Assessment of disease severity and prognosis

T.A. Medsger, Jr.[1], S. Bombardieri[2], L. Czirjak[3], R. Scorza[4], A. Della Rossa[2], W. Bencivelli[2]

[1]Division of Rheumatology and Clinical Immunology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; [2]Rheumatology Unit, Department of Internal Medicine, University of Pisa, Pisa, Italy; [3]Department of Immunology and Rheumatology, Hungarian Brothers of St. John of God and the University of Pécs, Pécs, Hungary; [4]Clinical Immunology and Allergology Unit, University of Milan and IRCCS, Ospedale Maggiore, Milan, Italy.

T.A. Medsger, Jr., Professor of Medicine; S. Bombardieri, Professor of Rheumatology; Laszlo Czirjak, Professor of Rheumatology; Raffaella Scorza, MD, Professor of Clinical Immunology and Allergology; Alessandra Della Rossa, MD; Walter Bencivelli , Associate Professor of Medical Physics.

Please address correspondence to: Prof. T.A. Medsger Jr., Division of Rheumatology and Clinical Immunology, University of Pittsburgh, School of Medicine, 3471 Fifth Avenue, Suite 502, Pittsburgh, PA 15213, USA.

Clin Exp Rheumatol 2003; 21 (Suppl. 29): S42-S46.

Key words: Systemic sclerosis, disease severity, disease prognosis.

ABSTRACT

The Subcommittee members initially agreed on the concepts of disease activity, damage and severity, defining severity as the total effect of disease on organ function. It was decided to start with the assessment of severity using the Medsger's severity scale. A revised version of this scale was constructed. The rationale for the exclusion of other variables was provided.

Introduction

The concepts of disease severity, damage and activity are difficult to define. There is no agreement among rheumatologists regarding the precise meaning of these terms. For the purpose of this conference, the Subcommittee on Disease Severity defined these conditions as follows:

1. Severity is the total effect of disease on organ function; it has both irreversible and reversible components.
2. Damage is that component of severity that is irreversible.
3. Activity is that component of severity that is reversible; activity may result in no or little damage in the future or be replaced completely by damage.

An important limitation of studies on systemic sclerosis (SSc) is the lack of a standardized method to determine the severity of disease, either in an individual organ system or globally. A disease severity scale would be extremely useful in assessing disease status at a given time (cross sectional) and in tracking the evolution of disease over time (longitudinal). Such a scale would also assist in developing stratification variables and in measuring treatment efficacy in clinical trials.

A severity scale should have: (1) face validity (makes sense) i.e., experts accept the grades within each organ system as clinically credible; (2) content validity, i.e., all variables that are considered important are included or are represented by other equally useful variables; and (3) construct validity, i.e., the scoring system parallels an independently ascertained severity measurement. It is also desirable to keep the total number of variables to a minimum and to retain only those variables that could be practically and feasibly collected in an academic medical center or office practice setting.

Initially, disease severity in SSc may be minimal and is likely to be mostly attributable to activity (reversible inflammation/edema) rather than damage (irreversible fibrosis). Later in the natural history of disease, activity and damage may contribute more equally to severity. However, in late stage disease the greatest component of severity will be damage (fibrosis) with little or no activity (inflammation). This concept is important since it directs our approach to therapeutic intervention. Anti-inflammatory or immunosuppressive treatment makes sense early in disease, whereas anti-fibrotic therapy (no agents currently available and effective) is more logical in late stage disease.

Disease prognosis refers to outcomes of the disease process on the host and includes mortality (survival) and morbidity (disability, limitations). The latter may take the form of deficits in the ability to perform activities of daily living, and require alterations in patient employment, lifestyle and psychosocial adjustment to illness.

