RHEUMATOLOGY

Rheumatology 2017;56:v53–v66
doi:10.1093/rheumatology/kex202

Assessment of skin involvement in systemic sclerosis

Gábor Kumánovics[1,*], Márta Péntek[2,*], Sangmee Bae[3], Daniela Opris[4], Dinesh Khanna[5], Daniel E. Furst[3] and László Czirják[1]

Abstract

Skin involvement in SSc is an important marker of disease activity, severity and prognosis, making the assessment of skin a key issue in SSc clinical research. We reviewed the published data assessing skin involvement in clinical trials and summarized the major conclusions important in SSc clinical research. A systematic literature review identified randomized controlled trials using skin outcomes in SSc. Analysis examined the validity of the different skin measures based on literature findings. Twenty-two randomized controlled trials were found. The average study duration was 10.2 (S.D. 4.5) months, mean (S.D.) sample size 32.4 (32.6) and 26.7 (27.8) in intervention and control arms, respectively. The 17-site modified Rodnan skin score is a fully validated primary outcome measure in diffuse cutaneous SSc. Skin histology seems to be an appropriate method for evaluation of skin thickness. These findings have important implications for clinical trial design targeting skin involvement in SSc.

Key words: scleroderma, systemic sclerosis, skin involvement, randomized controlled trials, outcome measuring, OMERACT filter, validation

Rheumatology key messages

- The modified Rodnan skin score is an appropriate instrument as a primary outcome measure in dcSSc.
- Disease progression rate of SSc before study entry may have significant impact on the results.
- Statistical challenges in the evaluation of treatments for small SSc subgroups should be considered.

REVIEW

Introduction

SSc is a multi-organ disease characterized by thickening, hardening and tightening of the skin. Skin thickening is caused by increased collagen and intercellular matrix formation in the dermis and by temporary oedema, probably caused by microvascular injury. Finally, in the end stage, the skin becomes thin, atrophic and often tightly tethered to the underlying tissue [1].

More extensive skin thickening coincides with more severe internal organ manifestation(s), poor prognosis and increased disability. The modified Rodnan skin score (mRSS), which uses palpation to estimate skin thickness, is currently considered the most appropriate technique for measuring skin involvement in SSc, at least in dcSSc [1]. Our aim was to analyse the evidence-based data on the skin assessment instruments in SSc and, in particular, their use in clinical trials as either primary or secondary end points.

Literature search

As part of a large international collaborative work [2], we performed a systematic literature review. PubMed was searched for the period between 1995 and 26 January 2010, using the Cochrane Highly Sensitive Search Strategy for identifying randomized trials (Table 1). The search produced 3865 titles. Two reviewers (D.E.F. and D.K.) reviewed the output, examining both titles and abstracts. Altogether, 138 studies dealing with skin involvement in SSc were selected. Some earlier papers (from

[1]Department of Rheumatology and Immunology, Medical School, University of Pécs, Pécs, [2]Department of Health Economics, Corvinus University of Budapest, Budapest, Hungary, [3]Department of Rheumatology, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA, [4]Internal Medicine and Rheumatology, Carol Davila University of Medicine and Pharmacy, Bucharest, Romania and [5]Department of Medicine, University of Michigan Scleroderma Program, University of Michigan, Ann Arbor, MI, USA

Submitted 6 July 2015; revised version accepted 12 April 2017

*Gábor Kumánovics and Márta Péntek contributed equally to this study

Correspondence to: László Czirják, University of Pécs, Medical School, Department of Immunology and Rheumatology, Akác u. 1. Pécs, Hungary, H-7632.
E-mail: czirjak.laszlo@pte.hu

Table 1 Search strategy used in PubMed to identify randomized trials in SSc

'Clinical Trial' [Publication Type] OR 'Randomized Controlled Trial' [Publication Type] OR 'randomized' [tiab] OR 'placebo' [tiab] OR 'drug therapy' [sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] OR 'Clinical Trials as Topic' [Mesh] OR 'Research Design' [Mesh] OR 'Epidemiologic Research Design' [Mesh] OR 'Epidemiologic Studies' [Mesh] OR 'research design' [text word] OR 'case control' [text word] OR 'cohort' [text word] OR 'cross sectional' [text word]) AND ('1995' [Publication Date]: '3000' [Publication Date]) NOT ('animals' [MeSH] NOT 'humans'[MeSH])

Table 2 Study disposition using our search strategy

Search		Result: number of articles
#1	Scleroderma OR 'systemic sclerosis'	19 818
#2	Skin	533 846
#3	(Random OR randomized OR tria) OR double-blind OR single-blind	1 037 188
#4	((#1) AND #2) AND #3	370
#5	((#1) AND #2) AND #3 Limits: Publication Date from 1 January 2009 to 1 March 2011	61
#6	Skin	339
#7	((#1) AND #2) AND #6	204
#8	((#1) AND #2) AND #6 Limits: Publication Date from 1 January 2009 to 1 March 2011	74
#9	Used Durometre	86
#10	((#1) AND #2) AND #9	13
#11	((#1) AND #2) AND #9 Limits: Publication Date from 1 January 2009 to 1 March 2011	1
#12	Used US	354 854
#13	((#1) AND #2) AND #12	139
#14	((#1) AND #2) AND #12 Limits: Publication Date from 1 January 2009 to 1 March 2011	21
#15	Used Plicometry	10
#16	((#1) AND #2) AND #15 Limits: Publication Date from 1 January 2009 to 1 March 2011	0
#17	Used HAQ	2423
#18	((#1) AND #2) AND #17	44
#19	((#1) AND #2) AND #18 Limits: Publication Date from 1 January 2009 to 1 March 2011	13
#20	Used survival OR mortality	1 071 107
#21	((#1) AND #2) AND #20	281
#22	((#1) AND #2) AND #21 Limits: Publication Date from 1 January 2009 to 1 March 2011	41
#23	Included gender OR ethnicity OR ethnic	305 517
#24	((#1) AND #2) AND #23	62
#25	((#1) AND #2) AND #24 Limits: Publication Date from 1 January 2009 to 1 March 2011	21

before 1995) and later (until March 2011) published data were also included (Table 2) if these were judged to be relevant (by consensus) [3, 4].

Of the 138 selected studies and including the additional search results discussing methodological issues based on the randomized controlled trials (RCTs), relevant data were extracted and analyzed by two independent reviewers (MP and DO). The quality of the RCTs was evaluated by the Jadad score [5]. All references cited in this article were categorized according to the level of evidence (Table 3) by two independent reviewers (M.P. and D.O.).

Table 3 Categories of evidence for evaluating credibility of articles included in bibliography

1A. One or more meta-analysis of randomized controlled trials
1B. One or more randomized controlled trial
2A. One or more controlled trial without randomization
2B. One or more quasi-experimental study
3. Descriptive studies (e.g. correlational, cohort, case–control)

RCTs with skin outcome

The main characteristics of the 22 RCTs finally selected are presented in Tables 4 and 5. The study duration was a mean (S.D.) of 10.2 (4.5; range 3–24) months. The mean (S.D.) number of patients in the active arms (summed) and control arm was 32.4 (32.6; range 6–137) and 26.7 (27.8; range 6–94), respectively. More than half of the studies (n = 14, 63.6%) were double blinded. Among the 12 RCTs with Jadad score $\geqslant 4$ [4], 10 were placebo controlled [6–15], one compared two different drug doses [16] and one had a crossover design [17]. Eleven applied skin outcome as primary end point.

