

RESEARCH ARTICLE

Open Access



# Skin model for improving the reliability of the modified Rodnan skin score for systemic sclerosis

Patnarin Pongkulkiat<sup>1</sup>, Bandit Thinkhamrop<sup>2</sup>, Ajanee Mahakkanukrauh<sup>1</sup>, Siraphop Suwannaroj<sup>1</sup> and Chingching Foocharoen<sup>1\*</sup> 

## Abstract

**Background:** The gold standard for skin thickness assessment in systemic sclerosis (SSc) is the modified Rodnan skin score (mRSS); however, inter- and intra-rater variation can arise due to subjective methods and inexperience. The study aimed to determine the inter- and intra-rater variability of mRSS assessment using a skin model.

**Methods:** A comparative study was conducted between January and December 2020 at Srinagarind Hospital, Khon Kaen University, Thailand. Thirty-six skin sites of 8 SSc patients underwent mRSS assessment: 4 times the first day and 1 time over the next 4 weeks by the same 10 raters. No skin model for mRSS assessment was used for the first two assessments, while one was used for the remaining three rounds of assessments. The Latin square design and Kappa statistic were used to determine inter- and intra-rater variability.

**Results:** The kappa agreement for inter-rater variability improved when the skin model was used (from 0.4 to 0.5; 25%). The improvement in inter-rater variability was seen in the non-expert group, for which the kappa agreement rose from 0.3 to 0.5 (a change of 66.7%). Intra-rater variability did not change (kappa remained at 0.9), and the long-term effect of using a skin model slightly decreased by week 4 ( $\Delta$ kappa 0.9–0.7).

**Conclusions:** Using a skin model could be used to improve inter-rater variation in mRSS assessment, especially in the non-expert group. The model should be considered a reference for mRSS assessment in clinical practice and health education.

**Keywords:** Scleroderma and related disorders, Skin, Clinical trials and methods, Primary care rheumatology, Study design

## Background

Systemic sclerosis (SSc) is a complex multisystem autoimmune connective tissue disease. SSc has been classified into two subtypes: (1) limited cutaneous systemic sclerosis (lcSSc) in which skin thickness is limited, presenting distal to the elbows and knees, with or without face

involvement, and (2) diffuse cutaneous systemic sclerosis (dcSSc) in which the extent of skin involvement presents above the elbows and knees, with or without face involvement [1]. The most common symptom and cause for concern among SSc patients is skin thickening [2].

The assessment of severity and extent of skin thickness is crucial as it is a surrogate marker of disease activity, severity, and prognosis as well as treatment responsiveness. The methods for skin thickness assessment thus need to be valid, reliable, precise, and practicable [3–6].

\*Correspondence: fching@kku.ac.th

<sup>1</sup> Department of Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand

Full list of author information is available at the end of the article



The skin biopsy validated, gold standard for skin thickness assessment for SSc is the modified Rodnan skin score (mRSS) [7–12]. The mRSS assesses skin thickness from 17 body sites: the face, chest, abdomen, arms, forearms, hands, fingers, thighs, legs, and feet. A score of 0 indicates normal skin thickness, 1 mild skin thickness, 2 moderate skin thickness, and 3 severe skin thickness with an inability to make skin folds between two fingers. The score is calculated by summing the rating from all 17 areas (range 0–51) [13, 14].

Although the mRSS has been validated at many centers, it has some limitations such as significant inter- and intra-rater variability due to (a) its subjective methodology, (b) physician inexperience, (c) significant differences between ethnic groups, (d) inaccuracies during the edematous and atrophic skin phase, and (e) lack of sensitivity in measuring minimal changes [7, 15–17].

To overcome these limitations, researchers have tried to develop new objective and quantitative methods for skin assessment. Mechanical devices and new imaging techniques include the durometer—a handheld device that measures skin hardness [18–20]; the plicometer—a medical device that measures skin folding [21]; the cutometer—a device that measures skin elasticity [22, 23]; the vesmeter—a computer-linked device that measures skin hardness, elasticity, and viscosity [24]; the twistometer—a device that measures skin rotation [25]; high-frequency ultrasound—an objective and quantitative tool that measures skin thickness [26, 27]; elastosonography—a tool that measures skin elasticity [28, 29]; shear wave elastography—a tool that measures skin thickness [30]; magnetic resonance imaging—a tool for demonstrating abnormalities of the skin and subcutaneous tissues [31, 32]; and, optical coherence tomography—a tool that identifies the microscopic features of the skin [33–35]. None of these techniques has matched the OMERACT standard of the mRSS for assessing the validity of outcomes; usually, because the techniques are not feasible in clinical practice due to time constraints, accessibility, dependence on trained experts, or lack of clarity defining what aspect of the skin to assess [8, 36].