A disease severity scale for SSc recently has been developed and internally tested (1). The authors identified nine organ systems and identified variables for each one which could be used for defining severity. They then used prospective data collection to determine the feasibility of certain variables and received feedback describing the association of each variable with mortality in an available comprehensive longitudinal SSc databank as a proxy for severity. After discussion, consensus was reached on each organ system and se-

Table I. Preliminary SSc severity scale (from reference 1)

Organ system	0 (normal)	1 (mild)	2 (moderate)	3 (severe)	4 (endstage)
1. General	Normal	Wt loss 5.0-9.9 kg; PCV 33.0-36.9%	Wt loss 10.0-14.9 kg; PCV 29.0-32.9%	Wt loss 15.0-19.9 kg; PCV 25.0-28.9%	Wt loss 20+ kg; PCV < 25.0%
2. Peripheral vascular	Normal	Raynaud's requiring vasodilators	Digital pitting scars	Digital tip ulcerations	Digital gangrene
3. Skin	TSS 0	TSS1-1-14	TSS 15-29	TSS 30-39	TSS 40+
4. Joint/tendon	FTP 0-0.09 cm	FTP 1.0-1.9 cm	FTP 2.0-3.9 cm	FTP 4.0-4.9 cm	FTP 5.0+ cm
5. Muscle	No proximal weakness	Proximal weakness, mild	Proximal weakness, moderate	Proximal weakness, severe	Proximal weakness, severe; ambulation aids required
6. GI tract	Normal	Distal esophageal hypoperistalsis; small bowel series abnormal	Distal esophageal aperistalsis; antibiotics required for bacterial overgrowth	Malabsorption syndrome; episodes of pseudo-obstruction	Hyperalimentation required
7. Lung	Normal	DLCO 70-80%; FVC 70-80%; basilar rales; fibrosis on radiograph	DLCO 50-69%; FVC 50-69%; mild pulmonary hypertension	DLCO < 50%; FVC < 50%; moderate-severe pulmonary hypertension	Oxygen required
8. Heart	Normal	EKG conduction defect; LVEF 45-49%	Arrhythmia; RVE plus LVE; LVEF 40-44%	hypertension LVEF < 40%	CHF; arrhythmia requiring Rx
9. Kidney	Normal	Serum creatinine 1.3-1.6 mg/dl; urine proteine 2+	Serum creatinine 1.7-2.9 mg/dl; urine proteine 3-4+	Serum creatinine 3.0+ mg/dl	Dialysis required

PCV: packed cell volume (hematocrit); TSS: total skin score; FTP: finger-to-palm distance in flection; DLCO: diffusing capacity for carbon monoxide % predicted; FVC: forced vital capacity, % predicted; EKG: electrocardiogram; LVEF: left ventricular ejection fraction; CHF: congestive heart fauilure; RVE: right ventricular enlargement; LVE: left ventricular enlargement.

verity scales were developed from 0 (no documented involvement) to 4 (endstage disease) for each organ system (Table I).

Thus far, only one article has been published in which the Medsger *et al.* severity scale has been utilized. One hundred Swedish SSc patients followed over 14 years had organ system involvement assessed using some (but not all) of the original scale items (2). In this report, as expected, a high severity score was shown to predict reduced survival. The primary determinants of poor survival were extensive skin thickening, ECG changes and reduced lung and renal function.

The original intent of the severity scale was that it should include the concept of improvement. Unlike the damage scale in systemic lupus erythematosus (3), where damage continues to accumulate without reversal, SSc severity can improve. This has been shown for skin thickening (4, 5), pulmonary function (6, 7) and renal disease (8, 9). In the Swedish study, organ dysfunction accumulated in the first five years of disease and remained stable/unchanged thereafter (2).

Candidate variables

The Subcommittee elected to use the Medsger *et al*. severity scale (1) as the basic document for discussion. For each organ system, the subcommittee thoroughly discussed the published 0-4 scale items and made recommendations for approval or modification.

Discussion

Identification of core set variables

1.*General System*. Calculation of weight loss in Kg should begin with the patient's baseline weight immediately prior to the onset of SSc. A percentage of total body weight loss from baseline should be used rather than absolute weight loss. 0 (normal) = < 5%; (1) mild = 5-10%; (2) moderate = 10-15%; (3) severe = 15-20%; and (4) endstage = > 20%. Hemoglobin is an acceptable alternative to packed cell volume (PCV or hematocrit),as follows: 0 (normal) = 12.3 Gm/dl or greater; 1 (mild) = 11.0-12.3; (2) moderate = 9.7-11.0; 3 (severe) = 8.3-9.7; and (4) endstage =<8.3 Gm/dl.

