Acta Derm Venereol (Stockh) 1989; 69: 41–45

Scoring Atopic Dermatitis: The Simpler the Better?

C. COSTA,[1] A. RILLIET,[1] M. NICOLET[2] and J.-H. SAURAT[1]

[1]*Clinique de Dermatologie,* [2]*Policlinique de médecine, Hôpital Cantonal Universitaire, Geneva, Switzerland*

Costa C, Rilliet A, Nicolet M, Saurat J-H. Scoring atopic dermatitis: the simpler the better? Acta Derm Venereol (Stockh) 1989; 69: 41–45.

Two scoring methods evaluating the severity of atopic dermatitis have been compared. One was simple, quick and compatible with a busy outpatients clinic. The other was more complicated and time consuming; it took into account most of the evaluable clinical signs of disease activity in each involved site. There was a highly significant correlation between the two methods of scoring, thus validating the simplest one; furthermore, the more complicated method was less reproducible than the simplest when used by two physicians on the same patient. This suggests that a simple and feasible scoring method is meaningful in keeping records at each visit in any patient with atopic dermatitis. Such records may then be used in retrospect for the evaluation of any new therapy. *Key words: Clinical evaluation; Scoring systems.* (Accepted July 12, 1988.)

J.-H. Saurat, Clinique de Dermatologie, Hôpital Cantonal Universitaire, 1211 Geneva 4, Switzerland.

There is no biological parameter that would allow an objective follow-up of atopic dermatitis (AD) in a given patient. At each visit AD patients seem "better" or "worse", but it is difficult in a long-term follow-up to keep objective records of their clinical course. The clinical record is usually unsuitable for individual evaluation of either the therapeutic regimen or multiple environmental factors. In the past it has been our policy to use, for all patients, a "score" that would allow such an evaluation as it gives a "number" reflecting the severity of the disease at each visit (1–2). We had designed a score aimed at the evaluation of AD in childhood since this represented the bulk of our practice at that time.

When we started to use this rather simplified scoring system in adult AD patients, we wondered if such a rough evaluation would be relevant to our purpose. We felt the lesions in adults to be less uniform than in young children so that an accurate score should take into account the severity of the dermatosis at each affected site. Therefore we designed a new scoring system, that was much more elaborate (and therefore time-consuming). It soon appeared that it would be unfair and perhaps unrealistic to set such a task for the medical staff unless this proves necessary. Thus, both scoring systems were evaluated in adult patients with AD. Since we think that a severity score is an important clinical problem in the care of AD patients and since we are not aware of any publication comparing two scoring systems in AD, we thought it was worthwhile to report our observations.

MATERIAL AND METHODS

Patients

14 consecutive patients (9 females, 5 males; aged 15 to 35 years, mean 23) from our AD clinic entered the study. They were evaluated with both scoring methods at each visit for a total of 100 visits. In order to check the reproducibility of each scoring method, 7 patients were examined simultaneously by two physicians over 10 visits. All the patients received standardized systemic and topical therapy according to the severity of the disease and no attempt was made to modify the treatment in relation to the scoring study results.

