ORIGINAL ARTICLE

Comparison of three methods for measuring psoriasis severity in clinical studies (Part 1 of 2): change during therapy in Psoriasis Area and Severity Index, Static Physician's Global Assessment and Lattice System Physician's Global Assessment

C. Chow,[1,†] M.J. Simpson,[1,2,†] T.A. Luger,[3] H. Chubb,[1] C.N. Ellis[1,*]

[1]Department of Dermatology, University of Michigan Medical School, Ann Arbor, Michigan, USA
[2]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, Michigan, USA
[3]Department of Dermatology, University of Münster, Münster, Germany
*Correspondence: C.N. Ellis. E-mail: cellis@umich.edu

Abstract

Background Accurate and reliable assessment of changes in psoriasis severity is critical in clinical trials of therapies.
Objective To compare Psoriasis Area and Severity Index (PASI), static Physician's Global Assessment (sPGA), and the Lattice System Physician's Global Assessment (LS-PGA) in a trial of systemic treatments for plaque psoriasis vulgaris and to assess whether they measure change in psoriasis induced by therapy.
Methods Patients were randomized to voclosporin or cyclosporine for 24 weeks (the '24-week-treatment' group, $n = 366$), or placebo for 12 weeks followed by voclosporin for 12 weeks (the 'initial-placebo' group, $n = 89$).
Results All scoring systems changed in concert and were sensitive enough to detect reductions in severity during placebo therapy as well as with active therapy ($P < 0.01$ for each measurement). At study onset, there were poorer correlations of sPGA with PASI ($r = 0.45$) and LS-PGA ($r = 0.39$) than between PASI and LS-PGA ($r = 0.68$). After therapy, all correlations were stronger, but sPGA continued to be less well correlated (with PASI, $r = 0.85$; with LS-PGA, $r = 0.79$) than LS-PGA with PASI ($r = 0.90$). Two- or three-step improvements in LS-PGA showed very good to excellent accuracy in corresponding to PASI-50 and PASI-75, respectively, and were more accurate than comparable changes in sPGA.
Conclusion PASI, sPGA and LS-PGA are responsive to the varying degrees of improvement in psoriasis induced by either placebo or active therapy. While the three systems capture similar information, each has different reasons for use in a clinical trial.
Received: 11 September 2014; Accepted: 10 March 2015

Introduction

In clinical trials, valid and repeatable measures of disease severity are essential in assessing efficacy of investigational treatments for regulatory approval and to compare results among experimental studies of new therapies.[1] Few studies have examined the validity and reliability of these measures, challenging our ability to make comparisons among different clinical trials.[1–9]

An ideal measurement system to assess disease severity in psoriasis should be objective, reproducible, easy to use and

†Both authors contributed equally to this work

clinically applicable.[1,4,10] Such a validated scoring system, if unchanging and widely adopted, would help compare studies and therapies.

Psoriasis Area and Severity Index (PASI, Fig. 1) was invented for use in a single clinical trial[11] and became the most frequently used measurement of psoriasis severity in clinical trials without any validation.[1,2] Different versions of PASI have been used, and errors in the formula have surfaced.[12] PASI is criticized for being time-consuming, complicated, lacking sensitivity and difficult to interpret.[13–15] Specific values of PASI (range, 0–72 units) have been translated into word-based severity terms;[16,17] however, there is no consensus on interpreting a specific value of PASI.

A static Physician's Global Assessment (sPGA) has many variations, including 5, 6, or 7-point scoring ranging from 'clear' to 'severe'.[14,18] The extent to which body surface area is involved (iBSA) is not considered in PGA scoring (Fig. 1).

The Lattice System Physician's Global Assessment (LS-PGA) was created in response to limitations in other measures.[19,20] Unlike most PGAs, it uses iBSA. The iBSA in the LS-PGA is segmented into ranges that, for ease of physician estimation, get larger as the amount of involvement increases. At low involvement, the patient's 'handprint' (representing one per cent of the patient's surface area) is used[21]; at higher ranges, intuitive anchor points are utilized, such as about a quarter of the body surface involved, about a third to a half and over a half (Fig. 1). Plaque elevation, erythema and scale severity are specifically defined. The scores for iBSA and plaque characteristics are weighted and combined in a computerized lattice algorithm[9] to determine the LS-PGA overall severity of psoriasis. The LS-PGA formula of weighting the clinical components of psoriasis has been validated in this clinical trial (see Ref. 22). The LS-PGA has one current version and its use is controlled to prevent changes

PASI

To obtain PASI in this clinical trial, investigators separately rated 4 regions of the body (head [h], trunk [t], and upper [u] and lower [l] extremities, not otherwise explained) for erythema (E), infiltration (I), desquamation (D), and body surface area involved with psoriasis (A).

