Localized Scleroderma Severity Index and Global Assessments: A Pilot Study of Outcome Instruments

THASCHAWEE ARKACHAISRI and SALLY PINO

ABSTRACT. ***Objective.*** To develop a disease outcome instrument measuring localized scleroderma (LS) severity and to determine its reliability.

Methods. Patients with LS were evaluated by 2 rheumatologists. The LS Severity Index (LoSSI) comprises the sum of 4 clinical skin scores measured at 14 cutaneous anatomic sites: extent of surface area (SA) affected, erythema score (ES), skin thickness (ST) score, and new lesion/extension (N/E). Physician and patient global assessments (GA) were recorded on a 100 mm visual analog scale.

Results. Twenty-two patients with LS had 66 visits, during which 91 lesions were assessed. Individual skin scores (SA, ES, ST) and LoSSI showed substantial interrater reliability (κ_w 0.77–0.83; percentage agreement 93.41%–96.70%). Intrarater variability was calculated using 26 anatomic sites and 9 pair-visits. Each skin score demonstrated excellent reliability (κ_w 0.56–0.80, percentage agreement 90.39%–94.23%). Physician GA showed substantial interrater correlation [0.72, 95% confidence interval (CI) 0.57, 0.87]. There was only fair correlation between physician and patient GA (0.27, 95% CI 0.00, 0.64). The standardized response means for LoSSI and physician GA were large (1.86 and 2.55) for those who improved after therapy.

Conclusion. LS clinical trials are impeded by the lack of reliable and reproducible outcome measures. We have developed the LoSSI to correct this deficiency. Our pilot study demonstrates that the LoSSI is reliable and reproducible in measuring LS severity and therapeutic effects and can be easily implemented into the clinical examination of patients with LS. Both LoSSI and physician GA were sensitive to clinical changes in patients with LS. A formal study should be conducted to validate these preliminary findings. (First Release Mar 1 2008; J Rheumatol 2008;35:650–7)

Key Indexing Terms:

LOCALIZED SCLERODERMA	OUTCOME MEASURE	DISEASE ACTIVITY
GLOBAL ASSESSMENT	SKIN SCORE	DISEASE SEVERITY

Localized scleroderma (LS) is an autoimmune disease primarily affecting cutaneous and subcutaneous tissue. It has long been known to be an entity separate from systemic sclerosis (SSc) as it rarely involves internal organs. The incidence of LS is estimated to be 2.7 cases per 100,000 persons at risk per year[1]. Children are afflicted 10 times more frequently with LS than with SSc[2]. Although LS is not a fatal disease, some children with LS experience cosmetic disfigurement, localized growth retardation, joint contractures, and related psychological disturbances[1].

Although its pathogenesis is unknown, it is apparent that LS includes an initial inflammatory phase followed by a late sclerotic and/or atrophic phases, which can involve deeper structures including muscle and bone[3]. Since no effective antifibrotic therapies are available, it is only during the early inflammatory phase that effective intervention may halt disease progression and prevent poor outcomes.

Accurate assessment of the effects of LS and the development of new therapies for this condition are limited by the lack of reliable and standardized outcome measures. Thermography and ultrasound in LS have been reported[4-6]. Although holding some promise, these methods require considerable time, expense, and operator experience. Skin biopsy accurately reflects tissue abnormalities but is subject to sampling bias and is inconvenient for longitudinal patient evaluation. Since LS mainly starts in the skin and changes of the skin and subcutis structures determine LS disease severity, if a reliable method to measure skin disease severity using semiquantitative methods could be found, it would facilitate clinical trials and inform treatment decisions and possibly help to improve outcomes.

We have developed a semiquantitative scoring method to minimize the above limitations. Our primary purpose was to determine interrater and intrarater reliability of this proposed system. A second goal was to assess the correlation of this measure with patient and physician global assessment (GA) in LS.

From the Division of Rheumatology, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA.

Supported by divisional funds.

T. Arkachaisri, MD, Assistant Professor of Pediatrics; S. Pino, MD, Clinical Fellow, University of Pittsburgh School of Medicine.

Address reprint requests to Dr. T. Arkachaisri, Division of Rheumatology, Children's Hospital of Pittsburgh, 3705 Fifth Ave., Room MT3887, Pittsburgh, PA 15213. E-mail: arkatx@chp.edu

Accepted for publication November 6, 2007.

MATERIALS AND METHODS

Patients with LS were recruited from the Scleroderma Clinic at Children's Hospital of Pittsburgh. Diagnosis and classification of LS was made based on that of Peterson, *et al*[1]. The University of Pittsburgh Institutional Review Board approved our study.

The LS Severity Index (LoSSI) is a scoring system based on the assessment of extent (surface area) and intensity (erythema, skin thickness) of LS lesions as well as new lesion development/existing lesion extension. The index includes the sum of the estimation of 4 separate scores, as follows.