Regarding the inclusion criteria of the 22 RCTs, age was not an inclusion criterion in 10 trials (45.5%). The analysis of age at onset in the European Scleroderma Trials and Research (EUSTAR) database (n = 8, 554) did not reveal any significant difference in mRSS between the late-onset

Table 4 Main characteristics of randomized controlled trials assessing skin outcome in SSc

Study	Phase	Design	Blinding	Jadad score	Study duration	Inclusion criteria: age	Inclusion criteria: SSc subtype	Inclusion criteria: disease duration, years	Inclusion criteria: skin manifestation severity	Inclusion criteria: skin manifestation progression
Domiciano (2011) [82]	NS	Prospective	Open	2	12 months	NS	NS (non-specific interstitial pneumonia confirmed by biopsy)	NS	NS	NS
Ostojic (2011) [83]	Pilot study	Prospective	Single	2	6 months	NS	dcSSc	<15 months	NS	Skin thickness progression rate ⩾12/year
Rubén *et al.* [6]	NS	Prospective	Double	4	6 months	>16	dcSSc and lcSSc	NS	NS	NS
Daoussis *et al.* [20]	1	Prospective	Single	2	1 year	NS	dcSSc	NS	NS	NS
Su (2009) [84]	1	Prospective	Single	3	48 weeks	18–70	dcSSc	⩽5 years (since first non-RP sign)	NS	NS
Khanna *et al.* [7]	3	Prospective	Double	5	24 weeks	18–70	dcSSc	Early phase dcSSc: ⩽5 years (since first non-RP sign)	Severe (mRSS ⩾ 20) or moderate (mRSS ⩾ 16 and truncal skin involvement)	mRSS varies ⩽5 points between screening and first treatment
Postlethwaite *et al.* [8]	2	Prospective	Double	5	12 months of treatment, follow-up at 15 months	>18	dcSSc	Early phase dcSSc: <3 years or late phase dcSSC: 3–10 years	mRSS ⩾16	mRSS stable during the 6 months preceding enrolment
Denton *et al.* [21]	1–2	Prospective	Double	3	18 weeks of treatment, follow-up 6–9 months	NS	dcSSc	Early stage dcSSc: 18 months (since first non-RP sign)	mRSS between 18 and 20	mRSS <3 between screening and baseline
Tashkin *et al.* [9]	NS	Prospective	Double	5	12 months of treatment, follow-up further 12 months	NS	NS	⩽7 years since first non-RP sign	NS	NS
Knobler *et al.* [10]	NS	Prospective	Double	5	12 months	NS	dcSSc	Recent onset: <2 years (from the time of first evidence of skin thickening)	Skin score between 18 and 55 at enrolment	NS
Genovese *et al.* [11]	NS	Prospective	Double	4	24 weeks	⩾18	dcSSc	NS (disease duration was measured as the time from initial physician diagnosis)	NS	NS
Scorza (2001) [78]	NS	Prospective	Single	3	12 months	⩾18	SSc with RP	NS	NS	NS

(continued)

Table 4 Continued

Study	Phase	Design	Blinding	Jadad score	Study duration	Inclusion criteria: age	Inclusion criteria: SSc subtype	Inclusion criteria: disease duration, years	Inclusion criteria: skin manifestation severity	Inclusion criteria: skin manifestation progression
Pope *et al.* [12]	NS	Prospective	Double	5	12 months	⩾18	dcSSc	<3 years, skin involvement within 3 years	UCLA skin score ⩾4 (maximum possible: 30)	NS
Seibold *et al.* [13]	NS	Prospective	Double	5	24 weeks	18–70	dcSSc	<5 years (since first non-RP sign)	mRSS >20 or mRSS >16 and truncal involvement	mRSS ⩽5 points variation between screening and first treatment
Enomoto *et al.* [58]	NS	Cross-over	Single	2	1 year of treatment (then change)	NS	NS	<5 years (since first symptom attributable to scleroderma)	NS	Progressive disease (increase in skin score within 3 months)
Filaci *et al.* [48]	NS	Prospective	Open	2	12 months	NS	NS	<2 years	NS	NS
Clements *et al.* [16]	NS	Prospective	Double	5	2 years	18–75	dcSSc (skin thickening proximal to the elbow and/or knee, with or without face and neck involvement)	⩽18 months (since first non-RP sign)	NS	NS
Black *et al.* [14]	NS	Prospective	Double	5	12 months	18–70	Proximal (to the wrist, ankle and neck) scleroderma	<3 years (since first non-RP sign)	NS	NS
Grasseger *et al.* [22]	NS	Prospective	Open	2	12 months (6 months follow-up)	NS	SSc Type I or II	NS	NS	Minimum 6 months of follow-up by a physician to exclude rapidly progressive cases
Della (1997) [90]	NS	Prospective	Open	2	12 months	NS	NS	NS	NS	NS
Van den Hoogen *et al.* [15]	NS	Prospective	Double	4	24 weeks (extension: 24 week observational)	⩾16	NS	<3 years (since first signs of skin thickening) or longer duration and progression of disease	NS	Progression of skin thickening; persistent digital ulceration in the past 6 months
Wilson *et al.* [17]	NS	Cross-over	Double	4	3 months	18–65	NS	NS	NS	Stable symptoms for 2 months

NS: not specified; mRSS: modified Rodnan skin score; NS: non-specific; UCLA: University of California, Los Angeles.

Table 5 Randomized controlled trials assessing skin outcome in SSc: interventions, end points, results and relevant *post hoc* analyses

Study	Active arm: intervention, sample size	Control arm: intervention, sample size	Skin manifestation as primary end point (outcome measure used)	Skin manifestation as secondary end point (outcome measure used)	Result on skin-related end points	Relevant *post hoc* analyses
Domiciano (2011) [82]	CYC i.v. with monthly infusions of 1 g/m^2/dose, + prednisolone (60 mg then reduced), n = 9	CYC i.v. with monthly infusions of 1 g/m^2/dose, n = 9	–	Change in mRSS	Tend to improve	–
Ostojic (2011) [83]	Antioxidants (α-tocopherol 400 IU/day and ascorbic acid 1000 mg/day) and CYC (500 mg/m^2 of body surface monthly), n = 6	CYC (500 mg/m^2 of body surface monthly), n = 7	mRSS, skin thickness progression rate (change in mRSS in a year)	–	No effect	–
Rubén *et al.* [6]	Oral ciprofloxacin (250 mg), n = 15	Placebo, n = 15	mRSS 17 sites	–	Tend to improve	–
Daoussis *et al.* [20]	Rituximab, n = 8	Continue previous drugs (no additional treatment), n = 6	mRSS	HAQ (0.2 point decrease), skin histology from affected and adjacent skin sites	Tend to improve	–
Su (2009) [84]	Rapamycine, n = 8	MTX, n = 9	–	mRSS, proportion of patients achieving minimal clinically important difference (for mRSS: ⩾5.3), HAQ, tendon friction rubs	No effect	–
Khanna *et al.* [7]	Recombinant human relaxin, two arms (10 and 25 μg/kg/day): n = 42, n = 95	Placebo, n = 94	mRSS 17 areas (mRSS training[a])	Oral aperture, maximal hand extension, HAQ, SF-36	No effect	Amjadi *et al.* [28]; Kaldas *et al.* [33]; Khanna *et al.* [30]
Postlethwaite *et al.* [8]	Oral Type I collagen, n = 83	Placebo, n = 85	Change of mRSS (17 areas) between baseline and 12 months (mRSS training[a])	–	No effect	Amjadi *et al.* [28]; Kaldas *et al.* [33]; Khanna (2007) [85]
Denton *et al.* [21]	CAT-192, three arms (0.5, 5 and 10 mg/kg): n = 11, n = 11, n = 10	Placebo, n = 11	–	Change of mRSS (17 sites) at weeks 12 and 24; proportion of patients with no change in mRSS (mRSS training[a]); HAQ, serum biomarkers, skin biopsy from affected skin sites	No effect	Merkel *et al.* [44]
Tashkin *et al.* [9]	CYC, n = 79	Placebo, n = 79	–	Skin thickening score (range 0–51)	Tend to improve	Tashkin, 2007 [86]
Knobler *et al.* [10]	Photopheresis, n = 27	Sham photopheresis, n = 37	Decrease in skin thickening score (20 areas) with	Joints with contractures	Tend to improve	–

