Validation of the Severity of Pruritus Scale (SPS) for the Assessment of Pruritus in Atopic Dermatitis (AD)

Gil Yosipovitch,[1] Eric Simpson,[2] Andrew G. Bushmakin,[3] Joseph C. Cappelleri,[3] Thomas Luger,[4] Sonja Ständer,[4] Wynnis Tom,[5] Katy Benjamin,[6] William C. Ports,[3] Anna M. Tallman,[7] Huaming Tan,[3] Robert A. Gerber[3]

[1]University of Miami, Miller School of Medicine, Miami, FL, USA; [2]Oregon Health and Science University, Portland, OR, USA; [3]Pfizer Inc., Groton, CT, USA; [4]University Hospital Münster, Münster, Germany; [5]Rady Children's Hospital-San Diego, San Diego, CA, USA; [6]ICON plc, Gaithersburg, MD, USA; [7]Pfizer Inc., New York, NY, USA

BACKGROUND

- AD is a chronic inflammatory skin disease characterized by the development of eczematous lesions[1]
- Pruritus is a significant feature of AD and is believed to be responsible for much of the burden associated with the disease[2]
- Crisaborole ointment is a nonsteroidal phosphodiesterase 4 inhibitor for the treatment of mild to moderate AD[3]
- The safety and efficacy of crisaborole was established in 2 Phase 3 clinical trials conducted in the United States (AD-301: NCT02118766; AD-302: NCT02118792)[4]
- Pruritus was assessed within these trials using the SPS, a 4-point rating scale, ranging from 0 (none: no itching) to 3 (severe, bothersome itching/scratching which disturbs sleep) that was adapted from the Atopic Dermatitis Severity Index[4] to quantify pruritus intensity within a 24-hour period (**Table 1**)

Table 1. Severity of Pruritus Scale

Instructions: Please think about your itching (or your child's itching if you are completing for your child) over the past 24 hours and choose the category that best describes it.		
Score	Grade	Definition
0	None	No Itching
1	Mild	Occasional, slight itching/scratching
2	Moderate	Constant or intermittent itching/scratching which is not disturbing sleep
3	Severe	Bothersome itching/scratching which is disturbing sleep

OBJECTIVES

- To assess the content validity of the SPS to ensure it is a clear and appropriate tool for the assessment of pruritus intensity
- To evaluate the psychometric properties (quantitative validation) of the SPS to ensure it is a valid and reliable measure of pruritus intensity

METHODS

Qualitative Content Validation of the SPS

- A combined concept elicitation and cognitive interview study was conducted to evaluate the content validity of the SPS in accordance with best practices[5-10]
- Eligible participants were aged ≥2 years, had a dermatologist confirmed diagnosis of AD within the preceding 12 months, and had experienced itching caused by AD
- Participants completed a brief background questionnaire and took part in a single one-on-one interview
  - For children aged 2-7 years, the caregiver completed the SPS and participated in the interview
- Because of the relative homogeneity of the US-only population and the single-item, single-concept nature of the SPS, a target sample size of 15 subjects was considered adequate

Quantitative Psychometric Validation of the SPS

Data Set

- Data from the 2 identically designed, multicenter, randomized, double-blind, vehicle-controlled Phase 3 clinical trials investigating the safety and efficacy of crisaborole were pooled to assess the psychometric properties of the SPS (AD-301: NCT02118766; AD-302: NCT02118792)
  - 1522 patients aged ≥2 years with mild to moderate AD were included in the intent-to-treat population
  - Efficacy was assessed using the Investigator's Static Global Assessment (ISGA; a 5-point rating scale measuring overall disease severity from clear [0] to severe [4]); the SPS; signs of AD (measured on a 4-point rating scale from none [0] to severe [3]); and health-related quality-of-life measures (Dermatology Life Quality Index [DLQI], Children's Dermatology Life Quality Index [CDLQI], Dermatitis Family Impact Questionnaire [DFI])

Test-Retest Reliability

- Assessed via an intraclass correlation coefficient (ICC) using all available SPS observations from stable subjects between baseline/day 1 and day 8 (stable group was defined as having no change on ISGA between baseline/day 1 and day 8)
  - ICC ≥0.70 was considered indicative of acceptable test-retest reliability; ICC ≥0.90 was considered indicative of excellent test-retest reliability[11,12]