(1) Surface area score (SA) — using 14 cutaneous surface anatomic sites [head, neck (chin to base of neck at the C7/C8 spinous process), chest (clavicle to anterior costal margin), abdomen (anterior costal margin to inguinal crease), upper back (C7/C8 spinous process to posterior costal margin), lower back (costal margin to gluteal crease), right and left arms (acromion process to cubital crease), right and left forearms/hands/fingers (cubital crease to finger tips), right and left buttocks/thighs (inguinal crease to popliteal crease), and right and left legs/feet/toes (popliteal crease to tip of toes)] and dividing into 3 segments, scores the extent of the LS (0: none; 1: ≤ 1/3; 2: > 1/3–2/3; and 3: > 2/3–3/3 of the anatomic site surface area).

(2) Erythema score (ES) — incorporating the severity of inflammation using the color of the lesion's edge (0: normal or postinflammatory hyper/hypopigmentation; 1: slight erythema/pink; 2: red/clearly erythema; and 3: dark red or marked erythema/violaceous).

(3) Skin thickness (ST) — adopting the modified Rodnan skin thickness system as performed by palpation or simple "pinching" of the skin (0: normal skin thickness and freely mobile; 1: mild increase of thickness, mobile; 2: moderate increase of thickness; impaired skin mobility or harder to pinch; 3: marked increase of thickness or no mobility of skin by palpation; degree of ST will be compared to patient's unaffected corresponding skin area on the contralateral side, or nearby ipsilateral side if symmetrical lesions affected. This will eliminate intersubject variability).

(4) New lesion/lesion extension (N/E) — any new lesion development and/or enlargement of an existing lesion within the past month will have a score of 3.

Each score is estimated on the most representative lesion of a given anatomic site. SA, ES, ST, and N/E scores are added together to obtain the LoSSI. The range of the LoSSI lies between 0 and 168.

Global assessment. At the time of examination, patients completed the overall and specific lesion GA and physician estimated the GA, using a 100 mm visual analog scale (VAS). Physician GA assessed overall LS severity. For patient GA, patients were asked to draw a vertical mark on the 0–100 mm line and respond to the question, "How do you feel about your LS disease overall for the past one month?", and patients responded to the question, "How do you feel about this specific lesion for the past one month?" for patient-specific lesion GA (we specified a lesion for each patient). This GA aimed to limit the patient attention to only 1 particular LS lesion in order to assess if the correlation between physician and patient GA could be improved. A 1-month time period was used to ascertain recall memory. The anchor at the 0 mm point was "absent" (example: no erythema, no skin thickening, no underlying tissue atrophy), and at the 100 mm point was "extremely severe" (example: severe skin thickening, total loss of underlying subcutaneous fat).

Other definitions. Patients were considered to have "active disease" if they had any one of the following characteristics: erythematous border lesion, enlargement of existing lesion, or new lesion development within the past month. Patients who had no aforementioned characteristics for at least 6 months were considered to have "inactive/stable disease."

Study design. LoSSI and GA were independently obtained by 2 examiners, a rheumatologist and rheumatologist in training, for interrater reliability assessment. Intrarater reliability assessment was obtained by the same examiners reevaluating the same patients 4–12 weeks after the initial examination. All patients included in this portion of the study had stable disease for at least 3 months.

Data analysis. Statistical analysis was performed using Stata v. 8.2 (Stata Corp., College Station, TX, USA). Descriptive statistics were used for demographic data. Interrater and intrarater agreement for the LoSSI were determined by calculating raw agreement and weighted kappa coefficients (κ_w), thus eliminating agreement by chance. κ_w penalizes a disagreement according to its seriousness[7]. Correlations between physician and patient/lesion GA were tested by intraclass correlation coefficient (ICC)[8].

Validity. Validity is defined as the extent to which any instrument actually measures what it is intended to measure. In LS, there is no "gold standard" against which to test the validity of LoSSI; thus it is impossible to assess criterion validity. For this reason, convergent construct validity was investigated. Construct validity is concerned with the extent to which a particular measure (in this case, the LoSSI) relates to other measures consistent with theoretically derived hypotheses concerning the concepts (or construct) that are being measured[9]. Spearman's rho correlation coefficient was used to assess LoSSI and physician GA correlation to demonstrate this validity. Interpretation of agreement followed the recommendation of Landis and Koch: 0.00–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial, and > 0.80 = almost perfect agreement[10].

Sensitivity to change/responsiveness. Clinical skin change as assessed by LoSSI and physician GA were recorded in 5 active LS patients and 5 stable/inactive LS patients from baseline, Week 0, to followup, Week 10–12. Standardized response mean (SRM) is the sensitivity statistic used to detect clinical change. SRM is computed as the mean change score (the change in score from baseline to followup) divided by the standard deviation of the change. This reflects the concept of a signal:noise ratio more effectively than the standardized effect size[11]. This method is similar to paired t-test; however, since it avoids the use of standard error of the mean as the denominator, it is less influenced by the sample size[11-14]. SRM is considered large (> 0.8), moderate (0.5–0.8), or small (0.2–0.5). 95% confidence interval (CI) for the SRM was calculated by the method described by Beaton, *et al*[11]. Wilcoxon signed-ranks test was used to demonstrate the difference between baseline and followup of LoSSI and physician GA.

