RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

Normal percentile reference curves for skin ultrasound thickness and stiffness at Rodnan sites

Tânia Santiago ,[1,2,3] Eduardo José Ferreira Santos ,[4] Mariana Luis,[1,2,3] Catarina Gaspar,[2] João Lima,[2] Francisco Caramelo,[2] Jose Antonio Pereira da Silva [1,2,3]

To cite: Santiago T, Santos EJF, Luis M, *et al*. Normal percentile reference curves for skin ultrasound thickness and stiffness at Rodnan sites. *RMD Open* 2022;**8**:e002577. doi:10.1136/rmdopen-2022-002577

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/rmdopen-2022-002577).

Received 13 July 2022
Accepted 26 September 2022

[1]Rheumatology, Centro Hospitalar e Universitario de Coimbra, Coimbra, Portugal
[2]Faculty of Medicine, University of Coimbra, Coimbra, Portugal
[3]Rheumatology, Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal
[4]Higher School of Nursing of Coimbra Health Sciences Research Unit Nursing, Coimbra, Coimbra, Portugal

Correspondence to
Dr Tânia Santiago; tlousasantiago@hotmail.com

ABSTRACT

Objectives Our primary objective was to establish preliminary normal reference curves for ultrasound–dermal thickness and skin stiffness in the 17 Rodnan skin sites, considering the effect of gender and age on these measures. As an exploratory objective, we investigated the effect of body mass index and the menopause on skin ultrasound measures.

Methods A cross-sectional study was conducted involving 140 healthy volunteers, aged 20–79 years. Recruitment was stratified by gender and age (10-year categories). Ultrasound–dermal thickness and skin stiffness were assessed by high-frequency ultrasound and shear-wave elastography, respectively, at the 17 Rodnan skin sites. Outcomes were evaluated through a mixed linear model, univariate and multivariate regressions. Normal reference curves were derived for both ultrasound measures in each skin site. An online calculator of the percentiles of skin ultrasound measures was developed.

Results Ultrasound–dermal thickness and stiffness measures were higher in men than women in all Rodnan skin sites (except in chest for ultrasound–dermal thickness). Age had also a significant impact in both ultrasound measures, but only in some skin sites. Gender and age percentile curves (97.5th, 95th, 75th, 50th, 25th, 5th, 2.5th) were plotted for each of the measures in each skin site.

Conclusions Gender and age are strongly associated with skin ultrasound parameters, imposing the need for gender-specific and age-specific reference values. Normal reference percentile curves are provided as a basis for future cooperative work to strengthen its evidence basis, representativeness and refinement regarding potentially influential factors.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Ultrasound is a reliable tool for skin assessment in systemic sclerosis (SSc), but factors influencing ultrasound measures are less well identified.

WHAT THIS STUDY ADDS

⇒ Gender and age affect ultrasound–dermal thickness and skin stiffness at Rodnan skin sites.
⇒ Gender-specific and age-specific normal percentile reference curves were derived for skin ultrasound measures at Rodnan sites.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Normal reference percentile curves for skin ultrasound measures are promising tools to support earlier diagnosis and refine sensitivity to time-induced or drug-induced changes.
⇒ This study provides a ground basis for international collaborations towards most robust and detailed normal referencing for SSc and other diseases affecting the skin.

INTRODUCTION

Systemic sclerosis (SSc) is an autoimmune rheumatic disease characterised by vascular damage and fibrosis of the skin and/or internal organs.[1] This disease is often associated with a high clinical burden and no current satisfactory treatment. Skin involvement is a hallmark of SSc and an important marker of disease activity, severity and prognosis.[2] The extent and rate of skin progression is associated with functional disability, visceral involvement and survival.[3–6] The modified Rodnan skin score (mRSS) is the current gold standard for the evaluation of the skin in SSc, both in clinical trials and in practice.[7 8] However, this score has considerable limitations, with emphasis on the low intrarater and inter-rater reproducibility and poor sensitivity to change, especially in the limited cutaneous form of the disease.[2] In clinical trials, the mRSS has shown a modest performance in discriminating drug versus placebo, despite improvement in other outcomes (lung function, disability index, American College of Rheumatology composite response index).[7–9] This raises the hypothesis that actual improvements in the skin may have been 'missed' because of mRSS limitations. This highlights the need for more sensitive and objective measures of skin involvement, not only support the evaluation and development of new treatments but also to facilitate the earlier diagnosis of SSc, allowing the initiation of therapy before irreversible damage of the skin and other organs is established.[10 11]

Increasing evidence suggests that ultrasound–dermal thickness and stiffness have good-to-excellent intra-reader and inter-reader reliability, and validity against mRSS and skin histological findings.[12–14] Previous studies have indicated that ultrasound skin thickness and stiffness in Rodnan sites of healthy people are affected by gender and age.[15–17] Some influence of body mass index (BMI)[16] and hormonal status[18] have also been suggested. Normal reference curves for skin ultrasound parameters, considering relevant influential factors, seem of pivotal importance to fully exploit the potential of these new techniques as reliable tools for diagnostic and follow-up purposes.

The aim of the present study was to provide normal percentile reference curves for ultrasound–dermal thickness and skin stiffness, in the 17 Rodnan skin sites, considering the effects of age and gender on these measures. The effects of BMI and the menopause were also investigated as exploratory outcomes. This is, hopefully, an important preliminary step towards the full exploitation of the potential of skin ultrasound as a tool for early diagnosis and sensitive follow-up of people with SSc.

METHODS

Study population and demographic measures

A cross-sectional study was conducted, aiming at including 140 healthy participants aged 20–79 years. Recruitment was stratified for age (10-year categories) and gender to guarantee a balanced distribution of these factors in the sample (online supplemental table S1a). Considering that SSc predominantly affects women, with a typical the age of onset 30–50 years, we have doubled the number of female participants in these age groups to increase statistical power. BMI and menopausal status (defined as years after age of menopause) were used as exploratory outcomes. To further support the study of the effect of menopause status, we included 10 premenopausal and 10 postmenopausal women in the group of 40–49 years of age.

Prospective participants were excluded if they had any of the following conditions: (1) diagnosis of any skin or connective tissue or rheumatic inflammatory disease, (2) history of insulin treated diabetes, (3) history of anticancer chemo/radiotherapy treatments, (4) exposure to organic solvents, (5) past glucocorticoid treatment for more than 4 months or (6) recent (<4 weeks) treatment with glucocorticoids, regardless of dose, clinical indication and timing. Participants were recruited among the hospital staff, patients' family members and university students.