Construct Validity

- Convergent validity was evaluated by calculation of Pearson correlations with the ISGA, quality-of-life instruments (DLQI, CDLQI, and DFI), and signs of AD
  - Evidence for convergent validity was based on a Pearson correlation ≥0.40 (correlations ≥0.50 were considered indicative of a strong association)
- Known-groups validity was assessed based on the difference in mean SPS scores between the "no disease group/clear" (ISGA = 0) and the "severe disease group" (ISGA = 4)
  - SPS scores as a function of ISGA were modeled using repeated-measures longitudinal analyses
  - The effect size was calculated as the difference in the mean divided by the baseline standard deviation (values of 0.20, 0.50, and 0.80 standard deviation units were considered small, medium, and large, respectively)

Ability to Detect Change

- Evaluated using a repeated-measures longitudinal mixed model to estimate the relationship between SPS and ISGA scores

Clinically Important Difference

- Estimated using a repeated-measures longitudinal model linked to a 1-cateogry difference on the ISGA

Clinically Important Response

- Estimated using a repeated-measures longitudinal model with the change in SPS score from baseline as the outcome and a newly created static global impression of change (SGIC) anchor as the predictor (SGIC was based on categorizing the change from baseline in ISGA scores as worse [–1], same [0], or better [1])

RESULTS

Qualitative Content Validation of the SPS

- 14 individuals participated in the content validation study (**Table 2**)
- 9 interviews were conducted in US English, and 5 were conducted in US Spanish

Table 2. Participant Characteristics

	Children 2-7 Years of Age n = 5	Children 8-11 Years of Age n = 4	Adolescents and Adults n = 5
Patients characteristics			
Sex Female	4	1	3
Language English	3	3	3
Age, mean (SD, range), years	5.0 (1.2, 3-6)	9.3 (1.5, 8-11)	27.8 (17.9, 17-59)
Race White Black	4 1	2 2	5 0
Ethnicity Hispanic/Latino	3	1	2
AD severity at interview (patient/caregiver reported) Almost clear Mild Moderate Severe	3 0 1 1	1 0 3 0	2 1 3 0
Currently receiving AD treatment (any)	5	4	5
Characteristics of caregivers of children 2-7 years of age (n = 5)			
Sex Female	4	—	—
Language English	3	—	—
Age, mean (SD, range), years	34.6 (4.5, 28-38)	—	—
Race White Black	4 1	— —	— —
Ethnicity Hispanic/Latino	3	—	—
Working status Homemaker Full-time work Part-time work	2 2 1	— — —	— — —
Education level High school Some college Master's degree	1 3 1	— — —	— — —

AD, atopic dermatitis; SD, standard deviation.

Concept Elicitation

- The most prevalent symptom was itch, with 79% (n = 11) of participants reporting it spontaneously and 21% (n = 3) reporting it after probing
- Other reported signs included change in skin color (50%, n = 7), dry skin (36%, n = 5), and change in skin texture (29%, n = 4)
- Concept saturation analysis showed that only 2 new concepts were reported by the fifth interview, and all concepts were reported by the ninth interview

Cognitive Interview

- All participants correctly interpreted the SPS instructions and found them easy to understand (**Table 3**)
- Most participants found the scale easy to complete and correctly interpreted the meaning of the questions and response options and the phrase "sleep disturbance" (**Table 3**)

Table 3. Summary of the Cognitive Interviewing Results

Question	Results Overview
Interpretation of SPS instructions	• 14 (100%) participants correctly interpreted the SPS instructions
Easy or difficult to understand the instructions	• 14 (100%) participants found the SPS instructions easy to understand
Easy or difficult to complete the scale	• 12 (86%) participants reported that the scale was easy to complete
Interpretation of item	• 13 (93%) participants correctly interpreted the meaning of the question – 1 child (age 8 years) did not interpret the meaning of the question correctly
Interpretation of response options	• 11 (79%) participants interpreted the response options consistently and in agreement with the provided definitions – 1 caregiver (US English) interpreted the meaning of each response option as frequency instead of severity – 1 child (age 8 years, US Spanish) defined "moderate" as "night" – 1 child (age 8 years, US English) did not differentiate between "mild" and "moderate"
Meaning of "sleep disturbance"	• 13 (93%) participants had no issues interpreting the terms "disturbing sleep" – 1 adult (US English) had difficulty interpreting the terms "disturbing sleep"
Recall period	• 11 (79%) participants correctly interpreted the 24-hour recall period – 1 adult reported on itch "in general" – 1 adult thought about itch severity "a few weeks ago" – 1 child (age 11) interpreted the question as frequency during the past week rather than the past 24 hours
How participant arrived at response	• No issues were raised regarding how patients arrived at their answers

SPS, Severity of Pruritus Scale.

