


RESEARCH

Open Access



# Durometry as an alternative tool to the modified Rodnan's skin score in the assessment of diffuse systemic sclerosis patients: a cross-sectional study

Marília de Fátima Cirioli de Oliveira<sup>1</sup>, Vanessa Cristina Leopoldo<sup>1</sup>, Karla Ribeiro Costa Pereira<sup>2</sup>, Daniela Aparecida de Moraes<sup>2</sup>, Juliana Bernardes Elias Dias<sup>2</sup>, Maynara Santana Gonçalves<sup>3</sup>, Leandra Náira Zambelli Ramalho<sup>4</sup>, Bruna Nogueira dos Santos<sup>1</sup>, Maria Carolina de Oliveira<sup>2</sup> and Renata Cristina Campos Pereira Silveira<sup>1\*</sup> 

## Abstract

**Background:** The reproducibility and reliability of the modified Rodnan's Skin Score (mRSS) are debated due to investigator-related subjectivity. Here, we evaluate if durometry correlates with mRSS in patients with diffuse systemic sclerosis (SSc).

**Methods:** This cross-sectional study was conducted from December 2018 to June 2019, including 58 diffuse SSc patients. Two certified researchers, blind to each other's scores, performed the mRSS, followed by durometry at 17 predefined skin sites. For durometry and mRSS, individual scores per skin site were registered. Durometry and mRSS results measured by each researcher, as well as scores from different researchers, were compared. Skin thickness measurements from forearm skin biopsies were available in a subset of the patients, for comparisons. Statistical analyses included Cohen's Kappa Coefficient, Intraclass Correlation Coefficient, Kendall's Coefficient and Spearman's test.

**Results:** Mean (standard deviation, SD) patient age was 44.8 (12.9) years, and 88% were female. Inter-rater agreement varied from 0.88 to 0.99 (Intraclass correlation coefficient) for durometry, and 0.54 to 0.79 (Cohen's Kappa coefficient) for mRSS, according to the specific evaluated sites. When data were compared with skin thickness assessed in forearm biopsies, durometry correlated better with skin thickness than mRSS.

**Conclusion:** Durometry may be considered as an alternative method to quantify skin involvement in patients with diffuse SSc. The strong inter-rater agreement suggests that the method may be useful for the assessment of patients by multiple researchers, as in clinical trials.

**Keywords:** Systemic sclerosis, Autoimmune diseases, Skin abnormalities, Modified Rodnan's score, Durometer

\* Correspondence: [recris@eerp.usp.br](mailto:recris@eerp.usp.br)

<sup>1</sup>Ribeirão Preto Nursing School, University of São Paulo, Av. Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, SP 14049-900, Brazil  
Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Background

Systemic sclerosis (SSc) is a chronic autoimmune connective tissue disease characterized by microvasculopathy, inflammatory phenomena and visceral and cutaneous fibrosis, associated with high morbidity and mortality [1, 2]. Skin involvement is the hallmark of SSc and affects more than 95% of the patients. Excess collagen and other extracellular matrix components and increased myofibroblasts in the dermis, epidermis and subcutaneous tissue lead to progressive thickening of the cutaneous tissue [3–5].

The modified Rodnan's Skin Score (mRSS) is the most commonly used noninvasive method to assess skin involvement in SSc [6]. However, the method is subject to evaluator skills and interpretation, and thus imposes limitations [7]. Despite adequate training and experience of health professionals in performing mRSS, the reproducibility among professionals is low [8]. Therefore, alternative methods to assess skin thickness are warranted [6, 9, 10]. Albeit invasive, skin biopsies are the gold-standard method to quantify cutaneous involvement, and have been used in the context of clinical research [10]. Biopsies have the limitation of assessing skin from only one location of the skin surface, usually the ventral area of a forearm, while mRSS enables a more global evaluation. Recently, durometry has been evaluated as an alternative tool for skin assessment, aiming to improve accuracy and consistency of skin hardness or stiffening evaluations [11–13].

## Methods

We aimed to compare mRSS and durometry results in diffuse SSc patients, and validate these measurements through comparisons with skin biopsies available in a subset of the patients.