Degree of severity in each body region	Value given
None	0
Slight	1
Moderate	2
Marked	3
Very marked	4

Surface area involved for each body region	Value given
No involvement	0
<10%	1
10-29%	2
30-49%	3
50-69%	4
70-89%	5
90-100%	6

The head, upper extremities, trunk, and lower extremities are assigned 10, 20, 30, and 40 percent, respectively, of the total body surface area. The final PASI calculation is the sum of:

$$0.1(E_h + I_h + D_h)A_h + 0.2(E_u + I_u + D_u)A_u + 0.3(E_t + I_t + D_t)A_t + 0.4(E_l + I_l + D_l)A_l.$$

The PASI ranges from 0 to 72 unlabeled units.

sPGA

To obtain the sPGA chosen for this clinical trial, investigators began with the interim steps of rating psoriasis plaque attributes of induration, erythema, and scaling averaged over the patient's entire body as shown.

	[a]Induration (averaged over all lesions)	[b]Erythema (averaged over all lesions)	[c]Scaling (averaged over all lesions)
0	No evidence of plaque elevation	No evidence of erythema; hyperpigmentation may be present	No evidence of scaling
1	Minimal plaque elevation (≈ 0.5 mm)	Faint erythema	Minimal, occasional fine scale over less than 5% of the lesions
2	Mild plaque elevation (≈ 1 mm)	Light red coloration	Mild, fine scale predominates
3	Moderate plaque elevation (≈ 1.5 mm)	Moderate red coloration	Moderate, coarse scale predominates
4	Marked plaque elevation (≈ 2 mm)	Bright red coloration	Marked, thick, non-tenacious scale predominates
5	Severe plaque elevation (≈ 2.5 mm or more)	Dusky to deep red coloration	Severe, very thick tenacious scale predominates

The investigator then decided upon the final sPGA based on the numbered choices shown on the next page.

Figure 1 Methods of evaluation of psoriasis severity. The choices made by the investigator, as would appear on a Case Report Form for each evaluation system for each patient in an investigational trial, are shown for Psoriasis Area and Severity Index (PASI), static Physician's Global Assessment (sPGA), and the Lattice System Physician's Global Assessment (LS-PGA).

Static Physician's Global Assessment Score (averaged over all lesions):
0 = clear, except for residual discoloration
1 = almost-clear, lesions have individual scores for induration[a] erythema[b] and scaling[c] (IES) of at least 1
2 = Lesions have individual scores for IES of at least 2
3 = Lesions have individual scores for IES of at least 3
4 = Lesions have individual scores for IES of at least 4
5 = Lesions have individual scores for IES of at least 5

Note: When required for comparison purposes, we converted the sPGA numerical range of 0 to 5 to clear = 0, almost-clear = 1, mild = 2, moderate = 3, marked = 4, and severe = 5.

LS-PGA

The investigator selects one score in each row that best matches the patient overall, as instructed.

1. INDICATE PERCENT BODY SURFACE INVOLVED (do not include areas with <u>only</u> post-inflammatory pigmentation). *The patient's handprint (palm with fingers and thumb held together) approximate a body surface area of 1%.*

❑ 0%   ❑ 1-3%   ❑ 4-9%   ❑ 10-20%   ❑ 21-29%   ❑ 30-50%   ❑ 51-100%

2. CHECK <u>ONE</u> *AND ONLY ONE* BOX IN **EACH** <u>ROW</u> BELOW. INDICATE AVERAGE QUALITIES ACROSS <u>ALL</u> INVOLVED AREAS.

Plaque Quality	NONE	MILD	MODERATE	MARKED
ELEVATION	❑ No elevation above normal skin *(Ignore scale in determining plaque elevation)*	❑ Slight elevation above surrounding normal skin; the edges are typically barely palpable and not seen *(Ignore scale in determining plaque elevation)*	❑ Visually apparent elevation in plaque edges; readily palpated edges with rounded or sloped edges *(Ignore scale in determining plaque elevation)*	❑ Visually obvious elevation in plaque edges; easily palpated with sharp edges that provide a discernible bump when sliding the finger from normal skin to the plaque *(Ignore scale in determining plaque elevation)*
ERYTHEMA	❑ Normal skin color or post-inflammatory color change; no erythema	❑ Pink color, including "residual hyperemia" ("end of therapy hyperemia")	❑ Light to medium red color	❑ Bright, full, or deep red color
SCALE	❑ No scales	❑ Fine scales looking like dust	❑ Individual scales discernible with yellow to silver color	❑ Coarse, thick scales, yellow to silver in color; plaque has rough surface; gives impression that scales could be lifted with fingernail

The selected ratings are used in an algorithm (Langley and Ellis, 2004) to obtain the overall LS-PGA severity score of clear, almost-clear, mild, mild to moderate, moderate, moderate to severe, severe, or very severe. As needed for purposes of analysis, we converted the LS-PGA ratings to numeric scores of 0 to 7, respectively. (LS-PGA Research Version 6.0, © and U.S. Pat. 7,955,260, DATAcquire, Inc. [info@datacquire.com]. All rights reserved. May not be reproduced in whole or part without permission; used here with permission.)

Figure 1 (*Continued*)

across different studies and their sponsors, thus allowing comparisons among clinical trials.