2. *Peripheral Vascular System*. The published scale was considered adequate to describe the spectrum of digital vascular ischemia.

3. *Skin System*. It was recommended that the modified Rodnan skin thickness scoring system (10) be retained without change.

4. *Joint/Tendon Systems*. Joint involvement is typically described as consisting of joint pain on motion, tenderness or swelling (synovitis or effusion). These features are amenable to description in diseases which affect only the articular structures, such as rheumatoid arthritis. However, in SSc, where skin, subcutaneous tissue, tendon sheaths and tendons themselves are all affected, and by both inflammation and fibrosis, the use of tender/swollen joint counts is less reliable and less reproducible. Palpable tendon friction rubs (11) are considered evidence of activity (fibrous tenosynovitis, tendinitis) rather than damage.

It is recognized that joint contractures can represent the end result of pathologic processes in all of the above tissues plus skeletal muscles. Considering the nearly uniform loss of motion of the finger joints, finger contracture was felt to be the best candidate variable to include. The published methods of the 3rd fingertip to distal palmar crease

Table II. Revised preliminary SSc severity scale

Organ system	0 (normal)	1 (mild)	2 (moderate)	3 (severe)	4 (endstage)
1. General	**Wt loss < 5%; PCV 37.0%+; Hb 12.3+ Gm/dl**	**Wt loss 5.0-9.9%; PCV 33.0-36.9% Hb 11.0-12.2 Gm/dl**	**Wt loss 10.0-14.9%; PCV 29.0-32.9% Hb 9.7-10.9 Gm/dl**	**Wt loss 15.0-19.9%; PCV 25.0-28.9% Hb 8.3-9.6 Gm/dl**	**Wt loss 20+ %; PCV < 25.0% Hb < 8.3 Gm/dl**
2. Peripheral vascular	**No Raynaud's; Raynaud's not requiring vasodilators**	Raynaud's requiring vasodilators	Digital pitting scars	Digital tip ulcerations	Digital gangrene
3. Skin	TSS 0	TSS 1-14	TSS 15-29	TSS 30-39	TSS 40+
4. Joint/tendon	FTP 0-0.09 cm	FTP 1.0-1.9 cm	FTP 2.0-3.9 cm	FTP 4.0-4.9 cm	FTP 5.0+ cm
5. Muscle	Normal proximal muscle strength	Proximal weakness, mild	Proximal weakness, moderate	Proximal weakness, severe	**Ambulation aids required**
6. GI tract	**Normal esophagram; normal small bowel series**	Distal esophageal hypoperistalsis; small bowel series abnormal	Antibiotics required for bacterial overgrowth	Malabsorption syndrome; episodes of pseudo-obstruction	Hyperalimentation required
7. Lung	**DLCO 80+%; FVC 80+%; No fibrosis on radiograph; sPAP < 35 mmHg**	**DLCO 70-79%; FVC 70-79%; basilar rales; Fibrosis on radiograph sPAP 35-49 mmHg**	**DLCO 50-69%; FVC 50-69%; sPAP 50-64 mmHg**	**DLCO < 50%; FVC < 50%; sPAP 65+ mmHg**	**Oxygen required**
8. Heart	**EKG normal; LVEF 50+%**	**EKG conduction defect; LVEF 45-49%**	**EKG arrhythmia; LVEF 40-44%**	**EKG arrhythmia requiring Rx; LVEF 30-40%**	**CHF; LVEF < 30%**
9. Kidney	**No Hx SRC with serum creatinine < 1.3 mg/dl**	**Hx SRC with serum creatinine < 1.5 mg/dl**	**Hx SRC with serum creatinine 1.5-2.4 mg/dl**	**Hx SRC with serum creatinine 2.5-5.0 mg/dl**	**Hx SRC with serum creatinine > 5.0 mg/dl or dialysis required**

Wt: weight; PCV: packed cell volume (hematocrit); Hb:hemoglobin; TSS:total skin thickness score; FTP: fingertip-to-palm distance in flexion; DLCO: diffusing capacity for carbon monoxide, % predicted; FVC: forced vital capacity, % predicted; sPAP:estimated pulmonary artery systolic pressure by Doppler echo; EKG:electrocardiogram; LVEF:left ventricular ejection fraction; Rx:treatment; CHF:congestive heart failure; Hx:history of; SRC:scleroderma renal crisis.
N.B. If two items are included for a severity grade, only one is required for the patient to be scored as having that severity level.

