2.5)	1.8 (1.5–2.2)	0.016
Echogenicity, grayscale 0–255, mean	62 (49–81)	69 (61–85)	0.107
SWE, kPa, mean	15.8 (9.4–25.9)	25.9 (19.5–29.4)	0.01

Patients with SSc were significantly older than controls and had a greater use of vasodilator therapy, which was expected. Statistical tests were performed using independent *t* test for continuous demographic data and chi-square test for categorical data. Mean SSc disease duration was 11.7 years. There were 26% of SSc (14 patients with SSc, including 10 with lcSSc) within 5 years of disease onset or transformation from lcSSc to dcSSc. Ultrasound demonstrated variations in skin thickness, echogenicity, and SWE in the disease group compared to controls at a number of regions of interest. All 3 ultrasound parameters were noted to be significantly different from controls at the finger. Statistical comparison between SSc and controls for ultrasound data was performed using Mann-Whitney *U* test. Anti-RNAP III: anti-RNA polymerase III antibody; anti-Scl70: antiscleroderma-70/anti-topoisomerase antibody; dcSSc: diffuse cutaneous systemic sclerosis; DMARD: disease-modifying antirheumatic drug; HFUS: high-frequency ultrasound; lcSSc: limited cutaneous systemic sclerosis; mRSS: modified Rodnan skin score; SSc: systemic sclerosis; SWE: shear wave elastography.
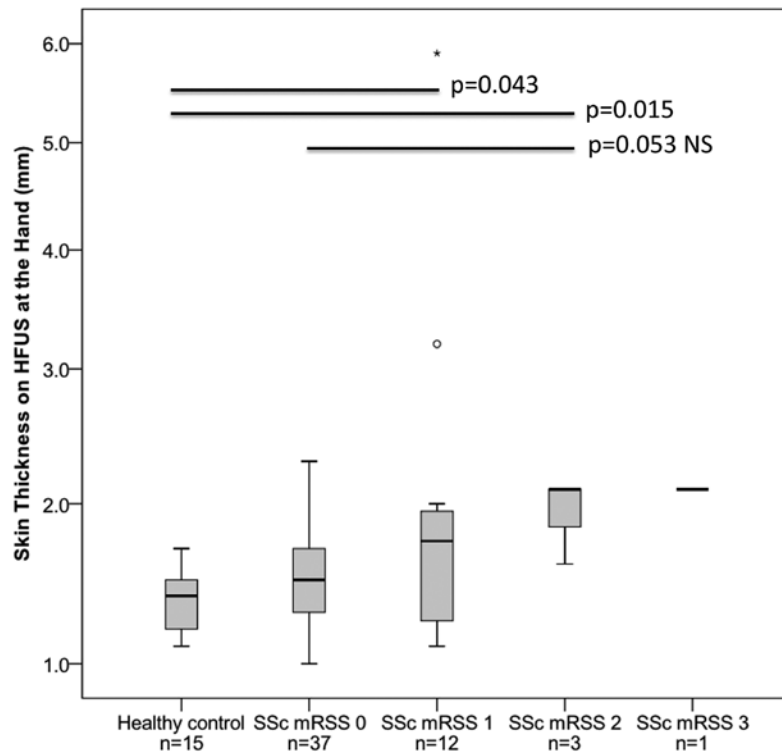
Figure 1. Objective HFUS assessment of skin thickness according to local mRSS at the hand. There were strong trends for objective hand skin thickness on HFUS to increase from normal HC skin across increasing skin scores in patients with SSc ($P = 0.034$, Kruskal-Wallis). *P* values shown in between 2 groups represent posthoc analysis (Dunn test). Circle and asterisk illustrate group outliers. HC: healthy controls; HFUS: high-frequency ultrasound; mRSS: modified Rodnan skin score; NS: not significant; SSc: systemic sclerosis.

Supplementary Table 1, available with the online version of this article). For example, the skin thickness at the forearm in patients with mRSS = 1 was significantly higher than in patients with mRSS = 0 ($P = 0.009$). These relationships are further illustrated by weak–moderate correlation of HFUS ST with local mRSS at the hand and forearm (Table 2), but not finger or abdomen. Similar correlation with ST is extended to the total mRSS (Table 2).

When mRSS = 0, the median ST at the forearm ($P = 0.046$) and abdomen ($P = 0.026$) was actually lower in patients with SSc compared with HCs, reflecting atrophic skin. Skin thickness at the hand and fingers was similar in patients with SSc and HCs when mRSS = 0 (Supplementary Table 1, available with the online version of this article).

