

CONCISE REPORT

The EUSTAR model for teaching and implementing the modified Rodnan skin score in systemic sclerosis

László Cziráj, Zoltán Nagy, Martin Aringer, Gabriela Riemekasten, Marco Matucci-Cerinic, Daniel E Furst, on behalf of EUSTAR

Ann Rheum Dis 2007;66:966–969. doi: 10.1136/ard.2006.066530

Objective: To evaluate the ability to teach scleroderma experts and young rheumatologists to perform the modified Rodnan skin score test.

Methods: Three international “teaching courses for teachers” were conducted with 6–9 experts who performed 3–9 skin score tests each. In addition, an international course for 90 young rheumatologists, in which 18 patients with systemic sclerosis (SSc) participated, was also organised. Finally, a local repeated training course for 5–9 rheumatologists was performed, in which 6–7 patients with SSc participated.

Results: When 6–9 scleroderma specialists investigated the patients, the intraclass correlation coefficient (ICC) showed “good” to “excellent” values (0.865 and 0.710, respectively). When 90 young rheumatologists were involved in one teaching course, the coefficient of variation (CV) was relatively satisfactory (35%) owing to the high number of investigators, and with a considerable within-patient SD value of 5.4.

Repeated teaching of 5–9 young rheumatologists in local courses clearly improved the consistency. The ICC increased from 0.496 to a “good” level of 0.722. The within-patient SD values for intraobserver variability ranged between 2.5 and 2.9. The intraobserver CV was about 20%.

Conclusions: This study strongly supports the need for standardisation among different centres when using skin scoring for clinical trials. The intraobserver variability and within-patient SD values can be significantly reduced by repeated teaching. For inexperienced rheumatologists, at least one repeated teaching course is needed.

Systemic sclerosis (SSc) is characterised by tightening and thickening of the skin, and involvement of internal organs. Particularly early in the disease, more extensive skin involvement coincides with more severe internal organ manifestation(s), poor prognosis^{1,2} and increased disability.^{3,4} Derived from the original Rodnan skin score,⁵ several modified and simplified skin score measuring methods have been developed.^{2,6,7} The modified Rodnan skin score (MRSS) uses clinical palpation to estimate the skin thickness. At present, the MRSS is considered the most appropriate and reproducible technique for measuring skin involvement in SSc.^{2,6,8–10} The method is easily used and fully validated.^{8,9}

The aims of this study were (1) to establish a method for teaching larger numbers of rheumatologists to appropriately perform the MRSS and (2) to evaluate the interobserver and intraobserver variabilities of skin scoring in a group of mixed patients with SSc with both limited and diffuse disease examined by different rheumatologists from centres all over Europe.

METHODS

The fully validated^{8,9} modified version of the Rodnan skin thickness score^{5,9} was used. A total of 17 skin sites were

evaluated, including the face, upper arms, forearms, dorsum of the hands, fingers, chest, abdomen, thighs, forearms and feet. The maximum total score is 51 (MRSS-51) and the grading is 0, normal skin; 1, thickened skin; 2, thickened and unable to pinch; and 3, thickened and unable to move.

Investigators

Teaching course for teachers

The 1st course (Berlin, June 2004), 2nd course (Budapest, January 2005) and 3rd course (Vienna, June 2005) were organised in a similar fashion. First, the “master teacher” (DEF) explained and demonstrated the technique of MRSS to the nine examiners, and examined each patient to establish the standard. Then, the examiners performed the investigations on the patients with SSc. The proportion of patients with diffuse cutaneous SSc was 66–100%, with a range of 33–80% of early cases (disease duration <3 years).

Teaching of a large number of young rheumatologists

A total of 90 young rheumatologists participated from all over Europe at the training in Budapest (January 2005), guided by the same master trainer (DEF). The master trainer gave a further demonstration of the technique to the nine tutors, followed by a discussion of details. Next, with the assistance of a patient with scleroderma, the master trainer demonstrated the details of the technique to the 87 “trainee” rheumatologists. Meanwhile, all tutors evaluated their patients, and after that, reasons for the recorded variability were once more discussed with the master trainer. Thereafter, the 87 trainee rheumatologists were divided into nine groups, each consisting of 5–6 students. Each trainee group had two patients for practice. Following a discussion with the tutor, each trainee investigated another two patients.

