

CONCISE REPORT

European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. III. Assessment of the construct validity of the preliminary activity criteria

G Valentini, W Bencivelli, S Bombardieri, S D'Angelo, A Della Rossa, A J Silman, C M Black, L Czirjak, H Nielsen, P G Vlachoyiannopoulos

Ann Rheum Dis 2003;62:901–903

Objective: To further assess the construct validity of the three European Scleroderma Study Group (EScSG) preliminary activity indices for systemic sclerosis (SSc): for SSc as a whole, for diffuse SSc (dcSSc), and for limited SSc (lcSSc).

Methods: 30/290 SSc clinical charts collected for the EScSG study used to develop activity criteria for SSc were selected and sent to four clinical experts in SSc. The experts ranked the charts from 1 to 30 (1=lowest activity, 30=highest activity). The relationships among the ranks given by each investigator and each of the three scores, and between any two of the ranks were investigated.

Results: A consistently significant correlation ($r_s=0.530-0.712$) was found between the ranks given by each of the four investigators and the index for the entire patient group. A similar level of agreement was detected between each couple of the four experts ($r_s=0.428-0.720$). Moreover, the ranks given in patients with an index >3 were significantly higher than those given for patients with an index ≤ 3 . This cut off point had previously been shown to best discriminate patients with active disease.

Conclusions: Of the originally developed activity indexes, the whole series index has been externally validated. The index comprises the first preliminary, but necessary, groundwork to improve the concept of disease activity in SSc, which is still ill defined. It can be used as a preliminary activity index in clinical investigational studies.

The European Scleroderma Study Group (EScSG) recently proposed three activity 10 point indices for systemic sclerosis (SSc): one for SSc as a whole (whole series index), one for diffuse cutaneous SSc (dcSSc), and one for limited cutaneous SSc (lcSSc).^{1,2} The indices were created after analysing the clinical charts of 290 patients with SSc from 19 European medical centres and identifying the set of items that was most significantly correlated with a subjectively defined disease activity score (on a scale from 0–10). The latter had been assigned to the patients under blinded conditions by three members of the protocol management committee and resulted in a reliable measure.²

In that study we assessed the construct validity of the indices by the jackknife technique³—that is, a statistical method in which one patient at a time is excluded and the significance of the data is evaluated by its dispersion. This approach, while statistically correct, is characterised by an intrinsic circularity: clinical investigators assign a score which is used to establish an index, which in turn is validated by comparing it with the assigned score.

Table 1 Proposed EScSG disease activity indices for SSc

Criteria	Whole series	dcSSc	lcSSc
TTS* >20	1.0		
Scleredema	0.5		0.5
Δ Skin \dagger	2.0	3.0	2.5
Digital necrosis	0.5		
Δ Vascular \dagger	0.5	2.0	1.0
Arthritis	0.5		1.0
Δ Articular/muscular \dagger		1.0	
\downarrow TLCO ($<80\%$ of predicted value)	0.5		
Δ Cardiopulmonary \dagger	2.0	4.0	1.5
ESR >30 mm/1st h	1.5		2.5
Hypocomplementaemia (C3 and/or C4)	1.0		1.0
Total maximum disease activity index	10.0	10.0	10.0

*Kahaleh total skin score⁴; \dagger deterioration, as evaluated by the patient, of the condition in the month preceding the assessment. TTS, total skin score; TLCO, carbon monoxide transfer factor; ESR, erythrocyte sedimentation rate.

We therefore decided to further test the validity of our indices by assessing the correlation of each of them with the activity ranks given by a different set of observers.

MATERIAL AND METHODS

Table 1 shows the three preliminary activity indices established in our previous study.² For the external validation of the indices, the protocol management members (GV, SB, SD'A, ADR) stratified the dataset in three subgroups with increasing activity (inactive, moderately active, and active/very active). Thirty charts (10 out of each group) were then randomly selected and sent to four experts (CMB, LC, HN, PGV) identified as representative of different countries and geographical areas and available to give their evaluation at the time the study was carried out. The four experts were requested to rank the charts under blinded conditions from 1 to 30 (least active disease to most active disease).

We then investigated the correlation between the ranks given by each expert and the respective activity indices for each patient using a Spearman rank correlation. In our previous work we had shown that a score of over 3 was the best single point discriminator between active and inactive disease. We then compared the median rank, for each expert, between those with and without active disease using the Mann-Whitney U test.