Since mRSS is a validated outcome in scleroderma, specialist rater training was needed to improve accuracy and reduce variability. Limitations of the mRSS skin assessment include; (a) training of the mRSS skin assessment needs experienced rheumatologist as a trainer, the training process might be affected if there are limited numbers of experienced rheumatologist; (b) the mRSS assessment method takes time and the trainee needs a learning curve, and; (c) subjective skin assessment according to the mRSS method needs recall memory, so it might affect the accuracy of data and causes a recall bias. The idea of using a skin model arose as a way to address these

limitations. Trainees can use this skin model, which has been validated by experienced rheumatologists, as a reference for skin thickness severity assessment without needing to take specialist skin assessment training from a rheumatologist and using recall memory after training. Moreover, nurses and/or healthcare workers can also perform skin assessments using the skin model as a reference. If the skin model can be validated, it would help ensure initial early disease severity assessments, early management, and early referral to specialists. In addition, using the skin model might save resources and provide better care for SSc patients.

The skin model has four grades of skin thickness, just like the mRSS (viz., 0, 1, 2, and 3 as validated by experienced rheumatologists) [14, 37]. The model is used as a nonviable trainer for inexperienced physicians who palpate each site of the patient's skin, compare it to the model, and score according to the mRSS assessment method. The study's objectives were to determine the inter- and intra-rater variability of the mRSS assessment after using the skin model. If the model achieved good agreement vis-à-vis both inter- and intra-rater variabilities for inexperienced or non-expert assessors, the model could be used as a reference of skin thickness assessment as per the mRSS assessment method for health education, routine clinical practice, and/or clinical trials.

## Methods

A descriptive, comparative, reliability study was conducted on eight Thai adult SSc patients, followed up at the Scleroderma Clinic, Srinagarind Hospital, Khon Kaen University, between January and December 2020.

The skin model was developed to mimic the four grades of disease according to the mRSS. The model comprises three layers—the base is made from polypropylene to maintain structural integrity; the subcutaneous layer from a synthetic sponge with or without polyurethane foam; and, the skin layer from raw, skin-colored, or synthetic rubber. The density of the polyurethane foam is set by the degree of skin elasticity being modeled: Grade 0—sponge alone; Grade 1—low-density polyurethane foam superimposed on a synthetic sponge; Grade 2—medium density polyurethane foam superimposed on a synthetic sponge; and Grade 3—high-density polyurethane foam superimposed on the synthetic sponge. According to the skin grading, the thickness of the skin layer is adjusted to be between 1 and 5 mm (Fig. 1). Four experienced rheumatologists independently validated the model. Each grade of the skin model was labeled after validation.

A Latin square experiment was used to determine the inter- and intra-rater variability of mRSS using the skin model. The sample size was calculated based on the kappa agreement to quantify the reliability among



**Fig. 1** Skin model

ten raters doing an mRSS assessment (range, 0–3). For example, a sample size of 36 skin sites (9 sites for each mRSS grading severity) with ten raters per subject would achieve 80% power to detect a kappa agreement of 0.80 under the alternative hypothesis when the kappa agreement under the null hypothesis is 0.49 (probability 0.25, 0.30, 0.30, 0.15) using an F-test with a significance level of 0.05.

All patients were over 18 years of age and met the 1980 American Rheumatism Association classification criteria of SSc [38] or 2013 ACR/EULAR Classification Criteria for Scleroderma [39]. Patients were excluded if they had any of the following conditions: overlap syndrome, a WHO functional class IV, needing oxygen therapy at rest, not available for skin assessment (i.e., post-amputation status), recent soft tissue or skin infection, coexisting disease including cancer, severe sepsis, and/or psychiatric or neurological problems.

Ten raters were included in the study, including two rheumatologists, three rheumatology fellows, three internal medicine residents, and two nurses. The types of raters were divided into expert raters (2 rheumatologists and 3 rheumatology fellows) and non-expert raters (3 general practitioners and 2 nurses). The 36 skin sites (9 sites for each level of mRSS severity grading ranging from 0 to 3) were selected and marked on 8 patients by consensus of 2 experienced rheumatologists.

