High frequency ultrasound assessment of skin involvement in systemic sclerosis: Repeatability and convergent validity with clinical assessment and dermal collagen

Authors

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

1. Department of Rheumatology Royal National Hospital for Rheumatic Diseases, Royal United Hospitals NHS Foundation Trust, Bath, BA1 3NG. United Kingdom.

2. Centre for Therapeutic Innovation & Department of Pharmacy and Pharmacology, University of Bath, Bath, BA2 7AY. United Kingdom.

3. Department of Respiratory Medicine, North Bristol NHS Trust, Bristol, BS10 5NB. United Kingdom.

4. Academic Respiratory Unit, School of Clinical Sciences, University of Bristol, Bristol, BS10 5NB. United Kingdom.

5. Department of Clinical Measurement, Royal National Hospital for Rheumatic Diseases, Royal United Hospitals NHS Foundation Trust, Bath, BA1 3NG. United Kingdom.

Word count main article text: 3499/3500

Corresponding Author

Dr John D Pauling

The Royal National Hospital for Rheumatic Diseases, (part of Royal United Hospitals NHS Foundation Trust), Combe Park, Bath. BA1 3NG. United Kingdom.

Email JohnPauling@nhs.net

Telephone: 0044 1225 821630

ABSTRACT

Objectives: The modified Rodnan skin score (mRSS) remains the preferred method for assessing skin involvement in systemic sclerosis (SSc). There are concerns regarding high inter-observer variability of the mRSS and negative clinical trials utilising the mRSS as the primary endpoint. High frequency ultrasound (HFUS) allows objective assessment of cutaneous fibrosis in SSc. We investigated the repeatability and convergent validity of HFUS with both mRSS and dermal collagen.

Methods Skin thickness (ST), echogenicity and novel Shear Wave Elastography (SWE) were assessed in 53 SSc patients and 15 controls at the finger, hand, forearm and abdomen. Intra-observer repeatability of HFUS was assessed using intra-class correlation coefficients (ICCs). The relationship between HFUS parameters with both mRSS (n=53) and dermal collagen (10 SSc patients and 10 controls) was assessed.

Results HFUS assessment of ST, echogenicity and SWE were related to local mRSS assessment. Subclinical abnormalities in ST, echogenicity and SWE were present in clinically uninvolved SSc skin compared to healthy controls. Changes in echogenicity and SWE were sometimes apparent despite normal ST. The ST and SWE correlated strongly with collagen quantification (rho 0.697 and 0.709 respectively). Intra-observer 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).

Conclusions ST and SWE have good convergent validity with dermal collagen deposition. HFUS abnormalities can be identified in clinically unaffected SSc skin. Intra-observer reproducibility of HFUS assessment is excellent. HFUS provides a reliable and valid alternative to mRSS for objective assessment of skin involvement in SSc.

Abstract word Count: 245/250

Key words

Collagen, Echogenicity, Elastography, Repeatability, Scleroderma, Systemic sclerosis, Ultrasound

Key messages

1. HFUS assessment of skin thickness and SWE has strong convergent validity with dermal collagen in SSc.
2. HFUS can detect sub-clinical changes in skin thickness, echogenicity and stiffness in clinically unaffected skin in SSc.
3. Assessment of skin involvement in SSc using HFUS has high intra-rater repeatability

INTRODUCTION

Aberrant tissue remodelling is a pathological hall mark of systemic sclerosis (SSc), resulting in fibrosis of the skin and other organs. Skin involvement in SSc is complex, comprising three distinct pathological phases [1, 2]. An early 'inflammatory' phase with cutaneous oedema [1, 3, 4] often manifesting as 'puffy fingers'. This evolves into a longer 'indurative' phase [1] where increased dermal collagen deposition results in thickened fibrotic skin [5]. The 'indurative' fibrotic phase typically plateaus clinically before entering an 'atrophic' stage characterised by natural regression of fibrosis [4]. At this stage, lesional skin may be thinner than healthy controls 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, non-invasive clinician assessment that correlates with histological grading of fibrosis on skin biopsy [7]. The mRSS is a useful tool in daily clinical practice that is associated with functional disability [8] and survival in SSc [9, 10]. It remains the preferred method for the assessment of skin involvement in clinical trials. The method is subjective and, whilst the intra-observer variability is acceptable [11], there is poor inter-observer agreement [11, 12], which has important implications for clinical trials. A number of recent trials of promising anti-fibrotic interventions have failed to demonstrate statistically significant improvements in mRSS despite encouraging pre-clinical data and an apparent improvement in composite clinical endpoints [13, 14]. The mRSS also lacks the sensitivity to differentiate between clinically indistinguishable pathological phases such as inflammatory oedema versus established fibrosis, which may also influence clinical trial findings [15].