With consideration for the influence of age on skin, ST correlated only with age in the control group at the forearm ROI ($r = -0.390$, $P = 0.001$), a site at which there was no significant difference between diagnostic groups (Table 1). There was no correlation between ST and age at any site for patients with SSc, or finger, hand, or abdomen in HCs. Additionally, there was no correlation between echogenicity or SWE with age for either HCs or patients with SSc.

Relationship between US assessment of skin stiffness and clinical assessment using mRSS. There was a linear relationship between

median skin stiffness (SWE) and local mRSS at the finger and hand ($P = 0.039$ and $P = 0.007$, respectively; Figure 2, Supplementary Table 1, available with the online version of this article). Correlation between SWE and local mRSS confirmed these relationships (Table 2). There was a similar trend for median SWE at the forearm and abdomen ($P = 0.062$ NS and $P = 0.035$, respectively; Supplementary Table 1). Correlation with local mRSS did not achieve significance (although numbers of patients with higher skin scores at these sites were low; Table 2). Correlation between SWE and total mRSS was not significant.

SWE values at both the abdomen ($P = 0.035$) and hand ($P = 0.023$; Figure 2) were significantly higher in patients with SSc with mRSS = 0 compared to HC. Further, when skin thickness was objectively normal on HFUS, there was still evidence of increased SWE at both the finger (71.2 kPa [49.0–80.8] vs 33.3 kPa [21.8 40.6]; $P < 0.001$) and hand (36.2 kPa [30.7–42.5] vs 27.0 kPa [15.9–40.4]; $P = 0.042$) compared to HCs. Taken together, these findings indicate that HFUS is capable of identifying aberrant tissue remodeling, even when the dermis is of objectively normal thickness and clinically normal to palpation. We noted that in patients with SSc with high local skin scores at the finger, it was sometimes difficult to obtain an accurate SWE reading on the first attempt. Iagnocco, *et al*, reported similar

Table 2. Correlation between clinical, histological, and ultrasound assessment of skin.

|	Skin Thickness, mm	HFUS Parameters — Echogenicity, Grayscale, Mean	SWE, kPa, Mean	mRSS — Local mRSS	Total mRSS
SSc					
Clinical assessment (n = 53)					
Local mRSS					
Finger	+0.115	+0.002	+0.290*	–	–
Hand	+0.322*	−0.137	+0.324*	–	–
Forearm	+0.410**	−0.232 (P = 0.094)	+0.236 (P = 0.088)	–	–
Abdomen	+0.243 (P = 0.079)	−0.084	+0.055	–	–
Total mRSS[b]					
Finger	−0.004	+0.128	+0.267 (P = 0.05)	–	–
Hand	+0.331*	−0.141	+0.209	–	–
Forearm	+0.357**	−0.337**	+0.154	–	–
Abdomen	+0.148	−0.163	−0.059	–	–
Histological parameters (n = 10)					
Forearm skin thickness	+0.483	−0.444	+0.321	+0.506	+0.407
Forearm collagen					
Area	+0.669*	−0.098	+0.717*	+0.650*	+0.497
Mean intensity	+0.215	−0.128	+0.730	+0.245	+0.366
Integrated density[a]	+0.697*	−0.255	+0.709*	+0.649*	+0.486
Sum of grayscale[b]	+0.697*	−0.255	+0.709*	+0.649*	+0.486
Overall cohort					
Histological parameters (n = 20)					
Forearm skin thickness	+0.417 (P = 0.067)	−0.571**	+0.229	–	–
Forearm collagen					
Area	+0.697**	−0.064	+0.307	–	–
Mean intensity	+0.046	−0.007	+0.182	–	–
Integrated density[a]	+0.735**	−0.136	+0.433 (P = 0.056)	–	–
Sum of grayscale[b]	+0.735**	−0.136	+0.433 (P = 0.056)	–	–

[a] Area × mean intensity of collagen staining. [b] Sum total of staining intensity. Spearman rank correlation coefficient (ρ) is demonstrated between mRSS and ultrasound parameters at 4 regions of interest (finger, hand, forearm, abdomen) for SSc (n = 53) and with histological assessment at the forearm in a subgroup of SSc biopsy donors (n = 10). Correlation between histological assessment of forearm skin and ultrasound is also shown for combined healthy controls and SSc forearm skin donors (n = 20). Significance is illustrated by * $P < 0.05$ and ** $P ≤ 0.01$. HFUS: high-frequency ultrasound; mRSS: modified Rodnan skin score; SSc: systemic sclerosis; SWE: shear wave elastography.

problems for finger elastography in all participants, including controls, and thus concluded that this may be due to interference from the closely underlying bone.[25] In our study, this practical difficulty did not occur with HCs or the patients with SSc with lower local mRSS. Thus, we concluded that it likely occurred due to the severity of pathology within the soft tissues rather than bony interference, which may limit its application for more severe skin lesions. Despite this, overall we still obtained high-quality reproducible data by this technique.