An experienced rheumatologist trained all raters on mRSS assessment before performing any assessments. Five assessments were performed. Each rater assessed 36 predetermined skin sites per session. The first and second assessments were performed about 30 min apart without using a skin model. The third and fourth assessments were performed 30 min apart using the skin model as the reference. The first four assessments were performed on day 1 of the study. The fifth assessment using the model was done 4 weeks later (Fig. 2). The assessments were independent and blinded from each rater.

### Statistical analysis

The demographic characteristics are presented using means ( $\pm$  SD) for the continuous data and numbers with percentages for the categorical data.

In the Latin square experiment, inter- and intra-rater variability of mRSS with and without using the skin model were assessed using the kappa statistic, a statistic used to measure inter- and intra-rater agreement for categorical items by eliminating agreement by chance. The kappa with its 95% confidence interval was estimated to demonstrate the level of agreement.

The first two assessments were used to investigate intra-rater reliability absent any interventions. The first assessment score was compared with the third assessment to investigate the intra-rater reliability under the rating with and without the skin model. The third and fourth assessments investigated intra-rater reliability under the rating using the skin model as the reference. The final kappa agreement comparing the score between the last two assessments 4 weeks apart was to evaluate the skin model's sustained improvement vis-à-vis inter-rater reliability. The kappa agreement was also evaluated to determine inter-rater reliability according to the experts in the mRSS assessment (expert vs. non-expert raters). The overall difference in kappa agreement between the first and fourth assessments with and without the skin model reflects the effect of using the skin model.

All data analyses were performed using STATA version 16.0 (StataCorp., College Station, TX, USA) and R program version 4.0.3.

The Human Research Ethics Committee of Khon Kaen University approved the study as per the Helsinki Declaration and the Good Clinical Practice Guidelines (HE621504). The experiment was explained to all participants who then signed informed consent before enrollment.