It is over 40 years since ultrasound was first proposed as an objective method for assessing skin involvement in SSc [16]. The emergence of high frequency ultrasound (HFUS, >15MHz)

has resulted in renewed interest in the use of ultrasound for assessing quantitative and qualitative changes in SSc skin, including skin thickness, stiffness (elastography) and oedema (low echogenicity) [3, 17-23]. Previous studies of ultrasound elastography used manual displacement of the skin (prone to procedural variation) whereas newer technology utilizes a 'pushing' pulse generated by the ultrasound transducer, providing a more consistent displacing waveform. The objective nature of HFUS overcomes limitations of the mRSS. HFUS also has better penetration than other promising non-invasive imaging modalities such as optical coherence tomography [24]. Nonetheless, studies have reported poor agreement between local mRSS assessment and objective skin thickness using HFUS [25]. The inter-relationship between skin thickness, echogenicity and elastography has not been fully explored. Furthermore, no previous studies have directly correlated ultrasound assessment with histological analysis of dermal collagen deposition. The aim of the present study was to further explore the potential of HFUS as a non-invasive tool for assessing SSc skin involvement. Specifically, we have examined the inter-relationship between HFUS parameters, intra-observer repeatability of HFUS and the relationship between HFUS parameters with both local mRSS assessment and histological analysis of dermal collagen content.

PATIENTS AND METHODS

Patients fulfilling the 2013 American College of Rheumatology/European league Against Rheumatism classification criteria [26] were enrolled from the SSc clinic at the Royal National Hospital for Rheumatic Diseases, Bath, UK. Exclusion criteria included the presence of diabetes, current pregnancy or current participation in a clinical drug trial. Previous or concurrent use of DMARDs and vasodilators was permitted. Healthy controls (HC) were recruited from members of staff and relatives of SSc participants. Healthy controls were excluded if they had current diabetes, pregnancy, a diagnosis of any inflammatory rheumatic disease or Raynaud's phenomenon. All participants underwent clinical and HFUS assessments. Skin biopsy was optional (see later).

Ethical approval

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

Clinical assessment

All study investigations were performed by the same observer (VF). Participants were asked to refrain from caffeine, smoking and alcohol for at least 4 hours prior to the assessment, which was for the benefit of concurrent vascular assessment that is not reported herein. A clinician case report form collected patient demographics, clinical phenotype and medication use. The local skin score [1] at each ultrasound region of interest (ROI) was assessed by a single observer (VF).

Ultrasound Image acquisition

Ultrasound imaging was undertaken on the non-dominant side (so usual daily activities were not affected by wound healing in those undergoing skin biopsy). The image capture protocol was adapted from previous studies [3, 25]. All images were obtained by a single observer (VF) using standardised settings on the same device (Toshiba APLIO A500). All assessments took place before 12pm to minimise the impact of diurnal variation in cutaneous fluid content. Shear-wave elastography (SWE) was assessed using a 14 MHz transducer, whereas skin thickness (ST) and echogenicity used a high frequency 18 MHz transducer. The transducers were applied perpendicular to the skin surface. The focus was set within the dermis at the following ROI: dorsal aspect of the proximal phalanx of the middle finger, dorsal hand just proximal to the 2nd and 3rd metacarpophalangeal joints, 7-10cm proximal to the wrist over the dorsal forearm, abdominal epigastrium midway between the xiphisternum and umbilicus.

A visible layer of gel was maintained between the transducer and the skin surface for ST and echogenicity to avoid artefactual variable applied pressure. For SWE, the transducer was applied perpendicular to the skin at each ROI using a 2cm solid gel pad standoff (Aquaflex, 04-02, Parker Laboratories Inc., New Jersey, USA) to allow focus within the correct depth of the tissues as well as a thin layer of transmission gel. Adequate but not excessive pressure was applied to maintain contact but avoid external compression of the skin, which would falsely increase elastography readings. The procedure was repeated separately at each ROI to assess for intra-observer variability of each ultrasound parameter.

Ultrasound image analysis

Electronic callipers within the APLIO A500 in-built software were used to assess ST (millimetres, mm). ST 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 SSc participants, the distance between the external surface of the epidermis down to the finger extensor tendon (clearly visible in all participants) was used to measure skin thickness in the finger, consistent with previous studies [27]. Dermal echogenicity was recorded as the mean 'brightness of Grey scale' (arbitrary scale 0-255) using Image J (https://imagej.nih.gov/ij/), across the whole width of the dermis (to the same depth used for ST measurement), such that low echogenicity suggests tissues oedema and increased echogenicity suggests fibrosis. For SWE, the APLIO A500 in-built software calculated the mean SWE (kPa) within an applied ROI (covering the depth of the dermis and crossing two parallel propagation lines).

Skin biopsy and histology

Optional skin biopsies were obtained from the forearm from the site of HFUS assessment in accordance with a purposive sampling framework that aimed to capture a mix of early/late and limited/diffuse skin changes. Anatomical site and biopsy size (4mm) were chosen to

minimize risk to participants in line with EUSTAR guidelines [28] as well as correspond with site of HFUS assessment. Participants on warfarin or direct oral anticoagulants were excluded from skin biopsy due to the risk of significant bleed in the absence of adrenaline use (a feature included for the benefit of a parallel vascular study). Biopsies were formalin fixed, paraffin embedded and sectioned at 4μm thickness. Skin tissue sections were deparaffinised and stained with Masson's trichrome (Sigma, Saint Louis, USA). Sections were imaged immediately using Leica microscope (CTR40000) and corresponding imaging software (LAS v4.3). Collagen staining was quantified using Image J (https://imagej.nih.gov/ij/) to calculate a mean intensity (Gray scale, 0-255) of blue colour across the tissue section, integrated density (mean intensity x blue pixel area) and total sum of the Gray scale (total sum of the intensity of each blue pixel).