Local courses to investigate the intraobserver variability

The first local teaching course (Pécs, April 2005) was organized as described above. Nine rheumatologists investigated 6 patients each a local teacher from the previous course. During the second local course (Pécs, June 2005) 5 examiners who had participated in the previous course in April, each investigated 7 SSc patients.

Statistical calculations

All calculations were made with SPSS/PC+ software. Box plot diagrams were constructed; the scores of the master trainer were treated as an external gold standard.

Abbreviations: ICC, intraclass correlation coefficient; MRSS, modified Rodnan skin score; SSc, systemic sclerosis

Interobserver variability

The coefficient of variation was calculated as described previously.⁷

Intraobserver variability

In April 2005, in Pécs, seven examiners performed the MRSS three times on each of six patients as described above.⁷

Intraclass correlation coefficient

The agreement among the examiners was expressed by the intraclass correlation coefficient (ICC).¹¹ A value of 0.4–0.6 was considered as moderate, 0.6–0.8 as good and >0.8 as excellent agreement.¹⁰

RESULTS

In fig 1, box plots of the three different measurements of the tutors (from Berlin, Budapest, and Vienna, respectively) are presented. In addition, the evaluations performed by European tutors were compared with the single evaluation of a US rheumatologist (DEF) with extensive experience in skin

scoring, who was invited to act as the master teacher for the whole course programme.

As shown in table 1, repeated teaching reduced substantially the coefficient of variation for interobserver variability both for this international group of experts (from 50% to 35%) and for Hungarian rheumatologists in a local repeated course (54% vs 32%). With regard to the ICC, both values were good or excellent for experts (0.865 and 0.710, respectively) (for definition of good or excellent ICC, see Methods). The local repeated study indicated that both the interobserver variability and ICC can be substantially improved by repeated teaching. The within-patient SD decreased from 54% to 32%. The ICC increased from 0.496 to the good level of 0.722. The coefficient of variation for intraobserver variability was close to 20% (table 1). When time constraints were placed on the time to complete the examinations as in the second course in Budapest, results were less consistent than in Berlin and Vienna.

For the Budapest course, the box plots of MRSS indicate larger interquartile range values (fig 1B) as compared with the values of the two other courses (fig 1A,C), which suggests better consistency.