From December 2018 to June 2019, 58 SSc patients, diagnosed according to the ACR/EULAR 2013 criteria [14], were included in this cross-sectional study, at the Ribeirão Preto Medical School, University of São Paulo. Subjects older than 16 years of age and presenting the diffuse form of SSc were enrolled. Patients were excluded if having amputation or absence of any limb, or skin abnormalities not related to SSc, which, in the opinion of the researchers, would have affected the evaluation. The Institutional Research Ethics Committee approved the study, under registration number 3.045.124. Subjects or their legal guardians, when patients were younger than 18 years of age, read and signed informed consent forms at study enrollment.

### Modified Rodnan's skin score and durometry assessments

Two researchers, blind to each other's scores, evaluated each patient for mRSS, followed by durometry, on the same date. Patients were comfortably placed on a bed, at

room temperature, with relaxed muscles. For mRSS, skin thickness was assessed as previously described [7]. Briefly, each of the 17 predetermined skin areas was scored from zero to three, according to severity of skin thickness, the final added score from all 17 areas ranging from zero to 51 points [6]. For durometry, a portable dial analog durometer (model 1600, type 00, Rex Gauge Company, Inc., Buffalo Grove, IL, USA) was used to measure skin hardness. The durometer (Fig. 1), an equipment of 2.5"×6.125" and 6 oz., is the international standard for the hardness measurement of rubber,



**Fig. 1** Durometer model 1600, type 00, Rex Gauge Company, Inc., Buffalo Grove, IL, USA

plastic and other non-metallic materials. The type 00, used in this study, is applicated in animal tissue. It has an indenter of 3/32" spherical and a spring force of 113 g. When the durometer presses the cutaneous tissue, it allows an assessment of the skin hardness through the resistance to indentation. Measurements are verified in Durometry units, with a continuous scale accurate to one decimal point [12]. Lower numbers indicate less resistance and softer materials. For each skin area, durometry was assessed three times, the mean value considered as final result. For each patient, mRSS and durometry were assessed by the same researcher and at the same areas comprised by mRSS [11], with some specific adjustments of sites to place the durometer, as described in Table 1.

### Skin biopsies

A subgroup of ten patients included in this study underwent forearm skin biopsies through a punch technique on the medial area of the left forearm, immediately after mRSS and durometry assessments. Biopsies were processed, fixated and stained with Hematoxylin-Eosin (HE). An experienced pathologist, blind to mRSS and durometry scores, evaluated the skin samples for quantification of epidermal plus dermal thickness in micrometers ( $\mu\text{m}$ ). Results were available for 9 samples, as one of them did not have enough quality to be measured.

### Statistical analysis

Data were analyzed in the following sequence. First, inter-investigator scores for mRSS and for durometry were evaluated for consistency. For mRSS, inter-rater agreement was analyzed by the Cohen's Kappa coefficient [15, 16], and for durometry, agreement was assessed by the Intraclass Correlation Coefficient [17]. Then, measurements from durometry and mRSS, for each investigator, were correlated through the Kendall's coefficient. For each of the correlation tests, a result of

zero indicates absence of agreement and one, total agreement. Finally, mRSS, durometry scores and skin thickness were correlated (Spearman's test), skin thickness considered as the golden-standard reference for skin involvement. The R software, version 3.4.1 [18] and SAS 9.2 [19] were used for statistical analyses. Significance was set at 0.05.

### Results

Fifty-eight diffuse SSc patients, 51 (88%) female, were included in this study. Mean (SD) age was 44.8(12.9) years and mean (SD) time from diagnosis was 22(42.3) months. All patients underwent complete mRSS and durometry evaluations. The mRSS ranged from 0 to 49, with a median value of 22.3 and all patients had visceral involvement, either gastrointestinal or pulmonary. Eighty-one percent had already been treated with one or more lines of immunosuppressant medications.

Table 2 shows the mean values for mRSS and durometry for investigators 1 and 2, respectively, at each evaluated skin site, as well as the inter-rater agreement evaluations. Inter-rater agreement ranged from 0.88 (finger) to 0.99 (abdomen) for durometry, and from 0.54 (thigh and abdomen) to 0.79 (foot) for mRSS.

Table 3 shows the intra-rater correlations between mRSS and durometry. For specific sites, such as face, fingers and legs, poor correlation was detected, with Kendall's coefficient ranging from 0.03 to 0.30.