The following demographic and clinical parameters were recorded at study entry: age, gender, BMI, age at menopause, years after menopause, dominant body side, smoking habits, past/current medications and professional activity.

Skin ultrasound measures

Skin ultrasound examination was performed using a Siemens ACUSON S2000 Ultrasound System HELX Evolution. During the examination, each participant was placed in supine and relaxed position with a pillow under the head. All ultrasound measures were performed in the morning (between 9:00 and 13:00 hours) in the same room, at stable temperature (between 22°C and 24°C), and after an acclimatisation period of 15 min.

Skin ultrasound evaluation was performed at the 17 sites of the mRSS. The probe was positioned perpendicular to the skin surface, in a longitudinal orientation, with the middle point of the probe footprint placed at the precise locations described by Moore *et al*.[19] An abundant gel interface was used to minimise probe pressure and avoid artefactual changes in the skin. Ultrasound measures were taken and reported according to the EULAR recommendations checklist for reporting studies using ultrasound.[20] The methodology used fully satisfied the World Scleroderma Foundation Recommendations for skin ultrasound studies in SSc, published later on.[21]

Dermal thickness

Dermal thickness was measured in ultrasound B-mode, with an 18 MHz, 57 mm wide, linear probe. All examinations were performed by the same sonographer (TS). To be accepted, every image was required to demonstrate: (1) adequate depiction of epidermis, dermis and subcutis, with distinct and parallel interfaces between them and (2) a gel film layer on the skin surface, to guarantee that no excessive pressure was being applied. Qualified images were collected and stored for later analyses (online supplemental figure S1, a).

Ultrasound–dermal thickness was determined on each single image, through a dedicated image viewer (DICOM), using an electronic calliper to measure the distance between the upper epidermis–dermis interface and the lower dermis–subcutis interface.[19] Separate measures were taken at the left, and right and in the middle of each image, and the average of these three values was considered for statistical analysis. All measures were expressed in millimetres and rounded to centesimals.

Image analysis was performed by one of four raters (rater 1, TS, experienced in skin ultrasound for>5 years; rater 2, ML, with 4 years' experience in musculoskeletal ultrasound; and rater 3 and 4, CG and JL, with no previous experience in ultrasound). The three latter raters received a 2-hour training session led by the first one (TS) before performing any study measurements. During image analysis, all raters were blinded for the participant's demographic characteristics.

In total, assessment of the 17 Rodnan sites took an average total time of 40 min per patient, equally distributed by image acquisition and image analysis.

Skin stiffness

Skin stiffness was evaluated through shear-wave elastography using virtual touch image quantification (VTIQ), with a 9 MHz, 40 mm wide, linear probe. The VTIQ output simultaneously displays a colour-coded tissue stiffness map as well as shear-wave velocity values (in m/s, up to 10 m/s) in distinct sampling gates of interest within each image. Higher shear-wave velocity values indicate greater tissue stiffness (online supplemental figure S1, b).

The probe was positioned as described above. The sonographer placed three sampling gates (2×2 mm) over the epidermis and dermis, at the left, centre and right of each image, and shear-wave velocity values were recorded. The average of these three measures was used for statistical analysis. Skin stiffness was expressed in metres per second (m/s) and all measures were given in centesimals.

Image acquisition and reading of shear-wave velocity values were performed simultaneously by the same examiner (TS). Each set of ultrasound examinations, at all 17 Rodnan skin sites, took about 30 min per patient.

Reliability

Intrarater reliability for stiffness was evaluated based on two ultrasound evaluations by the same sonographer (TS) within an interval of 1–4 weeks, in 20 healthy individuals (included in this study). Contextual and disease factors were stable in all duplicate evaluations.

Intrarater and inter-rater reliability of B-mode image analysis was calculated based on the measures of 4 raters, in 21 healthy individuals. After reliability analyses, agreement between raters was considered adequate and each of the raters proceeded with independently measuring different images.

Statistical analysis

Descriptive statistics (mean and 95% CIs) were used to describe the overall distribution of skin ultrasound measures in different skin sites.

Intrarater reliability for B-mode and shear-wave elastography (one rater) and intrarater and inter-rater reliability for B-mode image analysis (four raters) were evaluated using the intraclass correlation coefficient (ICC). Also, reliability between right and left sides of bilateral Rodnan sites (ie, upper arm, forearm, hand, finger, thigh, leg and foot) was assessed using ICC analysis. An ICC value lower than 0.4 was considered poor, 0.4–0.5 moderate, 0.5–0.7 good and 0.7–1.0 excellent.[22]

Initially, a mixed linear model was used to evaluate whether gender and age significantly influenced skin thickness and stiffness, regardless of body sites. The models included both random intercepts and random slopes using an unstructured covariance matrix. We chose this approach because linear mixed models are an extension of simple linear models which allow both fixed and random effects and are particularly used when there is non independence in the data, such as arises from a hierarchical structure.[23] Then, a more specific and in-depth analysis was performed, considering each site individually and carrying out univariate and multivariate associations between age (years) and gender (male/female).

Development of normal reference percentile curves

The percentile curves were determined by calculating the cumulative frequency of each ultrasound value considering the categories under analysis. Age was arranged in decades to ensure sufficient cases in each group to obtain the percentile values more robustly. Nonetheless, the minimum number of cases was, in some groups, limited to 10 when gender and age were considered simultaneously. Hence, a bootstrapping technique with 5000 samples was used to compute the accumulated frequencies in each group of gender and age. For each bootstrap sample the accumulated frequency was determined and the average for the 5000 samples was considered as the final accumulated frequencies.[24] Based on the frequencies obtained before, a linear regression was carried out and the values of stiffness and thickness corresponding to the 2.5%, 5%, 25%, 50%, 75%, 95% and 97.5% percentiles were computed. Finally, using the values of the percentiles for each age group a line was fitted and plotted to produce the different percentile curves. The regression models used were evaluated by the coefficient of determination.

Development of the online calculator

An online calculator was developed in Dash (Python) and Plotly to calculate percentile curves for skin ultrasound measures based on the evaluation of five skin Rodnan sites (face, forearm, hand, finger and foot) (https://skin-us-ts.herokuapp.com/).

Exploratory analyses

The sample for this study was stratified for age and gender and did not guarantee appropriate conditions to reliably assess the influence of BMI and menopause (mainly due to heterogeneity and small subgroup sampling). Even so, these were the object of exploratory analysis to investigate whether further studies are warranted.