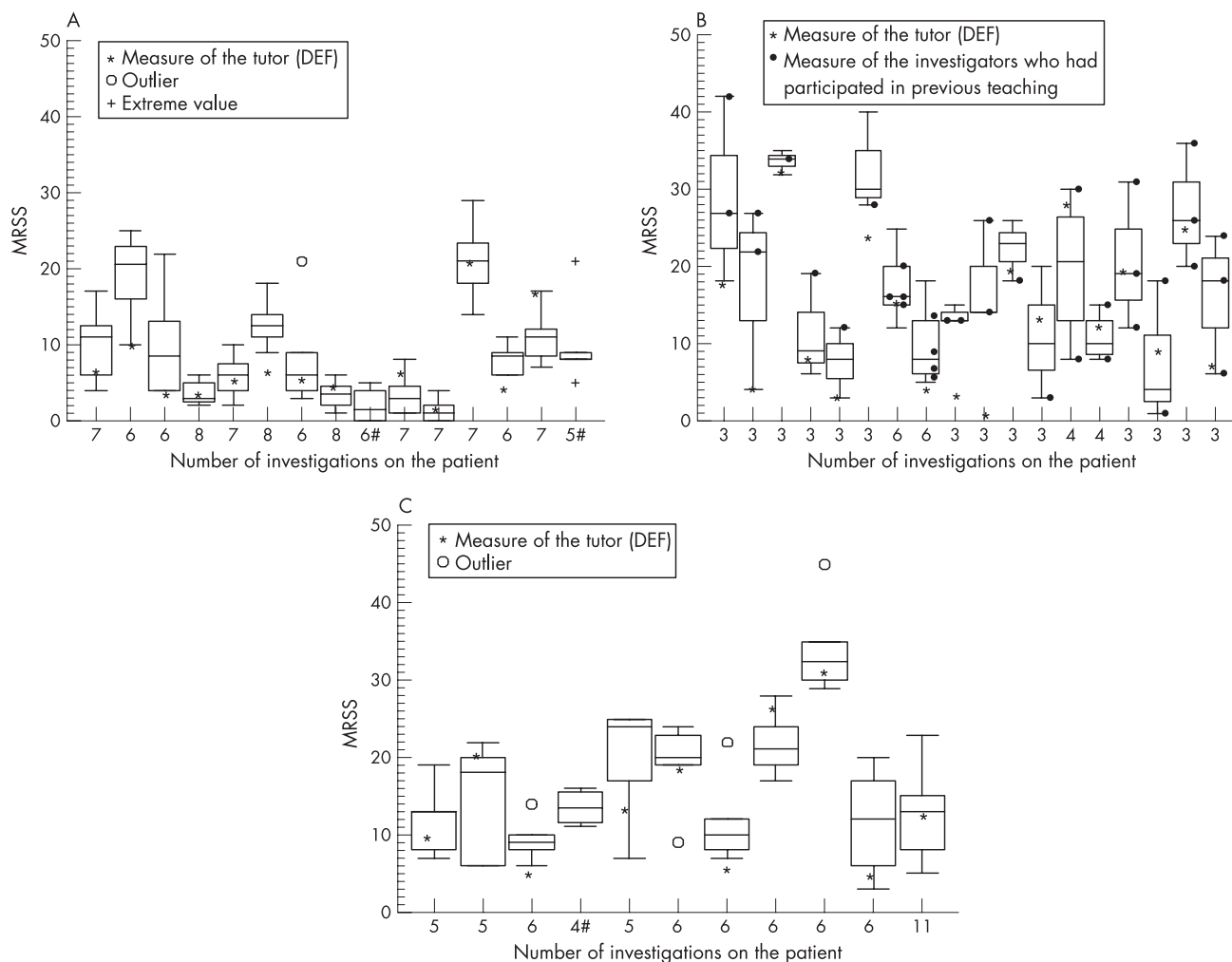


Figure 1 Box plots of modified Rodnan skin score (MRSS-51) observed for each patient. (A) Berlin course: the nine examiners were experts from five different European countries. The examiners performed 5–9 investigations on nine patients with SSs. (B) Budapest course: the nine examiners represented seven different countries. Five of the nine examiners had already participated in the previous course in Berlin. 3–6 examinations were performed on 18 patients with SSs. (C) Vienna course: the six examiners represented six different countries. Each of these examiners had also participated in at least one of the previous two courses, and the master teacher was the same (DEF). 4–6 examinations were performed on 11 patients with SSs. #The patient was not measured by the master teacher. For details, see Methods.

Table 1 Interobserver variability, and intraclass correlation coefficient as per modified Rodnan skin score (MRSS-51)

