

Standardization of modified Rodnan skin score in clinical trials of systemic sclerosis

Dinesh Khanna¹, Daniel E. Furst², Philip J. Clements², Yannick Allanore³, Lazlo Czirjak⁴, Oliver Distler⁵, Ivan Foeldvari,⁶ Masataka Kuwana⁷, Marco Mattucci-Cerinic⁸, Maureen Mayes⁹, Peter Merkel¹⁰, James Seibold¹¹, Christopher P. Denton¹² on behalf of the SCTC and the WSF

Affiliations

¹University of Michigan Scleroderma Program, Ann Arbor, MI, USA; ²UCLA, Los Angeles, CA, USA; ³Paris Descartes University, Cochin Hospital, Paris, France; ⁴Department of Immunology and Rheumatology, University of Pécs, Hungary; ⁵Division of Rheumatology, University Hospital Zurich, Zurich, Switzerland; ⁶Hamburger Zentrum für Kinder und Jugendrheumatologie, Hamburg, Germany; ⁷Department of Allergy and Rheumatology, Nippon Medical School Graduate School of Medicine, Keio University School of Medicine, Tokyo, Japan; ⁸ Department of Experimental & Clinical Medicine, Division of Rheumatology AOUC, University of Florence, Firenze, Italy; ⁹University of Texas Health Science Center Houston, TX, USA; ¹⁰University of Pennsylvania, Philadelphia, PA, USA; ¹¹Scleroderma Research Consultants, Litchfield, CT, USA; ¹²Centre for Rheumatology, Royal Free and University College London Medical School, London, UK.

Corresponding author

Dinesh Khanna, MD, MSc
Professor of Medicine
Director, University of Michigan Scleroderma Program
Division of Rheumatology/Dept. of Internal Medicine Suite 7C27
300 North Ingalls Street, SPC 5422
Ann Arbor, MI 48109
Email: khannad@med.umich.edu
Phone: 734.647.8173
Fax: 734.763.5761

CONFLICTS

Dr. Khanna has/had consultancy relationship with and/or has received research funding from Bayer, Bristol-Myers Squibb, Cytos, EMD Serono, Genentech/ Roche, Gilead, Glaxo SmithKline, Medac, Sanofi-Aventis/Genzyme. This work was funded in part by the NIH/NIAMS K24 grant.

Dr. Denton has had consultancy or speaker fees from Actelion, Bristol-Myers Squibb, GlaxoSmithKline, Bayer, Sanofi-Aventis, Boehringer Ingelheim, Genentech-Roche, CSL Behring and research grant funding from GlaxoSmithKline, Actelion, Novartis, CSL Behring.

Abstract

To be added

Introduction

The modified Rodnan skin score (mRSS) is a measure of skin thickness and is used as a primary or secondary outcome measure in clinical trials of systemic sclerosis (SSc). This article gives a brief history of mRSS development, outlines practical aspects of assessing mRSS in clinical trials, and provides recommendations for Phase 2 and 3 clinical trials where the measurement of skin thickness is the primary or a secondary outcome measure.

Natural history of skin involvement in SSc

SSc is subclassified into “diffuse cutaneous” or “limited cutaneous” based on the extent of skin involvement. In diffuse cutaneous SSc (dcSSc), it is generally accepted that the mRSS tends to worsen in early disease and decrease in late disease, although the time of peak involvement remains poorly defined. “Early” dcSSc is often defined as the period of rapidly and severely increasing induration (“thickening”) of the skin, which has been thought to peak 1.5-5 years after onset of the disease. Skin involvement in dcSSc usually goes through 3 phases: 1. An edematous phase that usually lasts 6-12 months, a fibrotic or indurative phase that lasts from 1-4 years or longer, and an atrophic phase, that lasts for rest of the patient’s life . (Figure xxx). The duration of different phases may differ in each individual patient. The time to peak MRSS in a cohort of dcSSc has been described previously and helps inform definitions of early skin disease (see: Nihtyanova SI, Denton CP. Current approaches to the management of early active diffuse scleroderma skin disease. *Rheum Dis Clin North Am.* 2008 Feb;34(1):161-79; viii. doi: 10.1016/j.rdc.2007.11.005.).