Table 4 shows correlations between skin thickness, mRSS and durometry scores for the nine patients with available skin biopsies (Table 4).

### Discussion

Skin involvement is an important characteristic of SSc patients and associates with prognosis [20]. Skin thickness is usually assessed by the mRSS, but limitations are associated to this method, mainly the evaluator-associated subjectivity, compromising its reproducibility

**Table 1** Specific sites to place the durometer at each evaluated skin area

Skin areas	Specifications for durometry
Fingers R and L	Third proximal phalange
Hands R and L	Between 2nd and 3rd metacarpals, 2 cm proximal to metacarpo-phalangeal joints
Forearms R and L	Medial dorsal area of the forearm
Arms R and L	Medial dorsal area of the arm
Face	Center of the forehead
Thorax R	Between 2nd e 3rd intercostal spaces, to the right
Abdomen R	5 cm to the right of the umbilical scar
Thigh R and L	Medial anterior area of the thigh
Legs R and L	Medial anterior area of the leg
Feet R and L	Between 1st and 2nd metatarsals, 2 cm proximal to the proximal phalange

R right, L left, cm centimeters

**Table 2** Inter-rater agreement scores for mRSS and durometry, at each evaluated skin area

Skin area	Durometry score agreement (investigator 1 vs investigator 2)		mRSS agreement (investigator 1 vs investigator 2)	
	ICC	95% CI	Kappa	95% CI
Face	0.89	(0.82; 0.94)	0.65	(0.50; 0.80)
Thorax	0.98	(0.97; 0.99)	0.60	(0.45; 0.74)
Abdomen	0.99	(0.98; 0.99)	0.54	(0.37; 0.71)
Arm R	0.91	(0.97; 0.99)	0.70	(0.58; 0.83)
Arm L	0.97	(0.94; 0.98)	0.56	(0.42; 0.70)
Forearm R	0.97	(0.95; 0.98)	0.75	(0.62; 0.87)
Forearm L	0.97	(0.95; 0.98)	0.71	(0.57; 0.84)
Hand R	0.95	(0.91; 0.97)	0.66	(0.52; 0.79)
Hand L	0.96	(0.94; 0.98)	0.72	(0.60; 0.84)
Finger R	0.90	(0.83; 0.94)	0.76	(0.64; 0.89)
Finger L	0.88	(0.79; 0.93)	0.76	(0.63; 0.89)
Thigh R	0.98	(0.97; 0.99)	0.57	(0.43; 0.71)
Thigh L	0.97	(0.95; 0.98)	0.54	(0.39; 0.69)
Leg R	0.93	(0.89; 0.96)	0.67	(0.53; 0.81)
Leg L	0.92	(0.87; 0.96)	0.66	(0.52; 0.79)
Foot R	0.94	(0.90; 0.97)	0.76	(0.64; 0.88)
Foot L	0.95	(0.92; 0.97)	0.79	(0.68; 0.90)

mRSS modified Rodnan's skin score, ICC Intraclass Correlation Coefficient, CI confidence interval, R right, L left, Kappa Weighted Cohen's Kappa coefficient

**Table 3** Intra-rater correlations between mRSS and durometry scores

Skin areas	Investigator 1 Kendall's coefficient ( <i>p</i> )	Investigator 2 Kendall's coefficient ( <i>p</i> )
Face	0.20 (0.06)	0.03 ( $\leq 0.80$ )
Thorax	0.50 ( $\leq 0.01$ )	0.36 ( $\leq 0.01$ )
Abdomen	0.53 ( $\leq 0.01$ )	0.45 ( $\leq 0.01$ )
Arm R	0.49 ( $\leq 0.01$ )	0.53 ( $\leq 0.01$ )
Arm L	0.50 ( $\leq 0.01$ )	0.43 ( $\leq 0.01$ )
Forearm R	0.62 ( $\leq 0.01$ )	0.57 ( $\leq 0.01$ )
Forearm L	0.62 ( $\leq 0.01$ )	0.55 ( $\leq 0.01$ )
Hand R	0.55 ( $\leq 0.01$ )	0.54 ( $\leq 0.01$ )
Hand L	0.49 ( $\leq 0.01$ )	0.51 ( $\leq 0.01$ )
Fingers R	0.22 (0.04)	0.28 ( $\leq 0.01$ )
Fingers L	0.34 ( $\leq 0.01$ )	0.40 ( $\leq 0.01$ )
Thigh R	0.54 ( $\leq 0.01$ )	0.46 ( $\leq 0.01$ )
Thigh L	0.54 ( $\leq 0.01$ )	0.37 ( $\leq 0.01$ )
Leg R	0.30 ( $\leq 0.01$ )	0.26 ( $\leq 0.01$ )
Leg L	0.29 ( $\leq 0.01$ )	0.22 ( $\leq 0.03$ )
Foot R	0.54 ( $\leq 0.01$ )	0.52 ( $\leq 0.01$ )
Foot L	0.44 ( $\leq 0.01$ )	0.57 ( $\leq 0.01$ )