Study	No of patients	No of patients per investigator	No of investigators	Interobserver variability			Intraclass correlation coefficient	
				Mean	Within-patient SD	Coefficient of variation %		
Tutors								
Berlin, 2004	15	8–15	9	8.6	4.2	50	0.865	
Budapest, 2005	18	6–8	9	18.1	7.5	41	0.530	
Vienna, 2005	11	9–11	6	16.3	5.7	34.6	0.710	
Students								
Budapest, 2005	18	2	90	15.4	5.4	35	0.378	
Repeated investigations								
Pécs, 2005 Apr	6	6	9	15.7	8.5	54	0.496	
Pécs, 2005 June	7	7	5	12	3.8	32	0.722	
Previous studies								
Harrison <i>et al</i> ²	12	12	3	–	–	–	0.67	
Brennan <i>et al</i> ⁶	12	12	6	18.3	4.6	25	0.87	
Silman <i>et al</i> ⁹	8	8	16	–	–	–	0.72	
Clements <i>et al</i> ³	20	5–6	23	17.7	4.6	25	0.92	
Clements <i>et al</i> ⁷	20	5–6	23	17.7	4.6	25	–	
Intraobserver variability								
	No of patients	No of patients per investigator	No of investigators	No of investigations	Overall mean	Overall within-patient SD	Coefficient of variation %	Mean of intraclass correlation coefficients for investigators
Pécs, 2005 April	6	6	9 (3)*	3	13.9	2.9	20	0.74
Pécs, 2005 June	7	7	5 (3)*	3	12.3	2.5	20.4	0.76
Previous studies								
Clements <i>et al</i> ⁷	55	3	21	3	20.7	2.45	11.8	–

Comparison with previous investigations is also depicted.

*The number of investigators who previously participated in the Budapest course is in parentheses.

When the 90 young rheumatologists performed the skin score measurements, variability was relatively good (35%) due to high numbers, but within-patient SD remained high (5.4; table 1) and the ICC value was low. Similar results were obtained when local repeated courses were organised in Pécs for young rheumatologists (table 1, fig 2). The ICC was good or excellent in all but one case (table 1). The intraobserver within-patient variability was relatively good from the beginning and remained stable thereafter.

DISCUSSION

The aim of this EUSTAR (European Scleroderma Trials and Research group) project was to unify European medical practice in skin score measuring. Overall, the results on MRSS-51 measurements were comparable to two previous independent studies, where the mean (within-patient SD) was 18.3 (4.6) and 17.7 (4.6), respectively.^{8,9} Our three investigations for teachers showed similar results (table 1), thereby confirming both that MRSS is an appropriate tool for skin score measuring and that it is possible to teach a large number of rheumatologists the appropriate skin score measuring technique in a short time.

A previous study had demonstrated that the interobserver variability of the MRSS-51 could be reduced by approximately 50%.⁷ The results of our repeat courses likewise demonstrated that the coefficient of variation for interobserver variability can be substantially reduced by repeated teaching (table 1). Although the ICC increased from 0.50 to the good level of 0.72 in the limited local experience with less experienced rheumatologists in Pécs, the ICC of the experts was very good from the beginning, but did not increase (table 1). This underlines the importance of a second training process for less experienced rheumatologists in particular.

It is encouraging that the coefficient of variation for intraobserver variability was around 20%, similar to that found

for joint swelling/tenderness counts in rheumatoid arthritis (table 2).¹⁴ Although a low ICC was observed even during the second repeated teaching course, the overall appropriateness of the two repeated courses appeared to be satisfactory (table 1). These findings coincide with previous experience.⁶

When, for the first EUSTAR/EULAR (European League Against Rheumatism) course, almost 100 young rheumatologists were trained, the value of the within-patient SD was relatively good (5.4) with low ICC value, mainly due to the high number of investigators. Interobserver variability after this single teaching course was still considerable, even though experienced experts trained the students. However, the other courses suggest that, following repeated training, interobserver variability could be substantially decreased. After completion, an overall value of 5.5 of within-patient SD for interobserver variability as well as good ICC were achieved in mixed patient populations. This observation is also in agreement with previous experience.⁷

Our study has certain limitations. Even repeated sessions of the experts did not bring them any closer to the main tutor with time (table 1). This may be a limitation of the teaching method, or it could reflect a systematic bias of the principal teacher or a systematic bias of the other experts. The definition of a gold standard remains a critical point in such teaching exercises.

On the other hand, this exercise demonstrates a significant strength—the feasibility of training a large group of relatively inexperienced rheumatologists and the feasibility of a follow-up which completes the training in an efficient and straightforward manner.