Statistical analysis

Statistical analyses were performed using SPSS v24. Patient demographics were compared using parametric tests including independent t-test and Chi-squared test.

Statistical comparison of ultrasound data between SSc (whole group) and HC utilized Mann-Whitney U (presented as group median [interquartile range]). Comparisons made with grades local mRSS utilized Kruskal-Wallis tests. Post-hoc analysis (Dunn test) was applied for comparison between 2 individual groups when K-W achieved significance.

Repeat assessments were used to calculate intra-class correlation coefficient (ICC) for intra-observer variability. SSc HFUS data for skin thickness at each site was sub-categorized according to atrophic skin (<mean +2 S.D. of controls), thickened skin (>mean + 2 S.D. of controls) or normal skin thickness. Multiple linear regression analysis for the 3 HFUS parameters to predict skin collagen content was performed using backwards exclusion (for echogenicity).

RESULTS

Participants

Fifty-three SSc patients were enrolled (45 limited cutaneous SSc [lcSSc] and 8 diffuse cutaneous SSc [dcSSc]) alongside fifteen healthy controls. A summary of the demographics and group median HFUS findings of the SSc participants and HCs is presented in Table 1. Across the 2 groups, patients with SSc were significantly younger and were more likely to be receiving vasodilator therapies. Across the whole cohort, skin thickness and SWE were generally higher at each ROI in SSc compared to controls (Table 1). Dermal echogenicity was lower at the finger but higher at the forearm in patients with SSc compared to controls.

Relationship between HFUS assessment of skin thickness and clinical assessment using mRSS

There was a linear relationship between HFUS median skin thickness and the local mRSS at the hand (p=0.034) with significantly higher skin thickness in patients with mRSS of 1 and 2 compared with healthy controls (1.75 [1.2-] and 2.1 [1.6-] versus 1.4 [1.2-1.5] respectively) (Figure 1). Similar trends were evident and the finger (p=0.137) and forearm (p=0.012). For example, the skin thickness at the forearm in patients with a mRSS of 1 was significantly higher than in patients with a skin score of 0 (1.6 [1.5-2.1] versus 1.4 [1.2-1.5], p=0.009). When the mRSS was 0, the median skin thickness at the forearm (1.4 [1.2-1.5] vs. 1.5 [1.3-1.7], p=0.046) and abdomen (1.8 [1.5-2.0] vs. 2.0 [1.9-2.5], p=0.026) was actually lower in SSc compared with healthy controls reflecting atrophic skin. Skin thickness at the hand and fingers was similar in SSc and controls when the skin score was 0.

Relationship between ultrasound assessment of skin stiffness and clinical assessment using mRSS

There was a linear relationship between skin stiffness (using SWE) and local mRSS at the finger and hand (p=0.039 and p=0.007 respectively) (Figure 2). There was a similar trend at the forearm and abdomen (p=0.062 and p=0.035 respectively) (although numbers of

patients with higher skin scores at these sites were low, data not shown). SWE values at both the abdomen (25.9 [19.6-29.2], n=51 versus 15.8 [9.4-25.9], p=0.026) and hand (36.6 [31.3-42.5], n=37 versus 27.0 [15.9-40.4], p=0.023) (Figure 2) were significantly higher in patients with SSc with a mRSS of 0 compared to healthy controls. Furthermore, when skin thickness was objectively normal on HFUS there was still evidence of increased SWE at both the finger (71.2 [49.0-80.8] versus 33.3 [21.8-40.6] p<0.001) and hand (36.2 [30.7-42.5] versus 27.0 [15.9-40.4], p=0.042) compared to healthy controls. Taken together, these findings indicate HFUS is capable of identifying aberrant tissue remodeling, even when the dermis is objectively normal thickness and clinically normal to palpation. We noted that in 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 problems for finger elastography in all participants including controls and thus concluded this may be due to interference from the closely underlying bone [22]. In our study, this practical difficulty did not occur with controls or those SSc with lower local mRSS. Thus, we conclude that it occurred due to the severity of pathology within the soft tissues rather than bony interference and may limit its application for more severe skin lesions. Despite this, overall we still obtained good 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 with local mRSS at any ROI. Echogenicity was, however, significantly lower at the finger in SSc with mRSS of 0 (52.0 [33.0-73.0], p=0.028), mRSS of 1 (47.0 [40.0-63.5], p<0.001) or mRSS of 2 (50.0 [46.0-60.0], p=0.023) compared to HC (67.0 [55.0-81.0]) reflecting more oedema (Figure 2). Similarly, echogenicity was significantly increased at the forearm in SSc with mRSS of 0 compared to HC (74.5 [64.8-87.3], n=42 versus 64 [57.0-71.0], p=0.001). Furthermore, there was a trend towards reduced echogenicity in SSc when ST was objectively normal on HFUS at the finger (52.0 [48.0-69.0], n=7 versus 67.0 [55.0-81.0], p=0.087). Taken together, these findings

suggest HFUS is able to demonstrate changes in skin oedema and fibrosis when skin thickness is otherwise objectively and clinically normal.

Contrary to expectations, there were no significant differences between echogenicity or SWE between early and late SSc (</>3 years since first non-RP symptom) or correlation with disease duration (data not reported). This may be due to a combination of small numbers in the early subgroup and few treatment naïve participants (both vasodilator and DMARDs).

