Measuring Atopic Eczema Severity Visually

Which Variables Are Most Important to Patients?

Carolyn R. Charman, BM, BCh, MRCP; Andrea J. Venn, PhD; Hywel Williams, MSc, PhD, FRCP; Michael Bigby, MD

Background: There is wide variation in the objective visual variables used to measure atopic eczema severity in clinical trials, making comparison and interpretation of results difficult.

Objective: To provide a rationale for simplifying and standardizing objective atopic eczema scoring by investigating which visual variables provide the best measure of disease severity from the patient’s perspective.

Setting: The dermatology outpatient department at the Queen’s Medical Centre, University Hospital in Nottingham, and 5 local general practices.

Patients: One hundred eighty individuals with atopic eczema.

Interventions: Clinical examination with scoring of 7 clinical signs and disease extent, followed by regression analyses of visual variable scores against a patient-rated measure of current disease severity.

Results: Objective measurements account for only a quarter of the variation in patient-rated disease severity. Three clinical signs were independent predictors of patient-rated disease severity: excoriations, erythema, and edema/papulation. Disease extent measurements do not reflect patient-rated disease severity in a linear manner, with mean severity scores increasing little above 30% body surface area involvement.

Conclusions: From the patient’s perspective, the measurement of 3 clinical signs—excoriations, erythema, and edema/papulation—provides as much information about current atopic eczema severity as more complex scoring systems that measure multiple clinical signs and disease extent. The simplicity of the Three Item Severity score, a previously published atopic eczema score based on measurement of these 3 clinical signs, makes it a suitable tool for research studies or clinical practice.

Arch Dermatol. 2005;141:1146-1151

MEASURING SKIN DISEASE severity is a central part of dermatologic practice. In the outpatient clinic and in therapeutic trials it is essential to have some method of recording disease activity to ensure that health care interventions are effective. In common with many skin diseases, there are no serologic markers that accurately reflect atopic eczema severity, and, therefore, measurements are primarily based on signs and symptoms. Although management revolves around symptom control, traditionally greater emphasis has been placed on visual assessments of clinical signs and disease extent by physicians, especially in the clinical trial setting. In a recent review of 93 randomized controlled clinical trials of atopic eczema therapy published between 1994 and 2001, scoring of clinical signs was the most widely used form of outcome measure (91% of trials), with disease extent recorded in 67% of trials. Symptoms were assessed in 86% of trials but were frequently combined with clinical signs in composite scoring systems, with relatively low weighting given to patient-rated symptoms compared with physician-rated clinical signs. Furthermore, quality of life was measured in only 3% of the trials. The most widely used clinical scoring system in this review was the SCORAD (SCORing Atopic Dermatitis) index, in which 60% of the total score is based on clinical signs and 20% is based on disease extent, with only 20% weighting for patient symptoms.

Given that objective scoring systems are playing such a central role in atopic eczema trials, it is essential that the visual variables...
being measured are important to patients and that they reflect current disease severity in a clinically meaningful way to practicing physicians. From the physician’s viewpoint, there has been a general lack of consensus as to which visual variables best reflect atopic eczema severity. Given that atopic eczema is one of the most common inflammatory diseases in dermatology, this lack of standardization in disease scoring has been astounding, with more than 50 different clinical scoring systems being identified in the 93 randomized controlled clinical trials published between 1994 and 2001 alone. These scoring systems comprised more than 30 different descriptions of clinical signs and more than 20 different methods of scoring disease extent, with widely variable weighting given to different visual variables across the randomized controlled clinical trials. In an era of patient-centered evidence-based medicine, such lack of standardization not only makes the comparison of results and the production of systematic reviews difficult but also challenges the meaningful interpretation of individual studies because researchers are selecting their own combination of signs in the absence of formal validity testing. Given the lack of consensus on which clinical signs are most important from the dermatologist’s perspective, the aim of this study is to investigate which visual variables provide the best measure of atopic eczema severity from the patient’s perspective, around which management ultimately revolves.