Although the LS-PGA was first described and tested for consistency in 2004[9], whether it reflects patients' responses to therapy was questioned.[23,24] We hypothesized that the LS-PGA would demonstrate change during treatment of psoriasis in a manner similar to PASI and sPGA. Proof of the sensitivity of the LS-PGA to change during therapy required the cooperation of a sponsor of a large clinical trial of an effective therapy. These data, now available for the first time, show that the LS-PGA demonstrates change in psoriasis severity during a clinical trial in concert with changes shown by PASI and sPGA. Another report[22] further assesses validity of all three psoriasis evaluation systems.

Patients, materials and methods

Study design

Data came from a phase III, randomized, multicenter, double-blind, placebo-controlled study performed in Canada, Germany and Poland that evaluated the efficacy of voclosporin and cyclosporine in the treatment of psoriasis. Men and women age ≥18 with chronic plaque psoriasis for at least 6 months and iBSA ≥10% were randomized 3 : 1 : 1 to divided-dose voclosporin

0.8 mg/kg/day, or cyclosporine 3.0 mg/kg/day, or placebo respectively. Patients in the voclosporin and cyclosporine arms were treated for 24 weeks; for purposes of this analysis, these patients were combined into a '24-week-treatment' group. Those in the 'initial-placebo' group received 12 weeks of placebo; by protocol, all then received 12 weeks of voclosporin. Patients with complete data were analysed.

This study (ClinicalTrials.gov NCT00408187) was approved by regulatory agencies in each country and by an Ethics Committee at each site. Study investigators adhered to the Guideline for Good Clinical Practice, which is based on the Declaration of Helsinki.[25]

Scoring of psoriasis

Severity was determined at weeks 0, 12 and 24 using PASI, sPGA and the LS-PGA (DATAcquire, Inc., Ann Arbor, MI, USA, version 6.0). Whenever possible, the same investigator at each site evaluated and scored the patients' severity of psoriasis through the trial using all three measurement systems. For sPGA, an overall score of 0 was *a priori* defined as 'clear' and 1 as 'almost-clear'; we labelled the final sPGA scores of 2, 3, 4 and 5 (Fig. 1) as 'mild', 'moderate', 'marked' and 'severe' respectively. We assigned LS-PGA scores of 0 'clear' through 7 'very severe'.

Statistical analysis

A *post hoc* analysis of the data that is being used in this study found powers in our various statistical tests above 80% to detect differences among the methods for measuring psoriasis. To compare demographical values, we used independent *t*-tests for continuous variables and Chi-square or Fisher's exact test for categorical variables. Paired *t*-tests were used to compare mean psoriasis severity over time by treatment assignment and assessment method, and independent *t*-tests were used to compare results between the two treatment assignments. For sPGA and LS-PGA, we calculated the prevalence ratio of subjects at better than mild and almost clear at weeks 12 and 24. We calculated pairwise Spearman's rank-correlation coefficients for all scoring systems at all three timepoints.

At weeks 12 and 24, patients who were 'responders' by PASI-50 were compared to 'responders' by a two-category reduction in sPGA or LS-PGA using Chi-square test. Similarly, Receiver Operating Characteristic (ROC) analysis and 95% confidence intervals (CIs) were used to compare accuracy of the 2-step and 3-step changes in sPGA and LS-PGA at weeks 12 and 24 using PASI-50 and PASI-75 respectively as reference standards. ROC areas under the curve (AUCs) range from 0.5–1.0; we use the ROC assessment of accuracy of 0.90–1.00 to be excellent; 0.80–0.89, very good; 0.70–0.79, good, 0.60–0.69, fair and 0.50–0.59, poor. We chose 2-step changes in LS-PGA and sPGA during treatment as our minimal clinically important difference (MCID) because 1-step changes in severity can occur with little real change in psoriasis when patient severity is near the border

between any two scores at pretherapy. For PASI, PASI-50 is regarded as the MCID for a useful therapy. Analyses used SAS version 9.2.1 (SAS Institute, Cary, NC, USA) and IBM SPSS Statistics for Windows version 22.0 (International Business Machines Corporation, Armonk, NY, USA).

Results

Demographics

There were 455 patients who had complete data; 366 patients were assigned to the 24-week-treatment group and 89 patients were assigned to the initial-placebo group. Both groups had similar initial characteristics (Table 1).

Disease severity

All scoring measures were sensitive enough to detect the effect caused by placebo therapy along with participating in a clinical trial; as shown in Table 2, the initial-placebo group improved by each scoring system after 12 weeks ($P < 0.01$ for each). All measures at 12 weeks showed that the 24-week-treatment group also improved from pretherapy ($P < 0.01$ for each) and that the 24-week-treatment group showed greater improvement than the initial-placebo group (Table 2). At week 24, after the initial-placebo group had received active treatment for 12 weeks, all severity scores showed further improvement and indicated no statistical difference in severity between the two assignment groups (Table 2).

At week 12, both sPGA and LS-PGA indicated that the 24-week-treatment group had a larger proportion of subjects who achieved a psoriasis score of mild or better (i.e. mild, almost-clear, or clear) and almost-clear or better (i.e. almost-clear or clear) compared to the initial-placebo group ($P < 0.0001$ for each, Table 3). At week 24, the LS-PGA indicated that the initial-placebo group had nearly caught up to the 24-week-treatment group and that the initial-placebo group and 24-week-treatment group were no longer statistically different (Table 2). However, the sPGA and PASI suggested that the initial-placebo group had surpassed the 24-week-treatment group overall (Table 2) and in achieving scores of mild or better ($P < 0.05$ for sPGA, Table 3); the reason for this is unknown.