(continued)

Table 5 Continued

Study	Active arm: intervention, sample size	Control arm: intervention, sample size	Skin manifestation as primary end point (outcome measure used)	Skin manifestation as secondary end point (outcome measure used)	Result on skin-related end points	Relevant *post hoc* analyses
			range of 0–66 (investigator training[a])			
Genovese *et al.* [11]	PVAC injection, two arms (15 and 50 mg injections), n = 6, n = 6	Placebo, n = 6	Change in mRSS at week 24	Post-baseline mRSS response, hand expansion, oral aperture, HAQ-DI, serum E-selectin and thrombomodulin	Tend to improve	–
Scorza (2001) [78]	Iloprost, n = 29	Nifedipine, n = 17	mRSS 17 areas	RP severity score	Tend to improve	–
Pope *et al.* [12]	MTX, n = 35	Placebo, n = 36	mRSS 26 areas, UCLA skin score 10 sites (investigator training[a])	Oral opening, grip strength, flexion index, HAQ-DI	Tend to improve	Sultan 2004 [87]; Johnson *et al.* [24]
Seibold *et al.* [13]	Recombinant human relaxin, two arms (25 or 100 μg/kg/day), n = 23, n = 26 (efficacy analysis: n = 21, n = 24)	Placebo, n = 19	mRSS 17 areas (investigator training[a])	Maximal oral aperture, hand extension	Significant improvement in the 25 mg/kg/day group	–
Enomoto *et al.* [58]	Photopheresis, n = 10	Cross-over, n = 9	Average change in four-scale skin score (74 areas) after 1 year, oral aperture, hand mobility	Blood tests, biopsy	Tend to improve	–
Filaci *et al.* [48]	Iloprost, n = 10	Iloprost + CYC, n = 10	Plicometry	Capillarmicroscopy, serum IL-6 concentration	Tend to improve	–
Clements *et al.* [16]	D-Pen 125 mg/day, n = 68	D-Pen 750–1000 mg/day, n = 66	Change of mRSS (17 areas), rate of responders (responder: ⩾25% lower score compared with baseline)	Active hand spread, fit closure, maximal oral aperture, HAQ-DI	Tend to improve. High D-Pen dose tend to harm	Amjadi *et al.* [28]; Khanna *et al.* [32]; Clements *et al.* [34]; Khanna, 2010 [79]; Sultan, 2004 [87]; Clements, 2001 [88]; Clements, 2004 [89]
Black *et al.* [14]	IFN-α, n = 19	Placebo, n = 17	mRSS (17 sites)	Skin biopsy, PINP and ICTP	No effect. Tend to harm	–
Grasseger *et al.* [22]	IFN-γ, n = 27	Control group, n = 17	RSS (15 areas), mouth aperture, grip strength		Tend to improve	–
Della (1997) [90]	Iloprost, n = 19	Nifedipine, n = 12	–	Rodnan score, nailfold capillaroscopy	Tend to improve	–
Van den Hoogen *et al.* [15]	MTX, n = 17 (at week 24: dose increase or switching to active treatment)	Placebo, n = 12	Total skin score (26 sites), global health VAS. Responders: improvement of ⩾ 30%	Hand extension, grip strength, maximal oral opening	Tend to improve	–
Wilson *et al.* [17]	Recombinant human tPA, n = 14	Placebo (cross-over)	RSS	–	Tend to improve	–

[a]Investigators underwent pre-study standardization training. HAQ-DI: HAQ-Disability Index; ICTP: carboxyterminal cross-linked telopeptide of type I collagen; mRSS: modified Rodnan skin score; PINP: blood amino-terminal propeptide; PVAC: therapeutic agent derived from delipidated, deglycolipidated *Mycobacterium vaccae*; RSS: Rodnan skin score; SF-36: short form 36 questionnaire; UCLA: University of California Los Angeles; VAS: visual analog scale.

group (onset $\geqslant$75 years of age) and the rest of the cases [18], indicating that an upper age limit does not seem to be a crucial inclusion criterion [19]. Subsetting of SSc (dcSSc or lcSSc) was clearly specified as a criterion in 13 (59.1%) trials, and a limit on disease duration was required in 14 trials (63.6%). Disease duration as an inclusion criterion varied between 15 months and 7 years (or was not specified), and several RCTs put a special focus on early SSc. Severity of skin manifestation was not a specific criterion in most of the trials (n = 16, 72.7%).

No specific criteria were applied with regard to the progression of skin manifestation in 13 (59.1%) trials, whereas stable disease was required in 6 (27.3%). Three studies (13.6%) involved patients with progressive disease; the most recent applied the skin thickness progression rate (STPR) for its determination.

The primary end point was the outcome of skin involvement in 17 trials (72.3%). Seven of the trials (31.8%) included biological agents [8, 11, 13, 14, 20–22]. No definitive conclusion can be reached from these data regarding the efficacy of biological drugs [23]. Improvement of skin symptoms as a primary outcome was observed with MTX in one study, although the difference compared with placebo was not statistically significant in that study [12]. Johnson *et al.* [24] pointed out that the study was underpowered to detect smaller but clinically important effects. On re-evaluation of the data with a Bayesian approach, they found favourable odds of beneficial treatment effects [25, 26].

From this analysis, the mRSS is the most widely applied measure to evaluate drug efficacy on skin involvement, and also to categorize patients by skin disease severity and progression. Also, small sample size, chosen to make the trial feasible using one or two centres, is a major challenge in RCTs targeting SSc subgroups (e.g. early dcSSc studies with other restrictions as well), so other clinical study designs and analyses deserve consideration to avoid statistically underpowered studies [25, 26].

mRSS

The mRSS is a 17-site assessment instrument to quantify the thickness of the skin and the extent of involvement in SSc [1]. The mRSS reflects disease activity and severity, and it is appropriate for assessment of skin, especially dcSSc, and early cases in particular [1, 27].