Inter-relationship between Ultrasound parameters

There was no correlation between echogenicity and either SWE or skin thickness. Only a small number of weak correlations between skin thickness and SWE were identified (at the hand Spearman's ρ=+0.454, p=0.001 and at the finger ρ=-0.366, p=0.007), although this may have been a consequence of multiple testing and were not considered relevant. The general lack of consistent relationship between the 3 parameters likely reflects the complex and non-linear evolution of SSc skin pathology. For example, an area of normal ST in SSc may reflect uninvolved skin, early oedematous change prior to thickening or regression of fibrotic skin.

Histological validation of ultrasound for assessing skin fibrosis

There were strong correlations between objective measurement of skin thickness on HFUS with each of the area, integrated density and sum of the Gray scale for collagen staining for both the overall cohort (SSc + HC) and SSc alone (Table 2). The same patterns of strong correlations were observed between SWE and collagen staining for SSc participants (Table 2 and 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. Multiple linear regression analysis confirmed skin thickness and SWE at the forearm as significant predictors of local collagen deposition (integrated density) in SSc $R^2 = 0.876$.

Reproducibility of High Frequency Ultrasound for Skin Assessment

Overall, reproducibility was very good; with very strong correlation between paired measurements for skin thickness (Intra-Class Correlation coefficient, 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

To our knowledge, we have demonstrated for the first time that skin thickness and SWE on ultrasound correlate with dermal collagen deposition, which provides significant convergent validation for its use in SSc skin assessment. Additionally, the excellent reproducibility that we have demonstrated like others before [18-23, 25] deems it reliable for such an application with appropriate training. Only one previous HFUS study has also taken anatomically paired skin biopsies and reported binary normal or abnormal skin based on thickening of collagen bundles [29], but no direct correlation with HFUS data was reported. Another study examined expression of ECM markers in patients examined by HFUS, but histological analysis was not performed [30]. This work reported a strong positive correlation between the rate of increase in HFUS skin thickness and echogenicity with fibroblast proteoglycan production [30]. 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.

Whitfield et al., [31] have reported on 4 gene expression profiles in SSc skin pertaining in part to the inflammatory and fibroproliferative phases that HFUS may be able to differentiate between, through examination of ST, SWE and echogenicity. Whilst the technology for efficient machine learning is advancing [31], gene expression profiling relies on invasive tissue sampling and is not practical to repeat throughout the disease course. We have

provided histological validation of ultrasound to reflect collagen deposition that occurs predominantly in the fibrotic phase. Paired study of gene profiling and ultrasound analysis may further advance the application of the latter as a non-invasive virtual biopsy to aid personalized precision medicine of the future.

We have demonstrated that changes in SSc skin quality can be detected by ultrasound when either clinically or objectively of normal thickness. Previous studies have demonstrated similar findings with skin thickness [3, 29, 32]; although we are the first to demonstrate such abnormalities using dermal echogenicity and SWE. There are two likely explanations for the disparity between detection of skin pathology by ultrasound and mRSS. The first is that ultrasound was able to identify sub-clinical skin pathology when skin was clinically felt to be uninvolved by the observer. The second is that mobile atrophic skin was scored as mRSS = 0 due to the lack of provision by the mRSS grading to document skin atrophy. Whatever the exact contribution of these 2 inter-related explanations, ultrasound appears to provide superior objective data regarding dermal pathology to the mRSS alone; with potentially important implications for SSc clinical trials. The incorporation of HFUS as a surrogate endpoint for skin involvement in SSc clinical trials may overcome issues around subjectivity and poor inter-rater reliability; potentially allowing treatment efficacy to be demonstrated more successfully in smaller single-blinded studies. HFUS is unlikely to surpass the mRSS in daily clinical practice, given the labour-intensive approach required for HFUS assessment at multiple sites.

Notably, neither SWE nor echogenicity consistently demonstrated significant changes in skin quality when skin thickness was normal. The lack of correlation between ultrasound parameters across the ROIs in our data, supports that one modality cannot substitute for another. This is in contrast to data from Hesselstrand et al., reporting a relationship between echogenicity and skin thickness [30], which may be due to shorter disease duration in their cohort. We therefore consider that all 3 ultrasound parameters provide complementary data.

Elastography may be additionally supportive by confirming the dermal interfaces and improving the accuracy and reproducibility of skin thickness measurement [20]. Larger multicenter longitudinal studies may be of benefit in further understanding the inter-relationship and clinical relevance of the different ultrasound parameters.

Notably the SSc group in our study had a statistically significantly older mean age, which could be considered to influence skin thickness and elasticity. However, normal values for skin thickness according to age has been described using 20MHz HFUS and shown only 0.04-0.09mm difference between 50-59 and 60-69 year olds at our ROIs [33]. We therefore concluded that whilst our SSc cohort was statistically significantly older than controls, the age difference was not felt to be clinically significant.

There are limitations to our study. We have demonstrated excellent intra-rater reproducibility but have not assessed inter-observer variability. We have assessed HFUS in a comparatively large number of SSc subjects compared to previous studies but SSc is a heterogeneous disease and larger studies would allow greater between-group comparisons to be made. Our study was cross-sectional in nature and larger multicenter longitudinal studies (including treatment naive early patients) shall be important in defining HFUS parameter evolution in SSc skin disease.