Correlation of scoring systems

In comparing the relationships of the three measures of psoriasis severity to each other, PASI and LS-PGA scores were the most highly correlated in both the 24-week-treatment group and the initial-placebo group at all timepoints (Table 4). sPGA had poorer correlations with PASI and LS-PGA, especially at pretherapy.

Change from pretherapy

Without regard to assignment to 24-week-treatment or initial-placebo, 299 subjects could be called responders at week 12 by

Table 1 Demographics and pretherapy psoriasis severity of study participants, stratified by treatment assignment. There were no significant differences in pretherapy characteristics between the 24-week-treatment group (patients who received active therapy throughout the study) and initial-placebo group (patients who received placebo for 12 weeks and active therapy for 12 weeks). Psoriasis Area and Severity Index has a range of 0 to 72

	Total (*n* = 455)	24-Week-Treatment (*n* = 366)	Initial-Placebo (*n* = 89)
Sex, *n* (%)			
Female	142 (31)	116 (32)	26 (29)
Male	313 (69)	250 (68)	63 (71)
Race/ethnicity, *n* (%)			
Caucasian	437 (96)	353 (97)	84 (94)
Asian	13 (3)	9 (3)	4 (5)
Black	1 (<1)	0 (0)	1 (1)
Hispanic	1 (<1)	1 (<1)	0 (0)
Native American	1 (<1)	1 (<1)	0 (0)
Other	2 (<1)	2 (<1)	0 (0)
Age, mean (SD)			
	43 (13)	43 (13)	45 (13)
Psoriasis Area and Severity Index, mean (SD)			
	18.4 (7.0)	18.5 (7.1)	17.9 (6.4)
Static Physician's Global Assessment, *n* (%)			
0	0 (0)	0 (0)	0 (0)
1	0 (0)	0 (0)	0 (0)
2	3 (<1)	3 (<1)	0 (0)
3	260 (57)	205 (56)	55 (62)
4	174 (38)	140 (38)	34 (38)
5	18 (4)	18 (5)	0 (0)
Lattice System Physician's Global Assessment, *n* (%)			
Clear	0 (0)	0 (0)	0 (0)
Almost-clear	0 (0)	0 (0)	0 (0)
Mild	0 (0)	0 (0)	0 (0)
Mild-to-moderate	0 (0)	0 (0)	0 (0)
Moderate	13 (3)	10 (3)	3 (3)
Moderate-to-severe	125 (28)	99 (27)	26 (29)
Severe	199 (44)	164 (45)	35 (39)
Very severe	118 (26)	93 (25)	25 (28)

SD, standard deviation.

achieving PASI-50. A 2-step improvement in LS-PGA was more likely to detect these PASI-50 responders than was a 2-step improvement in sPGA (256 [86%] vs. 203 [68%], respectively, *P* < 0.0001). This pattern persisted at week 24.

Using ROC analysis, a 2-step improvement at 12 weeks in LS-PGA was more likely to detect PASI-50 responders than was a 2-step improvement in sPGA (AUC 0.88 [95% CI, 0.84–0.91] vs. AUC 0.81 [95% CI, 0.77–0.85] respectively). After 24 weeks, the results were similar but closer to each other (AUC 0.82 [95% CI, 0.76–0.87] vs. AUC 0.79 [95% CI, 0.74–0.84] respectively). Similar results were found when 3-step changes were compared to PASI-75; at 12 weeks, the AUCs were 0.86 (0.82–0.91) for LS-PGA vs. 0.71 (0.66–0.76) for sPGA; at 24 weeks, 0.80 (0.76–0.85) for LS-PGA vs. 0.70 (0.65–0.75) for sPGA.

Discussion

This is the first report directly comparing PASI, sPGA and LS-PGA in a clinical trial. Previously, all three scoring systems correlated highly with one another when dermatologists evaluated patients with psoriasis on a single day.[9,14,26,27] With this analysis, we have now proved that PASI, sPGA and LS-PGA demonstrate change with active therapy, distinguish active from placebo therapy and track in parallel and correlate well overall (Tables 1–4) over the course of a typical psoriasis clinical trial. All three systems detected a wide range of change, including the modest change that occurs with placebo[28,29] and the greater effect of oral calcineurin therapy (Table 2). As expected, most patients in this trial who achieved final sPGA and LS-PGA scores of almost-clear or clear also achieved PASI-75 and PASI-90 (data not shown).