Relationship between HFUS assessment of dermal echogenicity and clinical assessment using mRSS. Unlike ST and SWE, there was not a linear relationship between median echogenicity and local mRSS at any ROI, reflected in a lack of correlation with local mRSS. Total mRSS only correlated with echogenicity at the forearm (r = −0.337, P = 0.014), which might have occurred by chance due to repeated testing. Echogenicity was, however, significantly lower at the finger in patients with SSc with mRSS = 0 (P = 0.028), mRSS = 1 (P < 0.001), or mRSS = 2 (P = 0.023) compared to HCs, suggesting more edema in the former (Figure 2). Further, there was a trend toward reduced echogenicity

in patients with SSc when ST was objectively normal on HFUS at the finger (52.0 [48.0–69.0], n = 7 vs 67.0 [55.0–81.0], P = 0.087). It is possible that this reflects tissue edema within this cohort of predominantly lcSSc. We were unable to identify differences in echogenicity (or SWE) between early and late SSc (</> 3 yrs since first non–Raynaud phenomenon symptom) or correlation with disease duration (data not reported). This may be due to small numbers of participants with early SSc and few treatment-naïve participants (both vasodilator and disease-modifying antirheumatic drugs). Further work is required to explore the potential relationship between HFUS echogenicity and the edematous phase of SSc. Nonetheless, our findings suggest the capacity of HFUS to identify changes in skin quality when skin thickness is otherwise objectively and clinically normal.

Interrelationship between US parameters. There was no correlation between echogenicity and either SWE or ST. Only a small number of weak correlations between ST and SWE were identified (hand: r = +0.454, P = 0.001; finger: r = −0.366, P = 0.007), although this might have been a consequence of multiple testing and was therefore not considered relevant. The general lack of consistent
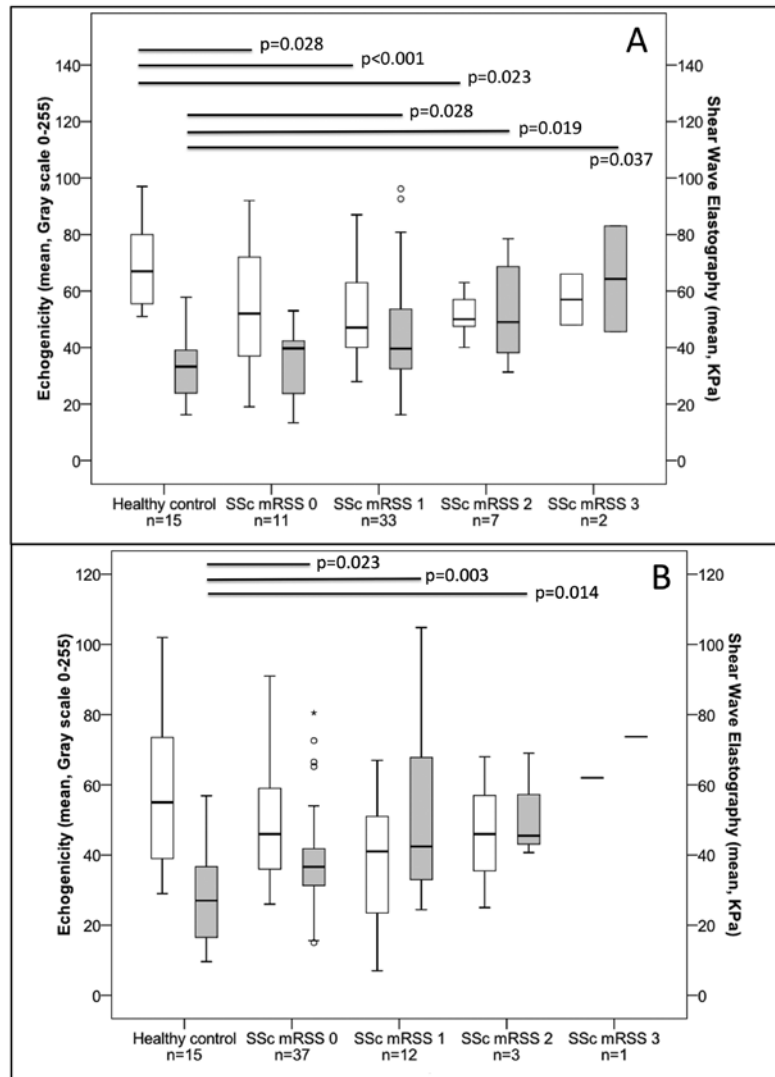
Figure 2. Ultrasound assessment of skin quality according to local mRSS at the finger and hand. Echogenicity (white) and SWE (gray) are shown for the (A) finger and (B) hand. Skin stiffness on SWE increased progressively at each region of interest from controls across increasing local skin scores in patients with SSc ($P = 0.039$ at the finger, $P = 0.007$ at the hand; Kruskal-Wallis). At the finger and hand, echogenicity tended to progressively reduce from HC across mRSS = 0 and mRSS = 1, suggesting cutaneous edema followed by an increase in echogenicity at higher mRSS grading ($P = 0.008$ at the finger, $P = 0.234$ at the hand; Kruskal-Wallis). Numbers for mRSS = 3 were small at both finger and hand. *P* values shown in figure for comparisons between 2 groups represent posthoc analysis (Dunn test). mRSS: modified Rodnan skin score; SSc: systemic sclerosis; SWE: shear wave elastography.