RESULTS

Patients. Twenty-two patients with LS with disease onset before age 16 years were included in this pilot study. There were 17 girls and 5 boys and the mean patient age at time of study was 12.6 ± 4.8 years. Ten had linear scleroderma (LiScl), 5 morphea (M), 3 subcutaneous morphea (SqM), and 4 mixed LiScl/M. The mean LS disease duration was 26.2 ± 23.8 months. The mean number of anatomic sites involved was 3.1 ± 2.1. Seven patients had no treatment (3 with inactive disease and 4 newly diagnosed) and 15 patients were receiving different therapy regimens [8 methotrexate (MTX), 3 systemic corticosteroids (CS) + MTX, and 4 topical CS + topical vitamin D].

Clinical skin scores. The distribution of SA, ES, ST, and LoSSI estimated by 2 examiners is shown in Figure 1. Ninety-one anatomic sites were assessed during 66 patient visits. Over 2/3 of our LS cohort had < 1/3 of surface area involved per anatomic site. The majority of lesions had no erythema; 21% of lesions had various degrees of skin erythema. For ST, 55% had no dermal thickening, and 25%, 19%, and < 1% of lesion areas assessed had mild, moderate, or severe skin thickening, respectively. About 70% of LS patients had LoSSI < 7, with scores of 6 or 2 in the majority of patients. No patient developed new lesions or enlargement of the existing lesions during the study period. Overall,

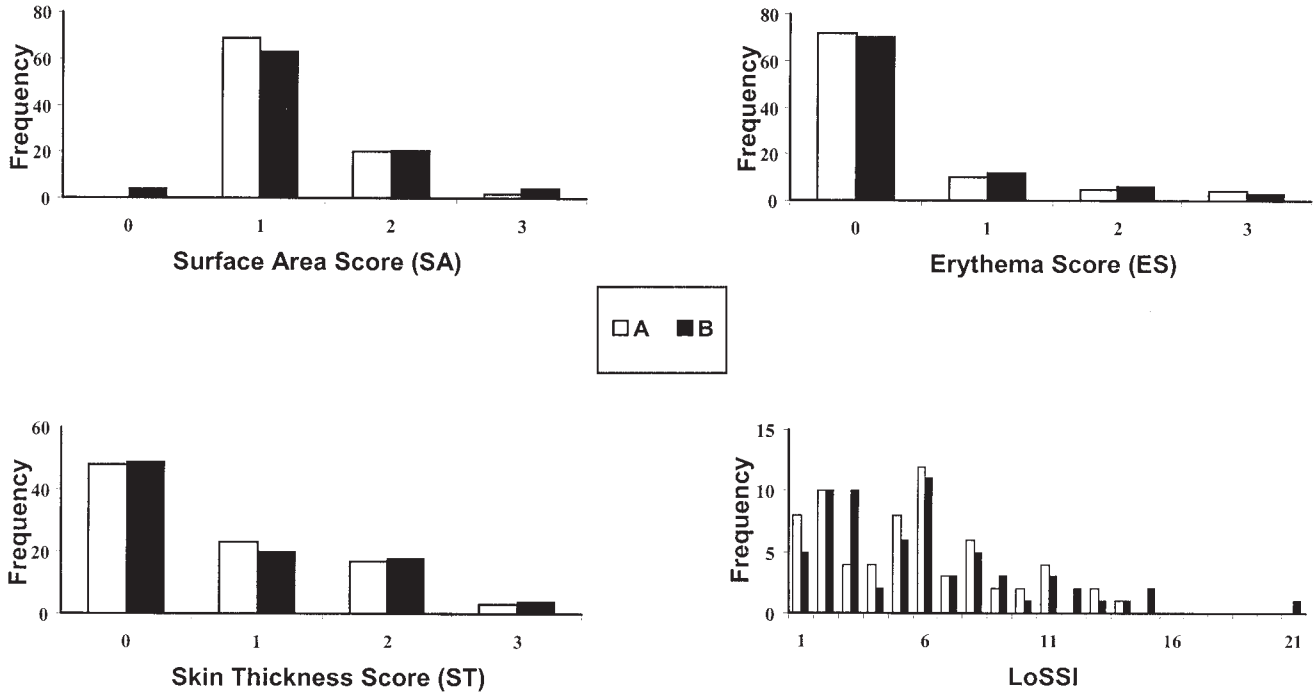

Figure 1. Clinical skin score distribution. Each skin score (surface area, erythema score, and skin thickness) was graded from 0 to 3, by 2 examiners (A and B) from 91 anatomic sites. LoSSI data were obtained from 66 patient-visits by 2 examiners (A and B).

as shown in Figure 1, the scores of the trainee were similar to those of the experienced rheumatologist.

Reliability assessment. Inter- and intrarater agreements for each clinical score are illustrated in Table 1. Ninety-one anatomic sites and 66 visits were studied for interrater reliability. LoSSI demonstrated an almost perfect interrater agreement (κ_w 0.83). When each domain was calculated separately, we found that all domains, SA, ES, and ST, showed substantial interrater agreement. For intrarater reliability assessment, 26 anatomic sites and 9 visit-pairs were studied. Both LoSSI and ST demonstrated substantial, and SA moderate intrarater agreement. We did not assess intrarater reliability of ES because all patients with erythema were treated after this evaluation with CS and none returned early enough for a repeat examination.