**METHODS**

**PATIENTS**

A total of 180 patients (age range, 1–67 years; median age, 11 years; 51% female) with atopic eczema were recruited for the study. Of these, 110 patients were recruited consecutively from the dermatology outpatient department at the University Hospital in Nottingham. Another 70 patients were recruited from 5 local general practices by writing to all patients listed with a diagnosis of atopic eczema and arranging local interviews in the general practice at a convenient time for the patients. The ethnicity of the 180 patients was as follows: 153 (85.0%) were white, 15 (8.3%) were Asian, and 12 (6.7%) were Afro-Caribbean. All the patients satisfied the UK Working Party’s refinement of the diagnostic criteria of Hanifin and Rajka for atopic eczema. The patients were stratified into a group of children younger than 16 years (n=76) and a group of children older than 16 years (n=68). The Children’s Dermatology Life Quality Index (CDLQI) for children aged 4 to 16 years (n=76). Itch is a dominant symptom in atopic eczema, and it was assessed on a visual analog scale ranging from 0 to 10 during the previous 3 days, per the SCORAD protocol. For young children, information about disease severity, itch, and quality of life was obtained from the child and the parents depending on the child’s age and understanding.

**ASSESSMENTS**

All the patients were examined by a single dermatologist (C.R.C.). For practical purposes, 7 clinical signs were recorded: erythema, edema/papulation, oozing/crusting, excoriations, lichenification, dryness, and cracking. We selected these clinical signs because they are measured in 3 of the most widely used and validated atopic eczema scoring systems currently available: the SCORAD index (erythema, edema/papulation, oozing/crusting, excoriations, lichenification, dryness), the Eczema Area and Severity Index (erythema, edema/papulation, excoriations, and lichenification), and the Six-Area, Six-Sign Atopic Dermatitis severity score (erythema, excoriations, oozing/crusting, lichenification, dryness, and cracking). The average intensity of each clinical sign was graded on a scale from 0 to 3 (0=absent, 1=mild, 2=moderate, and 3=severe) at a representative body site, per the SCORAD protocol. Dryness was assessed in noninflamed skin. Disease extent was measured using the rule of nines and was drawn directly on the SCORAD evaluation sheet for calculation, per the SCORAD protocol. The objective SCORAD index (possible score of 0–83) was calculated as $A/5 + B/2$, where $A$ is the percentage disease extent (0–100) and $B$ is the sum of 6 clinical signs (possible score of 0–18).

To assess current atopic eczema severity from the patient’s perspective, a variety of questions were pilot tested before the study. The question “How much bother has your eczema caused over the last 3 days?” (scale of 0–10) was found to be more acceptable and meaningful to patients than asking them to directly rate their disease severity. The bother caused by the eczema provides a good reflection of the overall impact of the disease and patients’ requirements for health care intervention. The term bother has previously been used successfully in the development of symptom-based outcome measures for eczema and asthma, and it was easily understood by all patients when pilot tested before the study. Alternative phrases, such as “How much trouble has your eczema caused?” may have provided similar information, but such terms have not, to our knowledge, been used previously in published outcome measure studies. Three days is the period for subjective assessment used in the SCORAD index, and it was used to minimize fluctuations in the visual variables under assessment. To provide construct validity for the measurement of bother, quality of life and itch were also measured. Quality of life was assessed using the Dermatology Life Quality Index (DLQI) for patients older than 16 years (n=68) and the Children’s Dermatology Life Quality Index (CDLQI) for children aged 4 to 16 years (n=76). Itch is a dominant symptom in atopic eczema, and it was assessed on a visual analog scale ranging from 0 to 10 during the previous 3 days, per the SCORAD protocol. For young children, information about disease severity, itch, and quality of life was obtained from the child and the parents depending on the child’s age and understanding.

**STATISTICAL ANALYSIS**

Statistical analysis was performed using a software program (SPSS for Windows version 11.0; SPSS Inc, Chicago, IL). The mean patient-rated disease severity score (on a scale from 0–10) was computed for each level of the clinical sign scores (grades 0–3) and was compared using 1-way analysis of variance. A test for trend across the levels was performed to assess whether each relation was linear. A scatterplot was used to assess the univariate relation between disease extent and patient-rated disease severity, and a correlation coefficient ($r$) was computed. Multiple linear regression analyses were performed to determine how much variability in patient-rated disease severity each variable explained and to determine which variables were independent predictors of patient-rated disease severity. Variables that did not display a linear relation with patient-rated disease severity were fitted as categorical factors, and coefficients were obtained for each level of the factor relative to the lowest. Residual plots were produced to check the model assumptions held. Correlation coefficients (Spearman or Pearson $r$) were calculated for patient-rated disease severity against total objective scores, quality of life, and itch.