relationships between the 3 parameters may reflect the complex and nonlinear evolution of SSc skin pathology. For example, an area of normal ST in SSc may reflect the early edematous change prior to overt thickening or regression of fibrotic skin.

Histological validation of US and mRSS for assessing skin fibrosis. There were strong correlations between objective measurements of ST on HFUS and each of the following: the area, integrated density, and sum of the grayscale for collagen staining for both the overall cohort (SSc + HC) and the SSc group alone (Table 2). The same patterns of strong correlations were observed between SWE and collagen staining for participants with SSc (Table 2, Figure 3). Of note, SWE associations did not achieve significance for the overall cohort ($P = 0.056$). There was no relationship between echogenicity and collagen content at the forearm, suggesting that increased echogenicity may reflect other features of skin fibrosis.

Local mRSS at the forearm (but not total mRSS) also correlated strongly with collagen quantification in patients with SSc (Table 2). Multiple linear regression analysis confirmed HFUS ST, SWE, and local mRSS at the forearm as significant predictors of local collagen deposition (integrated density) in SSc ($R^2 = 0.891$, $P = 0.001$).
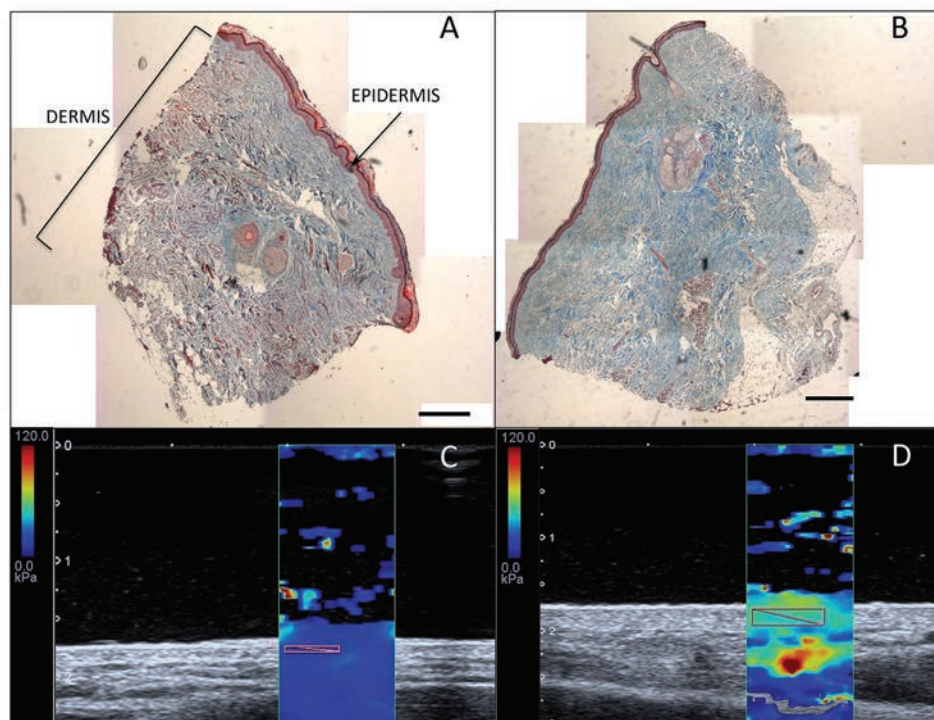
Figure 3. Assessment of skin fibrosis by Masson's Trichrome and elastography. (A, B) Histological collagen content (blue) of forearm skin biopsies are shown by Masson trichrome staining (C,D) with paired shear wave elastography (SWE) assessment at the same site for respective participants, demonstrating skin stiffness. A HC representative of the control group mean for integrated density of collagen staining (A) has low skin stiffness on SWE (C). The participant with SSc with the maximum integrated density of collagen staining for the SSc group shows visibly increased collagen staining in the dermis (B) with increased skin stiffness (D). Scales: trichrome scale bar (500 μm). SWE color scale: stiff (red) to soft (blue). Ultrasound white numeric scale: depth (cm). HC: healthy controls; SSc: systemic sclerosis.