At 18MHz, the dermo-subcutis junction at the proximal finger was not easily distinguishable which we hypothesised 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 which naturally has little fatty tissue even in healthy subjects. This is a feature also observed in other studies on OCT with reduced clarity of the dermo-epidermal junction and papillary-reticular dermis interface in the scleroderma disease state [34]. This is further reflected by previous reports of increased inter-observer variability in HFUS dermal thickness measurement at the finger

compared to other anatomical sites [18]. It has been suggested that higher frequencies of HFUS may provide better skin assessment than those in the lower end of the 'high frequency' range and particularly that they may have better sensitivity to identifying proximal skin thickening in lcSSc [32]. Whilst this seems logical, there are few studies to affirm it and therefore further studies are required to determine the optimum frequency for this application whilst balancing against the limited accessibility of very high frequency machines (>50MHz).

In conclusion, the present study has demonstrated excellent repeatability of HFUS and we have reported analyses of the convergent validity of different HFUS parametrers with both clinical assessment using the mRSS and dermal collagen content that suggests HFUS may have potential as a non-invasive virtual biopsy for the assessment of SSc-related skin pathology. We have demonstrated changes in the qualities of SSc skin at both clinically apparent lesional and non-lesional skin. Skin thickness and SWE stand out as highly reproducible parameters that reflect local collagen burden. Future studies comparing HFUS parameters with novel molecular subsets defined by paired gene profile signatures may help to support the use of HFUS in clinical trials, as well as personalizing medical interventions in SSc.

Funding

This work was supported by Scleroderma & Raynaud's UK.

Conflict of interests

Dr John Pauling has received speaker's honoraria and research grant support (>$10,000) from Actelion pharmaceuticals. Dr Pauling has undertaken consultancy work for Actelion pharmaceuticals and Boehringer Ingelheim.

Tables & Figures

Table 1. Demographic and Ultrasound data for SSc and healthy control groups. 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-squared for categorical data. Ultrasound demonstrated variations in skin thickness, echogenicity and Shear Wave Elastography 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. Abbreviations: anti-RNAPIII, anti-RNA Polymerase III antibody; anti-Scl-70, anti-Scleroderma-70/anti-Topoisomerase antibody; dcSSc, diffuse cutaneous systemic sclerosis; DMARD, disease modifying anti-rheumatic drug; IQR, interquartile range; kPa, kilopascal; lcSSc, limited cutaneous systemic sclerosis; mRSS, modified Rodnan skin score; S.D., standard deviation; SWE, Shear wave elastography.

	Healthy control (n=15)	Systemic sclerosis (n=53)	p value
Age, mean, (S.D.)	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 Caucasian	15 (100)	51 (96.2)	
Asian	0 (0)	2 (3.8)	0.445
Smoking status, n. (% of subgroup)			
Current	1 (6.7)	6 (11.3)	
Ex-smoker	3 (20)	11 (20.8)	
Non-smoker	11 (73.3)	36 (67.9)	0.861
Disease duration mean years, (S.D.)	-	11.7 (11.6)	-
Subgroup n. (% of SSc)			
LcSSc	-	45 (84.9)	-
DcSSc	-	8 (15.1)	-
Autoantibody n.			
Anti-centromere	-	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	-
High frequency ultrasound skin assessment			
Proximal Phalanx			
Skin thickness (mm; IQR)	2.9 (2.6-3.4)	3.4 (2.8-4.05)	0.047
Echogenicity (mean, Gray scale 0-255)	67 (55-81)	48 (40.5-63.5)	<0.001
SWE (mean, KPa)	33.3 (21.8-40.6)	39.8 (32.4-53.3)	0.02
Dorsal Hand			
Skin thickness (mm; IQR)	1.4 (1.2-1.5)	1.6 (1.3-1.8)	0.052
Echogenicity (mean, Gray scale 0-255)	55 (37-74)	46 (34.5-58)	0.112
SWE (mean, kPa)	27 (15.9-40.4)	38.9 (31.8-50.1)	0.003
Distal Forearm			
Skin thickness (mm; IQR)	1.5 (1.3-1.7)	1.4 (1.2-1.6)	0.183
Echogenicity (mean, Gray scale 0-255)	64 (57-71)	73 (63-84)	0.025
SWE (mean, kPa)	27.9 (21.1-41.1)	35.4 (28.6-43.2)	0.117
Abdomen			
Skin thickness (mm; IQR)	2.0 (1.9-2.5)	1.8 (1.5-2.2)	0.016
Echogenicity (mean, Gray scale 0-255)	62 (49-81)	69 (61-85)	0.107
SWE (mean, kPa)	15.8 (9.4-25.9)	25.9 (19.5-29.4)	0.01

Figure 1. Objective HFUS assessment of skin thickness according to local mRSS at the hand. There were strong trends for objective skin thickness on HFUS at the hand to increase from HC across the local skin scores in SSc (p=0.034, Kruskal-Wallis). P values shown in figure between 2 groups represent post-hoc analysis (Dunn test). Circle and asterisk illustrate group outliers.