Natural disease course and mRSS improvement in trials

The course of mRSS in patients with dcSSc was analysed based on pooled data of three large RCTs [7, 8, 16, 28]. Results of this *post hoc* analysis suggest that skin thickening of dcSSc patients recruited into clinical trials does not necessarily follow the same trend in natural history as previously reported. The improvement in mRSS independent of treatment group was also detected in a placebo-controlled CAT-192 RCT by Denton *et al.* [21]. The importance of considering the natural disease course in clinical trials was also highlighted in the scleroderma lung study [9]. Meta-analysis by Merkel *et al.* [29] confirmed that currently there are no variables that reliably identify groups of subjects whose mRSS will predictably increase or decrease during the course of a clinical trial. This meta-analysis suggested that early patients respond in the same manner as later patients (with active skin involvement) enrolled in trials that were pooled, so there may be innovative ways of enrolling active dcSSc patients, such as worsening skin involvement or elevated inflammatory markers. These findings have significant implications for clinical trial design in early dcSSc and challenge the feasibility of studies of the prevention of worsening.

Minimally important differences of mRSS and sensitivity to change

Minimally important difference (MID) is considered clinically meaningful in the case of mRSS, as change of >25 or 30%, based on surveys but not on statistical MIDs [16, 29–31]. In a D-Pen trial, statistical minimal differences were a change in mRSS of 3.2–5.3, or 15–25% of the baseline skin score of 21. This is smaller than the expert- and survey-defined MID >30% [16, 32]. MID estimates may depend on baseline scores. Moreover, it is questionable whether some areas have a stronger impact on the total score. According to Kaldas *et al.* [33], the chest, forearms and hands had moderate (0.50–0.74) effect size; however, the lower extremities, face, abdomen and fingers had only small effect size (0.16–0.49) in dcSSc [7, 8, 33].

Change in mRSS as a predictor of survival and overall morbidity

Progressive changes in skin thickness scores over a 2-year period were related to mortality and scleroderma renal crisis in a cohort of dcSSc patients who were being followed up as part of a D-Pen trial [16] in dcSSc [34]. Steen and Medsger [35] have also found that skin thickening has longer term (5–10 years) prognostic value in dcSSc [36]. Tyndall *et al.* [37] used the EUSTAR database (n = 5860). mRSS was an independent risk factor for mortality, with a hazard ratio of 1.20/10 score points. Hachulla *et al.* [38] analysed the French ItinérAIR-Sclérodermie SSc cohort (n = 546, 1347 patient-years) and confirmed the association of Rodnan skin score and increased mortality, with hazard ratio of 1.045 per one point.

With regard to the relationship between skin thickness and morbidity, the latent trajectory model, by Shand *et al*., identified three subgroups of patients demonstrating a similar trajectory of skin score changes, using unbiased mathematical analysis [31]. One group had low mRSS score (mean 20) at baseline and improved (33%) during the trial. The two other groups had high baseline mRSS score (mean 25 and 42, respectively); one included improvers (28%) the other non-improvers (worsening of 5%). The results confirm that the extent of skin disease is correlated with mortality, but its relationship with overall morbidity is more complex.

Table 6 Current state of validation of the skin assessment instruments by OMERACT filter

Skin assessment instruments	Face and content validity	Criterion validity	Construct validity	Discrimination	Responsiveness	Reliability	Feasibility
mRSS[a]	+	+	+	+	+	+	+
Durometry	+	+	+	±	+	+	±
Plicometry	+	±	–	+	±	±	±
Cutometry	+	+	–	+	–	+	±
US	+	+	+	+	–	+	±
Histology	+	+	+	+	+	±	–
Self-assessment questionnaire	+	+	+	±	–	+	+
Maximal oral aperture	+	–	+	+	±	–	±

[a]A teaching video course of the mRSS assessment (provided by D.E.F.) is available (http://video.edraspa.it/PublishingPoint/eustar.org/100kbps_part_1.wmv; http://video.edraspa.it/PublishingPoint/eustar.org/100kbps_part_2.wmv). mRSS: modified Rodnan skin score; +: fulfilling the criteria; –: there are no data or there are negative results; ±: there are not enough data or there is a remarkable difference between methods.

Disease progression measure and disease activity indexes based on mRSS

The STPR has been recently defined by Domsic *et al*. [39] as the mRSS at the first visit, divided by the duration of skin thickening (in years) by patient report. Rapid STPR was an independent predictor of early mortality, development of scleroderma renal crisis and severe cardiac disease [39].

The extent and change in skin involvement plays a large role in some disease activity indices [40, 41]. In the currently available, partially validated disease activity index (the European Scleroderma Study Group activity index, EScSG), skin involvement contributes 35% to the total index score. In a large consecutive SSc patient cohort, the mRSS was correlated with the EScSG activity index, both at baseline and at 1 year reinvestigation [41]. In another available partially validated activity index, the skin domain also plays an important role, with a representation of 29% [41].

Impact of gender, ethnicity and environmental factors on mRSS assessment

Nashid *et al.* [42] examined the baseline differences and course of mRSS, HAQ-Disability Index (HAQ-DI) and forced vital capacity percentage between men and women, and among three different ethnic groups. The course of skin thickness, functional disability and lung function was similar among genders and among ethnicities, even though there were several baseline differences between men and women and among the three ethnic groups. These findings confirm that there is no need to apply restrictions with regard to gender and ethnicity in RCTs studying skin manifestations of SSc [42].

In summary, the mRSS is a fully validated gold standard and widely used tool to assess skin involvement in dcSSc (Table 6). Its applicability is limited in late stage disease. Further research is required to improve our knowledge on its applicability in specific patient subgroups and clarify the current minor uncertainties. In future studies, both multicentric approaches, with larger sample size, and methods for cohort enrichment for dcSSc are required. Further research may be required for clarifying the potential differences in the natural disease course of skin involvement in patients with anti-toposoisomerase I *vs* patients with anti RNA polymerase III. Regarding mRSS assessment, minimizers and maximizers show a better performance with regard to the sensitivity to change.

Evaluation of late stage disease: tethering score

The University of California Los Angeles skin tethering score considers 10 skin regions scored from 0 to 3 for skin tethering, with a maximal score of 30. Inter- and intra-observer reliability of the University of California Los Angeles skin score has been quantified and found acceptable, and has demonstrated sensitivity to change [43]. It has been applied as a primary end point alongside the mRSS in an MTX trial [12], but it is rarely used.

Durometry, plicometry and elastometry

Different methods can be used to assess different skin properties; that is, mRSS for thickness, cutometry for skin elasticity and durometry for skin hardness [44, 45]. Unfortunately, these measurements are not performed at exactly the same body sites. Furthermore, these sites are too large for such instruments and they have been used in several different ways: sequential measurements at different regions of these sites (e.g. testing every 2 cm over the entire length of one arm); or only one measurement at a predefined specific region of these sites (e.g. periumbilical aspect of the abdomen). The consecutive assessment is very time consuming if it is performed for each body site. In the single measurement approach, one may miss involved skin, and the sensitivity to change is probably low because of sampling issues. These methods have not been adequately tested for use at the present time.

Durometry

Durometry is a validated method for the measurement of skin hardness in SSc [44] (Table 6). In a single centre and in a multicentre trial both intra- and inter-observer reproducibility were higher for durometry than for mRSS [44, 46]. Change in durometry scores was correlated with change in mRSS ($r = 0.70–0.77$). For each level of mRSS there was a wide and overlapping range of durometre readings, which makes it very difficult to assign a non-overlapping range of durometry scores to particular mRSS scores [46].