HFUS and histological ST shared a moderate but nonsignificant correlation (whole cohort [n = 20]: r = +0.417, *P* = 0.067; and SSc group only [n = 10]: r= +0.483, *P* = 0.157).

Variability of HFUS measurements. Overall, reproducibility was very good, with very strong correlation between paired measurements for ST (ICC 0.946–0.978) and SWE (ICC 0.953–0.973) and strong correlation for echogenicity (ICC 0.648–0.865; Table 3).

DISCUSSION

We have demonstrated for the first time, to our knowledge, that ST and SWE on US correlate strongly with dermal collagen deposition. Confirmation of excellent reproducibility reported in earlier studies[17,19,20,21,23,25,26] also suggests that the technique is reliable for such an application with appropriate training. Only 1 previous HFUS study undertook paired skin biopsies, which reported binary normal or abnormal skin based on thickening of collagen bundles[32] but made no direct correlation with HFUS

Table 3. Intraobserver variability in ultrasound parameters.

Region of Interest	ICC, mean (95% CI)		
	Skin Thickness, mm	Echogenicity, Grayscale	SWE, kPa
Proximal middle finger	0.946 (0.913–0.967)	0.782 (0.642–0.866)	0.954 (0.926–0.972)
Dorsal hand	0.970 (0.952–0.982)	0.744 (0.586–0.842)	0.953 (0.923–0.971)
Distal forearm	0.978 (0.965–0.987)	0.648 (0.432–0.783)	0.964 (0.941–0.978)
Abdomen	0.963 (0.940–0.977)	0.865 (0.781–0.917)	0.973 (0.955–0.983)

ICC demonstrating excellent reproducibility (intraobserver variability) for skin thickness and shear wave elastography. Reproducibility for echogenicity was good, but marginally less so than the other ultrasound parameters. ICC: intraclass correlation coefficient; SWE: shear wave elastography.

data. One other study reported an association between expression of extracellular matrix markers and increased HFUS ST and echogenicity, but histological analysis was not performed.[18] The lack of correlation between echogenicity and collagen staining in our data suggests that increased echogenicity may reflect other dermal components of fibrotic skin, such as perhaps fibrillin and elastin rather than collagen alone, but may also be influenced by the shorter disease duration in the former study.

Recent work has described distinct gene expression signatures in SSc skin pertaining to the inflammatory and fibroproliferative phases.[33] HFUS may be able to differentiate between such molecular subsets through examination of ST, SWE, and echogenicity. While the technology for efficient machine learning is advancing,[33] gene expression profiling relies on invasive tissue sampling and is not practical to repeat throughout the disease course. Further studies incorporating paired genetic profiling alongside HFUS analysis would be valuable in further developing HFUS as a noninvasive method for assessing SSc skin involvement.

The convergent validity between collagen quantification with ST and mRSS suggests that both noninvasive measures reflect skin fibrosis (at least at the examined forearm). The disparity between the strong correlation of collagen with SWE despite a nonsignificant relationship between SWE, and local mRSS may be a reflection of small numbers of patients with mRSS ≥ 2. This may also account for the nonsignificant relationship between HFUS and histological ST.

We identified HFUS abnormalities in SSc skin quality when ST was either clinically or objectively normal. Previous studies have demonstrated similar findings with ST[3,24,32] and more recently with SWE,[27,34] although we are the first to demonstrate such abnormalities using dermal echogenicity. The overlap and wide ranges of our HFUS data across mRSS groups has been demonstrated previously[3,23] and may be due in part to mixed cohort subtypes and disease duration, but speaks further to the discriminant validity of HFUS over mRSS. The incorporation of HFUS as a surrogate endpoint for skin involvement in SSc clinical trials may overcome issues around the subjectivity and poor interrater reliability of mRSS, potentially allowing treatment efficacy to be assessed in smaller single-blinded studies. Other noninvasive imaging modalities, such as optical coherence tomography, have been proposed as potentially attractive methods for objectively assessing skin involvement in SSc.[35,36,37] The major limitations of all such methods are the need for expensive equipment as well as labor-intensive image acquisition and analysis, particularly if assessment is required at multiple sites.