Physician GA had substantial interrater correlation between the 2 raters (0.72, 95% CI 0.57, 0.87). There was only a fair correlation between physician and patient GA (0.27, 95% CI 0.00, 0.64). If patients were asked about specific lesions, the correlation increased somewhat (0.38, 95% CI 0.04, 0.73).

Construct validity. To assess convergent construct validity of the LoSSI, we evaluated whether the LoSSI correlated with physician GA. Both scales from 29 patient visits were compared. The Spearman's correlation coefficient was 0.44, which is a moderate correlation (p = 0.016; Figure 2).

Sensitivity to change/responsiveness. Figure 3 shows temporal changes of the LoSSI and physician GA for 5 patients with active LS, 3 with LiScl (Patients A, B, C), 1 with mixed LiScl/M (Patient D), and 1 with SqM (Patient E). The mean

Table 1. Interrater and intrarater agreement of LoSSI.

Scores	Interrater Agreement*		Intrarater Agreement**	
	κ_w	% Agreement	κ_w	% Agreement
SA	0.77	96.34	0.36, 0.75	88.46, 92.31
ES	0.83	96.70	NA	NA
ST	0.79	93.41	0.80, 0.80	94.23, 94.23
LoSSI	0.83	95.56	0.83, 0.66	93.65, 88.89

* Calculated from 91 anatomic sites and 66 patient-visits. ** Calculated from 26 anatomic sites and 9 patient-visit-pairs, 4–12 weeks apart. SA: surface area score, ES: erythema score, ST: skin thickness score (14 anatomic sites), score 0–3; κ_w: weighted kappa coefficient (0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, 0.81–1.00 almost perfect agreement). NA: not available.

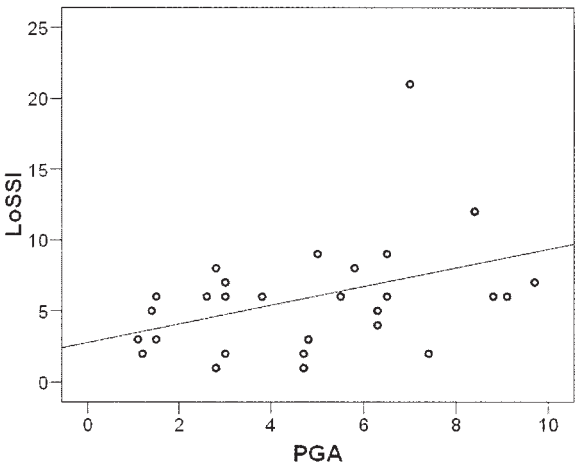

Figure 2. Correlation of the LoSSI and Physician Global Assessment (PGA) in 29 patient-visits. Spearman's coefficient 0.44, p = 0.016.

number of anatomic sites was 4.4 ± 2.1. The followup period after first evaluation was 10 to 36 weeks. All patients received medium to high dose of systemic CS therapy and subcutaneous MTX. Patient A had LiScl affecting the left upper and lower extremities, upper back, and lower abdomen. We first evaluated her LS disease during CS tapering, 4 months after treatment (MTX + CS) was begun by another physician (Week 0 point in Figure 3A was her first evaluation for LoSSI, which was at the 4-month point after her therapy was initiated). The first recorded LoSSI was 13. Four weeks later the hand ST score increased, giving a LoSSI of 15. At that time topical high-potency CS was added. Two months later, the ST improved and LoSSI declined to 9 and 8 at subsequent 8 and 16-week followup intervals, respectively. LoSSI and physician GA in 4 other patients also declined after therapy. There was slight LoSSI fluctuation in Patient B.

Table 2 gives responsiveness/sensitivity to change statistics of LoSSI and physician GA in patients with active and inactive LS. For the active LS patients [4 newly diagnosed (Patients B, C, D, E) and one with a flare (Patient A)], mean change and SRM of LoSSI were 5.80 ± 3.12 (p = 0.039) and 1.86 (95% CI 0.98, 2.73). These were compared to 5 patients with inactive or stable disease (3 LiScl, 1 M, and 1 mixed LiScl and M). The mean number of anatomic sites was 3.0 ± 2.3. The duration of disease was 49.3 ± 31.2 months. Mean change and SRM of the inactive LS group were 0.00 ± 0.71 and 0 (95% CI –0.80, 0.80). The mean change of physician GA for the active versus the inactive group was 47.00 ± 18.43 (p = 0.043) versus 1.20 ± 3.19 (p = 0.42), respectively, and for the SRM was 2.55 (95% CI 1.67, 3.43) versus 0.37 (95% CI –0.42, 1.18). The SRM of LoSSI and physician GA in the active LS group represents strong/large responsiveness, which indicates the ability of LoSSI and physician GA to detect change in those patients who had experienced change in their status. Further, the SRM of ES and ST were also large, 1.13 and 0.94, respectively, suggesting strong sensitivity to improvement after therapy. For these representative patients, SA did not change within the 12-week followup period (SRM = 0).