**RESULTS**

**CLINICAL SIGNS**

The distribution of patient-rated disease severity scores is shown in Figure 1. All 7 clinical signs were significantly associated with current disease severity from the patient’s perspective. Patient-rated disease severity increased in a linear manner with increasing severity of ery-
thema, edema/papulation, oozing/crusting, excoriations, lichenification, and cracking (Figure 2). Dryness was the only clinical sign that did not display a linear trend, with patient-rated disease severity increasing from dryness grade 0 to grade 2 but decreasing slightly in patients with grade 3 dryness.

The results of univariate regression analysis given in Table 1 demonstrate how much variability in patient-rated disease severity can be explained by each clinical sign separately. Individually, each clinical sign accounts for only a small proportion of the overall variation (adjusted $R^2$). Measurements of excoriations provided the best predictor of current disease severity from the patient’s perspective, accounting for almost a fifth (19.0%) of the variation in patient-rated disease severity scores. For every 1-point increase in erythema and excoriations, patient-rated disease severity increased on average by 1.5 and 1.25, respectively (effect estimate).

In a multiple linear regression model, all 7 clinical signs together predicted only a quarter of the variation in patient-rated disease severity (adjusted $R^2=0.25$). Three variables were identified as independent predictors of patient-rated disease severity with similar effect estimates after adjustment for age and sex.

DISEASE EXTENT

Patient-rated disease severity was positively correlated with disease extent, although Figure 3 suggests a nonlinear relation (Spearman $r=0.44$, $P<.001$). To investigate this correlation further, we categorized disease extent into 5 groups representing 0% to 10% ($n=56$), 11% to 20% ($n=69$), 21% to 30% ($n=22$), 31% to 50% ($n=17$), and 51%
or greater (n=16). Comparison of the means indicates a greater increase in patient-rated disease severity with disease extent across the lower values (0%-30% area involvement) than across the higher values (means, 3.64, 4.59, 6.05, 6.88, and 6.06, respectively; P<.001 for between-group effect). Of the patients with 20% or less skin involvement, almost a quarter (29/125:23.2%) rated the bother caused by their eczema as 7 or greater (on a scale from 0-10). The median age of these 29 patients with 0% to 20% skin involvement was 8.6 years (range, 2.6-51.8 years; 62% female), and the main body sites involved were the hands or feet (12/29:41%), the limbs (11/29:38%), and the head and neck (6/29:21%). In contrast, the hands or feet and head and neck were the main site of involvement in only 16% and 14% of the total sample, respectively.

Disease extent on its own explained 14.0% of the variability in patient-rated disease severity (adjusted R²=0.14; P<.001). However, when included in the multiple linear regression model with the clinical signs, disease extent added no statistically significant extra information about patient-rated disease severity vs measuring excoriations, edema/papulation, and erythema alone (adjusted R²=0.26 vs 0.25).

OBJECTIVE SCORAD INDEX

The sum of scores for excoriations, erythema, and edema/papulation (possible score of 0-9) was at least as closely correlated with patient-rated disease severity as the objective component of the SCORAD index (excoriations, erythema, edema/papulation, oozing/crusting, lichenification, dryness, and disease extent: possible score of 0-83) (Pearson r=0.51 and 0.49 and adjusted R²=0.25 and 0.23, respectively; P<.001).

QUALITY-OF-LIFE AND ITCH SCORES

Correlation between patient-rated disease severity (possible score of 0-10) and quality-of-life scores (possible score of 0-30) was high (Spearman r=0.65, n=86 for the CDLQI and Spearman r=0.64, n=54 for the DLQI; P<.001 for both). Similarly, correlation between patient-rated disease severity and itch (possible score of 0-10) was high (Spearman r=0.86, n=200; P<.001). The sum of scores for excoriations, erythema, and edema/papulation correlated as highly with the CDLQI as the more complex objective SCORAD (Spearman r=0.65 and 0.69, respectively; P<.001 for both), although correlation coefficients against the DLQI did not reach statistical significance (Spearman r=0.20 and 0.11; P=.16 and .42, respectively). Similarly, the sum of scores for excoriations, erythema, and edema/papulation correlated as highly with itch as the more complex objective SCORAD (Spearman r=0.48 and 0.49, respectively; P<.001 for both).