To this purpose, BMI was included in a second version of the mixed linear model. The effect of menopause was assessed through a dedicated mixed linear model considering only women and taking the number of years after menopause as fixed effects and, additionally, through a linear analysis restricted to women aged 40–59 years, taking into account each site individually, and age as a covariate.

Python V.3.9.7 and IBM SPSS were used for all the previous referred analysis. P value of <0.05 was considered as statistically significant.

RESULTS

A total of 140 caucasian participants, that is, 80 women, mean (SD) age of 47.2 (16.0) years, with a mean (SD) BMI of 25.4 (4.6) kg/m^2; and, 60 men, 49.5 (17.3) years, with a BMI of 26.1 (3.3) kg/m^2 were included in this study. Half of the women (n=40) were postmenopausal. Six participants (4.2%) were left-handed. The descriptive statistics of all measures are presented in online supplemental text 1.

Reliability

Details of reliability analysis are depicted in online supplemental text 2. The intrarater reliability of the full process (image acquisition and analysis) for B-mode evaluation ranged from 0.83 (finger) to 0.94 (abdomen), and for shear-wave elastography ranged from 0.83 (face) to 0.96 (leg), in a group of 20 participants (online supplemental table S2a).

In a previous study, we have reported an ICC for inter-rater reliability for B-mode and shear-wave elastography image acquisition and analysis (two independent raters, TS and MS) ranging from 0.78 (finger) to 0.88 (forearm), and 0.82 (finger) to 0.90 (chest), respectively, (detailed data not shown).

The ICC for the inter-rater reliability of the four raters for image analysis was 0.76 (p<0.001) in the first measurement and 0.69 (p<0.001) in the second (online supplemental table S2b). The ICC for intrarater reliability (test/retest) of raters 1, 2, 3 and 4 for B-mode image analysis (dermal thickness) ranged from 0.85 to 0.94 (p<0.001).

Effect of right versus left side

The agreement between values obtained in the left and right sides of each individual is high, as reflected by ICCs rated from good to excellent (online supplemental table S3). The lowest agreement was obtained for skin stiffness assessed at the upper arm (ICC 0.58, p<0.001).

Table 1 Univariate and multivariate regression analysis exploring the effects of age (years) and gender (male vs female), on ultrasound–dermal thickness (left) and skin stiffness (right), in all the sample (N=140)

		Dermal thickness				Skin stiffness			
		Univariate analysis		Multivariate analysis*		Univariate analysis		Multivariate analysis*	
Skin site	Variable	β value	P value	β value	P value	β value	P value	β value	P value
Face	Age	−0.322	**<0.0001**	−0.368	**<0.0001**	−0.039	0.649		
	Gender	−0.501	**<0.0001**	−0.533	**<0.0001**	−0.489	**<0.0001**	−0.494	**<0.0001**
Chest	Age	−0.014	0.874	–		0.053	0.534		
	Gender	0.046	0.587	–		−0.43	**<0.0001**	−0.428	**<0.0001**
Abdomen	Age	−0.066	0.448	–		−0.188	0.029	−0.212	0.009
	Gender	−0.334	**<0.0001**	−0.347	**<0.0001**	−0.34	**<0.0001**	−0.355	**<0.0001**
Upper arm	Age	−0.117	0.167	−0.153	0.038	−0.392	**<0.0001**	−0.485	**<0.0001**
	Gender	−0.502	**<0.0001**	−0.513	**<0.0001**	−0.279	0.001	−0.343	**<0.0001**
Forearm	Age	−0.223	0.009	−0.433	0.001	−0.269	0.001	−0.298	**<0.0001**
	Gender	−0.412	**<0.0001**	−0.433	**<0.0001**	−0.402	**<0.0001**	−0.423	**<0.0001**
Hand	Age	−0.217	0.01	−0.248	0.001	−0.231	0.006	−0.257	0.001
	Gender	−0.430	**<0.0001**	−0.448	**<0.0001**	−0.362	**<0.0001**	−0.380	**<0.0001**
Finger	Age	−0.110	0.194	–		−0.056	0.514	–	
	Gender	−0.393	**<0.0001**	−0.403	**<0.0001**	−0.31	**<0.0001**	−0.316	**<0.0001**
Thigh	Age	0.157	0.085	–		−0.486	**<0.0001**	−0.520	0.0001
	Gender	−0.381	**<0.0001**	−0.365	**<0.0001**	−0.272	0.001	−0.326	**<0.0001**
Leg	Age	−0.221	**<0.0001**	−0.264	0.001	0.079	0.354		
	Gender	−0.393	**<0.0001**	−0.420	**<0.0001**	−0.266	0.002	−0.261	0.002
Foot	Age	−0.031	0.718	–		0.087	0.087	−0.172	0.03
	Gender	−0.431	**<0.0001**	−0.438	**<0.0001**	−0.367	**<0.0001**	−0.379	**<0.0001**

All p values in bold are statistically significant.