Abbreviations: EScSG, European Scleroderma Study Group; dcSSc, diffuse cutaneous systemic sclerosis; lcSSc, limited cutaneous systemic sclerosis; SSc, systemic sclerosis

Table 2 Correlations between the rankings given by four investigators and the calculated index in the whole series, in patients with dcSSc, and in patients with lcSSc

Expert	Whole (n=30)	dcSSc (n=17)	lcSSc (n=13)	
HN	r_s	0.686	0.821	0.367
	p	<0.001	<0.001	>0.05
PGV	r_s	0.623	0.357	0.417
	p	<0.001	>0.05	>0.05
CMB	r_s	0.530	0.778	0.379
	p	0.003	<0.001	>0.05
LC	r_s	0.712	0.840	0.851
	p	<0.001	<0.001	<0.001

Table 3 Correlations between the rankings given by four investigators in 30 patients with SSc considered as a whole

Expert	HN	PGV	CMB	LC	
HN	r_s	1.000			
	p				
PGV	r_s	0.497	1.000		
	p	0.005			
CMB	r_s	0.563	0.428	1.000	
	p	0.001	0.018		
LC	r_s	0.720	0.430	0.720	1.000
	p	<0.001	0.018	<0.001	

RESULTS

Table 2 shows the correlations between the disease activity rankings given by each of the four experts and the three ESsSG indices. A significant correlation was found between the ranks given by each of the four external investigators and the derived index for the whole series. However, the correlations were not uniformly significant for the dcSSc index and very rarely significant for the lcSSc index. Therefore, we decided only to consider the whole series index for further evaluation.

Table 3 shows the correlations between the ranks given by the four experts for the 30 charts. This table clearly shows that even among experts in the field there is not a very high level of agreement in the assessment of disease activity.

Figure 1 shows the distribution of the ranks given by each expert in patients with a whole series index ≤ 3 and > 3 (that is, the value that had been found, in our previous paper, to discriminate between patients with low and high activity). The ranks in patients with high activity disease were in each case significantly higher than those given in patients with low activity.

DISCUSSION

We identified and validated three sets of criteria devoted to defining disease activity in SSc as whole, in dcSSc, and in lcSSc. Our first validation, however, was carried out by the jackknife approach,³ which is affected by an intrinsic circularity. In this report we provide the results of an external validation process. Our data show that the index derived for use in all patients is the only one which was validated by all the

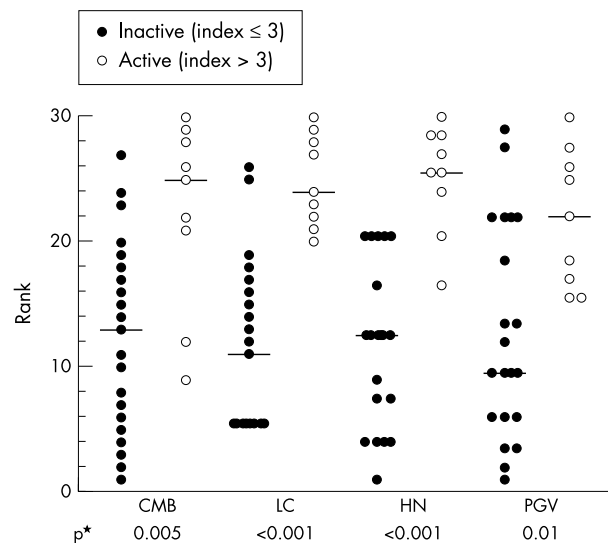


Figure 1 Distributions of the ranks given by each investigator in patients with active and inactive disease according to the whole series index.

external investigators. We are inclined to think that the unsatisfactory performance of the dcSSc index may be due to the small number of patients investigated, and the unsatisfactory performance of the lcSSc to the difficulties in identifying activity in patients whose disease has an indolent course. Therefore, further evaluation of these subgroup indices should be undertaken.

To define disease activity criteria for SSc, we used the same methodology previously used by two of us (SB and WB) to define both classification criteria for Sjögren's syndrome³ and activity criteria for systemic lupus erythematosus, which has been validated both internally⁶ and externally.⁷ Therefore we relied on real patients. Such methodology had the substantial advantage of being based on parameters which are currently evaluated in each centre, enhancing the feasibility of the developed index.