Measurement of skin thickness is used as a surrogate measure of disease severity and mortality in patients with dcSSc—an increase in skin thickening is generally associated with involvement of internal organs and increased mortality (Clements D Pen; Shand AR). In addition, the trajectory of worsening of mRSS is associated with worse outcomes, including higher mortality (Domsic ARD). The change in mRSS is associated with change in mortality, renal and cardiac outcomes (Clements PJ, Hurwitz EL, Wong WK et al. Skin thickening score as a predictor and correlate of outcome... A&R 2000. 43:2445-2454; Marco PJ, Weisman MH, Seibold JR, Furst DE et al. Predictors and outcomes of scleroderma renal crisis. A&R 2002. 46:2983-2989).. Some nuance was added by Shand L et al who showed that outcomes were particularly associated with the prior trajectory of skin thickening(Shand L, Lunt M, Nihtyanova S,Hoseini M et al. Relationship between change in skin score and disease outcome in diffuse cutaneous systemic sclerosis. A&R 2007. 56: 2422-2431) In addition, those cases that never achieved a high skin score had lower frequency of major organ based complications and those with high skin scores that failed to improve by 3 years of disease has significantly higher mortality, implying that attenuation of peak skin score or improvement may reflect better outcome. This point is noteworthy as we need to emphasize the potential benefit of including milder or early cases - both a preventative as well as an improvement strategy for clinical trials and future clinical practice. CD: I am sure Oliver will add something good here about the EUSTAR analysis and cohort enrichment.

Historical perspective

Phil: I did the above.this one is yours, ok? DF

mRSS as an outcome measure

mRSS meets the OMERACT filters of truth, discrimination, and feasibility. This has been detailed by Merkel et al and the readers are encouraged to review the OMERACT proceedings. Briefly, mRSS has the following characteristics.

- a. Feasible—easy to do with little equipment and requires little time
- b. Face Validity—captures skin thickness
- c. Content Validity— covers the affected body areas in SSc
- d. Construct validity—correlates with other measures of SSc such as durometer and ultrasound (ref)
- e. Criterion Validity—reflects gold standard-biopsy (ref)
- f. Sensitivity to change—analysis of multicenter clinical trials show change over time. For example, the MRSS was able to differentiate between treatment with methotrexate and placebo in early diffuse SSc and between cyclophosphamide and placebo in the diffuse SSc subset in the Scleroderma Lung Study-I. The MRSS also improved in the Autologous Stem Cell Transplant Study and showed a trend towards improvement for tocilizumab.

Conducting mRSS in clinical trials

Practical aspects

1. **Positioning of the patient:** This is an important part of assessing mRSS and should be done in a relaxed fashion. Figure xx shows proper positioning of the patient for conducting mRSS. We recommend that the patient be in sitting a relaxed position with arms by the side of the body when examining upper extremities, face, and chest. We recommend that the patient lie down on a bed while examining the abdomen and lower extremities, with the hips at 45 degrees flexion and dorsi flexion of the feet. The point is that the muscles underlying the areas of skin be relaxed or there is a danger of falsely elevated skin scores.
2. **Assessing skin thickness:** There are 2 techniques to accomplish this by palpation of the skin(Figure xx).
 - a. Index finger and thumb. This is preferred method where the skin is rolled gently to gauge thickness.
 - b. The other method is to use two thumbs to form a fold of skin between advancing lateral thumb borders and may be easier when skin overlies bone such as the fingers and the dorsum of the hands.
 - c. It is important to understand the relative distribution of fat and underlying structure as it can make the skin thickness feel differently in different anatomical areas.
 - d. Skin should be scored based on how the skins feels and not how it looks. For example, shiny skin, usually over the chest, doesn't mean that skin has increased thickness.

Global average vs. maximum score of an anatomic area—There are differences in scoring mRSS in clinical practice but we recommend standardizing mRSS measurement in clinical trials. We propose to average the score of a given anatomic area (Figure xx). This has been successfully used in previous clinical trials. In addition, there is theoretical disadvantage of using maximum score over an anatomic area as it may limit sensitivity to change. Consider an area over forearm [as an example] that remains a 3 but the skin improves overall. Using maximum score, this will continue to be scored as mRSS of 3 but the global average will change (improve) over time as it incorporates the whole area. One point for clarification is that if an area of skin in one of the 17 sites has grade 1 but no more severe thickening is in that site then the score of 1 is given – i.e. for global average maximizing applies in this situation.

3. **Scoring of each individual body area** (Figure xx)
 - a. mRSS =0 is defined as normal skin where the assessor can appreciate fine wrinkles and appreciable skin thickness appreciated

- b. mRSS =1 Appreciate skin thickness and easily able to make the skin fold between 2 fingers and may demonstrate fine wrinkles. mRSS=0 vs. 1 depends on the feel of the skin
 - c. mRSS =2 Appreciate skin thickness with difficulty in making skin folds and unable to appreciate wrinkles
 - d. mRSS =3 Appreciate skin thickness with inability to make easy skin folds between 2 fingers

- 4. **Same examiner should evaluate the patient during the trial:** Each outcome measure inherently has measurement variability. It is recommended that same assessor examine the patient during the duration of the trial. The inter-observer variability of mRSS is 4.6 units (coefficient of variation=25%) and the intra-observer variability is 2.5 units (coefficient of variation= 12%). Although the mRSS has been considered less than satisfactory outcome measure by rheumatology community, in a study assessing inter-observer variability of joint and tender joint count in rheumatoid arthritis, the coefficient of variation was 82% for swollen joints and 66% for tender joints (Scott et al Clinical Rheumatology 1996).