DISCUSSION

The findings from our pilot study demonstrate that a simple clinical skin scoring system, the LoSSI, is feasible and reliable in assessing LS disease severity. This system is suitable for both clinical trials and routine patient care. The results will require confirmation using a larger number of examiners.

LS skin lesions are often classified into 2 major stages. The active stage is characterized by the development of new lesions, increased induration and/or extension of the existing lesions, or a violaceous/erythematous border surrounding existing lesions. In the inactive stage, the number of lesions is unchanged and existing lesions show no change in size or become smaller, induration is unchanged or less, and erythema is absent. The development of a method to measure these LS stages would be important in quantifying the natural history of the disease and its response to therapeutic interventions. At present, there are no simple, feasible, or reliable measures of LS skin changes. Thermography is a noninvasive tool that detects differences in skin temperature. It was demonstrated to have high sensitivity (92%) and moderate specificity (68%) as compared to clinical description of LS skin lesions[6]. This measurement was proven to be reproducible between 2 investigators ($\kappa = 0.82$). Its disadvantages include the requirement for a temperature-controlled examination room and the need to acclimatize to the preset room temperature (usually taking 15–20 min). Also, atrophic lesions may cause false-positive results and thick subcutaneous fat tissues may give false-negative results. Many reports have shown that skin thickness and possibly skin echogenicity of LS lesions can be accurately measured by both 12–15 MHz and high frequency (> 15 MHz) ultrasound[5,15-18]. Both of these techniques are time-consuming and require extensive training and are thus not feasible for routine use.

We developed the LoSSI to assess the cardinal signs of skin changes during the active phase of LS — extent of disease (SA) and intensity of the lesion: inflammation (ES), skin thickness (ST), and new lesion development or extension of existing lesion (N/E). Surface area and skin thickness are partly damage measures; thus this tool is named the Localized Scleroderma Severity Index, LoSSI.

Surface area involvement in skin diseases has been the subject of many proposals for quantification. The rule of 9, "flat closed hand" technique, and schematic figure outlines were proven to be unreliable measures especially when used by untrained assessors who tended to overestimate areas affected[19-21]. Computerized image analysis scoring of the size of LS lesions was reported recently to be reliable, but its