Quantitative Psychometric Validation of the SPS

Test-Retest Reliability

- The ICC value for a single SPS measurement was estimated to be 0.54
- Reliability improved with the use of multiple SPS measurements
  - The average of 2 SPS measurements (representing average pruritus over 1 day) provided an ICC value of 0.70 (indicative of acceptable test-retest reliability)
  - The average of 14 SPS measurements (representing average pruritus over a 1-week period) provided an ICC value of 0.94 (indicative of excellent test-retest reliability)

Construct Validity

- Convergent validity was supported by strong correlation (Pearson correlation ≥0.50) between SPS scores and ISGA, DLQI, CDLQI, and DFI instruments, and correlations of ≥0.40 with 4 of 5 of the signs of AD at day 29 (**Table 4**)

Table 4. Correlation Between the SPS and All Instruments (Pearson correlation coefficient [r value][a])

	Baseline	Day 8	Day 15	Day 22	Day 29
ISGA	0.22	0.36	0.42	0.48	0.50
Erythema (redness)	0.16	0.30	0.38	0.39	0.42
Induration/ papulation	0.16	0.30	0.36	0.40	0.44
Exudation (oozing or crusting)	0.18	0.29	0.29	0.36	0.35
Excoriation (evidence of scratching)	0.25	0.36	0.42	0.45	0.41
Lichenification (epidermal thickening)	0.16	0.31	0.33	0.35	0.40
DLQI	0.46	—	—	—	0.59
CDLQI	0.47	—	—	—	0.58
DFI	0.38	—	—	—	0.53

CDLQI, Children's Dermatology Life Quality Index; DFI, Dermatitis Family Impact questionnaire; DLQI, Dermatology Life Quality Index; ISGA, Investigator's Static Global Assessment; SPS, Severity of Pruritus Scale.
[a]P < 0.0001 for all correlations.

- A 4-category difference in the ISGA (between the "no disease group/clear" and "severe disease group") was associated with a difference in SPS score of 0.80 (continuous anchor) and 0.87 (categorical anchor), indicating that the SPS can distinguish between groups known to be different (**Table 5**)

Table 5. Known-Groups Validity of the SPS in Relation to Scores on the ISGA

	Differences in Mean SPS Scores on ISGA Between the "Severe Disease Group" and the "No Disease Group"	
Data and Model	Difference (95% CI)	ES
Pooled studies (AD-301 and AD-302)[a]	0.80 (0.73-0.88)	1.03
Pooled studies (AD-301 and AD-302)[b]	0.87 (0.73-1.0)	1.12
Study AD-301[a]	0.75 (0.64-0.85)	0.96
Study AD-301[b]	0.89 (0.71-1.07)	1.14
Study AD-302[a]	0.86 (0.75-0.97)	1.10
Study AD-302[b]	0.84 (0.65-1.04)	1.08

ES, effect size; ISGA, Investigator's Static Global Assessment; SPS, Severity of Pruritus Scale.
ES calculated as difference/standard deviation of 0.78 (based on pooled data from post hoc analyses).
[a]ISGA as a continuous anchor.
[b]ISGA as a categorical anchor.

Ability to Detect Change

- The relationship between SPS scores and ISGA scores provides evidence of sensitivity to change over time (**Table 6**, **Figure 1**)