THE THREE ITEM SEVERITY SCORE

Three clinical signs provide the best measure of current disease severity from the patient’s perspective: excoriations, erythema, and edema/papulation. Excoriations provide a visual reflection of pruritus, and erythema and edema/papulation reflect acute inflammation. A simple scoring system comprising a measurement of these 3 variables (total score, 0-9) has previously been described in the literature as a simplified version of the SCORAD index called the Three Item Severity (TIS) score, although the 3 variables were originally chosen on clinical grounds rather than on patient assessment. Other clinical signs in this study included cracking and oozing/crusting, both of which were associated with high severity scores. However, severe cracking was uncommon in our patient population, and this clinical sign added no significant extra information about current disease severity. Similarly, oozing/crusting was sufficiently related to other acute variables in the regression analysis to make its measurement redundant. Lichenification is one of the more difficult clinical signs to grade accurately, but it is often included in clinical scoring systems to provide a measure of disease chronicity. How-

Figure 3. Relationship between disease extent and patient-rated disease severity scores (r=0.44).
however, lichenification can take several weeks to resolve, even after successful treatment has brought the acute inflammation and symptoms under control, and, therefore, does not provide an accurate measure of current disease activity, as demonstrated in this study. Dryness alone is associated with less bother than other variables and does not show a linear relationship with patient-rated disease severity, possibly reflecting increased emollient application in patients with the most severe symptoms or decreased dryness in acute exudative disease. Given that the median randomized controlled clinical trial duration in recent years was only 6 weeks,² the inclusion of these long-term variables to assess short-term changes in disease severity can make interpretation of results difficult and could potentially dilute any patient benefit detected.

DO WE NEED TO MEASURE DISEASE EXTENT?

Patient-rated disease severity increases as disease extent increases, but the relationship is nonlinear. A threshold effect is observed whereby mean severity scores show little increase above approximately 30% body surface area involvement. Whether the high severity scores seen in patients with limited disease reflect the limits of those patients’ experience is unclear from this cross-sectional study. Nevertheless, patients’ perceptions of their disease severity ultimately dictate their demand for treatment, regardless of their disease severity relative to other patients. Disease extent and intensity were highly correlated in this study, and from the patient’s perspective, disease extent measurements added no significant information to that obtained by measuring the 3 clinical signs mentioned herein because the latter reflected disease severity more closely. From a measurement perspective, accurate assessments of percentage of skin involvement are time-consuming and are subject to wide interobserver variation.⁴⁻¹² More important, although percentage of skin involvement may provide a tidy quantitative figure for analysis, the human body is not a uniform structure but consists of body sites of varying size and functional importance to patients. In clinical practice, the location of skin involvement is often more important to patients than is percentage of involvement per se, with involvement of small but functionally and cosmetically important areas, such as the hands, feet, and face, frequently being associated with significant morbidity. In this study, patients with limited disease and high patient-rated severity scores were more likely to have hand or foot eczema as the main site of involvement compared with the total sample of patients, reflecting the functional importance of these sites, particularly in young children.

IMPLICATIONS FOR MEASURING ATOPIC ECZEMA SEVERITY IN CLINICAL TRIALS AND DAILY PRACTICE

Although patient-based assessments, such as quality-of-life measurements,¹³,¹⁴ patient global assessments,⁷ and the patient-oriented eczema measure (POEM), provide the most pragmatic measure of atopic eczema severity, objective measures are likely to remain widely used because they can provide useful information when comparing patients across different cultural backgrounds, which may affect subjective reporting. In the clinical trial setting, researchers may want to use a validated scoring system, such as the SCORAD index or the Eczema Area and Severity Index, to provide detailed information about the disease process and to capture small changes in visible disease severity. However, additional information about visible skin changes does not necessarily translate into additional information about the illness experienced by the patient. When using such complex scoring systems it is often difficult to interpret whether a small change, such as 8 of a possible 103 points, is of real meaning to the patient rather than of academic interest only, or even simply a reflection of observer variation. It is also important to recognize that scoring systems that incorporate measurements of disease extent and skin dryness do not reflect patient-rated disease severity in a linear manner. Therefore, we suggest that researchers who use these composite scoring systems present the data for changes in individual variables (clinical signs and disease extent) alongside the total score to enable more meaningful interpretation by practicing physicians.