(FTP) measurement (12) and handspread (13) were considered. Both have limitations. The FTP however does have a normal value of 0.0 ^{cm}, even if the measurement technique is not standardized and the variability between examiners is large. Also, fingers which are so severely contracted that they are "fixed" in exaggerated flexion have a small FTP distance, which suggests only minimal impairment. Such hands are actually very poor from a functional standpoint since they have virtually no motion. The subcommittee members felt that a more precise/reliable measure of finger contracture would be most desirable, and thus issues a challenge for clinical investigators to develop such a measure. In the future, another alternative would be a patient-completed hand function questionnaire or practical test. However, for the present we recommend retaining the FTP measurement as described in the severity scale publication, without any changes.

5. *Skeletal Muscle System*. The subcommittee agreed to the use of the proximal muscle strength grading system administered by the examining physician, as currently recommended by the severity scale authors. However, this method will require detailed instructions for the examiner and validation before it is acceptable as a measure of skeletal muscle dysfunction.

6. *Gastrointestinal System.* The subcommittee felt that a radiographic method was preferable, supplemented by "clinical judgement", with the esophagus and small intestine being the primary sites for evaluation. Small intestinal radiographic abnormalities alone need to be distinguished from the clinical disorders which can result, i.e. bacterial overgrowth (hydrogen breath test) with diarrhea and the more severe manifestations of episodes of functional small intestinal obstruction (pseudo-obstruction) and frank malabsorption syndrome. The subcommittee recommended continuing to use the published severity scale for the gastrointestinal system, with better descriptions provided for the examiner to facilitate proper classification of the individual patient.

7. *Lung System*. The lung has two primary types of involvement, i.e. interstitial inflammation/fibrosis and pulmonary vascular disease. Each should be accommodated by the severity scale. Subcommittee members agreed that the spectrum of pulmonary function test results, particularly the FVC percent predicted (restrictive disease) and DLCO percent predicted (vascular disease) should be included. The DLCO should be corrected for alveolar volume (DLCO/VA).

8. *Heart System*. It was generally agreed that the most important abnormalities include those resulting from dysfunction of the conduction system and the left ventricular myocardium. Thus routine electrocardiogram and, if available, Holter monitor tests are needed to detect conduction system abnormalities and arrhythmias and echocardiography to quantitate LV contractile function. The published list of mild, moderate, severe and endstage manifestations might be modified in the future.

9. *Kidney System.* For renal disease, the Subcommittee felt that reorganization

of the severity scale was necessary. Because scleroderma renal crisis (SRC) is the most frequent and dominant renal manifestation, its presence or absence was considered to be the most important distinction between no involvement (no history of SRC) and the varying degrees of renal disease, grade 1-4 (history of SRC). It was recommended that these latter grades be associated with specific levels of serum creatinine or a simple method for estimation of creatinine clearance such as the Cockroft index (14). For example, (mild) could be a history of SRC and current serum creatinine < 1.5 mg/dl or creatinine clearance > 100 cc/min; 2 (moderate) = 1.5-2.4 or 60-99; 3 (severe) = 2.5-5.0 or 30-79; and 4 (end-stage) = dialysis, 5.0 mg/dl or greater or clearance < 30 cc/minute.

Rationale for the exclusion of other variables

1. *General system*. Serum albumin was considered as a possible variable to describe general organ dysfunction, as follows: 0 (normal) = > 3.5 gm/dl; 1 (mild) = 3.2-3.5; 2 (moderate) = 2.9-3.2; (3) severe = 2.6-2.9; and 4 (endsdtage) = < 2.6. It was not recommended as a substitute for either weight loss or hemoglobin. In addition, there was discussion concerning patient self-evaluation e.g. the HAQ disability index and physician global assessment (VAS) as reflecting the General System (15), but there was concern that the responses might reflect the most serious organ system involvement. For example, the patient with moderate to severe pulmonary fibrosis would answer this question based on the particular limitations imposed by his/her lung disease.