Plicometry

Plicometry is commonly used to measure the s.c. plica in obese individuals (Table 6). Nives Parodi *et al*. [47] performed measurements of plica thickness in only nine skin areas. High specificity (95–99%) and a high negative predictive value (95.5–100%) were found. Inter-observer variation was very low [47].

In a paper by Filaci *et al*. [48], plicometry could detect a significant improvement of skin involvement after CSA and iloprost treatment. Basso *et al*. [49] showed similar usefulness of plicometry. Although interesting, this methodology has not undergone formal validation, so its use in clinical trials is not recommended at present.

Elastometry

Skin elasticity can be measured with several different mechanical instruments. The cutometer lifts the skin into a measurement chamber by vacuum, followed by relaxation (Table 6). Skin elasticity was evaluated at 74 body areas, with high intra- and inter-observer agreement. The correlation coefficient between mRSS and elastometry was 0.67 [50]. Ishikawa *et al.* [51] investigated skin elasticity at only two regions. In this cross-sectional study, significantly less distension and retraction ability of the skin of patients with dcSSc was found compared with values of patients with lcSSc and normal controls [51].

Of the above, the durometer is the best validated for hardness and may be a useful instrument. All the other instruments need to be validated further to clarify whether they are also appropriate for this use.

Ultrasound

Most studies use a 10–25 MHz US probe. In clinical practice, this method has difficulties very similar to durometry; sequential measurements are time consuming, and in single or a few standardized sites, one may overlook positive cases, and the sensitivity to change is probably lower than that of the mRSS (Table 6).

Given that tissue thickness, echogenicity and vascularity vary with age and body site and show diurnal variation [3], accurate US evaluation of skin lesions is very difficult. Kaloudi *et al*. [52] used high-frequency US for measurement of skin thickness at two different sites on the second digit of the dominant limb. A highly significant correlation between the global mRSS and the local dermal thickness at the two examined sites ($P = 0.032$, $P = 0.021$) was detected [52].

Moore *et al.* [53] investigated dermal thickness at 17 sites of 39 patients with SSc. Intra-observer variability ranged from 0.55 to 0.96; the inter-observer variability ranged from 0.65 to 0.94. The inter-observer variation for the anterior chest was fair (0.84). Akesson *et al.* [54] investigated the usefulness of US in a longitudinal study. The measurements were performed at only five skin sites. Increasing echogenicity in patients with dcSSc was seen at 2 and 3 years in all areas except the forearm. Some years later, this workgroup compared skin assessment by mRSS and by high-frequency US in patients with early dcSSc [55]. The skin involvement of the chest could be detected earlier by US than by palpation. Kuhn *et al*. [56] investigated the effect of bosentan on skin fibrosis in patients with SSc. There were no significant differences noted in the results of US analysis, although patients with both dcSSc and lcSSc exhibited a statistically significant mean difference in the mRSS compared with normal. The authors explain this difference between results of US and mRSS by the difficultly of interpretation of the high-frequency US examination.

US could be a secondary or exploratory end point and will require highly experienced sites, although its pathological significance is not clear and it is also not clear what changes in the US denote. These and the minimally required, most relevant sites to be assessed by US need further research.

Skin histology for the assessment of skin involvement in SSc

Activity and damage

There is a good correlation between total mRSS and both wet and dry biopsy weights. The dry weight as a percentage of wet weight was constant both in lcSSc and dcSSc. There were no differences between the early and late disease subsets in lcSSc and dcSSc for either wet or dry biopsy weights [57], indicating that although mRSS is a very useful method for assessing disease activity, it is not totally independent from tissue damage and is likely to represent a severity score (combination of activity and damage).

Cell types and biomarkers

We found only 4 studies of 20 where analysis of skin biopsy specimens was performed for assessment of primary or secondary study end points.

Black *et al.* [14] examined the effect of IFN-α therapy on skin involvement in patients with early dcSSc. The median reduction of type I collagen secretion of IFN-α-treated patients was not significantly different from the placebo group. Enomoto *et al.* [58] investigated the effect of photopheresis on clinical, immunological and apoptosis (skin biopsy) parameters. The only significant observation was the induction of apoptosis in leucocytes by photopheresis. Denton *et al.* [21] examined consequences of recombinant human anti-TGF-1 antibody

(CAT-192) therapy in SSc. Real-time quantitative PCR analysis of fibrosis markers (COL1A1, COL3A1, TGFB1 and TGFB2) demonstrated that levels of mRNA for these markers were increased in affected skin obtained from patients with SSc, but there was no change during treatment.

Denton *et al*. [59] did not find clear benefit of infliximab treatment in dcSSc at 26 weeks, but it was associated with clinical stabilization and a rediction in two collagen markers.

Daoussis and Andonopoulos [60] investigated the effect of rituximab therapy. Administration of rituximab significantly reduced the number of B cells in three patients (who had clinical responses) but had no effect on the remaining three patients of the rituximab group (who did not respond) [60].

It should be commented that the techniques for obtaining and storing skin biopsies vary a great deal. Methods for biopsying skin have been published, should be included in clinical protocols and should be followed [61]. There is no standardized skin biopsy procedure [62], although skin histology seems to be an appropriate method for the evaluation of skin thickness in SSc (Table 6).

Questionnaires

Several new scleroderma-specific self-reported measures have been created in the last few years [63], but only one (the EScSG activity index) was developed originally for assessment of sclerodermatous skin lesions [41]. Most of the 138 studies reviewed used general (not organ-specific) patient self-assessment questionnaires. Out of the standard disability indexes and quality-of-life assessments, patients' visual analog scale (VAS) for global assessments of health was used in 17 studies.

At least partially skin-related and scleroderma-specific questionnaires (not counting questionnaires about RP and digital ulcers) were performed in only seven studies. These were the Scleroderma Health Assessment Questionnaire (SHAQ), the fully validated scleroderma UK functional score, which focuses on disability caused by skin tightness in the upper limb and proximal muscle weakness [59, 64], the self-rating VAS for perceived pain, stiffness and skin elasticity [65], a VAS about patient self-assessment of skin disease [44], the functional assessment [66], the patients' skin assessment score and physician's skin score [67], and functional discomfort as determined for each hand by SSc patients on a VAS [68].

In a recent manuscript about bosentan therapy for skin fibrosis, a VAS was used to evaluate several parameters, but the only significant result was seen for breathing problems [56]. The conclusion is that sufficiently validated self-rated VASs or skin self-assessment questionnaires are not currently available, and future work is required to clarify the role of these particular instruments in clinical trials and to assess the impact of skin involvement on SSc patients' quality of life.

Further assessments closely related to skin involvement

Maximal oral aperture, hand mobility, grip strength and tendon friction rubs were applied as secondary outcomes in several trials (Table 5). Wherever a detailed description was presented, variability could be found between the methods applied [7, 11–13, 15, 16, 22, 32, 69–80] (Table 6).

Discussion and conclusions

We present here the results of a systematic literature search for SSc studies of skin outcomes, for the period between 1999 and 2011. The design and end points of the available RCTs were analysed, and skin assessment tools were critically evaluated, based on literature findings. Most of the RCTs were placebo controlled in the past, although ethical issues may arise in today's world to prevent the use of placebo arms in RCTs. Other designs can be considered, such as comparing the test drug with the standard of care, adding the test treatment to the background standard of care or having a well-defined escape arm in the RCT.