Table 2 Mean psoriasis severity by treatment assignment and assessment method. The initial-placebo group crossed over to active therapy at week 12. The *P* value was obtained from testing the null hypothesis that the mean psoriasis severity is equal between the 24-week-treatment and initial-placebo groups

	Pretherapy	Week 12	Week 24
Psoriasis Area and Severity Index			
24-Week-Treatment	18.5	6.3*	5.4†
Initial-Placebo	17.9	14.3*	5.1†
P value	0.48	<0.01	0.63
Static Physician's Global Assessment			
24-Week-Treatment	4.0	1.8*	1.7†
Initial-Placebo	3.5	2.9*	1.6†
P value	0.19	<0.01	0.35
Lattice System Physician's Global Assessment			
24-Week-Treatment	5.9	3.2*	2.8†
Initial-Placebo	5.9	5.0*	3.0†
P value	0.94	<0.01	0.38

*$P < 0.01$ comparing mean psoriasis severity at week 12 compared to pretherapy within each assignment group.

†$P < 0.01$ comparing mean psoriasis severity at week 24 compared to pretherapy within each assignment group.

The correlation between PASI and LS-PGA was better at all timepoints than was sPGA with either of the other assessment systems. While perfect correlations were not expected (which would indicate that all three systems yield identical results), a high correlation was expected because the systems use similar

aspects of psoriasis assessments. Because iBSA is not a specified component of sPGA (but is used in the algorithms to compute PASI and the LS-PGA), this may account for the poorer correlation of sPGA with the other measures.

Not using iBSA removes an important characteristic that often influences how severe psoriasis is perceived to be. Because iBSA in theory is not considered in disease severity in sPGA, a patient with one small plaque should be scored the same as a patient with numerous similar plaques. However, in practice, physicians likely consider iBSA when determining sPGA. Indeed, iBSA was statistically more closely associated with sPGA score than was plaque elevation, erythema, or scale.[9] As a result of physicians' apparent innate use of iBSA in determining sPGA scores despite being instructed not to do so, exclusion of iBSA in sPGA may increase variability across investigator sites and among clinical trials.

This study used a version of sPGA that defined many psoriasis plaque qualities and included instructions on combining the plaque qualities into the final score (Fig. 1). In many versions of sPGA, psoriasis is defined simply as mild, moderate, or severe; some are even more vague.[9] Because sPGA definitions differ among studies, and because investigators participating in more than one study may have difficulty in adjusting their thresholds of severity for each, one must be cautious in directly comparing sPGA across clinical trials.

Standardized evaluation systems that are accurate, consistent and clinically meaningful and that allow comparisons among studies and across therapies are desired.[4,7,30] However, different

Table 3 Prevalence of subjects at or below various thresholds of sPGA and LS-PGA at week 12 and 24. The initial-placebo group crossed over to active treatment after week 12. The *P* values were obtained from testing the null hypothesis that the proportion of subjects who achieved the various outcomes between the two assignment groups is equal

Threshold*	24-Week-Treatment (*n* = 366) *n* (%)	Initial-Placebo (*n* = 89) *n* (%)	Prevalence ratio estimate (95% CI) of 24-Week-Treatment vs. Initial-Placebo†	*P*
Week 12				
sPGA≤almost-clear	150 (41)	7 (8)	5.2 (2.5, 10.7)	<0.0001
sPGA≤mild	264 (72)	21 (24)	3.1 (2.1, 4.5)	<0.0001
LS-PGA≤almost-clear	100 (27)	3 (3)	8.1 (2.6, 25.0)	<0.0001
LS-PGA≤mild	176 (48)	11 (12)	3.9 (2.2, 6.8)	<0.0001
Week 24				
sPGA≤almost-clear	161 (44)	41 (46)	1.0 (0.7, 1.2)	0.72
sPGA≤mild	268 (73)	74 (83)	0.9 (0.8, 1.0)‡	0.05‡
LS-PGA≤almost-clear	113 (31)	22 (25)	1.2 (0.8, 1.9)	0.25
LS-PGA≤mild	208 (57)	49 (55)	1.0 (0.8, 1.3)	0.76

≤almost-clear, almost-clear or clear; ≤mild, mild, almost-clear, or clear; CI, confidence interval; LS-PGA, Lattice System Physician's Global Assessment; sPGA, Static Physician's Global Assessment.

*An sPGA score of 2 was assigned the term 'mild'.

†A prevalence ratio 95% CI that includes 1 indicates no statistical difference between the two assignment groups.

‡The upper limit of the CI is <1 and $P < 0.05$; however, they appear as shown due to rounding. Among the initial-placebo patients, there was a higher percentage who achieved sPGA≤mild after receiving 12 weeks of active therapy compared to the patients in the 24-week-treatment group. This may have occurred due to chance (in view of the CI approaching 1) or due to our arbitrary assignment of the rubric 'mild' to level 2 of the sPGA. This occurrence was not seen with use of the LS-PGA.