2. *Peripheral vascular*. For the assessment of severity of Raynaud's phenomenon, an alternative could be the visual analog scale (VAS) from the Scleroderma Health Assessment Questionnaire (SHAQ) (15) with "mild" defined as >50% on the VAS. However, this approach is not commonly used. In addition, quantitation of capillary microscopic abnormalities was considered to be a variable with relevance to microvascular disease, but no scale for its use has been developed to date.

3. *Skin*. Ultrasound methods were considered but the equipment is expensive and not uniformly available and the results are operator-dependent and not yet standardized.

4. *Joint/tendon.* Hand spread was considered. However, it has no normal value, and thus a handspread distance cannot be related to a "baseline" measurement to determine the degree of abnormality attributable to the disease. The detection of joint contractures was believed by the Subcommittee to be the best available, but not ideal, measure to use.

5. *Muscle.* No other variable was considered.

6. *Gastrointestinal.* To assess esophageal involvement, a quantitative manometric method also would be helpful, but is not widely available. In addition, definitions would be necessary for "hypoperistalsis" and "mild to moderate manometric abnormalities" so that they could be distinguished from "aperistalsis" and "severe manometric abnormalities". The SSc HAQ GI VAS (15) was also discussed as a candidate variable, but it still needs to be validated.

7. *Lung*. Restrictive disease should be further characterized by HRCT (1.0 mm slices) and vascular disease by echo Doppler or right heart catheterization. An oxygen desaturation test could be also useful. However, these techniques are not feasible everywhere. A precise rearrangement of the published severity scale variables was not attempted during this conference. The 6-minute walk distance (16) could also be considered but its interpretation is difficult in patients with SSc who often have musculoskeletal factors limiting their ambulation. An SSc HAQ VAS for lung disease could also be included but dyspnea in SSc may sometimes be attributable to problems with other organ systems and is thus not lung-specific.

8. *Heart*. Pericardial disease was not discussed, but could also be incorporated in terms of the variables small, moderate or large effusion and pericardial tamponade. Stress echocardiogram and thallium perfusion studies or MUGA scans were discussed as candidate measures, but not incorporated as definite recommendations at this time.

9. *Kidney*. The evaluation of the glomerular filtration rate could be carried out by an isotopic technique, if an inexpensive and uniformly available procedure were to be established. This methodology is not feasible in many centers.

Mortality

Reports on survival should consider methods of calculating survival. The death date is the endpoint for all such calculations. The starting date can be any one of three dates: (1) date of disease onset (first symptom or finding attributed to SSc) as judged by the patient or the physician; (2) date of the first physician diagnosis of SSc; or (3) date of first enrollment in a research study. Cumulative survival from the starting date should be calculated using a standard method and taking into account patients lost to follow-up, such as the Kaplan-Meier method. All reports should include 5- and 10-year cumulative survival rates so that published studies can be directly compared with one another.

The cause of death should be recorded as accurately as possible, using the method described by Geirsson *et al.* (2) which includes the categories of: (1) definitely SSc-related (due to organ failure [specific organ]); (2) probably SSc-related (due to a complication caused or aggravated by SSc associated with organ injury or treatment); (3) possibly SSc-related (due to a manifestation reported to have increased prevalence in SSc, such as malignancy or suicide); and (4) unrelated to SSc, its organ involvement or treatment.

Morbidity

Important outcomes in addition to survival are morbidity and disability related to SSc involvement and treatment complications. We recommend the most widely used instrument, the disability index (DI) of the HAQ, a validated and reliable instrument (17). HAQ DI scores correlate well with the extent of skin thickening, loss of first closure, proximal muscle weakness and tendon friction rubs. The SHAQ "global" VAS is an alternative (15), and a variety of other validated instruments could be used. Several instruments have recently

been developed to measure hand dysfunction, but require observation of the patient and scoring by a trained professional (18-20).