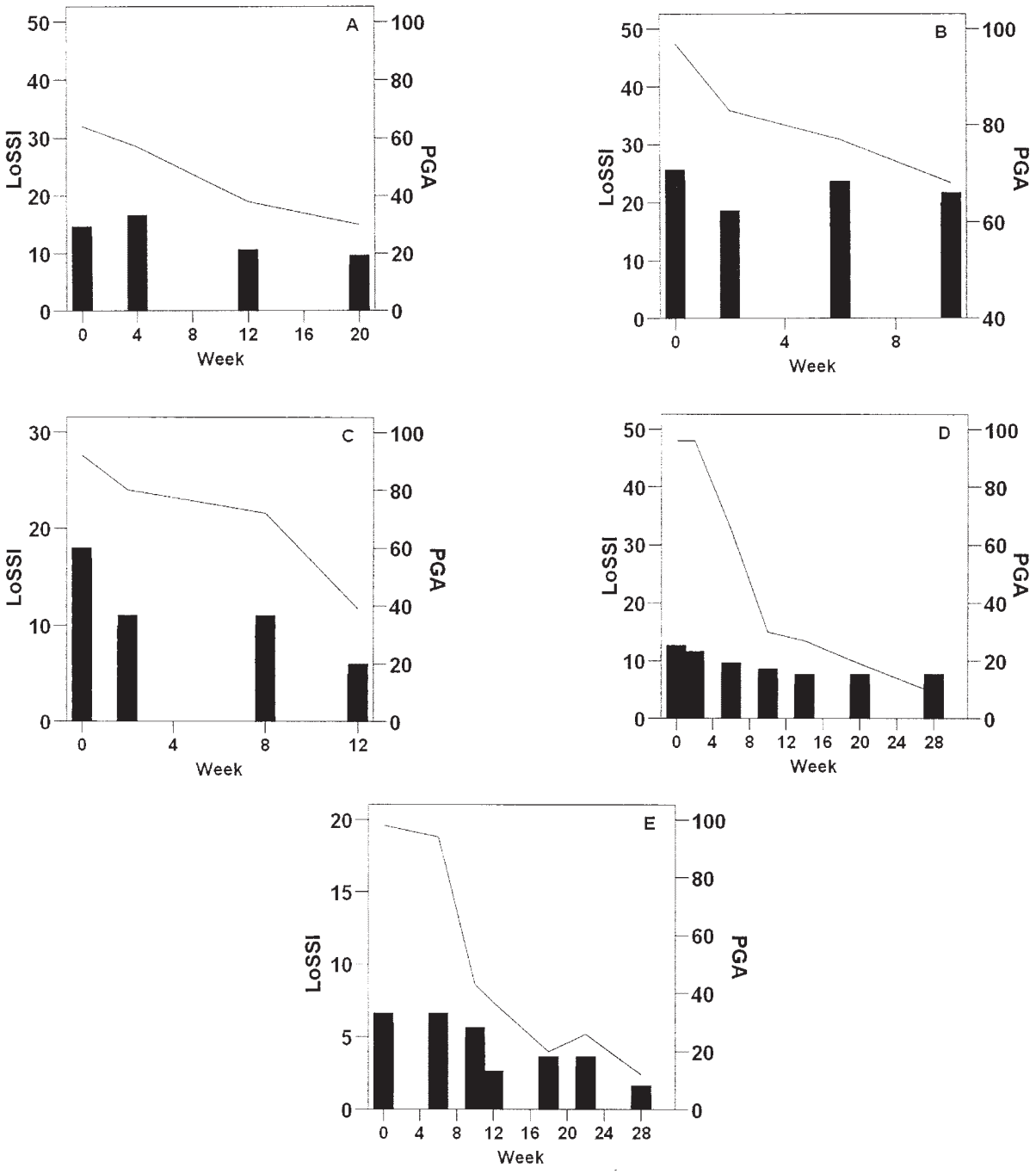

Figure 3. Longitudinal data of the LoSSI and Physician Global Assessment (PGA) in 5 representative LS patients at first visit and subsequent followup period from 10 to 36 weeks. Scales for the LoSSI (black bars) and PGA (lines) are shown at the left and right, respectively. LoSSI is sensitive to disease severity changes over time. Data shown for patients with linear scleroderma (A, B, C), mixed linear scleroderma and morphea (D), and subcutaneous morphea (E).

use is limited to small, circumscribed LS lesions[22]. It is inaccurate over curved body sites and intertriginous areas. This method is expensive, time-consuming, and impractical for clinical trials or busy clinical practices. Dividing the body into anatomic sites was shown to have less interrater variability[23]. Dividing the cutaneous surface area into 3 rather than more segments was done for 2 purposes: first, to reduce interrater variability as shown by others, and second, to be consistent with the other scores, which are recorded from 0 to 3[19-21]. The limitation of this approach would be its limited capacity of detecting a small extension of a few millimeters. This smaller extension will be recorded by the N/E domain (see below). The SA score was a highly reliable estimate of area involved in our LS patient cohort.

Table 2. Sensitivity to change of LoSSI and physician global assessment of severity (PGA).

Variable	Active, n = 5	Stable, n = 5
Duration of disease, mo (mean ± SD)	6.1 ± 3.5	49.3 ± 31.2
Number of anatomic sites (mean ± SD)	4.4 ± 2.1†	3.0 ± 2.3†
Medication	5 MTX + CS	1 N, 3 MTX, 1 Vit D
LoSSI		
Mean change (SD)	5.8 (3.12)	0 (0.71)
P for difference from baseline to followup††	0.039	1.00
Standardized response mean* (95% CI)	1.86 (0.98, 2.73)	0 (–0.80, 0.80)
PGA		
Mean change (SD)	47 (18.43)	1.20 (3.19)
P for difference from baseline to followup	0.043	0.42
Standardized response mean (95% CI)	2.55 (1.67, 3.43)	0.37 (–0.42, 1.18)

† p = 0.39; †† Wilcoxon signed-ranks test was used to assess score difference between baseline (Week 0) and followup (Week 10–12). * Calculated as mean change in LoSSI/PGA score from baseline to followup divided by standard deviation of the change. MTX: methotrexate; CS: corticosteroids; N: none; Vit D: topical vitamin D.

Erythema is one of the hallmarks of skin inflammation. The degree of erythema assessed visually was reported to correspond well to several objective measures including laser Doppler flowmeter, spectroradiometer, erythema meter, and chromameter[24,25]. The erythema scale has been incorporated in many skin severity measures and has been found to be reliable[26-31]. Wolkerstorfer, *et al* reported moderate interrater agreement of the ES in 20 children with atopic dermatitis ($\kappa = 0.52$)[30]. The ES from our pilot data showed almost perfect agreement between 2 raters ($\kappa_w = 0.83$). Since erythema often improves rapidly after systemic CS therapy, repeated assessment of ES will need to be carried out before such therapy is begun or within 24–48 h after its initiation to determine intrarater reliability.