- 5. **Teaching reduces variability of mRSS (Laszlo and Dan)** The mRSS is a validated outcome in diffuse SSc but, as in rheumatoid arthritis, training of the measure is important. In a 12 investigator teaching course, the ICC was 0.639, a very good reliability. On repeat training 7 months later, the ICC was 0.684, essentially the same number. This seems to indicate that repeated retraining over time is not necessary (at least over 7 months). On the other hand, the other hand, inter-investigator variability was high at both training sessions (standard deviation: 3-5-5.3) emphasizing the need to have the same investigator consistently measuring the same patient.

- 6. **Caveats in mRSS measurement**
 - a. The edematous phase of dcSSc is usually associated with non edematous pitting of the skin. mRSS may be difficult to assess and has been shown to be inaccurate with a correlation between an edema score and edema on biopsy of 0.069-0.169 (ibid Furst, Clements PJ, Steen VD et al). There is no agreed upon approach to this issue (although some score one lower than the apparent score if there is edema (3+ becomes 2+ etc, even including 1+ becoming 0))
 - b. Tethering or hidebinding is defined as involvement of underlying tissues [such as fascia and muscles] over an anatomic area [generally fingers and dorsum of the hands] . mRSS does not score tethering. A UCLA developed score is generally agreed to measure tethering rather than thickness or, at least, cannot differentiate tethering from thickening. It is not surprising that it is less sensitive than the mRSS, although it has changed in a trial of methotrexate for SSc (Sultan N, Pope JE, Clements

PJ et al. The HAQ is strongly predictive of good outcome in early diffuse scleroderma. Rheum(Oxford).2004.43:472-478.)

- c. Atrophic skin should be given a mRSS of 0 (Figure xx)

7. Other aspects

- a. Variability within different races/ countries and paediatric population Masa and Ivan to write and others add to this

Recommendations for assessing mRSS in clinical trials

Phase 3 trial focused on mRSS as primary outcome measure or key secondary outcome measure

Training is strongly encouraged for standardization of mRSS.
We recommend 2 phases for training.

- 2 Teaching phase: This should be conducted by an expert, following points should be made in 1- 2 patients with SSc
 1. Show proper technique of performing MRSS including keeping body parts in relaxed position and highlight most likely involved areas.
 2. mRSS should be performed based on the 'average' method , assigning individual areas with a score that is most representative of the area under examination (For example, if the distal forearm has a patch considered 2+ whereas the remainder of the forearm is 1+, then 1+ would be the most representative score. Alternatively, if the area rated 2+ was relatively extensive, then 2+ would be the recommended score for the forearm).
 3. Skin should be scored based on how the skins feels and not how it looks. Shiny skin, usually over the chest, doesn't mean that the skin is thick.
 4. Each site for the skin involvement (dorsum of the hand, forearm) should be scored from 0 to 3. Tethered skin areas, for purpose of skin scoring, should be scored as 0 as it is usually associated with atrophic skin.

This should be followed by an Evaluation phase where each assessor should preferably score 3 patients with SSc. These patients should be different than the training patients and preferably reflect the inclusion criteria for the proposed trial.

1. The experts should examine each patient and reach a consensus score.
2. Trainees should be given the mRSS sheet (Figure xx and hyperlink for downloadable copy) and asked to examine each patient.
3. Due to the measurement variability in mRSS, we recommend the following criteria for certification for the clinical trial.
 - a. if the trainee is within +/-5 units of the experts' score for all 3 patients, then the trainee has passed the Evaluation phase.
 - b. if the trainee is within +/-5 units of the experts' score for 2 of 3 patients, then the expert should review the disparity with the trainee and clarify concerns and

the trainee preferably be asked to review mRSS training video and then considered to have passed the Evaluation phase.

c. if the trainee is within +/-5 units of the experts' score for 1 of 3 patients, then the trainee should repeat the training, either on the same day or another venue.

All results should be documented for quality control and each trainee should receive a certificate at the end of training that should be part of the regulatory binder for the trial.

Phase 3 trial where mRSS is a secondary outcome measure or in a Phase 2 focused on mRSS as the primary outcome measure or key secondary outcome measure

It is encouraged to use the aforementioned plan for standardization of the mRSS but we acknowledge the costs associated with training each assessor. In lieu of this, consider training at the investigator meeting where the above points should be discussed and a video or live demonstration should be performed and the trainee should be provided enough time for clarifications. Even in this situation, at least one patient should be scored under the supervision of an expert so that clarity of the methods can be achieved. Each trainee should receive a certificate at the end of investigator meeting that should be part of the regulatory binder for the trial.

Finally, we recommend to discuss the above mentioned guidance with the regulatory agencies.