In the present study, as well as in our first study, the correlation between the activity items and the "gold standard" (that is, the activity score, from 1 to 10, given in blinded conditions by protocol management members³) was lower than that detected in other diseases such as systemic lupus erythematosus.⁶ This might suggest that the developed index has an intrinsic weakness. However, when a correlation between the ranks given by different experts was carried out (table 3), a similar level of agreement was detected. We therefore suggest that the low level of agreement found reflects the low level of concordance on the degree of disease activity in SSc. Defining activity in SSc is different from defining activity in other diseases such as rheumatoid arthritis and systemic lupus,⁶⁻¹⁰ because of the lack of clear flares and quiescent phases and its indolent course.

The presence of three Δ factors based on patient self assessment may cause some concern because of the lack of standardisation. However, such an approach has been accepted as valid and reliable in assessing disease activity in rheumatoid arthritis.¹⁰⁻¹¹

In conclusion, we have developed a validated index to assess disease activity in SSc. This index must still be considered preliminary. However, it should be viewed as the first necessary step towards developing a better agreement among different investigators in the assessment of disease activity in SSc. It can be used as a preliminary activity index in clinical investigational studies. Nevertheless, the search for other measures of disease activity in SSc must continue.

.....

Authors' affiliations

G Valentini, S D'Angelo, Second University of Naples, Italy
W Bencivelli, S Bombardieri, A Della Rossa, Clinical Immunology/Rheumatology Units, Department of Internal Medicine, University of Pisa, Italy
A J Silman, University of Manchester, UK
C M Black, Royal Free Hospital, London, UK
L Czirjak, University of Pécs, Hungary
H Nielsen, Rheumatology Unit, Herlev Hospital, Denmark
P G Vlachoyiannopoulos, National University, Athens, Greece

See also "European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. IV. Assessment of skin thickening by modified Rodnan skin score" in this issue of the journal p 904–5.

Correspondence to: Professor G Valentini, U.O di Reumatologia, Edificio 3, Via Pansini, 5 Naples, Naples, Italy; gabriele.valentini@unina2.it

Accepted 22 January 2003

REFERENCES

- 1 **Della Rossa A**, Valentini G, Bombardieri S, Bencivelli W, Silman AJ, D'Angelo S, *et al*. European multicentre study to define disease activity criteria for systemic sclerosis. I. Clinical and epidemiological features of 290 patients from 19 centres. *Ann Rheum Dis* 2001;60:585–91.
- 2 **Valentini G**, Della Rossa A, Bombardieri S, Bencivelli W, Silman AJ, D'Angelo S, *et al*. European multicentre study to define disease activity criteria for systemic sclerosis. II. Identification of disease activity variables and development of a preliminary activity index. *Ann Rheum Dis* 2001;60:592–8.
- 3 **Shao J**, Tu D. *The jackknife and the bootstrap*. New York: Springer, 1995.
- 4 **Kahaleh MB**, Sultany GL, Smith EA, Huffstutter JE, Loadholt CB, Le Roy EC. A modified scleroderma skin scoring method. *Clin Exp Rheumatol* 1986;4:367–9.
- 5 **Vitali C**, Bombardieri S, Moutsopoulos HM, Balestrieri G, Bencivelli W, Bernstein RM, *et al*. Preliminary criteria for the classification of Sjögren's syndrome. Results of a prospective concerted action supported by the European Community. *Arthritis Rheum* 1993;36:340–7.
- 6 **Vitali C**, Bencivelli W, Isenberg DA, Smolen JS, Snaith ML, Sciuto M, and the European Consensus Study Group for disease activity in SLE. Disease activity in systemic lupus erythematosus: report of the Consensus Study Group of the European Workshop for Rheumatology Research. II. Identification of the variables indicative of disease activity and their use in the development of an activity score. *Clin Exp Rheumatol* 1992;10:541–7.
- 7 **Ward MM**, Marx AS, Barry NN. Comparison of the validity and sensitivity to change of 5 activity indices in systemic lupus erythematosus. *J Rheumatol* 2000;27:664–70.
- 8 **Liang MH**, Sucher SA, Roberts WN, Esdaile JM. Measurement of systemic lupus erythematosus activity in clinical research. *Arthritis Rheum* 1988;31:817–25.
- 9 **Scott DL**, Panayi GS, van Riel PL, Smolen J, van de Putte LB. Disease activity in rheumatoid arthritis: preliminary report of the Consensus Study Group of the European Workshop for Rheumatology Research. *Clin Exp Rheumatol* 1992;10:521–5.
- 10 **Felson DT**, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, *et al*. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993;36:729–40.
- 11 **Boers M**, Tugwell P. OMERACT conference questionnaire results. OMERACT Committee. *J Rheumatol* 1993;20:552–4.