Depending on the study design, researchers not wanting to use a complex scoring system can record as much clinically meaningful information from the patient’s perspective by using the TIS score, the validity of which has been confirmed in this study. Because the TIS score is derived from elements of the SCORAD index, it is not surprising that it has showed high correlation with the SCORAD index in the original description of the index (r = 0.86)¹⁵ and in the present study (r = 0.82). The 3 clinical signs are graded as in the SCORAD index, and standardization of scoring can be improved by using the definitions and photographs in the SCORAD protocol (available at http://adserver.sante.univ-nantes.fr).³ Reliability data for the 3 clinical signs in the TIS score have been documented previously during validation of the SCORAD index and vary among studies, although the measurement of edema and papulation is most likely to vary between or within observers.¹³,¹⁵,¹⁶ In the original description of the TIS score, the total score showed fair interobserver agreement (k = 0.58).¹³ The 3 clinical signs have shown sensitivity to change in clinical trials using the SCORAD index,¹⁰ although further trials are needed to confirm the sensitivity of the TIS score in different patient populations. The TIS score was originally described for routine clinical practice and for prescreening for clinical trials, but it has been used successfully as a primary outcome measure in a recent randomized, double-blind, parallel-group study,¹⁸ of topical corticosteroid use, using a score of 4 or higher (of a possible 9) to indicate a relapse and 1 or lower to indicate remission. In addition to its potential use for screening and monitoring in clinical trials, the simplicity of the TIS score makes it suitable for use in the outpatient clinic or general practice, particularly for clinical audit.

The limitations of all forms of visual scoring systems should be recognized. Visual assessments reflect a single snapshot in time and do not comprehensively capture the chronic, fluctuating nature of the atopic eczema. Furthermore, accurate grading of clinical signs and disease extent can be time-consuming and subject to wide interobserver
variation, which could potentially overshadow small changes in disease severity.16,17 The assessment of erythema can be particularly difficult in black skin, despite being one of the most widely measured clinical signs.19

**CONCLUSIONS**

Objective clinical scoring systems have a defined but limited role in the assessment and monitoring of disease severity in children and adults with atopic eczema. From the patient’s perspective, the current array of visual measurements is not necessary to provide a clinically meaningful reflection of disease severity, and visual scoring can be simplified without losing important information. From the physician’s perspective, standardization and simplification of scoring will allow improved interpretation and comparison of disease outcomes. In a disease dominated by symptoms, visual assessments should be used to supplement but not replace patient-based symptom measures20,21 and quality-of-life scores,13,14 which have remained underused in recent years and can ultimately provide more valuable information about day-to-day disease severity from the patient’s perspective, around which management revolves.

**Accepted for Publication:** March 22, 2005.  
**Correspondence:** Carolyn R. Charman, BM, BCh, MRCP, Department of Dermatology, Queen’s Medical Centre, Nottingham NG7 2UH, England (carolyn.charman@nottingham.ac.uk).

**Author Contributions:** Study concept and design: Charman and Williams. Acquisition of data: Charman. Analysis and interpretation of data: Charman, Venn, and Williams. Drafting of the manuscript: Charman. Critical revision of the manuscript for important intellectual content: Charman, Venn, and Williams. Statistical analysis: Charman and Venn. Obtained funding: Charman and Williams. Administrative, technical, and material support: Charman. Study supervision: Williams.

**Financial Disclosure:** None.

**Funding/Support:** This work was supported by a Health Services Research Training Fellowship funded by the former National Health Service Research and Development Board, Trent Region, Sheffield, England.

**Role of the Sponsor:** The sponsor had no role in the design and conduct of the study; in the collection, analysis, and interpretation of data; or in the preparation, review, or approval of the manuscript.

**Disclaimer:** Dr Charman had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Acknowledgment:** We are indebted to the patients and the outpatient reception staff who helped with this study and to the Trent Focus Group and general practices for their time and administrative assistance.

---

the risk-benefit ratio. This case illustrates the benefits of thalidomide therapy for patients with NL, and further studies are needed to understand the full potential of this unique drug.

Accepted for Publication: April 28, 2005.
Correspondence: Tarun Kukreja, MD, Center for Dermatological and Cosmetic Surgery, Washington University, 969 Mason Rd, Suite 200, St Louis, MO 63141 (tkukreja@gmail.com).
Financial Disclosure: None.

REFERENCES