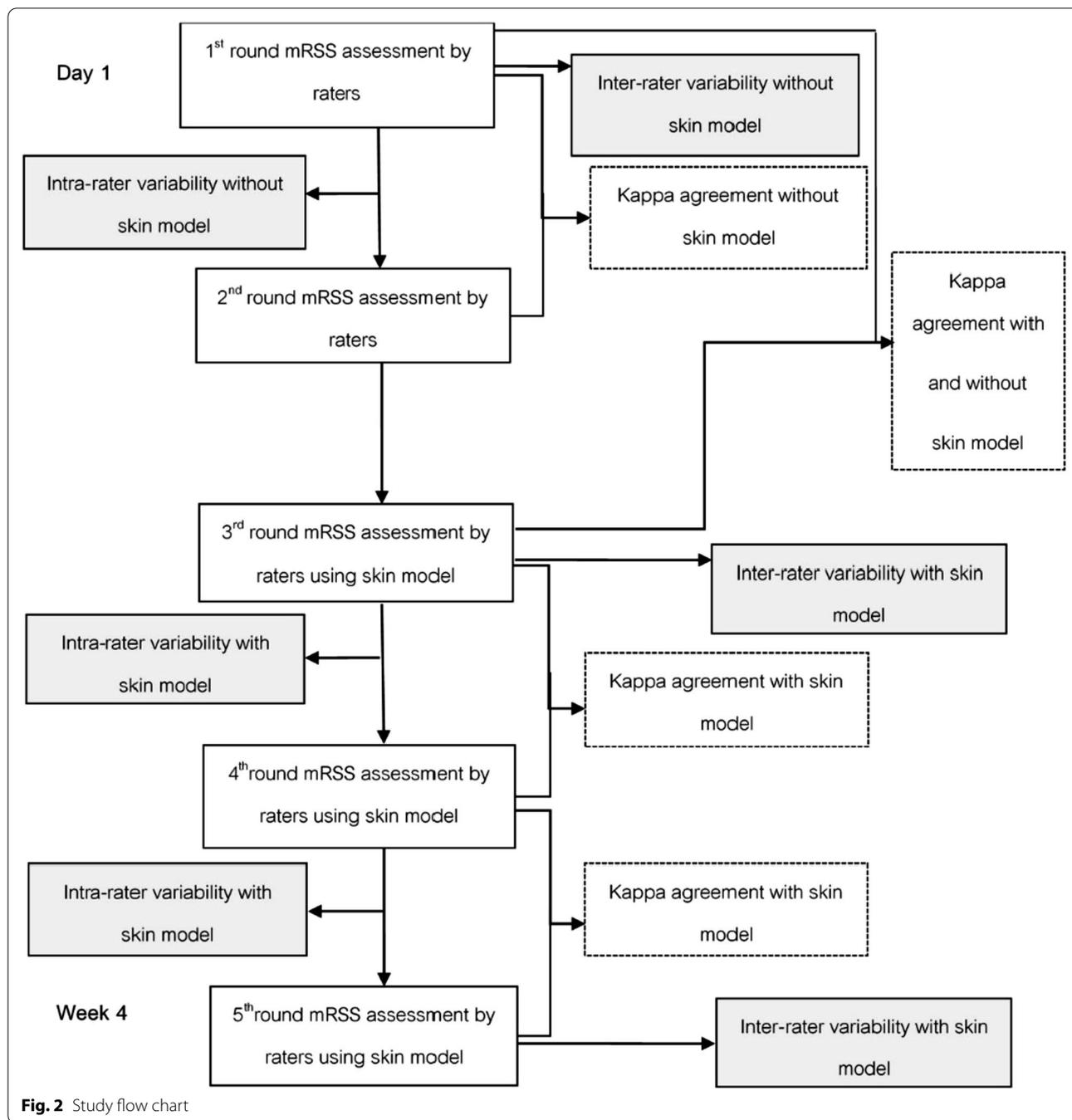


Fig. 2 Study flow chart

## Results

### Patients and raters baseline characteristics

The study included 8 SSc patients (5 males). The mean age was  $59.0 \pm 3.5$  years. Both SSc subsets were included (7 dcSSc and 1 lcSSc). The mean disease duration was  $5.8 \pm 3.9$  years. The ten raters included 7 females whose mean age was  $33.7 \pm 7.7$  years (Table 1).

### Skin score assessment

The respective individual kappa agreement for the intra-rater variability analysis with and without the skin model was 0.9 and 0.9 (Table 2.); however, the intra-rater agreement variability using the skin model at week 4 decreased slightly to 0.7 (Table 2).

**Table 1** Demographics of patients and raters presented as number and percentage unless specified otherwise

Characteristic	Number N (%)
<i>A. Patient characteristics</i>	
	N = 8
Age in years, mean ± SD (min–max)	59.0 ± 3.5 (55–64)
Duration of disease in years, mean ± SD (min–max)	5.8 ± 3.9 (0.7–13.9)
Sex female	3 (37.5)
Male	5 (62.5)
Subset dcSSc	7 (87.5)
lcSSc	1 (12.5)
<i>B. Rater characteristics</i>	
	N = 10
Age in years, mean ± SD (min–max)	33.7 ± 7.7 (25–48)
Sex: female	7 (70)
Male	3 (30)
Type of raters	
Rheumatologists	2 (20)
Fellowship in rheumatology	3 (30)
Internists	3 (30)
Nurse	2 (20)

SD standard deviation

The respective overall kappa agreement for inter-rater variability analysis with and without the skin model was 0.5 and 0.4, respectively (25% difference). Notwithstanding, the kappa agreement for inter-rater variability analysis in the non-expert group was improved by 67% when using the skin model compared to not using the skin model (kappa = 0.3 vs. 0.5), while the kappa agreement was comparable in the expert group with and without the skin model (kappa = 0.5 vs. 0.5). (Table 3.)

The overall and individual skin thickness agreement in day 1 and week 4 are presented in Additional file 1:

Table S1, Additional file 2: Table S2, Additional file 3: Table S3, Additional file 4: Table S4, Additional file 5: Table S5, Additional file 6: Table S6 and Additional file 7: Table S7.

### Discussion

Skin thickness is the most troubling clinical symptom among SSc patients. In SSc, the extent and severity of skin thickness are also associated with the severity of internal organ involvements, poor prognosis, morbidity, and mortality [2–6]. The accuracy and validity of mRSS, the gold standard for skin thickness assessment, are essential [7, 8, 10]. Although previous studies revealed that the mRSS was a valid and reliable test for skin thickness assessment, there is significant inter- and intra-rater variation, particularly when inexperienced assessors perform it [10, 15, 16]. A similar finding was shown in a study by Foocharoen et al. conducted among Thai SSc patients [17]. To overcome this limitation, we developed a skin model to use as a reference for mRSS assessment to reduce inter- and intra-rater variability in mRSS assessment.