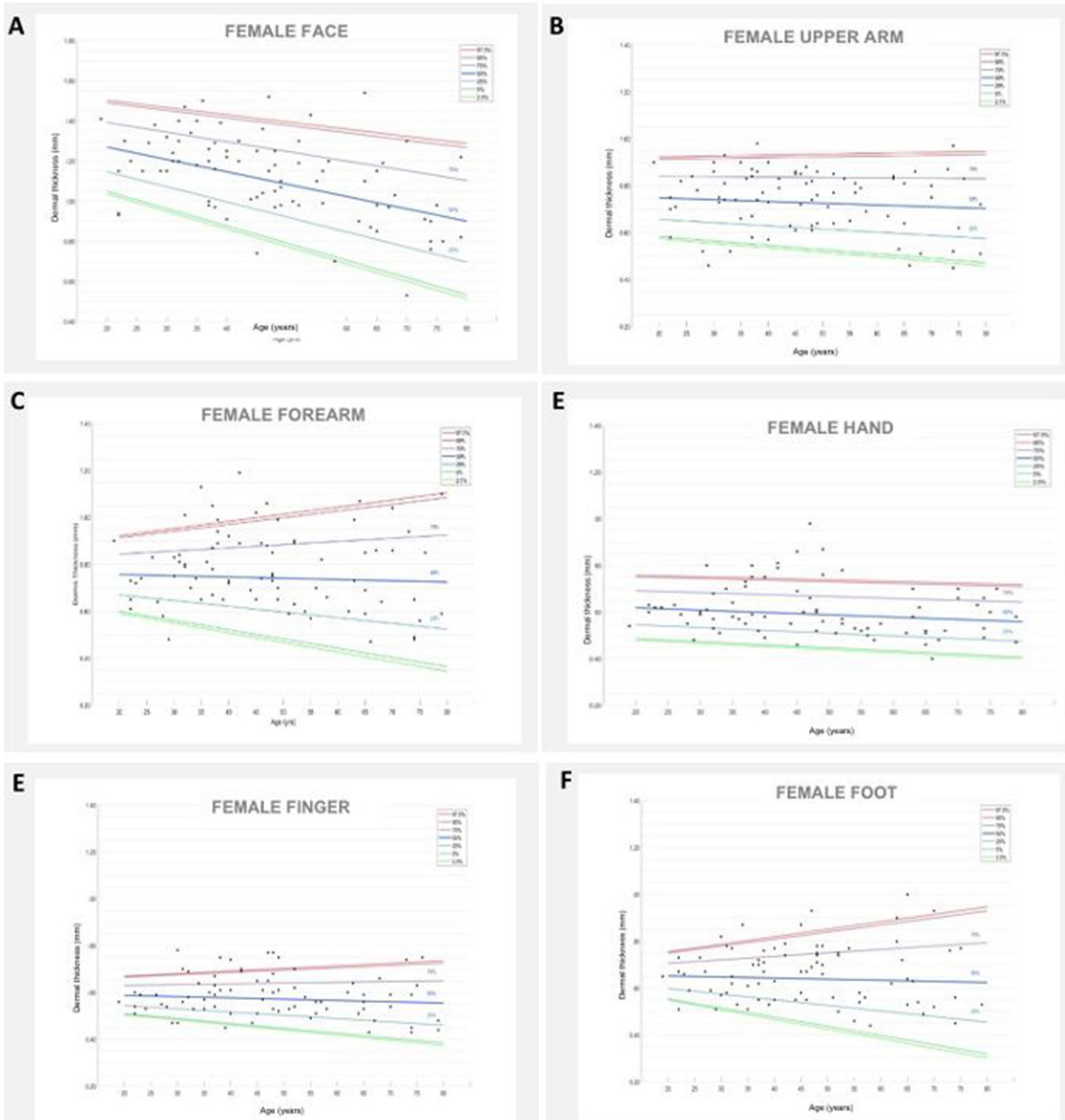

Figure 1 Ultrasound–dermal thickness in females percentile curves. (A) Face; (B) upper arm; (C) forearm; (D) hand; (E) finger; (F) foot.

In addition, we haven't found any significant difference between the dominant versus non-dominant side for dermal thickness in the forearm dominant versus forearm non-dominant (0.87±0.26 vs 0.84±0.19, p=0.07), hand dominant versus hand non-dominant (0.66±0.14 vs 0.65±0.14, p=0.499) and finger dominant versus finger non-dominant (0.63±0.14 vs 0.62±0.12, p=0.661); and, for shear-wave velocity values in the forearm dominant versus forearm non-dominant (1.54±0.02 vs 1.56±0.06, p=0.184), hand dominant versus hand non-dominant sites (1.79±0.30 vs 1.78±0.31, p=0.924) and finger dominant versus finger non-dominant (1.76±0.25 vs 1.75±0.29, p=0.189).

Therefore, the mean value of the two sides was calculated, and used in all subsequent analysis. Hereafter, we present the bilateral sites as a single one.

Effects of gender and age on ultrasound skin measures

Gender and age showed statistical significance in the mixed model for dermal thickness (p<0.001 and p=0.008, respectively) and for skin stiffness (p<0.001 and p<0.001, respectively).

Men had significantly higher ultrasound–dermal thickness and shear-wave velocity values than women in all the skin sites evaluated, except the chest, in univariate and multivariate analyses (table 1). The standardised beta coefficient values ranged from β=−0.533, p=<0.0001 (face) to β=−0.347, p=<0.0001 (abdomen), for dermal thickness; and β=−0.494, p=0.0001 (face) to β=−0.261, p=0.002 (leg), for skin stiffness. The effect of gender, based on the magnitude of the effect size (ie, beta weights), was considered clinically relevant for all skin sites and both ultrasound measures, except for dermal thickness in the chest.

Considering the overall population, age was significantly associated with ultrasound–dermal thickness, in multivariate analyses, independently of gender, in the face, upper arm, forearm, hand, thigh and leg. The same was observed for stiffness in the abdomen, upper arm, forearm, hand, thigh and foot (table 1).

Further, in women, there was a statistically significant association between age and dermal thickness in the face and leg; and, with shear-wave velocity values in the abdomen, upper arm, forearm and thigh (online supplemental table S4). In men, age was significantly associated with ultrasound–dermal thickness in the face, forearm and hand; and, with skin stiffness in upper arm, forearm, hand, thigh and foot (online supplemental table S4). In general, a higher age was associated with thinner and softer skin.

Gender and age normal reference percentile curves

The percentile curves for ultrasound–dermal thickness for women and men are presented in figures 1 and 2, respectively.

Figures 3 and 4 present the percentile curves for stiffness in women and men, respectively. Only the curves for the face, upper arm, forearm, hand, finger and foot are depicted here. Individual figures for all sites, together with equations for percentile lines, are presented in online supplemental text 6—figures. They can also be found at www. http://www.reumatologiachuc.pt.

Online calculator

A dedicated online calculator is available at https://skin-us-ts.herokuapp.com. The website layout is presented in figure 5. The calculator is split into two tabs corresponding to dermal thickness and skin stiffness. The researcher is invited to insert gender, age and each of the