In conclusion, in view of a significant amount of observer variability, teaching courses should be repeated at least twice for teaching inexperienced rheumatologists.

ACKNOWLEDGEMENTS

Besides the authors of the paper, the following colleagues participated in the teaching course for teachers: Phil Clements (US), Christopher

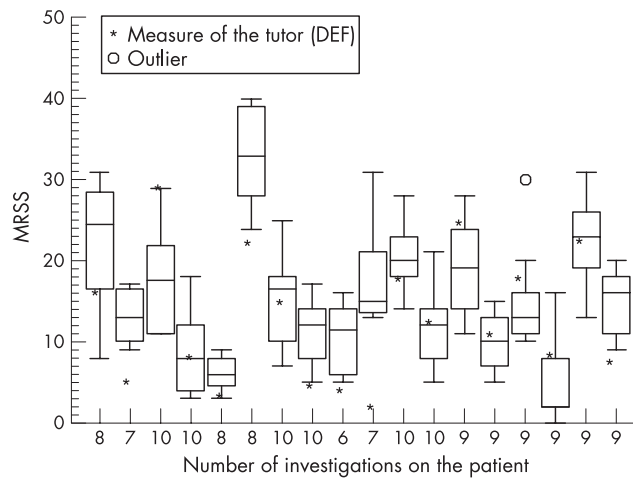


Figure 2 Box plot of modified Rodnan skin scores (MRSS-51) of the 90 students in the Budapest course. Ninety young rheumatologists participated in the training from all over Europe in Budapest (January 2005). The master teacher was the same person (DEF) who guided the teaching courses of the teachers. For further details, see Methods.

Denton (UK), Ivan Foeldvari, Ulf Mueller-Ladner (Germany), Alan Tyndall (Switzerland), Gabriele Valentini (Italy) and Frederick Wigley (US).

The participants of the local repeated courses were: Valéria Dudics, Gábor Kumánovics, Csaba G Kiss, Zoltán Nagy, Nóra Nusser, Gábor Sütő, Sándor Szántó, Zoltán Szekanez, Gabriella Szűcs, Sándor Szántó and Cecília Varjú (Hungary).

Participants of the European training course for young rheumatologists were: Stummvoll G (Austria); Culo MI, Marasovic Krstulovic D, Novak S, Radic M, Soldo JD (Croatia); Sorensen LJ, Sondergaard K, Strauss GI, Ullman S (Denmark); Tuvik P (Estonia), Assous N, Ilie D, Launay D (France); Clarenbach R, Foeldvari I, Gensch K, Grundt B, Hanitsch LG, Himsel A, Müller A, Olski TM, Seidel M, Sunderkoetter C, Süß A, Walzer K (Germany); Kampakis G, Mamoulaki M, Siakka P (Greece); Lee Ka Wing G (Hong Kong); Dudics V, Orbán I, Szűcs G, Varjú C (Hungary); Cavazzana I, Chialà A, Comina D, Frassi M, Guiducci S, Iannone F, Lo Monaco A, Maglione W, Miniati I, Pieropan S, Ruocco L, Tiso F (Italy); Jae-Bum J (Korea); Pileckyte M (Lithuania); Coleiro B (Malta); Grandaunet BH, Midtvedt Ø (Norway); Kotulska A, Kowal-Bielecka O, Krasowska DM (Poland); Figueira R, Cordeiro F, Oliveira R, Henriques MJ, Pinto SP, Resende C, Sequeira G (Portugal); Damjanovska Rajcevska L (Macedonia); Dragomir D, Isa ML, Micu M, Nicola M, Rednic S (Romania); Nevskaya T (Russia); Stamenkovic B, Ostojic P, Zlatanovic M (Serbia); Lukáčová O (Slovak Republic); Sipek DA (Slovenia); Banegil Espinosa I, de la Peña G, Lefebvre P, Joven B, Nuño Nuño L, Ortega de la O Mdc, Rodríguez Rubio S (Spain); Tikly M (South Africa); Chizzolini C, Oehri M, Ribl C, von Muehlenen I (Switzerland); Vonk M (The Netherlands); Dass S, Prabu A, Solanki K (UK).