We found that the overall inter-rater variability (kappa agreement) of the mRSS assessment when using the skin model improved from 0.4 to 0.5—a 25% improvement. In the non-expert group, where the kappa agreement had been 66.7%, agreement rose from 0.3 to 0.5. By comparison, in the expert group, the inter-rater agreement did not improve (before 0.5 and after 0.5). Our results agree with previous studies where mRSS assessment accuracy improved with experience [15, 16, 40, 41]. The inter-rater variability for mRSS assessment using the skin model was good compared to raters who took a standardized mRSS training course [40]. The EUSTAR study (European League Against Rheumatism Scleroderma Trials

**Table 2** Kappa agreement for intra-rater variability of skin thickness scoring agreement

	Agreement (%)	Overall kappa	95%CI
Assessment without skin model (round 1 and 2)	96.1	0.9	0.82–0.98
Assessment with skin model (round 3 and 4)	95.4	0.9	0.82–0.98
Assessment with skin model (round 4 and 5)	89.8	0.7	0.62–0.78

CI confidence interval

**Table 3** kappa agreement for inter-rater variability analysis in overall raters, expert-raters, and non-expert raters

	Overall kappa	95%CI	Kappa of expert	95%CI	Kappa of non-expert	95%CI
Assessment without skin model	0.4	0.39–0.41	0.5	0.47–0.53	0.3	0.27–0.33
Assessment skin model	0.5	0.49–0.51	0.5	0.47–0.53	0.5	0.47–0.53
Percentage of difference	25.0%	–	0%	–	66.7%	–

CI confidence interval

and Research) reported that repeated mRSS assessment training courses decrease inter-rater variability. The ICC (intra-class correlation coefficient) agreement increased from 0.5 to 0.7 after two training courses in less experienced rheumatologists, while for experts it was very good at the outset and did not increase [42]. The authors also found that intra-rater variability was relatively good before mRSS assessment training and remained stable thereafter [42]. Another report reported a significant difference in the training effect according to physicians' professional seniority. The training effect was more pronounced in students than in senior staff [40]. The results of our study likewise showed good intra-rater agreement that did not change when using the skin model. In addition, using a skin model can save time, no need to recall memory, no need learning curve, and the mRSS assessment training by the specialist. Our results suggest that the skin model might be used as a reference for mRSS assessment instead of a health education training course.

Intra-rater variability doing the mRSS assessment with or without the skin model was comparable on day 1 (pre-/post kappa were both 0.9), perhaps because of a post-training effect lending confidence on how to assess skin thickness using mRSS. After evaluating the long-term effects on intra-rater variability, the kappa agreement was slightly dropped from 0.9 on the first day of evaluation to 0.7 by week 4, even though skin thickness had hardly progressed over the 4-week study period. The results suggest that the time interval might affect intra-rater variability. Although there was a slight increase in long-term intra-rater variability, the inter-rater variability was not much improved. The findings suggest a decline in rater confidence in only 1 round of skin assessment, week 4, without training or orientation. Further study is suggested to evaluate intra-rater variability over a longer duration to confirm the long-term effect on intra-rater variability when using the skin model. We propose doing further study, including internal and external validation of the skin model vis-à-vis implementation as a reference for mRSS assessment in daily practice and/or clinical trials.

Our study had some limitations, including (1) we are a single-center trial, so further investigation is needed of the external validity of the test, and (2) we enrolled only Asians whose skin thickness may be different from other ethnic groups [10, 43]. Notwithstanding, the reliability of mRSS assessment does not depend on the patients' skin rather it depends upon the assessors. Irrespective of the patient's nationality or ethnicity, the skin assessment is unlikely to influence the findings. Although a small number of patients (8 patients with 36 skin sites) were included in determining inter- and intra-rater variability of mRSS using the skin model, we are confident

that we have included an adequate sample size according to the method of sample size calculation. Further study for external validation of the skin model should include a larger sample size.