Table 4 Spearman's rank-correlation coefficients (*r*) between Psoriasis Area and Severity Index (PASI), Static Physician's Global Assessment (sPGA) and Lattice System Physician's Global Assessment (LS-PGA) at pretherapy, week 12 and week 24. The 24-week-treatment group is the patients who received active therapy throughout the study; the initial-placebo group is the patients who received placebo for 12 weeks followed by active therapy for 12 weeks

	PASI and sPGA		PASI and LS-PGA		LS-PGA and sPGA	
	24-Week-Treatment (*n* = 366)	Initial-Placebo (*n* = 89)	24-Week-Treatment (*n* = 366)	Initial-Placebo (*n* = 89)	24-Week-Treatment (*n* = 366)	Initial-Placebo (*n* = 89)
Pretherapy	0.45	0.23	0.68	0.63	0.39	0.26
Week 12	0.83	0.75	0.91	0.83	0.76	0.59
Week 24	0.85	0.72	0.90	0.86	0.79	0.58

P < 0.05 for all *r* values (testing the null hypothesis that *r* = 0).

versions of PASI have been used in various clinical trials, and PASI scoring lacks clear definitions. PASI has significant intrarater and interrater variation that is affected by the experience level of the evaluator[9] and can be difficult and time-consuming to do, especially with limited experience. In PASI, the iBSA must be determined separately for each segment of the body, a provision that may induce errors, particularly at small degrees of involvement when the handprint method[21] cannot be used because the patient's handprint covers 1% of the entire body, not 1% of the segment being scored.

Despite its wide use, the numerical PASI score has no intuitive meaning to patients or physicians[7,9,13,14,18,31–33] and lacks clinical relevance for many dermatologists, in part because it is rarely used in clinical practice, particularly in the USA. Although PASI is non-linear, treatment success for individual patients in clinical trials of investigational therapies for psoriasis has been PASI- 50 or PASI-75. This has been problematic, because some patients with improvements less than these cutoffs have clinically meaningful success or achieve satisfaction with their therapy, while others still have extensive psoriasis or lack satisfaction after achieving even PASI-75.[18,26,34–36]

For a subject to achieve treatment success in a clinical trial, the U.S. Food and Drug Administration (FDA) expects the patient's psoriasis severity to reach clear or almost-clear, a threshold that is not defined by PASI number or by a specific reduction of PASI score. The FDA and the European Medicines Agency (EMA) do not support the use of PASI alone to gauge psoriasis severity for approval of new therapies in phase III trials and suggest use of a validated, standardized scoring system that produces word scores; the EMA guidance specifically mentions the LS-PGA or an sPGA.[37]

sPGA scores have clinical context regarding severity and are typically quick and easy to perform. Unfortunately, multiple versions of sPGAs exist, most of which lack clear definitions; furthermore, sPGAs are designed not to take iBSA into account. A newer measure, the LS-PGA, is a standard method for evaluating psoriasis severity; provides clear definitions of severity of each component of psoriasis and intuitive estimations of iBSA; is quick and easy to perform by checking only four boxes

(Fig. 1); yields a clinically relevant score; and has reproducible results.[9,14]

To be used with confidence, a new system must be shown to detect the effect of treatment; in this large clinical trial, the LS-PGA proved sensitive to changes in disease severity during therapy in concert with PASI and an sPGA. Furthermore, a two- or three-category reduction in the LS-PGA was more often associated with PASI-50 or PASI-75 respectively than was sPGA at both evaluation timepoints, and the confidence range of accuracy for the LS-PGA was very good to excellent. Because two and three-category reductions in the LS-PGA so consistently reflect PASI-50 or PASI-75, such changes in LS-PGA may be used to assess clinically significant improvement in psoriasis.

This research studies the evaluation of psoriasis severity in clinical trials; it does not fully assess the efficacy of voclosporin or any other treatment; drug efficacy was used solely to test the scoring systems. Therefore, the group of patients treated with calcineurin inhibitors meets all requirements for this research. The trial sponsor did not identify to us which patients in the 24-week-treatment group received voclosporin and which received cyclosporine. Even with that identification, to evaluate psoriasis rating systems we would have performed the exact same combined group analyses that are presented in this report. Additional analysis of subgroups receiving the two similarly active calcineurin inhibitors would only result in smaller numbers in each group as compared to the combined group.

Our data indicate that all three severity measures used in this psoriasis clinical trial, namely the LS-PGA, PASI and sPGA, may be used to monitor change induced by placebo or active therapy. In theory, any one of the measurements would do, so how do designers of clinical trials select the evaluation systems to be used? PASI is often chosen for use in a new clinical trial to allow comparison to data from prior trials that reported PASI. However, in comparing various trials, the severity of patients' psoriasis before therapy may be important when attempting to compare patients' percentage changes (e.g. PASI-75) across studies because a 75 per cent reduction from a large extent of psoriasis may be a different result from the same per cent reduction if starting from a much lesser extent of involvement.

To meet regulatory requirements, a severity measure that provides a word-based result (e.g. severe, mild, almost-clear, etc.) is desired. A version of the sPGA typically is chosen, usually with the concurrence of the FDA or EMA. However, the selected sPGA may differ from previous studies, may be poorly defined, and usually does not incorporate extent of psoriasis. Alternatively, the well-defined LS-PGA includes measurement of psoriasis extent and may be used to provide a word-based result that can be compared across studies. However, it is not yet widely used and awaits greater familiarity. In practice, and as in the trial reported here, two or more severity systems are commonly employed to satisfy various interests in the clinical trial.