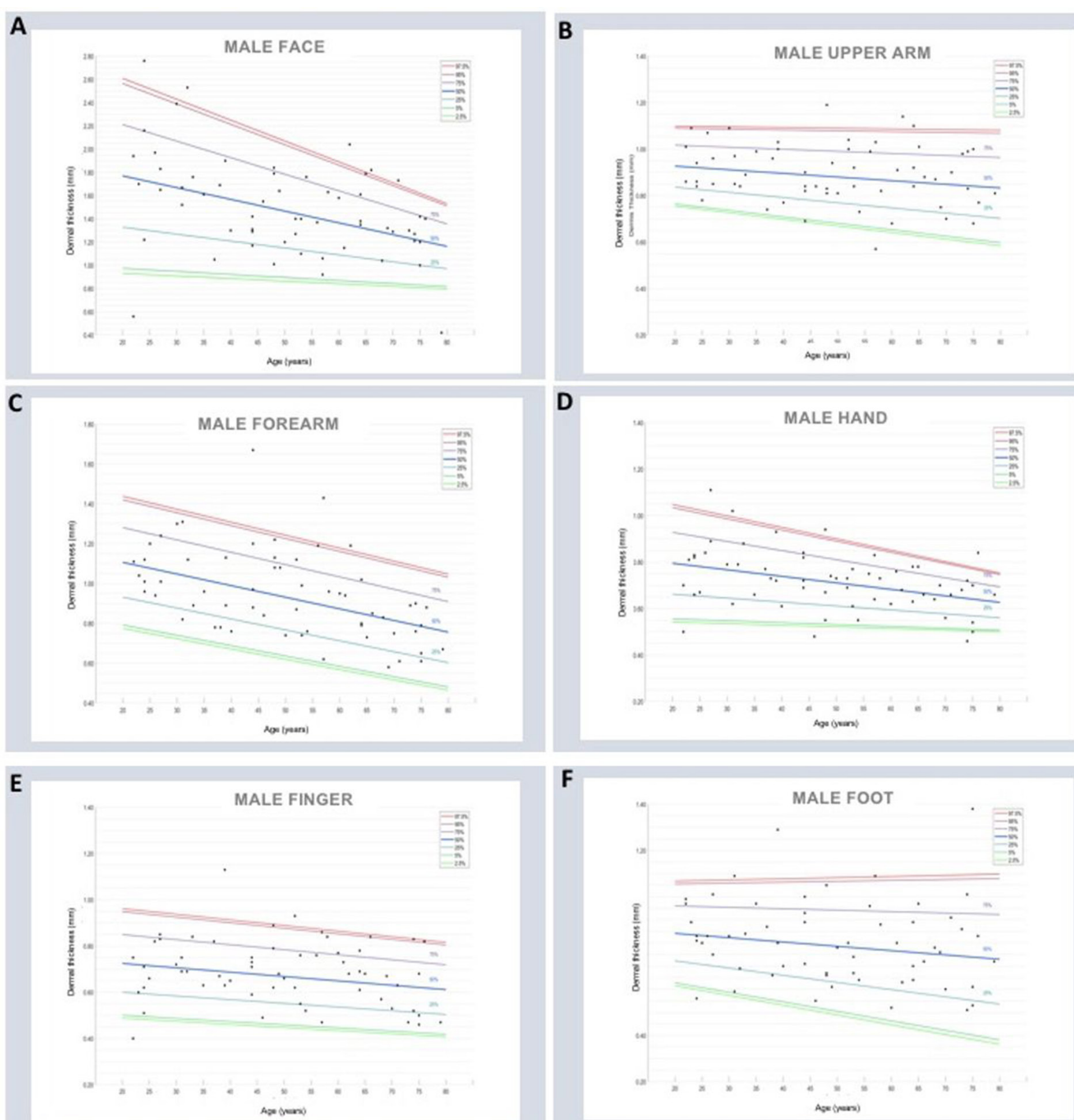

Figure 2 Ultrasound–dermal thickness in males percentile curves. (A) Face; (B) upper arm; (C) forearm; (D) hand; (E) finger; (F) foot.

 Santiago T, *et al*. *RMD Open* 2022;**8**:e002577. doi:10.1136/rmdopen-2022-002577

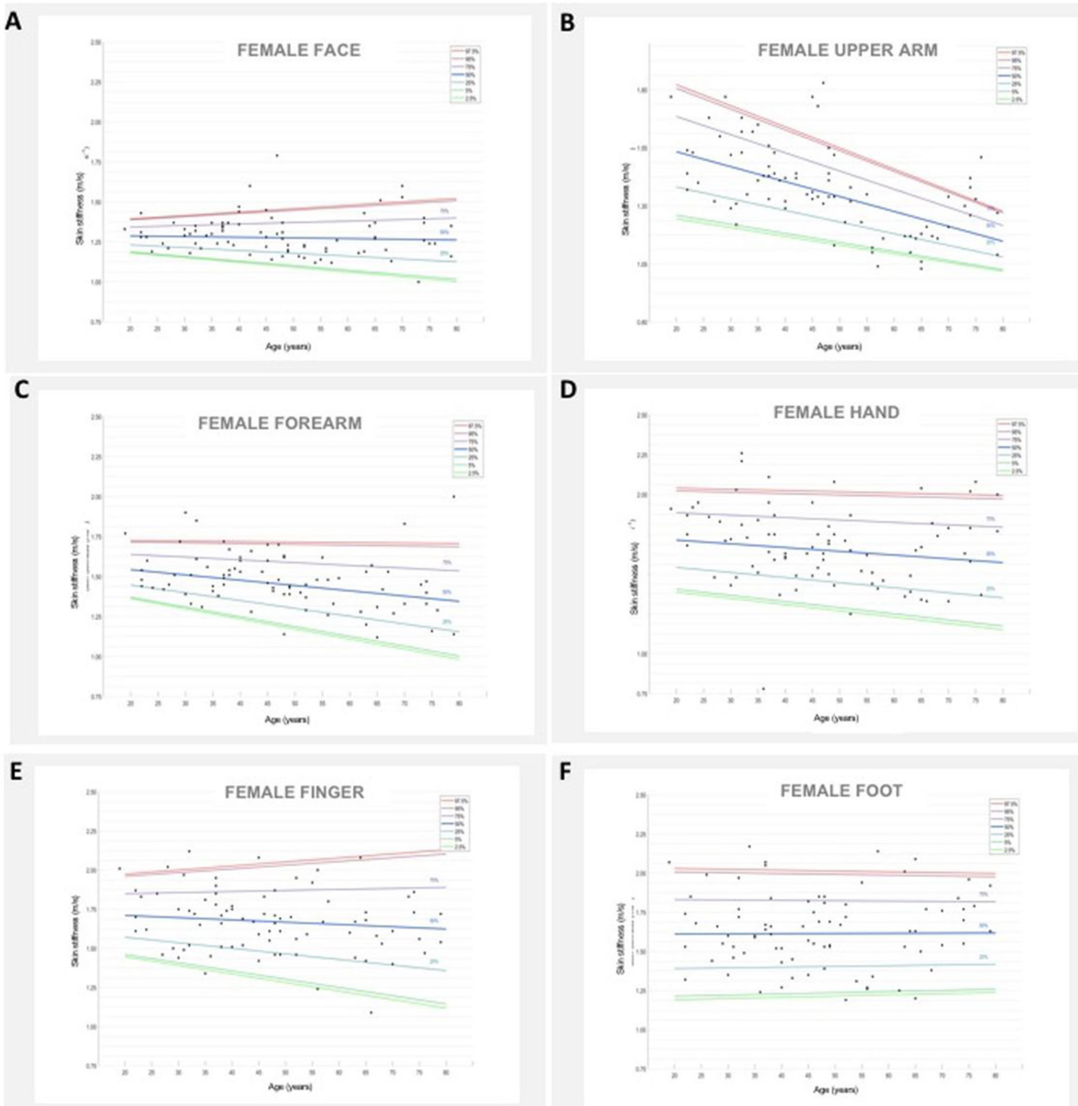

Figure 3 Ultrasound skin stiffness in females percentile curves. (A) face; (B) upper arm; (C) forearm; (D) hand; (E) finger; (F) foot.

skin ultrasound measures obtained at five skin sites (face, forearm, hand, finger and foot). A radar chart showing these values and the percentile for each site is automatically drawn, allowing the clinician to export the results obtained.