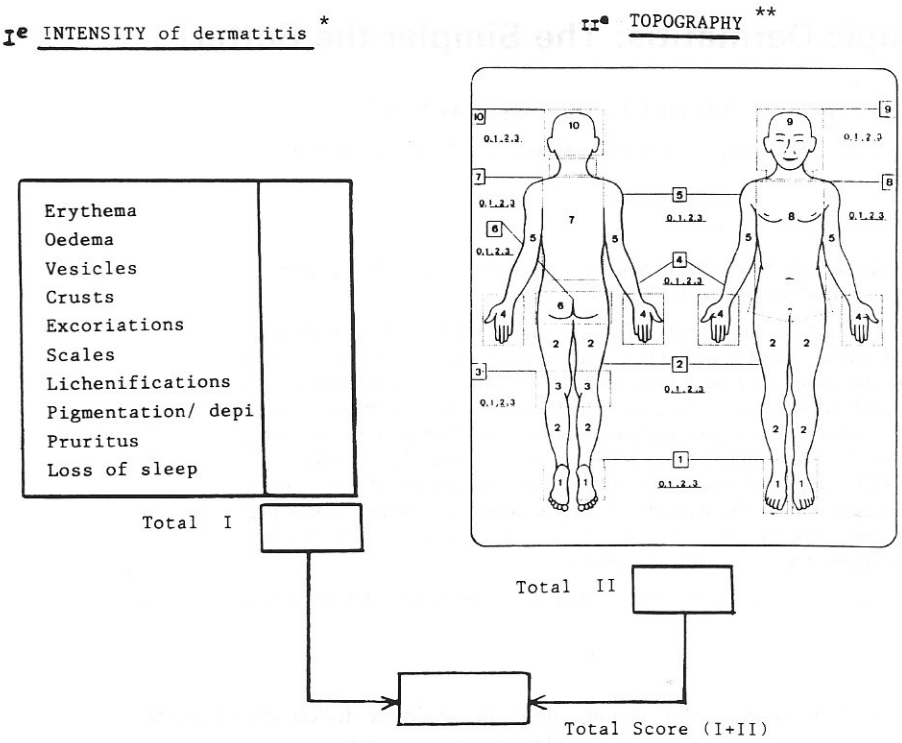

Fig. 1. Simple scoring system. (I) Intensity of dermatitis: each item scored 0 to 7 (see text) was evaluated in the most severely affected areas. (II) Topography: quotation 0 to 3 for the extent in each of the 10 areas.

Scoring systems

Simple scoring system (Fig. 1). This method consists of scoring ten "severity" criteria and ten topographic sites. The "severity" criteria are listed in Fig. 1; each was scored from 0 (no lesion) to 7 (extremely severe). The most severely affected areas were deliberately chosen for evaluation with each criterion. For the "topography" item, each of the following areas was scored from 0 to 3 according to the extent of the involvement: five symmetrical areas: feet, knees, legs, hands, arms (one value for both) and five non symmetrical areas: face, scalp, buttock, anterior and posterior aspects of the trunk.

More elaborate scoring system (Fig. 2). This method was developed in order to include in the score the maximum number of clinical variables, taking into account the fact that, in many patients, there was a difference in the severity of the disease from site to site.

—"Severity" criteria

The number of these was reduced from 10 to 5 by grouping objective signs that are usually linked at least clinically (see Fig. 2): 1) Erythema and oedema; 2) Vesicles, pustules and crusts; 3) Excoriations and cracking; 4) Scaling and dryness; 5) Lichenification. The grading scale of each criterion was reduced from 0 to 7 to 0 to 3. Indeed, retrospective analysis of grading 0 to 7 in the "simple scoring system" indicated that the highest quotations were in fact never used. Further, we quoted pruritus and loss of sleep as subjective criteria to be distinguished for the objective ones; each of these two criteria was given a score scale sufficient to contribute significantly to the final score.

—"Topography" item

Rather than just quoting the extent of the dermatitis whatever its severity form site to site, each site was graded separately for both severity and extent. The body was divided into 20 areas (see Fig. 2), so it would be possible to calculate the percentage of involvement of each particular area. Since each area does not represent a similar surface, each was given a relative correcting factor (for example, if