In 1979, Rodnan, *et al* first reported the use of a semiquantitative skin palpation technique to assess skin thickness in patients with SSc[32]. Skin thickness using this method was shown to correlate well with histologic estimation of the amount of skin collagen deposition and wet weight of skin biopsies from the same anatomic site[32]. The original Rodnan method assessed cutaneous thickening in 26 body surface areas and graded the degree of skin thickening from 0 (none) to 4 (severe). A modified Rodnan Skin Score (mRSS) was developed and validated using a scale of 0–3 in 17 surface areas, eliminating those sites found to have the greatest interrater variability[28]. The mRSS was demonstrated to be both accurate and reliable. The coefficient of variance (CV) of 12% for intrarater variability and CV 25% and ICC 0.74–0.87 for interrater variability were better than 2 methods used to assess joints in patients with rheumatoid arthritis, the Ritchie index, and joint counts (CV 37% and 43%, respectively)[33-36]. The mRSS has been used as a primary outcome measure in SSc clinical trials[37]. There are no reports of this method applied to LS lesions. Assessing ST in children, Foeldvari and Wierk found that healthy children had mean mRSS of 13.92 and this range would be expected for mild SSc in adults[38]. This skin score also correlated with Tanner stage and age of the children[38]. In order to apply this system in LS, we proposed to assess ST compared to patient's unaffected corresponding skin area on the contralateral side, or nearby ipsilateral side if symmetrical lesions were present. The patient will serve as their own control regardless of the Tanner stage and maturity, thus eliminating intersubject variability. Although assessing ST requires training, repeated use was shown to improve reliability in a recent international study[36]. Our pilot data showed substantial inter- and intrarater agreement of ST score in the assessment of individual LS lesions ($\kappa_w = 0.79$ and 0.80, respectively) comparable to that previously reported in SSc[36].

New lesion development can occur during the course of LS and often will prompt physicians to initiate or change therapy. We are aware that existing lesions may enlarge without extending into an adjacent anatomic site. Therefore, we included "new lesion" development or "extension of existing lesion" (N/E) as a variable in the LoSSI and have given it the highest score of 3. The sensitivity of this variable can be judged only in a longitudinal study. We also demonstrated that changes in LoSSI fluctuate with disease severity and response to therapy. In our pilot study, despite small sample size, using SRM to assess sensitivity to change, both the LoSSI and physician GA demonstrated high sensitivity to change/improvement after therapy. Indeed, we do not expect the change in SA in such a short followup period (10–12 weeks); thus change in the LoSSI in our representative patients was parallel to that of ES and/or ST, as demonstrated by large SRM of ES and ST. However, the LoSSI's sensitivity to change over time, as either the result of the natural history of the disease or response to therapy, will need to be examined in larger prospective longitudinal studies.

Overall, the LoSSI is sensitive, as demonstrated by large SRM, and possesses excellent inter- and intrarater reliability. Today, GA is an integral part of almost all rheumatolog-

ic disease outcome measures. We report here for the first time results using a 100 mm VAS to assess physician, patient, and patient-lesion GA (where patients were instructed to report on a specific lesion). Although our GA incorporated both activity (extent, erythema, induration, new lesion development and extension of existing lesion) and damage (extent, skin thickening, and atrophy of soft tissue) features, physician GA revealed excellent interrater correlation, but only fair reliability was shown between physician and patient GA. This result was not unexpected, as physicians and patients often have different perceptions regarding disease severity. As expected, this discrepancy was reduced somewhat by narrowing patient focus to a specific cutaneous lesion. The moderate LoSSI correlation with physician GA suggests that changes in physician GA may be influenced by response to therapy and may need to be recorded together.

LS clinical trials are hampered by the lack of effective, reliable outcome measurements. For this reason, we have proposed a simple, semiquantitative method, the LoSSI. This study is considered to be the first step toward the development of LS core set outcome measures. Our proof-of-concept study demonstrates that the LoSSI is a simple and reliable instrument for measuring LS disease severity. It can be easily incorporated into the examination of patients with LS. GA could also be used as part of a composite outcome measure in LS clinical trials. The severity score may have different values in different LS subtypes; thus further validation in various types of LS is needed. A larger study should be conducted to confirm and extend these initial findings.

ACKNOWLEDGMENT

We thank Dr. Thomas A. Medsger, Jr for helpful comments and critical review of the manuscript, and Molly Vogt, PhD, for her statistical analysis assistance.