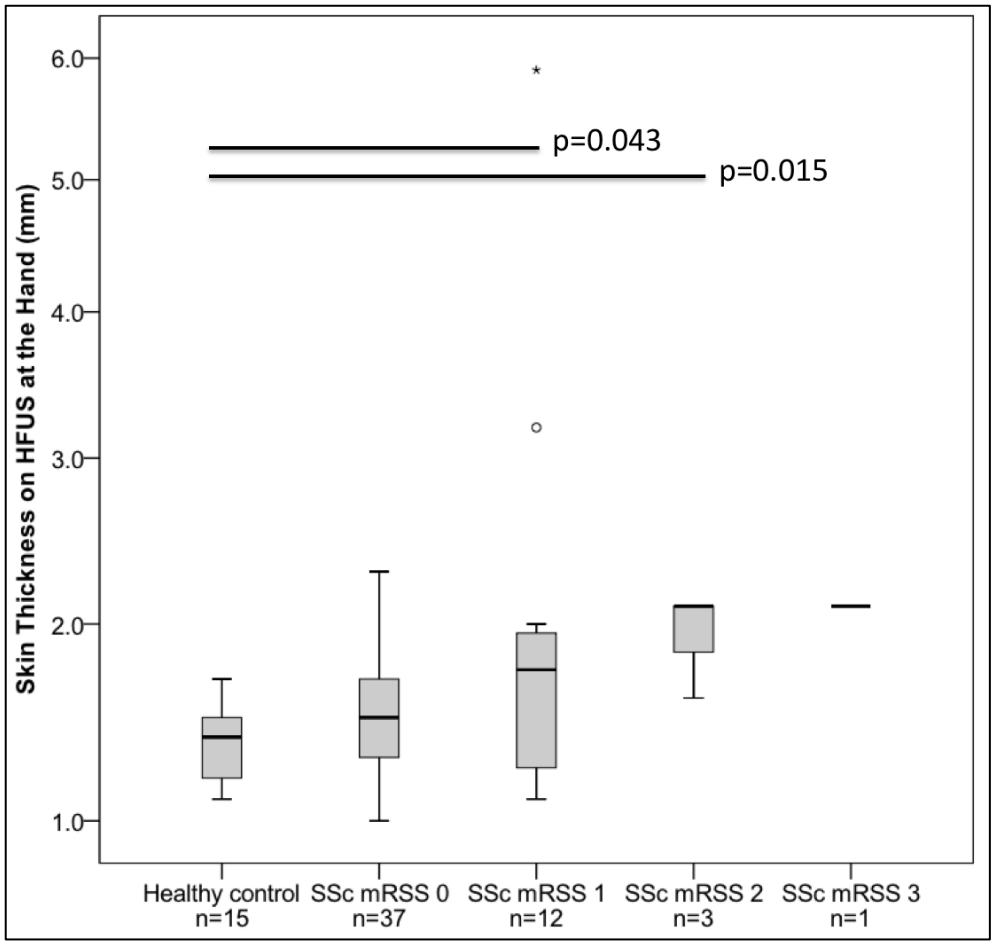

Figure 2. Ultrasound assessment of skin quality according to local mRSS at the finger and hand. Echogenicity (white) and SWE (grey) are shown for the finger (A) and hand (B). Skin stiffness on SWE increased progressively at each region of interest from controls across increasing local skin scores in 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 =1 reflecting cutaneous oedema followed by an increase in echogenicity at higher mRSS grading (p=0.008 at the finger, p=0.234 at the hand; Kruskal-Wallis). P values shown in figure for comparisons between 2 groups represent post-hoc analysis (Dunn test).