Exploratory analysis: effects of BMI and menopause on ultrasound skin measures

For exploratory purposes, BMI was reintroduced in the linear mixed model. The results showed that, in the overall population with all skin sites considered, the correlation between BMI and dermal thickness was statistically significant (p<0.001) but of little magnitude (r=0.0113). Regarding skin stiffness the magnitude was even smaller (r=−0.0036) and did not reach statistical significance (p=0.130).

Years after menopause was not associated with statistically significant effects on either dermal thickness or skin stiffness (p=0.117 and p=0.597, respectively), in the dedicated mixed linear model constructed to this purpose.

In the subgroup of women aged 40–59 years, considering each skin site individually, no statistically significant impact of years after menopause was observed on dermal thickness or skin stiffness, independently of ageing (online supplemental table S5a and S5b).

DISCUSSION

Based on observations made in heathy individuals, we found that ultrasound–dermal thickness and stiffness measures were higher in men than women, in all Rodnan skin sites (except in chest for ultrasound–dermal thickness); and, age had also a significant impact in both ultrasound measures, but only in some skin sites. This study presents a preliminary proposal of gender and age percentile reference curves of ultrasound–dermal thickness and skin stiffness for all 17 Rodnan skin sites. This is, to our knowledge, the first study addressing reference values for these ultrasound measures.

The results confirm that both gender and age are strongly associated with skin ultrasound parameters, imposing the need for gender-specific and age-specific reference values. An important gender effect was found on both ultrasound measures in all 17 Rodnan skin sites (except for dermal thickness in the chest). This confirms and extends observations made in previous ultrasound studies.[15–17] Seidenari *et al* showed, with a 20 MHz probe, that skin (epidermis and dermis) was thicker in men (27–31 years and over 60 years) than women, in five of the six skin sites evaluated.[15] Another study evaluating ultrasound thickness (epidermis plus dermis) using a 20MHz probe, at 2 skin sites of 100 normal individuals, found that thickness was higher in men than women, although not to a statistically significant degree.[17] Skin elasticity values were evaluated in a single shear-wave elastography study including 90 normal individuals, which showed that man had significantly higher values than women at the right dorsal forearm, chest and abdomen.[16]

A significant effect of age on skin ultrasound thickness and stiffness was also found. In general, a higher age was associated with thinner and softer skin, considering all sites considered in the overall population, and in more than half of the skin sites evaluated. This effect was more pronounced in men. A previous study showed that skin was thinner in men and women aged 60+ versus aged 27–31 years, in the six skin sites evaluated.[15] Another group reported that skin thickness tends to increase in the 21–40-year category and decrease after 60–70 years of age.[17] In a shear-wave elastography study of age-related skin changes, it was found that values were significantly lower in individuals older than 50 years compared with the 20–50 year group, at the finger and forearm, but not in the chest or abdomen.[16]

Direct comparisons between our results and those observed in previous studies are hindered by several