Good-quality evidence has confirmed the validity of the mRSS as an assessment tool for skin involvement in dcSSc; its applicability is limited to early stages of the disease. Given that there is significant inter-observer variability in mRSS assessment, patients should be evaluated by the same investigator throughout a study, and a careful teaching course for mRSS assessment should be strongly considered before starting the study [81]. The skin change trajectory before the study start may enrich a study for higher sensitivity to change.

Other skin measures, such as elastometry, durometry and US, are not fully validated, although they may be used in an exploratory capacity or may be incorporated to validate these methods. Furthermore, especially for local treatment, skin biopsy is appropriate and, in certain studies, a necessary tool for skin assessment.

It is important to note that new classification criteria of SSc have been established in 2013 that will undoubtedly lead to changes in clinical trial designs [91, 92]. The new criteria enhance the detection and involvement of early and very early SSc patients in RCTs. This new tool, however, might increase the involvement of overlap cases that on the one hand, deserve attention when setting up the inclusion an exclusion criteria and, on the other hand, hamper comparability with former trials [93]. Valuable systematic reviews have been published in specific areas, such as evidence on therapeutic options [94], the role of ultrasonography [95], antibodies [96] and quality-of-life assessment tools [97]. These basically seem to confirm, albeit update and refine, our main observations. Analyses of the EUSTAR database provided further insights that should be considered when designing future studies [98, 99].

Owing to the multi-organ involvement and heterogeneity of active SSc, one should consider the use of composite outcomes, such as the Combined Response Index of SSc and patient-reported outcomes [100].

Supplement

This paper is part of the supplement titled Points to consider: systemic sclerosis and was funded by an unrestricted educational grant from EULAR.

Funding: No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: D.K. is supported by the National Institutes of Health/National Institute of Arthritis and Musculoskeletal and Skin Diseases K24 AR063120 and has received investigator-initiated grants and acts as a consultant to Actelion, Bristol-Myers Squibb, Bayer, Corbus, Cytori, ChemoMab, GlaxoSmithKline, Genentech/Roche and Sanofi-Aventis. D.E.F. has received grant/research support from AbbVie, Actelion, Amgen, Bristol-Myers Squibb, Corbus, National Institutes of Health, Novartis, Pfizer and Roche/Genetic and consulting fees from AbbVie, Actelion, Amgen, Bristol-Myers Squibb, Cytori, Novartis, Pfizer and Roche/Genentech. All other authors have declared no conflicts of interest.

References

1 Clements PJ, Medsger TA Jr, Feghali CA. Cutaneous involvement in systemic sclerosis. In: Systemic sclerosis (eds: Clements PJ, Furst DE), 2nd edition. Lippincott Williams and Wilkins, 2004. pp. 129–50.

2 Khanna D, Furst DE, Allanore Y *et al*. Twenty-two points to consider for clinical trials in systemic sclerosis, based on EULAR standards. Rheumatology 2015;54:144–51.

3 Gniadecka M, Serup J, Søndergaard J. Age-related diurnal changes of dermal oedema: evaluation by high-frequency ultrasound. Br J Dermatol 1994;131:849–55.

4 Fleischmajer R, Perlish JS, Shaw KV, Pirozzi DJ. Skin capillary changes in early systemic scleroderma. Electron microscopy and ''in vitro'' autoradiography with tritiated thymidine. Arch Dermatol 1976;112:1553–7.

5 Jadad AR, Moore RA, Carroll D *et al*. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1–12.

6 Rubén EC, Manuel VR, Agustín OR *et al*. Ciprofloxacin utility as antifibrotic in the skin of patients with scleroderma. J Dermatol 2010;37:323–9.

7 Khanna D, Clements PJ, Furst DE *et al*. Recombinant human relaxin in the treatment of systemic sclerosis with diffuse cutaneous involvement: a randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2009;60:1102–11.

8 Postlethwaite AE, Wong WK, Clements P *et al*. A multicenter, randomized, double-blind, placebo-controlled trial of oral type I collagen treatment in patients with diffuse cutaneous systemic sclerosis: I. Oral type I collagen does not improve skin in all patients, but may improve skin in late-phase disease. Arthritis Rheum 2008;58:1810–22.

9 Tashkin DP, Elashoff R, Clements PJ *et al*. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354:2655–66.

10 Knobler RM, French LE, Kim Y *et al*. A randomized, double-blind, placebo-controlled trial of photopheresis in systemic sclerosis. J Am Acad Dermatol 2006;54:793–9.

11 Genovese MC, Chakravarty EF, Boyle DL *et al*. A randomized, blinded, parallel group, placebo controlled pilot study evaluating the effect of PVAC treatment in patients with diffuse systemic sclerosis. J Rheumatol 2005;32:2345–50.

12 Pope JE, Bellamy N, Seibold JR *et al*. A randomized, controlled trial of methotrexate versus placebo in early diffuse scleroderma. Arthritis Rheum 2001;44:1351–8.

13 Seibold JR, Korn JH, Simms R *et al*. Recombinant human relaxin in the treatment of scleroderma. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2000;132:871–9.

14 Black CM, Silman AJ, Herrick AI *et al*. Interferon-α does not improve outcome at one year in patients with diffuse cutaneous scleroderma: results of a randomized, double-blind, placebo-controlled trial. Arthritis Rheum 1999;42:299–305.

15 van den Hoogen FH, Boerbooms AM, Swaak AJ *et al*. Comparison of methotrexate with placebo in the treatment of systemic sclerosis: a 24 week randomized double-blind trial, followed by a 24 week observational trial. Br J Rheumatol 1996;35:364–72.

16 Clements PJ, Seibold JR, Furst DE *et al*. High-dose versus low-dose d-penicillamine in early diffuse systemic sclerosis trial: lessons learned. Semin Arthritis Rheum 2004;33:249–63.

17 Wilson D, Edworthy SM, Hart DA, Fritzler MJ. The safety and efficacy of low-dose tissue plasminogen activator in the treatment of systemic sclerosis. J Dermatol 1995;22:637–42.

18 Hügle T, Schuetz P, Daikeler T *et al*. Late-onset systemic sclerosis—a systematic survey of the EULAR scleroderma trials and research group database. Rheumatology 2011;50:161–5.

19 Weng HH, Ranganath VK, Oh M *et al*. Differences in presentation of younger and older systemic sclerosis patients in clinical trials. Clin Exp Rheumatol 2010;28(5 Suppl 62):S10–4.

20 Daoussis D, Liossis SN, Tsamandas AC *et al*. Experience with rituximab in scleroderma: results from a 1-year, proof-of-principle study. Rheumatology 2010;49:271–80.

21 Denton CP, Merkel PA, Furst DE *et al*. Recombinant human anti-transforming growth factor β1 antibody therapy in systemic sclerosis: a multicenter, randomized, placebo-controlled phase I/II trial of CAT-192. Arthritis Rheum 2007;56:323–33.

22 Grassegger A, Schuler G, Hessenberger G *et al*. Interferon-gamma in the treatment of systemic sclerosis: a randomized controlled multicentre trial. Br J Dermatol 1998;139:639–48.

23 Phumethum V, Jamal S, Johnson SR. Biologic therapy for systemic sclerosis: a systematic review. J Rheumatol 2011;38:289–96.

24 Johnson SR, Feldman BM, Pope JE, Tomlinson GA. Shifting our thinking about uncommon disease trials: the case of methotrexate in scleroderma. J Rheumatol 2009;36:323–9.