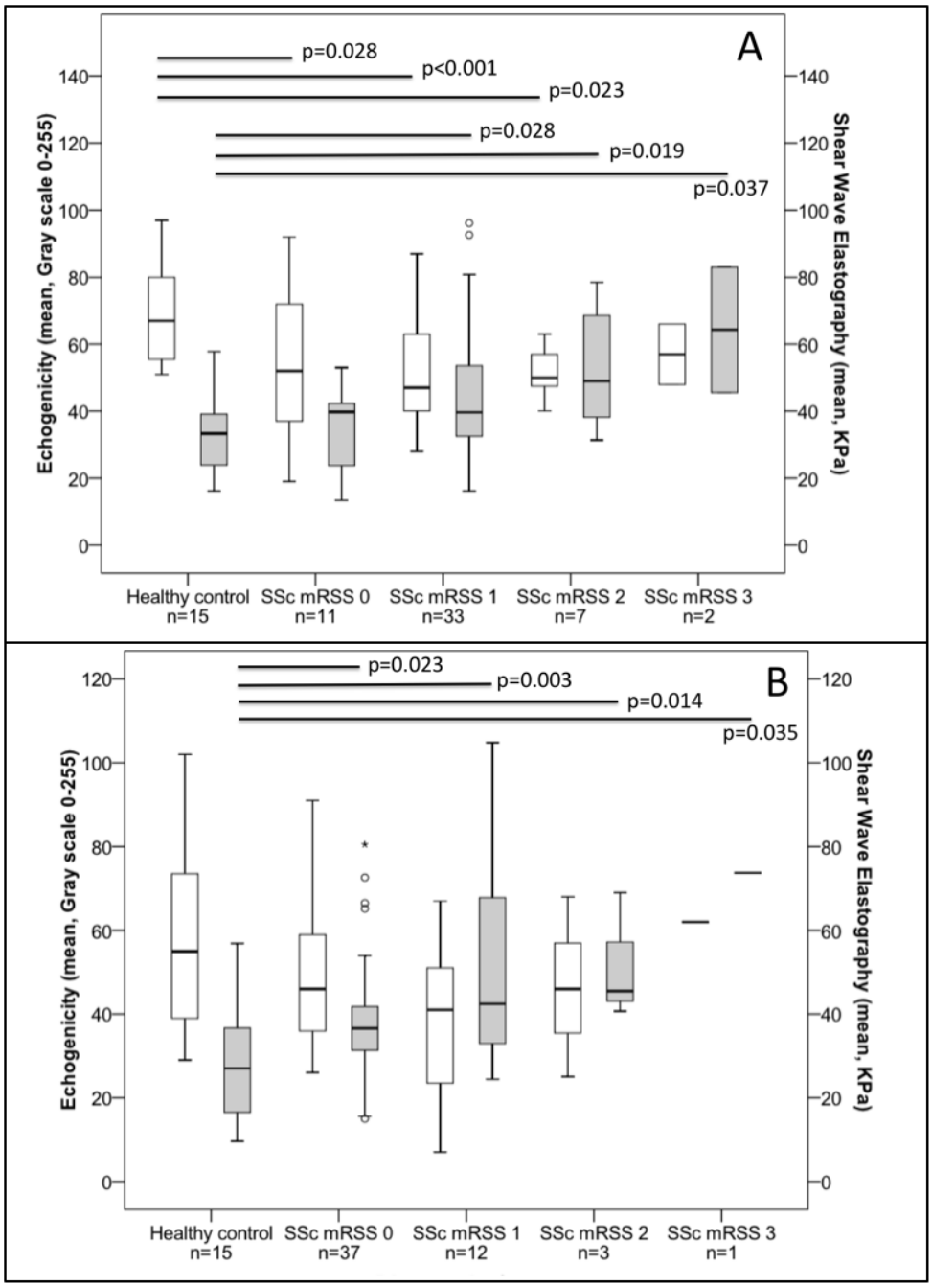

Table 2. Correlation between Dermal Collagen and Ultrasound parameters. Strong positive correlations (Spearman's rank correlation coefficient, ρ) are observed at the forearm between skin thickness and each of the area of collagen staining, the integrated density (area x mean intensity of collagen staining) and sum of the Gray scale (sum total of staining intensity) across the overall study cohort and SSc alone. Similar strong correlations are observed within the SSc group with collagen staining. Significance is illustrated by p<0.05* and p</=0.001**. Abbreviation: [a]Shear Wave Elastography, SWE.

	Skin thickness (mm)	**Echogenicity (mean, Gray scale)**	**SWE[a] (mean, kPa)**
Overall Cohort (Systemic sclerosis + Healthy controls, n=20)			
Area	**+0.697****	-0.064	+0.307
Mean Intensity	+0.046	-0.007	+0.182
Integrated Density	**+0.735****	-0.136	+0.433 (p=0.056)
Sum of Gray scale	**+0.735****	-0.136	+0.433 (p=0.056)
Systemic sclerosis (n=10)			
Area	**+0.669***	-0.098	**+0.717***
Mean Intensity	+0.215	-0.128	+0.73
Integrated Density	**+0.697***	-0.255	**+0.709***
Sum of Gray scale	**+0.697***	-0.255	**+0.709***

Figure 3 Assessment of skin fibrosis by Masson's Trichrome and Elastography. Histological collagen content (blue) of forearm skin biopsies are shown by Masson's Trichrome staining (A, B) with paired Shear Wave Elastography (SWE) assessment at the same site for respective participants (C, D respectively) demonstrating skin stiffness. A healthy control representative of the control group mean for integrated density of collagen staining (A) has low skin stiffness on SWE (C). SSc participant 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, 500μm. SWE colour scale, stiff (red) to soft (blue). Ultrasound white numeric scale, depth (cm).