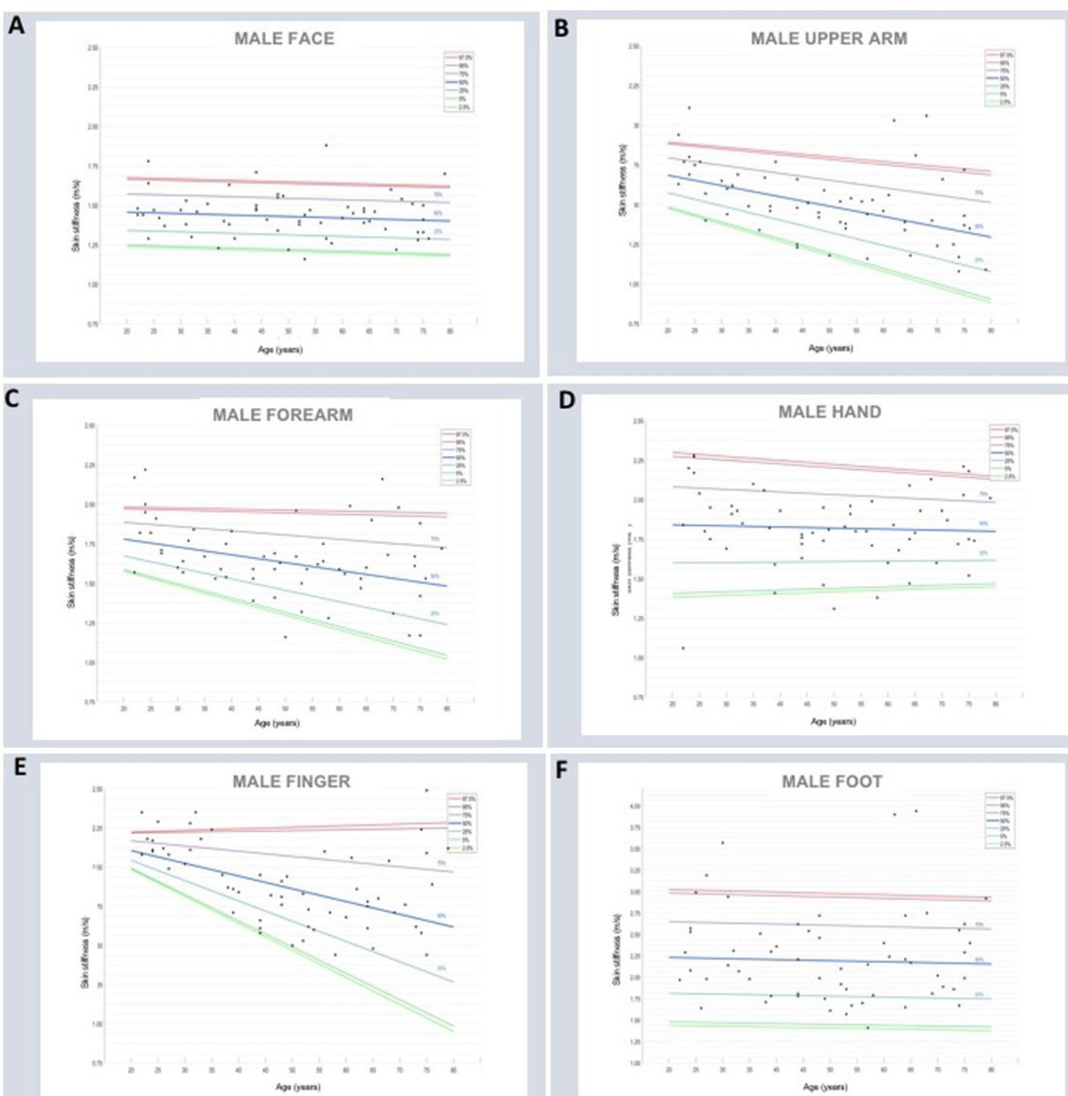

Figure 4 Ultrasound skin stiffness in males percentile curves. (A) face; (B) upper arm; (C) forearm; (D) hand; (E) finger; (F) foot.

factors, with emphasis on the differences in the skin sites assessed and in the ultrasound measurement techniques (eg, equipment and probe frequency, units reported and different software used in elastography). Furthermore, no previous study developed reference values for the whole age range, so we can only compare subgroup mean values. Our study seems to be unique in the recruitment stratified for age and gender, to obtain a sample in which these factors are balanced.

Our study confirms that ultrasound evaluation of all 17 Rodnan scores is reproducible and feasible in a reasonable amount of time. We have found a good-to-excellent intraobserver reliability for B-mode and elastography full process of acquisition and image analysis, and this was also extensive to intrarater and inter-rater B-mode image analysis (by four raters). This may suggest that, in the future, saved ultrasound images may be centralised and analysed provided that a minimum of training is provided. This was also confirmed in a previous study that showed good-to-excellent inter-rater reliability of the full process and is in line with data reported in the literature.[25] Regarding feasibility, we found a strong concordance between the evaluation of the two body sides. This aspect is relevant, as it might raise the rational to only evaluate one side of the body, on the bilateral symmetric Rodnan sites. Nonetheless, this finding should be further confirmed in the SSc population. Feasibility and time reduction of ultrasound evaluation (image acquisition and analysis) is one of the key aspects prioritised in our previous research agenda,[14] and the results of this study definitely contribute to shed a light in this topic.

Our proposal of percentile curves is supported by a robust methodology. We had rigorous eligibility criteria which considered a large array of factors suspected of affecting skin properties. The ultrasound evaluation was fully standardised and in accordance to international recommendations. We exclusively evaluated the dermal thickness (instead of total skin thickness) and used a 18MHz probe, which is a recommended frequency and usually available in standard ultrasound equipment for use in clinical practice.[14] We performed a comprehensive evaluation of all 17 Rodnan skin sites using well-defined landmarks. The evaluation of the sites was conducted by a well-trained sonographer and image assessors, with consistent results as reflected by good-to-excellent intrarater and inter-rater reliability. Finally, we employed robust statistical methodology.

Some limitations of the present study should also be considered. Factors potentially influencing skin ultrasound measures such cumulative sun exposure in sun-exposed areas and manual work occupation (eg, effect on in hand and finger) were not taken into consideration. For shear-wave elastography we have used a transducer with a frequency of 9Mhz, which is lower than the recommended for evaluating the skin.[20] Although, we acknowledge, that high-frequency probes for shear-wave elastography are not yet widely available.[14] Finally, and most important, this a single-centre study with a very homogeneous population (eg, same racial/ethnic background) and relatively small sample size: these results do not necessarily apply to other populations and should not be generalised before further study.

These should be considered as preliminary results which will, hopefully, provide a basis to support multinational cooperation towards a wider and more refined body of evidence. Cooperative studies to enlarge the

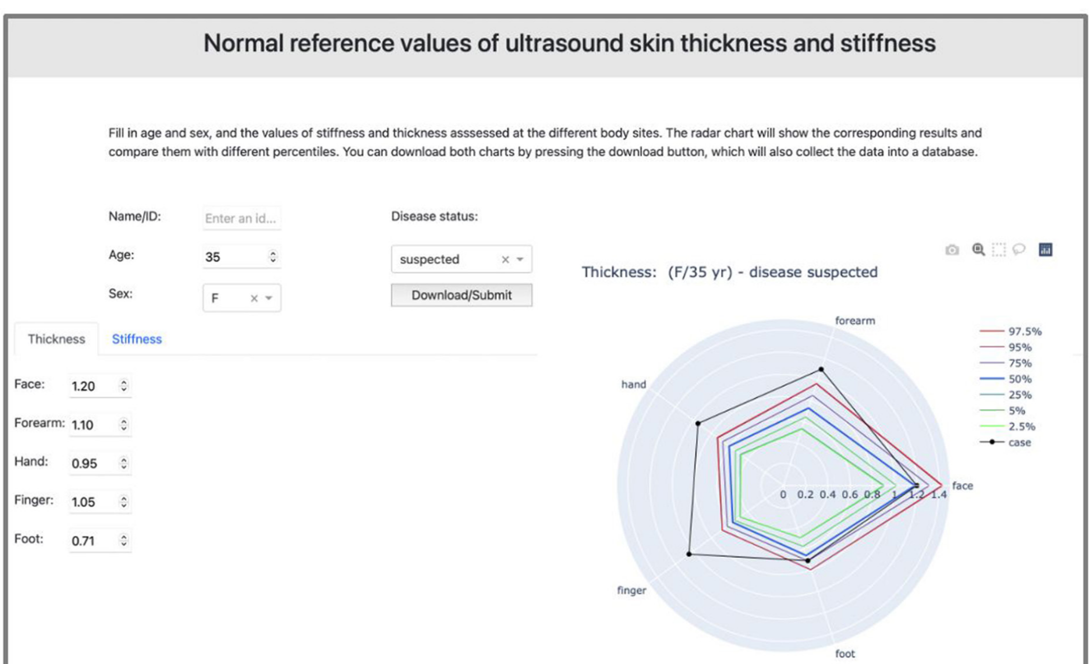

Figure 5 Compact layout of the online calculator of percentile value for ultrasound skin measures. Freely available at https://skin-us-ts.herokuapp.com.

sample size and its geographical diversity are most welcome. Studies assessing the influence of cultural and racial factors, among others, on the normal reference values for skin ultrasound are undoubtedly necessary to take this first step to full fruition. The relevance of BMI and the menopause seem to warrant further evaluation. The reliability of the assessment of only one side of the body and/or a reduced number of skin sites needs to be further investigated. The potential applications of percentile curves in classification criteria, early diagnosis and treatment assessments are inspiring prospects that can only be envisaged after extensive additional research.