R right, L left

and reliability. Here, we show that durometry may be used as an alternative method to quantify skin involvement. Durometry scores are reproducible between investigators and directly reflect the severity of skin involvement assessed by mRSS and skin biopsies.

We evaluated the inter-rater agreement for mRSS and durometry scores. Correlation coefficients between investigators were higher for durometry than for mRSS, which is in accordance with Merkel et al. [13] and with Kissin et al. [11]. Durometry, therefore, is more reproducible between different researchers than mRSS, on

**Table 4** Correlations between durometry and mRSS versus skin thickness (biopsies)

	Skin Thickness (forearm)	
	Total scores <i>r</i> ( <i>p</i> )	Forearm scores <i>r</i> ( <i>p</i> )
mRSS #1	0.95 ( $< 0.001$ )	0.85 (0.005)
mRSS #2	0.63 (0.08)	-0.13 (0.75)
Durometry #1	0.86 (0.005)	0.80 (0.014)
Durometry #2	0.88 (0.003)	0.85 (0.006)

In the left column, skin thickness from the forearm is correlated with total mRSS and durometry scores. On the right column, forearm skin thickness is correlated with forearm mRSS and durometry. Analysis by Spearman's test. mRSS: modified Rodnan's skin score; #1 and #2 indicate investigator 1 and 2, respectively

transversal evaluations. Specifically in the abdomen and thighs, we detected an inter-rater agreement for mRSS that was lower than average for the remaining skin sites, indicating that reproducibility of mRSS is uneven across evaluated areas. In our perspective, this is as an opportunity to use durometry as a complementary tool, enabling better assessment of the areas where mRSS is not sufficiently accurate.

When each investigator's mRSS and durometry scores were compared (intra-rater evaluations, Table 3), weak correlation indices were observed in areas where bone tissue was more prominent to the skin surface. These results may reflect the smaller amounts of subcutaneous tissue at specific body sites, such as the face, fingers, and legs. Durometry measures the hardness of a material; therefore underlying bone may be misevaluated as thick skin. Moon et al. [21] have already described similar difficulties for leg evaluations, and Kissin et al. [11] have suggested, for future applications, modifications in sites for durometry, aiming to further increase the accuracy of the method. For further studies, the durometer should be placed in areas with less underlying bone, thereby replacing fingers, forehead and anterior areas of the legs for other sites.

The use of durometry, complementary to mRSS assessments, should be considered in the follow-up of SSc patients. First, the device does not require certification or extended learning time to be used, as it consists of a simple place-and-measure procedure. The method is reproducible between evaluators and may be used interchangeably on patient follow-up when the same investigator is not available [13]. This may be a valuable tool to simplify the logistics of clinical trials [11]. Moreover, in our patient cohort, durometry scores, both global and next to the site of skin biopsy, correlated better to skin thickness evaluated by histopathology, than mRSS. We believe that these discrepant results reflect the lack of reliability of mRSS across different evaluators and reinforce the need of alternative (or complementary) skin assessment strategies in SSc patients. Finally, durometry measurements are provided as a continuous scale, accurate to decimal units, and may evidence more subtle changes in skin involvement than mRSS. This hypothesis should be further evaluated in longer, prospective studies.

Nine patients from this study coincidentally underwent forearm skin biopsies for clinical purposes, on the same date that durometry and mRSS were assessed. The skin thickness of the fragment enabled comparisons to durometry and mRSS values, as a golden-standard reference. In our study, we did not evaluate myofibroblasts or hyalinized collagen, but Kissin et al. [12] showed positive correlations between mRSS and durometry with both collagen and myofibroblast content. In healthy individuals, conversely, analyses of skin biopsy results did not show

myofibroblasts and the amount of collagen in the control sample was lower compared to the diseased samples.