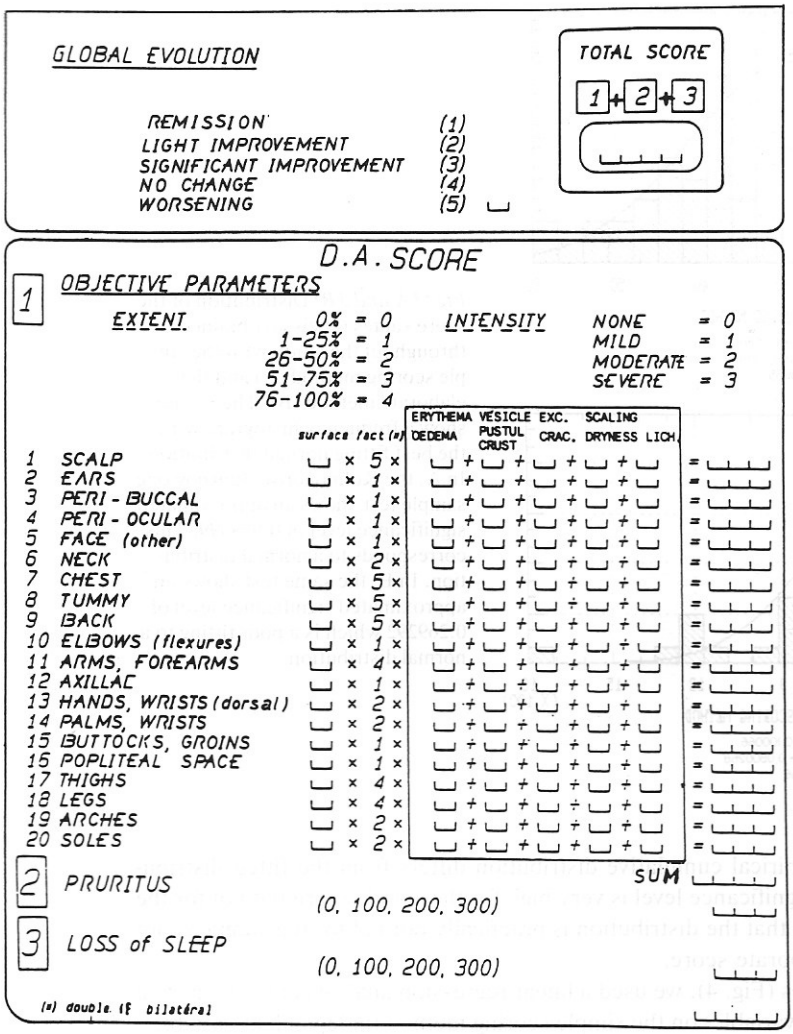

Fig. 2. More elaborate scoring system. Extent: quotation 0% to 100% for each of the 20 areas. Intensity: number of severity items reduced from 10 to 5; quotation reduced from 0 to 7 to 0 to 3.

the involvement is 100% of the surface of one ear, this would correspond to an "extent" value of 4, then 4 is multiplied by the correcting factor 1. When 100% of the scalp is affected, the same value of 4 for the extent is multiplied by the correcting factor 5, etc.).

—Scoring procedure

1) Each line (Fig. 2) is filled as follows:
   - calculation of the percentage of the involved surface: 0%=0, 1–25%=1, 26–50%=2, 51–75%=3, 76–100%=4
   - multiplication by the relative surface factor (correcting factor)
   - multiplication by the sum of the severity criteria of the particular site.
2) The score of all lines (namely all sites) is summed up
3) Addition of pruritus and loss of sleep gives the final score.

Needless to say such a procedure is time-consuming and needs a calculator.

RESULTS

Fig. 3 A and 3 B show histogram plots of the scores with the best fitting normal distribution superimposed on them. The Kolmogorov-Smirnov one sample test has been used to

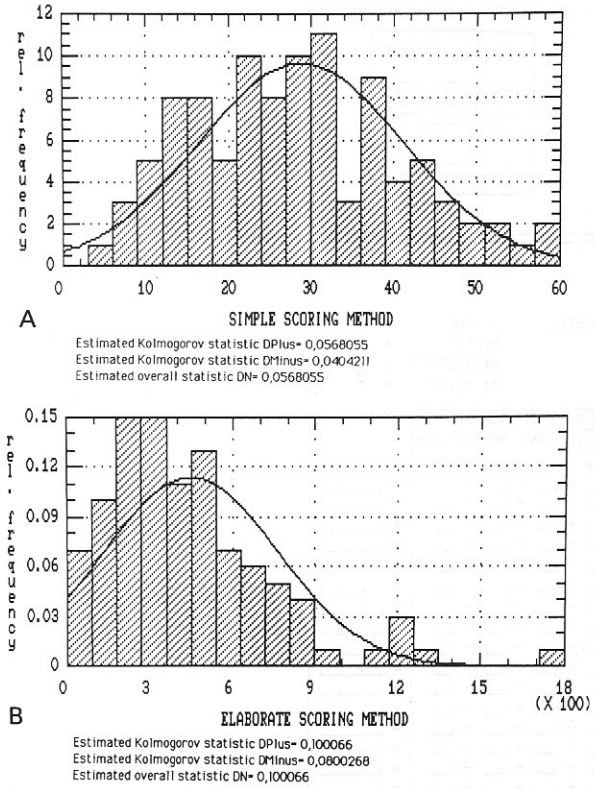

Fig. 3 A and 3 B: Distribution of the score values (abcissa) obtained throughout the study with the simple scoring method (A) and the elaborate method (B). The figures show a frequency histogram with the best fitting normal distribution. In A, the Kolmogorov-Smirnov one sample test shows an approximated significance level of 0.999999. This corresponds to a normal distribution. In B, the same test shows an approximated significance level of 0.269292 which is a poor fitting to a normal distribution.

measure how much the empirical cumulative distribution differs from the fitted distribution. One can see that the significance level is very high for the simple score but not for the elaborate one. We conclude that the distribution is practically normal for the simple score but not so good for the elaborate score.