References

1. MEDSGER TA, JR., SILMAN AJ, STEHEN VD, *et al.*: A disease severity scale for systemic sclerosis: development and testing. *J Rheumatol* 1999; 26: 2159-67.
2. GEIRSSON AJ, WOLLHEIM FA, AKESSON A: Disease severity of 100 patients with systemic sclerosis over a period of 14 years: using a modified Medsger scale. *Ann Rheum Dis* 2001; 60: 1117-22.
3. GLADMAN D, GINZLER E, GOLDSMITH C, *et al.*: The development and initial validation of the systemic lupus international collaborating clinics / American College of Rheumatology damage index for systemic lupus erythematosus. *Arthritis Rheum* 1996; 39: 363-9.
4. STEHEN VD, MEDSGER TA, JR.: Improvement in skin thickening in systemic sclerosis associated with improved survival. *Arthritis Rheum* 2001; 44: 2828-35.
5. BLACK C, DIEPPE PK, HUSKISSON T: Progressive systemic sclerosis. *Ann Rheum Dis* 1986; 45:384-88.
6. SILVER RM, WARRICK JH, KINSELLA MB, *et al.*: Cyclophosphamide and low dose prednisone therapy in patients with systemic sclerosis (scleroderma) with interstitial lung disease. J.Rheumatol. 1993; 20: 838-44.
7. STEEN VD, OWENS G, REDMOND C, *et al.*: The effect of D-penicillamine on pulmonary findings in progressive systemic sclerosis. *Arthritis Rheum* 1985; 28: 882-8.
8. STEEN VD, COSTANTINO JP, SHAPIRO AP, *et al.*: Outcome of renal crisis in systemic sclerosis: relation to availability of angiotensin convertine enzyme (ACE) inhibition. *Ann Intern Med* 1990; 113:352-57.
9. STEEN VD, MEDSGER TA, JR.:Long term outcome of scleroderma renal crisis:Ann. Intern. Med. 2000; 133: 600-3
10. CLEMENTS PJ, LACHENBRUCH PA, SEIBOLD JR, *et al.*: Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in Systemic Sclerosis. *Rheumatol* 1995; 22:1281-5.
11. STEEN VD, MEDSGER TA, JR.: The palpable tendon friction rub: an important physical examination finding in patients with systemic sclerosis. *Arthritis Rheum* 1997; 40: 1146-151.
12. CLEMENTS PJ, FURST DE, WONG WK, e*t al.*: High-dose versus low-dose D-penicillamine in early diffuse systemic sclerosis:analisys of two-year, double-blind, randomized, controlled clinical trial. *Arthritis Rheum* 1999; 42: 1194-203.
13. CLEMENTS PJ, WONG WK, NURWITZ EL *et al.*: Correlates of the disability index of the health assessment questionnaire. *Arthritis Rheum* 1999; 42: 2372-80.
14. COCKCROFT DW, GULT MH: Prediction of creatinine clearance from serum creatinine. *Nephron* 1976; 16: 31-41.
15. STEEN VD, MEDSGER TA, JR.: The value of health assessment questionnaire and special patient generated scales to demonstrate change in patients with systemic sclerosis over time. *Arthritis Rheum* 1997; 40: 1984-91.
16. BADESCH DB, TAPSON VF, MCGOON MD, *et al.*: Continuous intravenous epoprostenol for pulmonary hypertension due to scleroderma spectrum of disease. *Ann Intern Med* 2000; 132: 425-34.
17. FRIES JF, SPITZ P, KRAINES RG, *et al.*: Measurement of patients outcome in arthritis. Arthritis Rheum 1980; 23: 137-45.
18. SANDQVIST G, EKLUND M: Hand mobility in scleroderma (HAMIS) Test: The reliability of a novel hand function test. *Arthritis Care and Research* 2000; 13: 369-74.
19. SANDQVIST G, EKLUND M: Validity of HAMIS:A test of hand mobility in scleroderma. *Arthritis Care and Research* 2000; 13: 382-7.
20. POOLE JL,GALLEGOS M,O'LINE S: Reliability and validity of the arthritis hand function test in adults with Systemic Sclerosis (scleroderma). *Arthritis Care and Research* 2000; 13: 69-73.