Table 6. Ability to Detect Change in the SPS in Relation to ISGA Scores

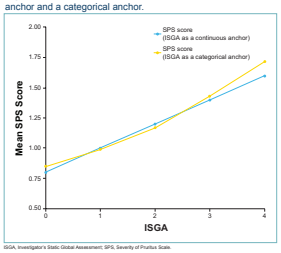
ISGA Score	Mean SPS Score (95% CI) (ISGA continuous predictor)	Mean SPS Score (95% CI) (ISGA categorical predictor)
Pooled studies AD-301 and AD-302		
0	0.79 (0.75-0.84)	0.84 (0.79-0.90)
1	0.99 (0.96-1.03)	0.99 (0.95-1.03)
2	1.19 (1.16-1.23)	1.16 (1.13-1.20)
3	1.40 (1.36-1.44)	1.42 (1.38-1.47)
4	1.60 (1.54-1.65)	1.71 (1.59-1.83)
Study AD-301		
0	0.82 (0.75-0.89)	0.88 (0.80-0.96)
1	1.01 (0.95-1.06)	0.99 (0.94-1.05)
2	1.19 (1.14-1.24)	1.17 (1.11-1.22)
3	1.38 (1.32-1.44)	1.40 (1.34-1.47)
4	1.57 (1.49-1.64)	1.76 (1.59-1.93)
Study AD-302		
0	0.77 (0.70-0.83)	0.81 (0.72-0.89)
1	0.98 (0.93-1.03)	0.98 (0.92-1.04)
2	1.20 (1.15-1.24)	1.16 (1.11-1.21)
3	1.41 (1.36-1.47)	1.44 (1.39-1.50)
4	1.63 (1.55-1.70)	1.65 (1.47-1.83)

ISGA, Investigator's Static Global Assessment; SPS, Severity of Pruritus Scale.

Clinically Important Difference

- The estimated clinically important difference for SPS was 0.20 (95% CI, 0.18-0.22)
  - The close functional relationship when using the ISGA as a categorical variable and continuous predictor supports the linearity assumption of the relationship between ISGA and SPS scores (**Figure 1**)

Figure 1. SPS score as a function of the ISGA as a continuous anchor and a categorical anchor.

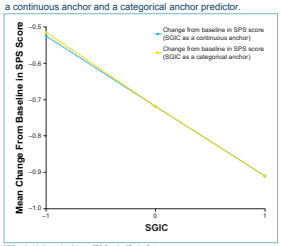


ISGA, Investigator's Static Global Assessment; SPS, Severity of Pruritus Scale.

Clinically Important Response

- A 1-category difference in SGIC corresponded to 0.19 points (95% CI, 0.16-0.22) in SPS score
  - Using the anchor as a categorical variable provided a close functional relationship to the results using the anchor as a continuous predictor, supporting the linearity assumption for the relationship between change from baseline of the SPS score and SGIC score (**Figure 2**)
- The responder definition was estimated as a decrease of 0.19 points from baseline in SPS score, linked to a 1-category difference between the SGIC categories

Figure 2. SPS change score as a function of the SGIC as a continuous anchor and a categorical anchor predictor.



SGIC, static global impression of change; SPS, Severity of Pruritus Scale.

CONCLUSIONS

- **The results of the qualitative content validation analysis confirm that itch is a significant symptom of AD and that the SPS is easy to understand and to complete in both US English and US Spanish**
- **The results of the quantitative analysis confirm the validity of the SPS, demonstrating that it has acceptable test-retest reliability provided ≥2 SPS measurements are used, has good convergent and known-group validity, and has an ability to adequately detect change**
- **A clinically important difference and a clinically important response were also identified, which could be used in future investigations that use the SPS to assess pruritus severity**

REFERENCES

1. Weidinger S, Novak N. Lancet. 2016;387(10023):1109-1122.
2. Eichenfield LF et al. J Am Acad Dermatol. 2014;70(2):338-351.
3. Eucrisa [crisaborole] ointment, 2%, for topical use [prescribing information]. Palo Alto, CA: Anacor Pharmaceuticals, Inc.; 2016.
4. Paller AS et al. J Am Acad Dermatol. 2016;75(3):494-503.e6.
5. Van Leent EJ et al. Arch Dermatol. 1998;134(7):805-809.
6. Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. December 2009. https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm193282.pdf. Accessed September 20, 2017.
7. Patrick DL et al. Value Health. 2011;14(8):978-988.
8. Patrick DL et al. Value Health. 2011;14(8):967-977.
9. Brod M et al. Qual Life Res. 2009;18(9):1263-1278.
10. Lasch KE et al. Qual Life Res. 2010;19(8):1087-1096.
11. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York: McGraw-Hill; 1994.
12. Reeve BB et al. Qual Life Res. 2013;22(8):1889-1905.

ACKNOWLEDGMENTS

Editorial/medical writing support under the guidance of the authors was provided by Jemimah Walker, PhD, and Corey Mandel, PhD, at ApotheCom, San Francisco, CA, USA, and was funded by Pfizer Inc., New York, NY, USA, in accordance with Good Publication Practice (GPP3) guidelines (*Ann Intern Med*. 2015;163:461-464).

Presented at the 2017 Fall Clinical Dermatology Conference; October 12-15, 2017; Las Vegas, NV