To compare the two scores (Fig. 4), we used a linear regression analysis of the elaborate scoring method (dependent variable) on the simple scoring method (independent variable). Both in the analysis of the variance and in the slope or in the intercept, we obtained a highly significant correlation between the two scores ($r=0.9002$, $p<0.0001$). To test how reproducible the two scores were between two examinators for the same patient, we used the Wilcoxon signed rank test (the samples are paired, two examinators for the same

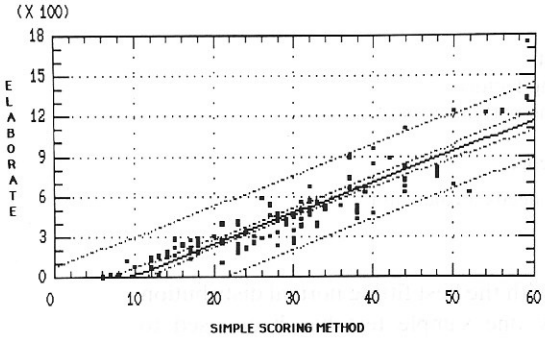

Fig. 4. Linear regression analysis of the elaborate scoring method (dependent variable) on the simple scoring method (independent variable). The analysis of the variance shows a F-ratio=418.8 which corresponds to a high significance level ($p<0.000001$). The slope has an estimate of 22.9233 (standard error: 1.12) with a T value=20.4655 ($p<0.000001$). The intercept value is −210.055 (standard error: 35.04) with a T value=5.99 ($p<0.000001$).

patient). If the results of both examinators are approximately the same, the test would not be significant which is the case with the simple score, but not with the more elaborate one ($p<0.05$, $n=10$).

DISCUSSION

The simple scoring method we have been using for many years in more than 500 patients proved to be easy to perform, quick and reproducible between different trained physicians (1, 2). The present study confirms its reproducibility. However, the selection of the "severity" criteria as well as the mode of evaluation of the extent of the dermatitis in this simple scoring method may not stand critical analysis.

Therefore when we tried to design a more elaborate mode of scoring, it was difficult to select, within the many lesions that may occur in AD patients, the ones that reflect severity in a given lesional site. The ones that have been finally used reflect a continuum in the inflammation process, and we found it useless to quote separately oedema and erythema, or vesicles and crusts, etc, because they usually go togheter. Such a limited number of lesional descriptive terms in the "more elaborate scoring method", 5 compared with 7 in the "simple method" did not apparently interfere with the final result because there was an extremely good statistical correlation between the two methods.

This study also suggests that a detailed scoring of each involved site seems useless because only the evaluation of the most severely affected sites as in the "simple scoring method" gave similar results. As both scoring systems appear comparable, it seems wise to use the simplest; this is further supported by the fact that the more elaborate method was less reproducible when two physicians' results were compared.

Scoring AD is an important factor in the care of such patients; however, two different situations must be considered. One is the need in a single patient of such a score to keep records of the evolution of his disease; the score does not need to be very much elaborate but should give better retrospective information than just a quotation "worse or better" in the files. We think the "simple scoring system" even reduced to 5 criteria for lesional analysis is adequate for that purpose. Whether such a simplified method is usable for research purposes in order to detect subtle clinical modifications, for instance during drug trials, remains to be established. However, it seems likely because the most important criterion in such a trial would be the overall comparison of the state of the disease over many months (or years) before and during the trial; this criterion could only be obtained when a reliable score has been used throughout.

REFERENCES

1. Queille-Roussel C, Raynaud F, Saurat, JH. A prospective computerized study of 500 cases of atopic dermatitis in childhood. I. Initial analysis of 250 parameters. Acta Derm Venereol (Stockh) 1985; Suppl 114: 87–92.
2. Higaki Y, Hauser C, Rilliet A, Saurat JH. Increased in vitro cell-mediated immune response to staphylococcal antigens in atopic dermatitis. J Am Acad Dermatol 1986; 15: 1204–1209.