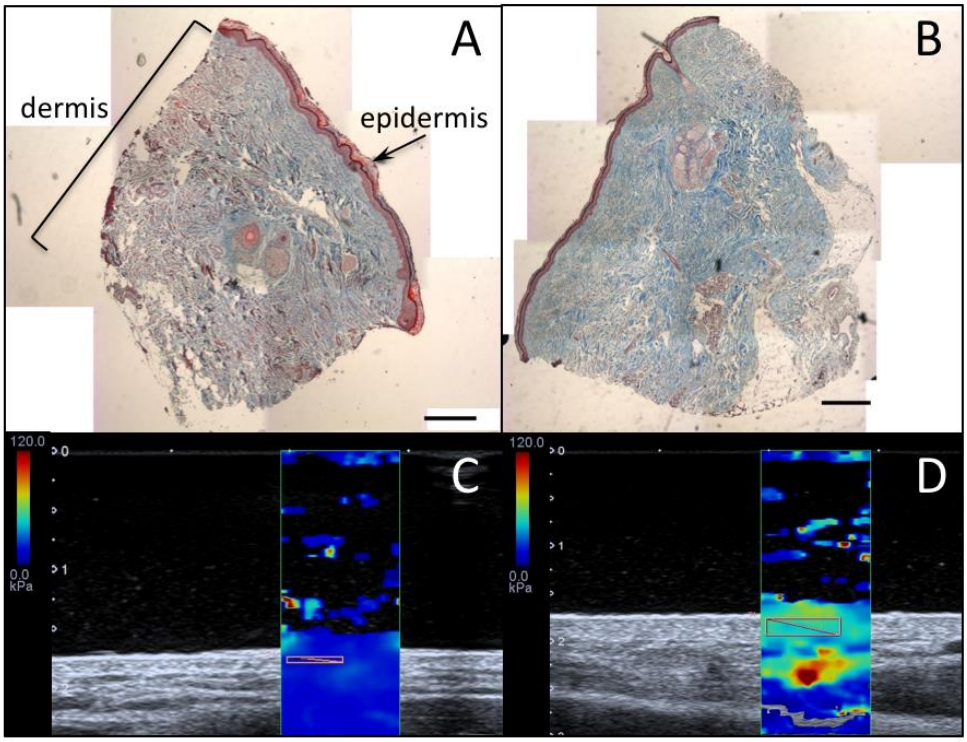

Table 3. Intra-observer variability in Ultrasound parameters. Intra-Class Correlation coefficient demonstrating excellent reproducibility (intra-observer variability) for skin thickness and Shear Wave Elastography. Reproducibility for echogenicity was good, but marginally less so than the other ultrasound parameters. [a]Shear Wave Elastography, SWE.

Region of interest	**Intra-class Correlation Coefficient (95% CI)**		
	Skin thickness (mm)	**Echogenicity (mean, Gray scale)**	**SWE[a] (mean, KPa)**
Proximal Middle Finger	0.946	0.782	0.954
	(0.913-0.967)	(0.642-0.866)	(0.926-0.972)
Dorsal Hand	0.970	0.744	0.953
	(0.952-0.982)	(0.586-0.842)	(0.923-0.971)
Distal Forearm	0.978	0.648	0.964
	(0.965-0.987)	(0.432-0.783	(0.941-0.978)
Abdomen	0.963	0.865	0.973
	(0.940-0.977)	(0.781-0.917)	(0.955-0.983)

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. Journal of scleroderma and related disorders. 2017;2(1):11-8.
2. Rodnan GP, Lipinski E, Luksick J. Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma. Arthritis and rheumatism. 1979;22(2):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(1):84-7.
4. Hesselstrand R, Carlestam J, Wildt M, Sandqvist G, Andreasson K. High frequency ultrasound of skin involvement in systemic sclerosis - a follow-up study. Arthritis research & therapy. 2015;17:329.
5. Hayes RL, Rodnan GP. The ultrastructure of skin in progressive systemic sclerosis (scleroderma). I. Dermal collagen fibers. The American journal of pathology. 1971;63(3):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. British journal of rheumatology. 1992;31(7):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(5):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 and rheumatism. 1997;40(11):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 and rheumatism. 1990;33(8):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. Annals of the rheumatic diseases. 2011;70(1):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. The Journal of rheumatology. 1995;22(7):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. The Journal of rheumatology. 1993;20(11):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). Annals of the rheumatic diseases. 2018;77(2):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. Respiratory medicine. 2017;122 Suppl 1:S14-s7.
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. The Journal of investigative dermatology. 2013;133(8):1979-89.
16. Alexander H, Miller DL. Determining skin thickness with pulsed ultra sound. The Journal of investigative dermatology. 1979;72(1):17-9.
17. 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 research and technology : official journal of International Society for Bioengineering and the Skin. 2013;19(1):e78-84.
18. Akesson A, Hesselstrand R, Scheja A, Wildt M. Longitudinal development of skin involvement and reliability of high frequency ultrasound in systemic sclerosis. Annals of the rheumatic diseases. 2004;63(7):791-6.
19. Scheja A, Akesson A. Comparison of high frequency (20 MHz) ultrasound and palpation for the assessment of skin involvement in systemic sclerosis (scleroderma). Clinical and experimental rheumatology. 1997;15(3):283-8.
20. 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. Clinical and experimental rheumatology. 2011;29(6):926-32.
21. 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. Annals of the rheumatic diseases. 2014;73(1):247-51.
22. 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. The Journal of rheumatology. 2010;37(8):1688-91.
23. Cannao 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. Journal of ultrasound. 2014;17(4):265-9.

24. Babalola O, Mamalis A, Lev-Tov H, Jagdeo J. Optical coherence tomography (OCT) of collagen in normal skin and skin fibrosis. Archives of dermatological research. 2014;306(1):1-9.
25. 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. Annals of the rheumatic diseases. 2010;69(6):1140-3.
26. 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 and rheumatism. 2013;65(11):2737-47.
27. 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 reumatologica portuguesa. 2017.
28. Denton C, Garcia P, Shiwen X. EUSTAR guidelines for preparation and storage of biologic samples in systemic sclerosis: Explant cultures of dermal fibroblasts [Available from: http://www.eustar.org/index.php?module=ContentExpress&func=display&ceid=75&bid=19&btitle=eustar&meid=59.
29. Ihn H, Shimozuma M, Fujimoto M, Sato S, Kikuchi K, Igarashi A, et al. Ultrasound measurement of skin thickness in systemic sclerosis. British journal of rheumatology. 1995;34(6):535-8.
30. 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. Clinical and experimental rheumatology. 2002;20(3):301-8.
31. 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 & rheumatology. 2019.
32. 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. Annals of the rheumatic diseases. 2017;76(Suppl 2):66-7.
33. 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 research and technology : official journal of International Society for Bioengineering and the Skin. 2016;22(3):263-75.
34. 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. Annals of the rheumatic diseases. 2013;72(11):1845-51.