The strengths of our study are that (1) we included both dcSSc and lcSSc, both sexes, various ages, various disease durations, and various sites of skin assessment in SSc which can represent general SSc patient and can be generalized; (2) both expert and non-expert raters were included, so the results can be implemented in daily practice and/or in clinical trial(s) if experienced rheumatologists or specialists are unavailable; (3) the skin assessment was performed 4 times on day 1 and one more time during week 4 (i.e., 2 times without the skin model and 3 times with the skin model). We thus have confidence that we have investigated both inter- and intra-rater variability; and (4) we used materials that are versatile, low-cost, available, and nontoxic. The textiles are like human skin, so the skin model can be easily produced, is harmless and long-lasting.

## Conclusion

The skin model improves inter-rater reliability of mRSS assessment, especially in the non-expert group. This finding suggests that the skin model might be helpful as a reference for mRSS assessment training and in clinical practice. However, further study is needed to assess the validity and the sustained effect in a larger population.

## Abbreviations

SSc: Systemic sclerosis; lcSSc: Limited cutaneous systemic sclerosis; dcSSc: Diffuse cutaneous systemic sclerosis; mRSS: Modified Rodnan skin score.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41927-022-00262-2>.

**Additional file 1.** Overall skin thickness scoring agreement without skin model on day 1.

**Additional file 2.** Overall skin thickness scoring agreement with and without skin model on day 1.

**Additional file 3.** Overall skin thickness scoring agreement with skin model on day 1.

**Additional file 4.** Overall skin thickness scoring agreement with skin model between round 4 and 5.

**Additional file 5.** Individual skin thickness scoring agreement without skin model (1st and 2nd round).

**Additional file 6.** Individual skin thickness scoring agreement with skin model (3rd and 4th round).

**Additional file 7.** Individual skin thickness scoring agreement with skin model (4th and 5th round).

### Acknowledgements

The authors thank (a) the Scleroderma Research Group and Faculty of Medicine, Khon Kaen University, for its support, (b) the Thai Rheumatism Association for its support, and (c) Mr. Bryan Roderick Hamman under the aegis of the Publication Clinic Khon Kaen University, Thailand for assistance with the English-language presentation. The manuscript was presented as poster presentation in the Pacific League of Associations for Rheumatology (APLAR) Congress, 28–31 August 2021, <https://onlinelibrary.wiley.com/doi/10.1111/1756-185X.14200>.

### Authors' contributions

PP designed the study, did the data collection, and drafted the manuscript. BT designed the study and proofread the manuscript. AM and SS did data collection, edited, and proofread the manuscript. CF designed the study, did data collection, data analysis, proofed, and approved the manuscript. All authors read and approved the final manuscript.

### Funding

Research on Faculty of Medicine, Khon Kaen University has received funding support from National Science, Research and Innovation Fund. The funder had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

### Availability of data and materials

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

### Declarations

#### Ethics approval and consent to participate

The Human Research Ethics Committee of Khon Kaen University reviewed and approved the study as per the Helsinki Declaration and the Good Clinical Practice Guidelines (HE621504). Furthermore, all eligible patients signed informed consent before enrollment.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Medicine, Faculty of Medicine, Khon Kaen University, Khon Kaen 40002, Thailand. <sup>2</sup>Department of Epidemiology and Biostatistics, Faculty of Public Health, Khon Kaen University, Khon Kaen 40002, Thailand.