The lack of correlation between US parameters across the ROI in our data is in contrast to Hesselstrand, et al, who reported a relationship between echogenicity and ST,[18] which may be due to shorter disease duration and proportionally more dcSSc in their cohort. Larger studies are needed for more sensitive subgroup analyses to further explore the relationship between HFUS parameters at different anatomic sites and within different subgroups of SSc. Elastography may be additionally supportive by confirming the dermal interfaces and improving the accuracy and reproducibility of ST measurement.[21]

There are limitations to our study. Notably, the SSc group in our study had a statistically significantly older mean age, which could be considered to have an influence on skin thickness and elasticity. However, normal values for ST according to age have been described using 20 MHz HFUS and shown only 0.04–0.09 mm difference between 50–59 and 60–69 year age groups at our ROI.[38]. Additionally, SWE appears to reduce with aging skin,[34] in contrast to the higher values in our older SSc group. We acknowledge that this is a weakness of the study, but do not feel that it lead to false-positive results. We have demonstrated excellent intraobserver reliability but further work should examine interobserver reliability that may be optimized using defined anatomical landmarks or transient skin surface markings. Additionally, the assessor was not truly blinded to the mRSS scores when performing HFUS, which may influence manual placement of the electronic calipers for ST measurement. We have assessed HFUS in a comparatively large number of subjects with SSc compared to previous studies, but with a minority of patients with early SSc and dcSSc in a single-center, cross-sectional study. SSc is a heterogeneous disease and larger longitudinal studies with greater numbers of treatment-naïve patients with early SSc and dcSSc will be important in defining HFUS parameter evolution, in allowing greater between-group comparisons, and in considering the application for clinical trials. It was surprising that our cohort of largely established patients with SSc demonstrated reduced finger echogenicity. While transition from the edematous to fibrotic phase is generally considered an early event in SSc, the natural history and timing of such transitions has not been fully described and may differ between limited and diffuse subsets. Further longitudinal studies are needed to interrogate this.

US has some limitations as a skin-assessment tool. Ultrasonography requires focused learning of both basic physics and procedural skill. As such, ultrasonography would require standardized assessor certification and standardized protocols if it were to be applied as an outcome in SSc clinical trials. Nonetheless, US is increasingly part of rheumatologists' training,[39,40,41] and acquiring the skills to assess skin is arguably easier than assessing the synovium or major organs.

At 18 MHz, the dermo-subcutis junction at the proximal finger was not easily distinguishable, which we hypothesized may be due to a combination of sclerodermatous changes in the dermis creating a similar echo to the hypodermis, thus reducing the echo interface, as well as pathological subcutaneous fat atrophy in the disease group at a site that naturally has little fatty tissue, even in healthy subjects. Similar issues have been reported in studies using optical coherence tomography, with reduced clarity of the dermoepidermal junction and papillary-reticular dermis interface in the scleroderma disease state.[37] This is further reflected by previous reports of increased interobserver variability in HFUS dermal thickness measurement at the finger compared to other anatomical sites.[19] Higher frequencies of HFUS may provide better skin assessment than those in the lower end of the high-frequency range and improve sensitivity for identifying proximal skin thickening in lcSSc,[24] although additional work is necessary to determine the optimal frequency.

SWE of skin reports a mean quantitative value across the region of interest, which does not reflect the heterogenicity within that tissue, potentially increasing error. The HFUS technology utilized in this study calculates an SD for SWE across the region of interest, but its meaning is easily lost by group analysis under statistical interrogation. This value is perhaps more useful when paired with a visual assessment of the regional color map (Figure 3), to make a subjective interpretation of tissue heterogeneity. Similar reflections are also relevant for echogenicity. Additionally, the relationship between shear wave propagation and the Young modulus makes some assumptions about tissue properties,[29] the exploration of which is beyond the aim of this study. However, our results, combined with recent studies examining pathological skin,[27,42] suggest SWE can be considered a reasonable reflection of fibrotic change.

In conclusion, the observed correlations with collagen quantification support further investigation of HFUS as an alternative to the mRSS. Skin thickness and SWE are highly reproducible parameters and have good convergent validity with local collagen burden. Further, HFUS identifies changes in SSc skin quality at both clinically lesional and nonlesional skin, which may be of benefit in clinical trials. Further studies are needed to include larger numbers of patients with early SSc and dcSSc, with biopsies from additional anatomical locations. Comparison between HFUS parameters and paired genenetic signatures may further support the development of HFUS as a noninvasive tool for assessing SSc skin involvement.

ONLINE SUPPLEMENT
Supplementary material accompanies the online version of this article.