REFERENCES

1. Peterson LS, Nelson AM, Su WP, Mason T, O'Fallon WM, Gabriel SE. The epidemiology of morphea (localized scleroderma) in Olmsted County 1960-1993. J Rheumatol 1997;24:73-80.
2. Murray KJ, Laxer RM. Scleroderma in children and adolescents. Rheum Dis Clin North Am 2002;28:603-24.
3. Torres JE, Sanchez JL. Histopathologic differentiation between localized and systemic scleroderma. Am J Dermatopathol 1998;20:242-5.
4. Birdi N, Shore A, Rush P, Laxer RM, Silverman ED, Krafchik B. Childhood linear scleroderma: a possible role of thermography for evaluation. J Rheumatol 1992;19:968-73.
5. Hoffmann K, El-Gammal S, Gerbaulet U, Schatz H, Altmeyer P. Examination of circumscribed scleroderma using 20-MHz B-scan ultrasound. In: Altmeyer P, El-Gammal S, Hoffmann K, editors. Ultrasound in dermatology. Berlin, Heidelberg: Springer-Verlag; 1992:231-43.
6. Martini G, Murray KJ, Howell KJ, et al. Juvenile-onset localized scleroderma activity detection by infrared thermography. Rheumatology Oxford 2002;41:1178-82.
7. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-68.
8. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989;45:255-68.
9. Carmines EG, Zeller RA. Reliability and validity assessment. Thousand Oaks: Sage Publications, Inc.; 1979:17-27.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
11. Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol 1997;50:79-93.
12. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992;30:917-25.
13. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-2.
14. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-42.
15. Cosnes A, Anglade MC, Revuz J, Radier C. Thirteen-megahertz ultrasound probe: its role in diagnosing localized scleroderma. Br J Dermatol 2003;148:724-9.
16. Seidenari S, Conti A, Pepe P, Giannetti A. Quantitative description of echographic images of morphea plaques as assessed by computerized image analysis on 20 MHz B-scan recordings. Acta Derm Venereol 1995;75:442-5.
17. Serup J. Ten years' experience with high-frequency ultrasound examination of the skin: development and refinement of technique and equipment. In: Altmeyer K, el-Gammal S, Hoffmann K, editors. Ultrasound in dermatology. Berlin, Heidelberg: Springer-Verlag; 1992:41-54.
18. Szymanska E, Nowicki A, Mlosek K, et al. Skin imaging with high frequency ultrasound — preliminary results. Eur J Ultrasound 2000;12:9-16.
19. Charman CR, Venn AJ, Williams HC. Measurement of body surface area involvement in atopic eczema: an impossible task? Br J Dermatol 1999;140:109-11.
20. Ramsay B, Lawrence CM. Measurement of involved surface area in patients with psoriasis. Br J Dermatol 1991;124:565-70.
21. Tiling-Grosse S, Rees J. Assessment of area of involvement in skin disease: a study using schematic figure outlines. Br J Dermatol 1993;128:69-74.
22. Zulian F, Meneghesso D, Grisan E, et al. A new computerized method for the assessment of skin lesions in localized scleroderma. Rheumatology Oxford 2007;46:856-60.
23. Fredriksson T, Pettersson U. Severe psoriasis — oral therapy with a new retinoid. Dermatologica 1978;157:238-44.
24. Lahti A, Kopola H, Harila A, Myllyla R, Hannuksela M. Assessment of skin erythema by eye, laser Doppler flowmeter, spectroradiometer, two-channel erythema meter and Minolta chroma meter. Arch Dermatol Res 1993;285:278-82.
25. Serup J, Agner T. Colorimetric quantification of erythema — a comparison of two colorimeters (Lange Micro Color and Minolta Chroma Meter CR-200) with a clinical scoring scheme and laser-Doppler flowmetry. Clin Exp Dermatol 1990;15:267-72.
26. Albrecht J, Taylor L, Berlin JA, et al. The CLASI (Cutaneous Lupus Erythematosus Disease Area and Severity Index): an outcome instrument for cutaneous lupus erythematosus. J Invest Dermatol 2005;125:889-94.
27. Ashcroft DM, Wan Po AL, Williams HC, Griffiths CE. Clinical measures of disease severity and outcome in psoriasis: a critical appraisal of their quality. Br J Dermatol 1999;141:185-91.
28. Charman CR, Venn AJ, Williams H. Measuring atopic eczema severity visually: which variables are most important to patients? Arch Dermatol 2005;141:1146-51; discussion 1151.
29. Chuh AA, Chan HH. Effect on quality of life in patients with

pityriasis rosea: is it associated with rash severity? Int J Dermatol 2005;44:372-7.
30. Wolkerstorfer A, de Waard van der Spek FB, Glazenburg EJ, Mulder PG, Oranje AP. Scoring the severity of atopic dermatitis: three item severity score as a rough system for daily practice and as a pre-screening tool for studies. Acta Derm Venereol 1999;79:356-9.
31. Ben-Gashir MA, Hay RJ. Reliance on erythema scores may mask severe atopic dermatitis in black children compared with their white counterparts. Br J Dermatol 2002;147:920-5.
32. Rodnan GP, Lipinski E, Luksick J. Skin thickness and collagen content in progressive systemic sclerosis and localized scleroderma. Arthritis Rheum 1979;22:130-40.
33. Brennan P, Silman A, Black C, et al. Reliability of skin involvement measures in scleroderma. The UK Scleroderma Study Group. Br J Rheumatol 1992;31:457-60.
34. Clements P, Lachenbruch P, Siebold J, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281-5.
35. Clements PJ, Lachenbruch PA, Seibold JR, et al. Skin thickness score in systemic sclerosis: an assessment of interobserver variability in 3 independent studies. J Rheumatol 1993;20:1892-6.
36. Czirjak L, Nagy Z, Aringer M, Riemekasten G, Matucci-Cerinic M, Furst DE. The EUSTAR model for teaching and implementing the modified Rodnan skin score in systemic sclerosis. Ann Rheum Dis 2007;66:966-9.
37. Merkel PA, Clements PJ, Reveille JD, Suarez-Almazor ME, Valentini G, Furst DE. Current status of outcome measure development for clinical trials in systemic sclerosis. Report from OMERACT 6. J Rheumatol 2003;30:1630-47.
38. Foeldvari I, Wierk A. Healthy children have a significantly increased skin score assessed with the modified Rodnan skin score. Rheumatology Oxford 2006;45:76-8.