Acknowledgements

References

1 Spuls PI, Lecluse LL, Poulsen ML, Bos JD, Stern RS, Nijsten T. How good are clinical severity and outcome measures for psoriasis?: quantitative evaluation in a systematic review. *J Invest Dermatol* 2010; **130**: 933–943.

2 Naldi L, Svensson A, Diepgen T *et al.* Randomized clinical trials for psoriasis 1977-2000: the EDEN survey. *J Invest Dermatol* 2003; **120**: 738–741.

3 Marks R, Barton SP, Shuttleworth D, Finlay AY. Assessment of disease progress in psoriasis. *Arch Dermatol* 1989; **125**: 235–240.

4 Bigby M, Gadenne AS. Understanding and evaluating clinical trials. *J Am Acad Dermatol* 1996; **34**: 555–590.

5 McKenna KE, Stern RS. The outcomes movement and new measures of the severity of psoriasis. *J Am Acad Dermatol* 1996; **34**: 534–538.

6 van de Kerkhof PC. The Psoriasis Area and Severity Index and alternative approaches for the assessment of severity: persisting areas of confusion. *Br J Dermatol* 1997; **137**: 661–662.

7 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**: 185–191.

8 Chalmers RJ, Griffiths CE. Resetting the research agenda for psoriasis. *J Invest Dermatol* 2003; **120**: ix–x.

9 Langley RG, Ellis CN. Evaluating psoriasis with Psoriasis Area and Severity Index, Psoriasis Global Assessment, and Lattice System–Physician's Global Assessment. *J Am Acad Dermatol* 2004; **51**: 563–569.

10 Jemec GB, Wulf HC. The applicability of clinical scoring systems: SCORAD and PASI in psoriasis and atopic dermatitis. *Acta Derm Venereol* 1997; **77**: 392–393.

11 Fredriksson T, Pettersson U. Severe psoriasis—oral therapy with a new retinoid. *Dermatologica* 1978; **157**: 238–244.

12 Finlay AY, Gibson SL. Error in the original description of the psoriasis area and severity index. *Br J Dermatol* 1998; **139**: 534.

13 Harari M, Shani J, Hristakieva E, Stanimirovic A, Seidl W, Burdo A. Clinical evaluation of a more rapid and sensitive Psoriasis Assessment Severity Score (PASS), and its comparison with the classic method of Psoriasis Area and Severity Index (PASI), before and after climatotherapy at the Dead-Sea. *Int J Dermatol* 2000; **39**: 913–918.

14 Berth-Jones J, Grotzinger K, Rainville C *et al.* A study examining inter- and intrarater reliability of three scales for measuring severity of psoriasis: Psoriasis Area and Severity Index, Physician's Global Assessment and Lattice System–Physician's Global Assessment. *Br J Dermatol* 2006; **155**: 707–713.

15 Puig L. Shortcomings of PASI75 and practical calculation of PASI area component. *J Am Acad Dermatol* 2013; **68**: 180–181.

16 Schmitt J, Wozel G. The Psoriasis Area and Severity Index is the adequate criterion to define severity in chronic plaque-type psoriasis. *Dermatology* 2005; **210**: 194–199.

17 Feldman SR. A quantitative definition of severe psoriasis for use in clinical trials. *J Dermatolog Treat* 2004; **15**: 27–29.

18 Feldman SR, Krueger GG. Psoriasis assessment tools in clinical trials. *Ann Rheum Dis* 2005; **64**(Suppl 2): ii65–ii68; discussion ii69-73.

19 McGuire JJr, Mindel J, Rosenberg W *et al.* Dermatologic and Ophthalmic Drugs Advisory Committee 49th meeting, Open Session [Internet]. U.S. Food and Drug Administration; March 1998. URL http://www.fda.gov/ohrms/dockets/ac/98/transcpt/3402t2.pdf (last accessed: 6 February 2015).

20 Jensen JD, Fujita M, Dellavalle RP. Validation of psoriasis clinical severity and outcome measures: searching for a gold standard. *Arch Dermatol* 2010; **147**: 95–98.

21 Long CC, Finlay AY, Averill RW. The rule of hand: 4 Hand Areas = 2 FTU = 1 g. *Arch Dermatol* 1992; **128**: 1129–1130.

22 Simpson MJ, Chow C, Morgenstern H, Luger TA, Ellis CN. Comparison of three methods for measuring psoriasis severity in clinical studies (Part 2 of 2): Use of quality of life to assess construct validity of the Lattice System Physician's Global Assessment, Psoriasis Area and Severity Index, and Static Physician's Global Assessment. JEADV 2015; in press.

23 Paul C, Gourraud P-A, Bronsard V *et al.* Evidence-based recommendations to assess psoriasis severity: systematic literature review and expert opinion of a panel of dermatologists. *J Eur Acad Dermatol Venereol* 2010; **24**(Suppl 2): 2–9.

24 Puzenat E, Bronsard V, Prey S *et al.* What are the best outcome measures for assessing plaque psoriasis severity? A systematic review of the literature. *J Eur Acad Dermatol Venereol* 2010; **24**(Suppl 2): 10–16.