REFERENCES
1. Khanna D, Furst DE, Clements PJ, Allanore Y, Baron M, Czirjak L, et al. Standardization of the modified Rodnan skin score for use in clinical trials of systemic sclerosis. J Scleroderma Relat Disord 2017;2:11-8.
2. Rodnan GP, Lipinski E, Luksick J. Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma. Arthritis Rheum 1979;22:130-40.
3. Hesselstrand R, Scheja A, Wildt M, Akesson A. High-frequency ultrasound of skin involvement in systemic sclerosis reflects oedema, extension and severity in early disease. Rheumatology 2008;47:84-7.
4. Hesselstrand R, Carlestam J, Wildt M, Sandqvist G, Andréasson K. High frequency ultrasound of skin involvement in systemic sclerosis - a follow-up study. Arthritis Res Ther 2015;17:329.
5. Hayes RL, Rodnan GP. The ultrastructure of skin in progressive systemic sclerosis (scleroderma). I. Dermal collagen fibers. Am J Pathol 1971;63:433-42.
6. Brennan P, Silman A, Black C, Bernstein R, Coppock J, Maddison P, et al. Reliability of skin involvement measures in scleroderma. The UK Scleroderma Study Group. Br J Rheumatol 1992;31:457-60.
7. Verrecchia F, Laboureau J, Verola O, Roos N, Porcher R, Bruneval P, et al. Skin involvement in scleroderma--where histological and clinical scores meet. Rheumatology 2007;46:833-41.
8. Steen VD, Medsger TA Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. Arthritis Rheum 1997;40:1984-91.
9. Clements PJ, Lachenbruch PA, Ng SC, Simmons M, Sterz M, Furst DE. Skin score. A semiquantitative measure of cutaneous involvement that improves prediction of prognosis in systemic sclerosis. Arthritis Rheum 1990;33:1256-63.
10. Domsic RT, Rodriguez-Reyna T, Lucas M, Fertig N, Medsger TA Jr. Skin thickness progression rate: a predictor of mortality and early internal organ involvement in diffuse scleroderma. Ann Rheum Dis 2011;70:104-9.
11. Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281-5.
12. Clements PJ, Lachenbruch PA, Seibold JR, Zee B, Steen VD, Brennan P, et al. Skin thickness score in systemic sclerosis: An assessment of interobserver variability in 3 independent studies. J Rheumatol 1993;20:1892-6.
13. Khanna D, Denton CP, Lin CJF, van Laar JM, Frech TM, Anderson ME, et al. Safety and efficacy of subcutaneous tocilizumab in systemic sclerosis: Results from the open-label period of a phase II randomised controlled trial (faSScinate). Ann Rheum Dis 2018;77:212-20.
14. Distler O, Pope J, Denton C, Allanore Y, Matucci-Cerinic M, de Oliveira Pena J, et al. RISE-SSc: riociguat in diffuse cutaneous systemic sclerosis. Respir Med 2017;122 Suppl 1:S14-S17.
15. Hinchcliff M, Huang CC, Wood TA, Matthew Mahoney J, Martyanov V, Bhattacharyya S, et al. Molecular signatures in skin associated with clinical improvement during mycophenolate treatment in systemic sclerosis. J Invest Dermatol 2013; 133:1979-89.
16. Alexander H, Miller DL. Determining skin thickness with pulsed ultra sound. J Invest Dermatol 1979;72:17-9.
17. Scheja A, Akesson A. Comparison of high frequency (20 MHz) ultrasound and palpation for the assessment of skin involvement in systemic sclerosis (scleroderma). Clin Exp Rheumatol 1997; 15:283-8.
18. Hesselstrand R, Westergren-Thorsson G, Scheja A, Wildt M, Akesson A. The association between changes in skin echogenicity and the fibroblast production of biglycan and versican in systemic sclerosis. Clin Exp Rheumatol 2002;20:301-8.
19. Akesson A, Hesselstrand R, Scheja A, Wildt M. Longitudinal development of skin involvement and reliability of high frequency ultrasound in systemic sclerosis. Ann Rheum Dis 2004;63:791-6.
20. Kaloudi O, Bandinelli F, Filippucci E, Conforti ML, Miniati I, Guiducci S, et al. High frequency ultrasound measurement of digital dermal thickness in systemic sclerosis. Ann Rheum Dis 2010;69:1140-3.
21. Di Geso L, Filippucci E, Girolimetti R, Tardella M, Gutierrez M, De Angelis R, et al. Reliability of ultrasound measurements of dermal thickness at digits in systemic sclerosis: role of elastosonography. Clin Exp Rheumatol 2011;29:926-32.
22. Sedky MM, Fawzy SM, El Baki NA, El Eishi NH, El Bohy Ael M. Systemic sclerosis: an ultrasonographic study of skin and subcutaneous tissue in relation to clinical findings. Skin Res Technol 2013;19:e78-84.
23. Sulli A, Ruaro B, Alessandri E, Pizzorni C, Cimmino MA, Zampogna G, et al. Correlations between nailfold microangiopathy severity, finger dermal thickness and fingertip blood perfusion in systemic sclerosis patients. Ann Rheum Dis 2014;73:247-51.
24. Ruaro B, Sulli A, Paolino S, Pizzorni C, Ghio M, Alessandri E, et al. OP0037 evaluation of skin involvement in systemic sclerosis patients by using two ultrasound transducers with different frequency. Ann Rheum Dis 2017;76 Suppl 2:66-7.
25. Iagnocco A, Kaloudi O, Perella C, Bandinelli F, Riccieri V, Vasile M, et al. Ultrasound elastography assessment of skin involvement in systemic sclerosis: lights and shadows. J Rheumatol 2010; 37:1688-91.