25 European Medicines Agency, Comittie for Medicinal Products for Human use (CHMP): Guideline in clinical trials in small populations. 2006. Doc. Ref. CHMP/EWP/83561/2005.

26 Korn EL, McShane LM, Freidlin B. Statistical challenges in the evaluation of treatments for small patient populations. Sci Transl Med 2013;5:178sr173.

27 Czirják L, Foeldvari I, Müller-Ladner U. Skin involvement in systemic sclerosis. Rheumatology 2008;47 (Suppl 5):v44–5.

28 Amjadi S, Maranian P, Furst DE *et al*. Course of the modified Rodnan skin thickness score in systemic sclerosis clinical trials: analysis of three large multicenter, double-blind, randomized controlled trials. Arthritis Rheum 2009;60:2490–8.

29 Merkel PA, Silliman NP, Clements PJ *et al*. Patterns and predictors of change in outcome measures in clinical trials in scleroderma: an individual patient meta-analysis of 629 subjects with diffuse cutaneous systemic sclerosis. Arthritis Rheum 2012;64:3420–9.

30 Khanna D, Furst DE, Clements PJ *et al*. Responsiveness of the SF-36 and the Health Assessment Questionnaire Disability Index in a systemic sclerosis clinical trial. J Rheumatol 2005;32:832–40.

31 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.

32 Khanna D, Furst DE, Hays RD *et al*. Minimally important difference in diffuse systemic sclerosis: results from the D-penicillamine study. Ann Rheum Dis 2006;65:1325–9.

33 Kaldas M, Khanna PP, Furst DE *et al*. Sensitivity to change of the modified Rodnan skin score in diffuse systemic sclerosis—assessment of individual body sites in two large randomized controlled trials. Rheumatology 2009;48:1143–6.

34 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.

35 Steen VD, Medsger TA Jr. Improvement in skin thickening in systemic sclerosis associated with improved survival. Arthritis Rheum 2001;44:2828–35.

36 Krieg T, Takehara K. Skin disease: a cardinal feature of systemic sclerosis. Rheumatology 2009;48 (Suppl 3):iii14–8.

37 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.

38 Hachulla E, Carpentier P, Gressin V *et al*. Risk factors for death and the 3-year survival of patients with systemic sclerosis: the French ItinérAIR-Sclérodermie study. Rheumatology 2009;48:304–8.

39 Domsic RT, Rodriguez-Reyna T, Lucas M, Fertig N, Medsger TA Jr. Skin thickness progression rate: a predictor of mortality and early internal organ involvement in diffuse scleroderma. Ann Rheum Dis 2011;70:104–9.

40 Valentini G, Della Rossa A, Bombardieri S *et al*. European multicentre study to define disease activity criteria for systemic sclerosis. II. Identification of disease activity variables and development of preliminary activity indexes. Ann Rheum Dis 2001;60:592–8.

41 Minier T, Nagy Z, Balint Z *et al*. Construct validity evaluation of the European Scleroderma Study Group activity index, and investigation of possible new disease activity markers in systemic sclerosis. Rheumatology 2010;49:1133–45.

42 Nashid M, Khanna PP, Furst DE *et al*. Gender and ethnicity differences in patients with diffuse systemic sclerosis—analysis from three large randomized clinical trials. Rheumatology 2011;50:335–42.

43 Merkel PA, Clements PJ, Reveille JD *et al*. Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. J Rheumatol 2003;30:1630–47.

44 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.

45 Aghassi D, Monoson T, Braverman I. Reproducible measurements to quantify cutaneous involvement in scleroderma. Arch Dermatol 1995;131:1160–6.

46 Kissin EY, Schiller AM, Gelbard RB *et al*. Durometry for the assessment of skin disease in systemic sclerosis. Arthritis Rheum 2006;55:603–9.

47 Nives Parodi M, 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.

48 Filaci G, Cutolo M, Scudeletti M *et al*. Cyclosporin A and iloprost treatment of systemic sclerosis: clinical results and interleukin-6 serum changes after 12 months of therapy. Rheumatology 1999;38:992–6.

49 Basso M, Filaci G, Cutolo M *et al*. Long-term treatment of patients affected by systemic sclerosis with cyclosporin A. Ann Ital Med Int 2001;16:233–9.

50 Enomoto DN, Mekkes JR, Bossuyt PM, Hoekzema R, Bos JD. Quantification of cutaneous sclerosis with a skin elasticity meter in patients with generalized scleroderma. J Am Acad Dermatol 1996;35:381–7.

51 Ishikawa T, Tamura T. Measurement of skin elastic properties with a new suction device (II): systemic sclerosis. J Dermatol 1996;23:165–8.

52 Kaloudi O, Bandinelli F, Filippucci E *et al*. High frequency ultrasound measurement of digital dermal thickness in systemic sclerosis. Ann Rheum Dis 2010;69:1140–3.

53 Moore TL, Lunt M, McManus B, Anderson ME, Herrick AL. Seventeen-point dermal ultrasound scoring system—a reliable measure of skin thickness in patients with systemic sclerosis. Rheumatology 2003;42:1559–63.

54 Akesson A, Hesselstrand R, Scheja A, Wildt M. Longitudinal development of skin involvement and reliability of high frequency ultrasound in systemic sclerosis. Ann Rheum Dis 2004;63:791–6.

55 Hesselstrand R, Scheja A, Wildt M, Akesson A. High-frequency ultrasound of skin involvement in systemic sclerosis reflects oedema, extension and severity in early disease. Rheumatology 2008;47:84–7.

56 Kuhn A, Haust M, Ruland V *et al*. Effect of bosentan on skin fibrosis in patients with systemic sclerosis: a prospective, open-label, non-comparative trial. Rheumatology 2010;49:1336–45.

57 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.

58 Enomoto DN, Mekkes JR, Bossuyt PM *et al*. Treatment of patients with systemic sclerosis with extracorporeal photochemotherapy (photopheresis). J Am Acad Dermatol 1999;41:915–22.

59 Denton CP, Engelhart M, Tvede N *et al*. An open-label pilot study of infliximab therapy in diffuse cutaneous systemic sclerosis. Ann Rheum Dis 2009;68:1433–9.

60 Daoussis D, Andonopoulos AP. Rituximab in the treatment of systemic sclerosis-associated interstitial lung disease: comment on the article by Yoo. Rheumatol Int 2011;31:841–2.

61 Verrechia F, Laboureau J, Verola O *et al*. Skin involvement in scleroderma – where histological and clinical scores meet. Rheumatology 2007;46:833–41.

62 Beyer C, Distler JHW, Allanore Y, *et al*. The EUSTAR biobanking group. EUSTAR biobanking: Recommendations for the collection, storage, and distribution of biospecimens in scleroderma research. Ann Rheum Dis 2011;70:1178–82.

63 Ruof J, Bruhlmann P, Michel BA, *et al*. Development and validation of a self-administered systemic sclerosis questionnaire (SySQ). Rheumatology (Oxford) 1999;38:535–42.

64 Serednicka K, Smyth AE, Black CM *et al*. Using a self-reported functional score to assess disease progression in systemic sclerosis. Rheumatology 2007;46;1107–10.

65 Sandqvist G, Akesson A, Eklund M. Evaluation of paraffin bath treatment in patients with systemic sclerosis. Disabil Rehabil 2004;26:981–7.