This study was limited to diffuse SSc patients as a strategy to homogenize the sample of patients. We found interesting that at specific skin areas, mRSS and durometry correlated well with each other, with a few already described exceptions in areas with scarce subcutaneous tissue. In addition, the skin thickness correlated well with mRSS and durometry in both global and forearm (site of biopsy) evaluations. These results suggest that the findings may be expanded to limited forms of SSc, still to be evaluated in the future.

An interesting future addition would be to compare mRSS and durometry measurements with those from skin ultrasound, which is another available tool for skin assessment. Ultrasound requires a specialist and is not as readily available as the other tools, but may contribute to the field. Ultrasound has a low intra and inter-observer variability, but correlates only low-to-moderately with mRSS [11, 22, 23]. In a prior study, ultrasound has shown moderate correlation with durometry and with total mRSS, but when specific skin sites were analyzed, there was poor correlation in the fingers [11]. Aspects such as to standardize imaging and best skin sites to be measured are still to be defined in ultrasonography. In the future, more than alternatives to the already established mRSS, these methods may be combined in the assessment of the scleroderma skin as a composite measure.

When compared to previous reports, the specific sites of durometry were slightly different in our study. The decision reflected our initial belief that durometry sites should be similar to those used in mRSS and that patients should be evaluated in anterior areas of the body. We also chose areas that we considered more stable to place the durometer, instead of sites subject to fat accumulation. For the facial evaluations, we used the center of the forehead, which has proven inadequate due to the great divergence between mRSS and hardness measured by the durometer. We therefore agree that alternative approaches, using the cheeks, instead of the forehead, may be more reliable [21]. A study that used the calves, instead of the anterior surface of the legs, also yielded non-reliable results [21]. The durometry of the fingers, however, still stands as a challenge. Our results show that it is not reliable to place the durometer on the hard dorsal surface of the phalanges. The alternative of placing it on the palmar surface of the fingers, however, does not add benefit, as part of the patients have claw-like contractures in the hands. For future studies, investigators should consider excluding the fingers from durometry assessments. Whether this may affect the overall evaluation of a patient, especially those with limited SSc or with initial manifestations of the disease, remains unknown.

## Conclusions

We were able to show that for single point transversal analyses of skin involvement, the durometry may reproduce, with possible advantages, the results obtained with the mRSS. The inter-rater reliability of the durometry may prove it useful for assessment of large numbers of patients in single or multicenter trials. For future studies, we seek to investigate whether durometry is also reliable in longitudinal evaluations of patients, with multiple assessments over time.

## Abbreviations

SSc: Systemic sclerosis; mRSS: modified Rodnan's Skin Score; ACR/ EULAR: American College of Rheumatology/ European League Against Rheumatism; HE: Hematoxylin-Eosin;  $\mu\text{m}$ : micrometers; SD: Standard deviation

## Acknowledgements

We thank Coordination of Superior Level Staff Improvement (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES) for the support. We are grateful to Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, where the study was conducted.

## Authors' contributions

M.F.C.O. conceived the original idea, collected and analysed the data, and wrote the manuscript. D.A.M. and J.B.E.D. collected and analysed the data. M.S.G. and L.N.Z.R. helped on data analysis. V.C.L. and B.N.S. helped on data analysis and contributed to the final version of the manuscript. K.R.C.P. worked on the final review of the manuscript. M.C.O. and R.C.C.P.S. designed and supervised the study and worked on the final review of the manuscript. All author(s) read and approved the final manuscript.

## Funding

Not applicable.

## Availability of data and materials

All data generated or analysed during this study are included in this published article.

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

## Ethics approval and consent to participate

The Institutional Research Ethics Committee approved the study, under registration number 3.045.124.