Received: 3 October 2021 Accepted: 22 March 2022

Published online: 02 June 2022

### References

- LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. *J Rheumatol*. 1988;15:202–5.
- Chularojanamontri L, Sethabuttra P, Kulthanan K, Manapajon A. Dermatology life quality index in Thai patients with systemic sclerosis: a cross-sectional study. *Indian J Dermatol Venereol Leprol*. 2011;77:683. <https://doi.org/10.4103/0378-6323.86481>.
- 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. [https://doi.org/10.1002/1529-0131\(200011\)43:11%3c2445::AID-ANR11%3e3.0.CO;2-Q](https://doi.org/10.1002/1529-0131(200011)43:11%3c2445::AID-ANR11%3e3.0.CO;2-Q).
- Steen VD, Medsger TA. Improvement in skin thickening in systemic sclerosis associated with improved survival. *Arthritis Rheum*. 2001;44:2828–35. [https://doi.org/10.1002/1529-0131\(200112\)44:12%3c2828::aid-art470%3e3.0.co;2-u](https://doi.org/10.1002/1529-0131(200112)44:12%3c2828::aid-art470%3e3.0.co;2-u).
- Shand L, Lunt M, Nihtyanova S, et al. Relationship between change in skin score and disease outcome in diffuse cutaneous systemic sclerosis: application of a latent linear trajectory model. *Arthritis Rheum*. 2007;56:2422–31. <https://doi.org/10.1002/art.22721>.
- Tyndall AJ, Bannert B, Vonk M, et al. Causes and risk factors for death in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. *Ann Rheum Dis*. 2010;69:1809–15. <https://doi.org/10.1136/ard.2009.114264>.
- Pope JE, Bellamy N. Outcome measurement in scleroderma clinical trials. *Semin Arthritis Rheum*. 1993;23:22–33.
- Furst DE. Outcome measures in rheumatologic clinical trials and systemic sclerosis. *Rheumatology (Oxford)*. 2008;47:v29–30. <https://doi.org/10.1093/rheumatology/ken269>.
- Kumánovics G, Péntek M, Bae S, et al. Assessment of skin involvement in systemic sclerosis. *Rheumatology (Oxford)*. 2017;56:v53–66. <https://doi.org/10.1093/rheumatology/kex202>.
- 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. <https://doi.org/10.5301/jsrd.5000231>.
- 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.
- Verccechia F, Laboureaux J, Verola O, et al. Skin involvement in scleroderma—where histological and clinical scores meet. *Rheumatology (Oxford)*. 2007;46:833–41. <https://doi.org/10.1093/rheumatology/kel451>.
- Brennan P, Silman A, Black C, et al. Reliability of skin involvement measures in scleroderma. *Rheumatology (Oxford)*. 1992;31:457–60. <https://doi.org/10.1093/rheumatology/31.7.457>.
- Black CM. Measurement of skin involvement in scleroderma. *J Rheumatol*. 1995;22:1217–9.
- Clements PJ, Lachenbruch PA, Seibold JR, et al. Skin thickness score in systemic sclerosis: an assessment of interobserver variability in 3 independent studies. *J Rheumatol*. 1993;20:1892–6.
- Clements P, Lachenbruch P, Siebold J, et al. Inter- and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol*. 1995;22:1281–5.
- Foocharoen C, Thinkhamrop B, Mahakkanukrauh A, et al. Inter- and intra-observer reliability of modified Rodnan skin score assessment in Thai systemic sclerosis patients: a validation for multicenter scleroderma cohort study. *J Med Assoc Thai*. 2015;98:1082–8.
- Falanga V, Bucalo B. Use of a durometer to assess skin hardness. *J Am Acad Dermatol*. 1993;29:47–51. [https://doi.org/10.1016/0190-9622\(93\)70150-r](https://doi.org/10.1016/0190-9622(93)70150-r).
- Kissin EY, Schiller AM, Gelbard RB, et al. Durometry for the assessment of skin disease in systemic sclerosis. *Arthritis Care Res*. 2006;55:603–9. <https://doi.org/10.1002/art.22093>.
- 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. <https://doi.org/10.1002/art.23564>.
- Parodi MN, Castagneto C, Filaci G, et al. Plicometer skin test: a new technique for the evaluation of cutaneous involvement in systemic sclerosis. *Br J Rheumatol*. 1997;36:244–50. <https://doi.org/10.1093/rheumatology/36.2.244>.
- Enomoto DN, Mekkes JR, Bossuyt PM, et al. Quantification of cutaneous sclerosis with a skin elasticity meter in patients with generalized scleroderma. *J Am Acad Dermatol*. 1996;35:381–7. [https://doi.org/10.1016/s0190-9622\(96\)90601-5](https://doi.org/10.1016/s0190-9622(96)90601-5).
- Ishikawa T, Tamura T. Measurement of skin elastic properties with a new suction device (II): systemic sclerosis. *J Dermatol*. 1996;23:165–8. <https://doi.org/10.1111/j.1346-8138.1996.tb03992.x>.
- Kuwahara Y, Shima Y, Shirayama D, et al. Quantification of hardness, elasticity and viscosity of the skin of patients with systemic sclerosis using a novel sensing device (Vesmeter): a proposal for a new outcome measurement procedure. *Rheumatology*. 2008;47:1018–24. <https://doi.org/10.1093/rheumatology/ken145>.
- Czirjak L, Foeldvari I, Muller-Ladner U. Skin involvement in systemic sclerosis. *Rheumatology*. 2008;47:v44–5. <https://doi.org/10.1093/rheumatology/ken309>.
- Ch'ng SS, Roddy J, Keen HI. A systematic review of ultrasonography as an outcome measure of skin involvement in systemic sclerosis. *Int J Rheum Dis*. 2013;16:264–72. <https://doi.org/10.1111/1756-185X.12106>.