26. Cannaò PM, Vinci V, Caviggioli F, Klinger M, Orlandi D, Sardanelli F, et al. Technical feasibility of real-time elastography to assess the peri-oral region in patients affected by systemic sclerosis. J Ultrasound 2014;17:265-9.

27. Sobolewski P, Maślińska M, Zakrzewski J, Paluch Ł, Szymańska E, Walecka I. Applicability of shear wave elastography for the evaluation of skin strain in systemic sclerosis. Rheumatol Int 2020;40:737-45.

28. DeJong HM, Abbott S, Zelesco M, Kennedy BF, Ziman MR, Wood FM. The validity and reliability of using ultrasound elastography to measure cutaneous stiffness, a systematic review. Int J Burns Trauma 2017;7:124-41.

29. Lee SY, Cardones AR, Doherty J, Nightingale K, Palmeri M. Preliminary results on the feasibility of using ARFI/SWEI to assess cutaneous sclerotic diseases. Ultrasound Med Biol 2015;41:2806-19.

30. van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737-47.

31. Sousa-Neves J, Cerqueira M, Santos-Faria D, Afonso C, Teixeira F. Ultrasound assessment of skin thickness performed on fingers: A tool for estimating overall severity of skin disease in systemic sclerosis patients? Acta Reumatol Port 2017;42:339-40.

32. Ihn H, Shimozuma M, Fujimoto M, Sato S, Kikuchi K, Igarashi A, et al. Ultrasound measurement of skin thickness in systemic sclerosis. Br J Rheumatol 1995;34:535-8.

33. Franks JM, Martyanov V, Cai G, Wang Y, Li Z, Wood TA, et al. A machine learning classifier for assigning individual patients with systemic sclerosis to intrinsic molecular subsets. Arthritis Rheumatol 2019;71:1701-10.

34. Santiago T, Santiago M, Coutinho M, Salvador MJ, Da Silva JAP. How much of skin improvement over time in systemic sclerosis is due to normal ageing? A prospective study with shear-wave elastography. Arthritis Res Ther 2020;22:50.

35. Pires NSM, Dantas AT, Duarte AL, Amaral MM, Fernandes LO, Dias TJC, et al. Optical coherence tomography as a method for quantitative skin evaluation in systemic sclerosis. Ann Rheum Dis 2018;77:465-6.

36. Liu CH, Assassi S, Theodore S, Smith C, Schill A, Singh M, et al. Translational optical coherence elastography for assessment of systemic sclerosis. J Biophotonics 2019;12:e201900236.

37. Abignano G, Aydin SZ, Castillo-Gallego C, Liakouli V, Woods D, Meekings A, et al. Virtual skin biopsy by optical coherence tomography: The first quantitative imaging biomarker for scleroderma. Ann Rheum Dis 2013;72:1845-51.

38. Nedelec B, Forget NJ, Hurtubise T, Cimino S, de Muszka F, Legault A, et al. Skin characteristics: normative data for elasticity, erythema, melanin, and thickness at 16 different anatomical locations. Skin Res Technol 2016;22:263-75.

39. Kang T, Emery P, Wakefield RJ. A brief history of ultrasound in rheumatology: where we are now. Clin Exp Rheumatol 2014; 32:S7-11.

40. de Miguel E, Andreu JL, Naredo E, Möller I; Grupo de Trabajo de Ecografía de la Sociedad Española de Reumatología (ECOSER). Ultrasound in rheumatology: where are we and where are we going? Reumatol Clin 2014;10:6-9.

41. Naredo E, D'Agostino MA, Conaghan PG, Backhaus M, Balint P, Bruyn GA, et al. Current state of musculoskeletal ultrasound training and implementation in Europe: results of a survey of experts and scientific societies. Rheumatology 2010;49:2438-43.

42. DeJong H, Abbott S, Zelesco M, Spilsbury K, Martin L, Sanderson R, et al. A novel, reliable protocol to objectively assess scar stiffness using shear wave elastography. Ultrasound Med Biol 2020; 46:1614-29.