Authors' affiliations

László Cziráj, Zoltán Nagy, Department of Immunology and Rheumatology, University of Pécs, Pécs, Hungary

Martin Aringer, Department of Rheumatology, Internal Medicine III, Medical University of Vienna, Vienna, Austria

Gabriela Riemekasten, Department of Rheumatology and Clinical Immunology, Charité University Hospital, Berlin, Germany

Marco Matucci-Cerinic, Department of Medicine, Division of Rheumatology, University of Florence, Florence, Italy

Daniel E Furst, Rheumatology Division, Geffen School of Medicine, University of California-Los Angeles, Los Angeles, California, USA

Funding: This work was supported by a EULAR educational grant provided for EUSTAR.

Competing interests: None declared.

EUSTAR is the EULAR Scleroderma Trials and Research group. Besides the authors, further participants in the skin score measurements are listed in the Acknowledgements.

Correspondence to: Professor L Cziráj, Department of Immunology and Rheumatology, University of Pécs, H-7621 Pécs, Irgalmasok u. 1, Hungary; laszlo.cziraj@aok.pte.hu

Accepted 24 December 2006

Published Online First 18 January 2007

REFERENCES

- Clements PJ, Hurwitz EL, Wong WK, Seibold JR, Mayes M, White B, *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.
- Clements PJ, Lachenbruch PA, Ng SW, Simmons M, Sterz M, Furst DE. Skin score. A semiquantitative measure of cutaneous involvement that improves prediction of prognosis in systemic sclerosis. *Arthritis Rheum* 1990;**33**:1256-63.
- Clements PJ, Wong WK, Hurwitz EL, Furst DE, Mayes M, White B, *et al.* The disability index of the health assessment questionnaire is a predictor and correlate outcome in the high dose versus low dose penicillamine is systemic sclerosis trial. *Arthritis Rheum* 2001;**44**:653-61.
- Sultan N, Pope JE, Clements PJ, for the Scleroderma trials Study Group. The health assessment questionnaire (HAQ) is strongly predictive of a good outcome in early diffuse scleroderma: results from an analysis of two randomized controlled trials in early diffuse scleroderma. *Rheumatology* 2004;**43**:472-8.
- Steen VD, Medsger TA Jr, Rodnan GP. D-penicillamine therapy in progressive systemic sclerosis (scleroderma): a retrospective analysis. *Ann Intern Med* 1982;**97**:652-9.
- Pope JE, Baron M, Bellamy N, Campbell J, Carrette S, Chalmers I, *et al.* Variability of skin scores and clinical measurements in scleroderma. *J Rheumatol* 1995;**22**:1271-6.
- Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, *et al.* Inter- and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. *J Rheumatol* 1995;**22**:1281-5.
- Clements PJ, Lachenbruch PA, Seibold JR, Zee B, Steen VD, Brennan P, *et al.* An assessment of inter-observer variability in 3 independent studies. *J Rheumatol* 1993;**20**:1892-6.
- Brennan P, Silman A, Black C, Bernstein R, Coppock J, Maddison J, *et al.* Reliability of skin involvement measures in scleroderma. *Br J Rheumatol* 1992;**31**:457-60.
- Silman AJ, Harrison M, Brennan P. Is it possible to reduce the observer variability in skin score assessment of scleroderma? *J Rheumatol* 1995;**22**:1277-80.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420-8.
- Harrison A, Lusk J, Corkill M. Reliability of skin score in scleroderma (comment). *Br J Rheumatol* 1993;**32**:170.
- Clements PJ, Medsger TA Jr, Feghali CA. Systemic sclerosis. In: Clements PJ, Furst DE, eds. Philadelphia: Lippincott Williams & Wilkins, 2004:129-50.
- Van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;**41**:1845-50.