27. Bendeck SE, Jacobe HT. Ultrasound as an outcome measure to assess disease activity in disorders of skin thickening: an example of the use of radiologic techniques to assess skin disease. *Dermatol Ther.* 2007;20:86–92. <https://doi.org/10.1111/j.1529-8019.2007.00116.x>.
28. Iagnocco A, Kaloudi O, Perella C, et al. Ultrasound elastography assessment of skin involvement in systemic sclerosis: lights and shadows. *J Rheumatol.* 2010;37:1688–91. <https://doi.org/10.3899/jrheum.090974>.
29. Di Geso L, Filippucci E, Girolimetti R, et al. Reliability of ultrasound measurements of dermal thickness at digits in systemic sclerosis: role of elastosonography. *Clin Exp Rheumatol.* 2011;29:926–32.
30. Wakhlu A, Chowdhury A, Mohindra N, et al. Assessment of extent of skin involvement in scleroderma using shear wave elastography. *Indian J Rheumatol.* 2017;12:194–8. [https://doi.org/10.4103/injr.injr\\_41\\_17](https://doi.org/10.4103/injr.injr_41_17).
31. Chapin R, Hant FN. Imaging of scleroderma. *Rheum Dis Clin.* 2013;39:515–46. <https://doi.org/10.1016/j.rdc.2013.02.017>.
32. Kang T, Abignano G, Lettieri G, et al. Skin imaging in systemic sclerosis. *Eur J Rheumatol.* 2014;1:111–6. <https://doi.org/10.5152/eurjrheumatol.2014.036>.
33. Mogensen M, Thrane L, Joergensen TM, et al. Optical coherence tomography for imaging of skin and skin diseases. *Semin Cutan Med Surg.* 2009;28:196–202. <https://doi.org/10.1016/j.sder.2009.07.002>.
34. Marschall S, Sander B, Mogensen M, et al. Optical coherence tomography—current technology and applications in clinical and biomedical research. *Anal Bioanal Chem.* 2011;400:2699–720. <https://doi.org/10.1007/s00216-011-5008-1>.
35. Abignano G, Aydin SZ, Castillo-Gallego C, et al. Virtual skin biopsy by optical coherence tomography: the first quantitative imaging biomarker for scleroderma. *Ann Rheum Dis.* 2013;72:1845–51. <https://doi.org/10.1136/annrheumdis-2012-202682>.
36. Khanna D, Merkel P. Outcome measures in systemic sclerosis: an update on instruments and current research. *Curr Rheumatol Rep.* 2007;9:151–7. <https://doi.org/10.1007/s11926-007-0010-5>.
37. Rodnan GP, Lipinski E, Luksick J. Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma. *Arthritis Rheum.* 1979;22:130–40. <https://doi.org/10.1002/art.1780220205>.
38. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum.* 1980;23:581–90.
39. van den Hoogen F, Khanna D, Fransen J, et al. Classification criteria for systemic sclerosis: an ACR-EULAR collaborative initiative. *Arthritis Rheum.* 2013;65:2737–47. <https://doi.org/10.1002/art.38098>.
40. Park JW, Ahn GY, Kim J-W, et al. Impact of EUSTAR standardized training on accuracy of modified Rodnan skin score in patients with systemic sclerosis. *Int J Rheum Dis.* 2019;22:96–102. <https://doi.org/10.1111/1756-185X.13433>.
41. 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. <https://doi.org/10.1136/ard.2006.066530>.
42. Pope JE, Baron M, Bellamy N, et al. Variability of skin scores and clinical measurements in scleroderma. *J Rheumatol.* 1995;22:1271–6.
43. Perera A, Fertig N, Lucas M, et al. Clinical subsets, skin thickness progression rate, and serum antibody levels in systemic sclerosis patients with anti-topoisomerase I antibody. *Arthritis Rheum.* 2007;56:2740–6. <https://doi.org/10.1002/art.22747>.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