In summary, gender-specific and age-specific normal reference percentile curves on the 17 Rodnan skin sites were derived for each of the skin ultrasound measures, from healthy individuals of a restricted geographical and cultural background. This study aims to provide a basis for cooperative studies aimed at testing and consolidating its validity and usefulness for clinical practice and research.

Twitter Eduardo José Ferreira Santos @EduardoJFSantos

Contributors TS and JAPdS designed the study. TS, ML, CG and JL collected the data. TS, EJFS and FC analysed the data and critically interpreted the results. TS prepared the first version of the manuscript. All authors reviewed the draft versions and gave their approval of the final version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was obtained from the Ethics Committee of Centro Hospitalar e Universitário de Coimbra (CHUC-118-17).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data sets used and/or analysed during the current study are available from the corresponding author on reasonable request. Data are available upon reasonable request.

ORCID iDs
Tânia Santiago http://orcid.org/0000-0002-1562-4022
Eduardo José Ferreira Santos http://orcid.org/0000-0003-0557-2377
Jose Antonio Pereira da Silva http://orcid.org/0000-0002-2782-6780

REFERENCES

1. Allanore Y, Simms R, Distler O. Systemic sclerosis. *Nat Rev Dis Primers* 2015;04:15002.
2. Khanna D, Furst DE, Clements PJ, *et al*. Standardization of the modified Rodnan skin score for use in clinical trials of systemic sclerosis. *J Scleroderma Relat Disord* 2017;2:11–18.
3. Domsic RT, Rodriguez-Reyna T, Lucas M, *et al*. Skin thickness progression rate: a predictor of mortality and early internal organ involvement in diffuse scleroderma. *Ann Rheum Dis* 2011;70:104–9.
4. Steen VD, Medsger TA. 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.
5. Ruaro B, Sulli A, Pizzorni C, *et al*. Correlations between blood perfusion and dermal thickness in different skin areas of systemic sclerosis patients. *Microvasc Res* 2018;115:28–33.
6. Hughes M, Huang S, Alegre-Sancho JJ, *et al*. Late skin fibrosis in systemic sclerosis: a study from the EUSTAR cohort. *Rheumatology* 2022. doi:10.1093/rheumatology/keac363. [Epub ahead of print: 22 Jun 2022].
7. Khanna D, Lin CJF, Furst DE, *et al*. Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Respir Med* 2020;8:963–74.
8. Khanna D, Spino C, Johnson S, *et al*. Abatacept in early diffuse cutaneous systemic sclerosis: results of a phase II investigator-initiated, multicenter, double-blind, randomized, placebo-controlled trial. *Arthritis Rheumatol* 2020;72:125–36.
9. Elhai M, Boubaya M, Distler O, *et al*. Outcomes of patients with systemic sclerosis treated with rituximab in contemporary practice: a prospective cohort study. *Ann Rheum Dis* 2019;78:979–87.
10. van den Hoogen F, Khanna D, Fransen J, *et al*. 2013 classification criteria for systemic sclerosis: an American College of rheumatology/European League against rheumatism collaborative initiative. *Ann Rheum Dis* 2013;72:1747–55.
11. Bellando-Randone S, Del Galdo F, Lepri G, *et al*. Progression of patients with Raynaud's phenomenon to systemic sclerosis: a five-year analysis of the European Scleroderma Trial and Research group multicentre, longitudinal registry study for Very Early Diagnosis of Systemic Sclerosis (VEDOSS). *Lancet Rheumatol* 2021;3:e834–43.
12. Santiago T, Santiago M, Ruaro B, *et al*. Ultrasonography for the assessment of skin in systemic sclerosis: a systematic review. *Arthritis Care Res* 2019;71:563–74.
13. Michael Hughes CB, Cuomo G, Sedie AD. The role of ultrasound in systemic sclerosis: on the cutting edge to foster clinical and research advancement. *J Scleroderma Relat Disord* 2020.
14. Santiago T, Santos E, Ruaro B, *et al*. Ultrasound and elastography in the assessment of skin involvement in systemic sclerosis: A systematic literature review focusing on validation and standardization - WSF Skin Ultrasound Group. *Semin Arthritis Rheum* 2022;52:151954.
15. Seidenari S, Pagnoni A, Di Nardo A, *et al*. Echographic evaluation with image analysis of normal skin: variations according to age and sex. *Skin Pharmacol* 1994;7:201–9.
16. Yang Y, Wang L, Yan F, *et al*. Determination of normal skin elasticity by using real-time shear wave elastography. *J Ultrasound Med* 2018;37:2507–16.
17. Kozarova A, Kozar M, Minarikova E, *et al*. Identification of the age related skin changes using high-frequency ultrasound. *Acta Medica Martiniana* 2017;17:15–20.
18. Eisenbeiss C, Welzel J, Schmeller W. The influence of female sex hormones on skin thickness: evaluation using 20 MHz sonography. *Br J Dermatol* 1998;139:462–7.
19. Moore TL, Lunt M, McManus B, *et al*. Seventeen-point dermal ultrasound scoring system--a reliable measure of skin thickness in patients with systemic sclerosis. *Rheumatology* 2003;42:1559–63.
20. Costantino F, Carmona L, Boers M, *et al*. EULAR recommendations for the reporting of ultrasound studies in rheumatic and musculoskeletal diseases (RMDs). *Ann Rheum Dis* 2021;80:840–7.
21. Santiago T, Santos EJF, Ruaro B, *et al*. Recommendations for the execution and reporting of skin ultrasound in systemic sclerosis: an international collaboration under the WSF skin ultrasound group. *RMD Open* 2022;8:e002371.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
23. Galecki AB T. *Linear mixed-effects models using R. Springer*, 2013.
24. Hall HFPvdB MJ, Fernando RC, Mynett AE. *The construction of confidence intervals for frequency analysis using resampling techniques*. European Geosciences Union: Hydrology and Earth System Sciences Discussions, 2004.
25. Vanhaecke A, Cutolo M, Heeman L, *et al*. High frequency ultrasonography: reliable tool to measure skin fibrosis in SSC? A systematic literature review and additional pilot study. *Rheumatology* 2021;61:42–52.