66 Stratton RJ, Wilson H, Black CM. Pilot study of anti-thymocyte globulin plus mycophenolate mofetil in recent-onset diffuse scleroderma. Rheumatology (Oxford) 2001;40:84–8.

67 Muellegger RR, Hofer A, Salmhofer W, *et al*. Extended extracorporeal photochemotherapy with extracorporeal administration of 8-methoxypsoralen in systemic sclerosis. An Austrian single-center study. Photodermatol Photoimmunol Photomed 2000;16:216–23.

68 Durand F, Staumont D, Bonnevalle A, *et al*. Ultraviolet A1 phototherapy for treatment of acrosclerosis in systemic sclerosis: controlled study with half-side comparison analysis. Photodermatol Photoimmunol Photomed 2007;23:215–21.

69 Sandqvist G, Eklund M. Hand Mobility in Scleroderma (HAMIS) test: the reliability of a novel hand function test. Arthritis Care Res 2000;13:369–74.

70 Del Rosso A, Maddali-Bongi S, Sigismondi F, *et al*. The Italian version of the Hand Mobility in Scleroderma (HAMIS) test: evidence for its validity and reliability. Clin Exp Rheumatol 2010;28(5 Suppl 62):S42–7.

71 Sandqvist G, Hesselstrand R, Eberhardt K. A longitudinal follow-up of hand involvement and activities of daily living in early systemic sclerosis. Scand J Rheumatol 2009;38:304–10.

72 Roberts-Thomson AJ, Massy-Westropp N, Smith MD, *et al*. The use of the hand anatomic index to assess deformity and impaired function in systemic sclerosis. Rheumatol Int 2006;26:439–44.

73 Roberts-Thomson AJ, Englert H, Ahern MJ, *et al*. A modified hand anatomic index to assesss hand deformity in scleroderma. Rheumatol Int 2009;29:847–8.

74 Medsger TA Jr, Silman AJ, Steen VD. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26:2159–67.

75 Medsger TA Jr, Bombardieri S, Czirjak L, *et al*. Assessment of disease severity and prognosis. Clin Exp Rheumatol 2003;21:S42–S46.

76 Torok KS, Baker NA, Lucas M *et al*. Reliability and validity of the delta finger-to-palm (FTP), a new measure of finger range of motion in systemic sclerosis. Clin Exp Rheumatol 2010;28(2 Suppl 58):S28–36.

77 Ingegnoli F, Boracchi P, Ambrogi F, *et al*. Hand impairment in systemic sclerosis: association of different hand indices with organ involvement. Scand J Rheumatol 2010;39:393–7.

78 Scorza R, Caronni M, Mascagni B, *et al*. Effects of long-term cyclic iloprost therapy in systemic sclerosis with Raynaud's phenomenon. A randomized, controlled study. Clin Exp Rheumatol 2001;19:503–8.

79 Khanna PP, Furst DE, Clements PJ, *et al*. D-Penicillamine Investigators. Tendon friction rubs in early diffuse systemic sclerosis: prevalence, characteristics and longitudinal changes in a randomized controlled trial. Rheumatology (Oxford) 2010;49:955–9.

80 Khanna D, Lovell DJ, Giannini E, *et al*. Scleroderma Clinical Trials Consortium co-authors. Development of a provisional core set of response measures for clinical trials of systemic sclerosis. Ann Rheum Dis 2008;67:703–9.

81 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.

82 Domiciano DS, Bonfá E, Borges CT *et al*. A long-term prospective randomized controlled study of non-specific interstitial pneumonia (NSIP) treatment in scleroderma. Clin Rheumatol 2011;30:223–9.

83 Ostojic P, Damjanov N. Effects of micronutrient antioxidants (alpha-tocopherol and ascorbic acid) on skin thickening and lung function in patients with early diffuse systemic sclerosis. Rheumatol Int 2011;31:1051–4.

84 Su TI, Khanna D, Furst DE *et al*. Rapamycin versus methotrexate in early diffuse systemic sclerosis: results from a randomized, single-blind pilot study. Arthritis Rheum 2009;60:3821–30.

85 Khanna D, Furst DE, Wong WK *et al*. Reliability, validity, and minimally important differences of the SF-6D in systemic sclerosis. Qual Life Res 2007;16:1083–92.

86 Tashkin DP, Elashoff R, Clements PJ *et al*. Effects of 1-year treatment with cyclophosphamide on outcomes at 2 years in scleroderma lung disease. Am J Respir Crit Care Med 2007;176:1026–34.

87 Sultan N, Pope JE, Clements PJ. The health assessment questionnaire (HAQ) is strongly predictive of good outcome in early diffuse scleroderma: results from an analysis of two randomized controlled trials in early diffuse scleroderma. Rheumatology (Oxford) 2004;43:472–8.

88 Clements PJ, Wong WK, Hurwitz EL *et al*. The Disability Index of the Health Assessment Questionnaire is a predictor and correlate of outcome in the high-dose versus low-dose penicillamine in systemic sclerosis trial. Arthritis Rheum 2001;44:653–61.

89 Clements PJ, Seibold JR, Furst DE *et al*. High-dose versus low-dose D-penicillamine in early diffuse systemic sclerosis trial: lessons learned. Semin Arthritis Rheum 2004;33:249–63.

90 Della BS, Molteni M, Mascagni B *et al*. Cytokine production in scleroderma patients: effects of therapy with either iloprost or nifedipine. Clin Exp Rheumatol 1997;15:135–41.

91 van den Hoogen F, Khanna D, Fransen J *et al*. 2013 classification criteria for systemic sclerosis: an American college of rheumatology/European league against rheumatism collaborative initiative. Ann Rheum Dis 2013;72:1747–55.

92 Pope JE, Johnson SR. New Classification Criteria for Systemic Sclerosis (Scleroderma). Rheum Dis Clin North Am 2015;41:383–98.

93 Baron M, Pope JE, Van Den Hoogen F *et al*. Specificity of Systemic Sclerosis Classification Criteria. J Rheumatol 2015;42:2512.

94 Volkmann ER, Furst DE. Management of Systemic Sclerosis-Related Skin Disease: A Review of Existing and Experimental Therapeutic Approaches. Rheum Dis Clin North Am 2015;41:399–417.

95 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.

96 Domsic RT. Scleroderma: the role of serum autoantibodies in defining specific clinical phenotypes and organ system involvement. Curr Opin Rheumatol 2014;26:646–52.

97 Almeida C, Almeida I, Vasconcelos C. A review on quality of life in systemic sclerosis. Autoimmun Rev 2015 published on 23 July 2015. doi: 10.1016/j.autrev.2015.07.012.

98 Maurer B, Graf N, Michel BA *et al*. Prediction of worsening of skin fibrosis in patients with diffuse cutaneous systemic sclerosis using the EUSTAR database. Ann Rheum Dis 2015;74:1124–31.

99 Wirz EG, Jaeger VK, Allanore Y *et al*. Incidence and predictors of cutaneous manifestations during the early course of systemic sclerosis: a 10-year longitudinal study from the EUSTAR database. Ann Rheum Dis 2015 published on 31 July 2015. doi: 10.1136/annrheumdis-2015-207271.

100 Khanna D, Berrocal VJ, Giannini EH *et al*. The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. Arthritis Rheum 2016;68:299–311.