## Consent for publication

Subjects or their legal guardians, when patients were younger than 18 years of age, read and signed informed consent forms at study enrollment.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Ribeirão Preto Nursing School, University of São Paulo, Av. Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, SP 14049-900, Brazil. <sup>2</sup>Department of Internal Medicine, Ribeirão Preto Medical School, University of São Paulo, Av. Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, SP 14049-900, Brazil. <sup>3</sup>Biological Scientist, Basic and Applied Immunology Program, Ribeirão Preto Medical School, University of São Paulo, Av. Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, SP 14049-900, Brazil. <sup>4</sup>Pathology and Legal Medicine Department, Ribeirão Preto Medical School, University of São Paulo, Av. Bandeirantes, 3900, Monte Alegre, Ribeirão Preto, SP 14049-900, Brazil.

Received: 5 June 2020 Accepted: 9 September 2020

Published online: 21 September 2020

## References

- Allanore Y, Simms R, Distler O, et al. Systemic sclerosis. *Nat Rev Dis Primers*. 2015;1(15002):1–12.
- Asano Y. Systemic sclerosis. *J Dermatol*. 2018;45:128–38.
- Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. *N Engl J Med*. 2009;360:1989–2003.
- Pattanaik D, Brown M, Postlethwaite BC, et al. Pathogenesis of systemic sclerosis. *Front Immunol*. 2015;6:1–40.
- Ferrelli C, Gasparini G, Parodi A, et al. Cutaneous manifestations of scleroderma and scleroderma-like disorders: a comprehensive review. *Clin Rev Allergy Immunol*. 2017;53:306–36.
- 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;2:11–8.
- Furst DE, Clements PJ, Steen VD, et al. The modified rodnan skin score is an accurate reflection of skin biopsy thickness in systemic sclerosis. *J Rheumatol*. 1998;25:84–8.
- Czirják L, Nagy Z, Aringer M, et al. The EUSTAR model for teaching and implementing the modified Rodnan skin score in systemic sclerosis. *Ann Rheum Dis*. 2007;66:966–9.
- Clements PJ. Measuring disease activity and severity in scleroderma. *Curr Opin Rheumatol*. 1995;7:517–21.
- Ziemek J, Man A, Hinchcliff M, et al. The relationship between skin symptoms and the scleroderma modification of the health assessment questionnaire, the modified Rodnan skin score, and skin pathology in patients with systemic sclerosis. *Rheumatology (Oxford)*. 2016;55:911–7.
- Kissin EY, Schiller AM, Gelbard RB, et al. Durometry for the assessment of skin disease in systemic sclerosis. *Arthritis Rheum*. 2006;55:603–9.
- Kissin EY, Schiller AM, Gelbard RB, et al. Myofibroblasts and hyalinized collagen as markers of skin disease in systemic sclerosis. *Arthritis Rheum*. 2006;54:3655–60.
- Merkel PA, Silliman NP, Denton CP, et al. Validity, reliability, and feasibility of durometer measurements of scleroderma skin disease in a multicenter treatment trial. *Arthritis Rheum*. 2008;59:699–705.
- Johnson SR. New ACR EULAR guidelines for systemic sclerosis classification. *Curr Rheumatol Rep*. 2015;17:32.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
- Agresti A. *Categorical data analysis*. New York: Wiley; 1990.
- Snedecor GW, Cochran WG. *Statistical methods*. Ames: The Iowa State University Press; 1972.
- R Core Team. R: a language and environment for statistical computing. Available at: <https://www.R-project.org/>. Accessed 28 Aug 2019.
- SAS. *The SAS system for Windows*. Release 9.2. SAS Inst., Cary: NC; 2011.
- Clements PJ, Hurwitz EL, Wong WK, et al. Skin thickness score as a predictor and correlate of outcome in systemic sclerosis: high-dose versus low-dose penicillamine trial. *Arthritis Rheum*. 2000;43:2445–54.
- Moon KW, Song R, Kim JH, et al. The correlation between durometer score and modified Rodnan skin score in systemic sclerosis. *Rheumatol Int*. 2012; 32:2465–70.
- Li H, Furst DE, Jin H, et al. High-frequency ultrasound of the skin in systemic sclerosis: an exploratory study to examine correlation with disease activity and to define the minimally detectable difference. *Arthritis Res Ther*. 2018; 20(1):181.
- Hesselstrand R, Carlestam J, Wildt M, et al. High frequency ultrasound of skin involvement in systemic sclerosis - a follow-up study. *Arthritis Res Ther*. 2015;17:329.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.