25 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Guideline for good clinical practice [Internet]. 1996. European Medicines Agency. URL http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf (last accessed: 6 February 2015).

26 Chandran V, Gottlieb A, Cook RJ *et al.* International multicenter psoriasis and psoriatic arthritis reliability trial for the assessment of skin, joints, nails, and dactylitis. *Arthritis Rheum* 2009; **61**: 1235–1242.

27 Wittkowski KM, Leonardi C, Gottlieb A *et al.* Clinical symptoms of skin, nails, and joints manifest independently in patients with concomitant psoriasis and psoriatic arthritis. *PLoS ONE* 2011; **6**: e20279.

28 Ellis CN, Barker JN, Haig AE, Parker CA, Daly S, Jayawardene DA. Placebo response in two long-term randomized psoriasis studies that were negative for rosiglitazone. *Am J Clin Dermatol* 2007; **8**: 93–102.

29 Lamel SA, Myer KA, Younes N, Zhou JA, Maibach H, Maibach HI. Placebo response in relation to clinical trial design: a systematic review and meta-analysis of randomized controlled trials for determining biologic efficacy in psoriasis treatment. *Arch Dermatol Res* 2012; **304**: 707–717.

30 Chren MM. Giving "scale" new meaning in dermatology: measurement matters. *Arch Dermatol* 2000; **136**: 788–790.

31 Jacobson CC, Kimball AB. Rethinking the Psoriasis Area and Severity Index: the impact of area should be increased. *Br J Dermatol* 2004; **151**: 381–387.

32 Robinson A, Kardos M, Kimball AB. Physician Global Assessment (PGA) and Psoriasis Area and Severity Index (PASI): Why do both? A systematic analysis of randomized controlled trials of biologic agents for moderate to severe plaque psoriasis. *J Am Acad Dermatol* 2012; **66**: 369–375.

33 Naldi L. Scoring and monitoring the severity of psoriasis. What is the preferred method? What is the ideal method? Is PASI passé? Facts and controversies. *Clin Dermatol* 2010; **28**: 67–72.

34 Krueger GG, Feldman SR, Camisa C *et al.* Two considerations for patients with psoriasis and their clinicians: what defines mild, moderate,

and severe psoriasis? What constitutes a clinically significant improvement when treating psoriasis? *J Am Acad Dermatol* 2000; **43**: 281–285.

35 Carlin CS, Feldman SR, Krueger JG, Menter A, Krueger GG. A 50% reduction in the Psoriasis Area and Severity Index (PASI 50) is a clinically significant endpoint in the assessment of psoriasis. *J Am Acad Dermatol* 2004; **50**: 859–866.

36 Schäfer I, Hacker J, Rustenbach SJ, Radtke M, Franzke N, Augustin M. Concordance of the Psoriasis Area and Severity Index (PASI) and patient-reported outcomes in psoriasis treatment. *Eur J Dermatol* 2010; **20**: 62–67.

37 European Medicines Agency, Committee for Medicinal Products for Human Use (CHMP). Guideline on clinical investigation of medicinal products indicated for the treatment of psoriasis. 2004. URL www.e-ma.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003329.pdf (last accessed: 6 February 2015).

Appendix

Study Investigators.

Canada: Lorne Albrecht, Surrey BC; Robert Bissonnette, Montreal QC; David Gratton, Montreal QC; Wayne Gulliver, St. John's NL; Aditya Gupta, London ON; Vincent Ho, Vancouver BC; Rod Kunynetz, Barrie ON; Ian Landells, St. John's NL; Richard Langley, Halifax NS; Charles Lynde, Markham ON; Kim Papp, Waterloo ON; Yves Poulin, Quebec City, QC; Michael Robern, Ottawa ON; Leslie Rosoph, North Bay ON; Norman Wasel, Edmonton AB. Germany: Peter Altmeyer, Bochum; Thorsten Auer, Bochum; Matthias Augustin, Hamburg; Beate Bräu, Gießen; Thomas Dirschka, Wuppertal; Yvonne Frambach, Lübeck; Beatrice Gerlach, Dresden; Petra Hauptmann, Salzwedel; Hans-Joachim Lüdcke, Potsdam; Thomas A. Luger, Münster; Karl-Gustav Meyer, Berlin; Ulrich Mrowietz, Kiel; Georg Popp, Augsburg; Henrik Prés, Berlin; Kristian Reich, Hamburg; Rudolf Schopf, Mainz; Michael Sebastian, Mahlow; Jan C. Simon, Leipzig; Margrit Simon, Berlin; Wolfram Sterry, Berlin; Diamant Thaci, Frankfurt; Gottfried Wozel, Dresden. Poland: Maria Blaszczyk-Kostanecka, Warszawa; Grazyna Chodorowska, Lublin; Jerzy Jarzab, Zabrze; Andrzej Kaszuba, Lodz; Andrzej Langner, Iwonicz Zdroj; Slawomir Majewski, Warszawa; Romuald Maleszka, Szczecin; Waldemar Placek, Bydgoszcz; Wojciech Silny, Poznan; Dorota Stajno-Danikiewicz, Gliwice; Anna Wojas-Pelc, Krakow.