



Comparison of Patient-Oriented Eczema Measure and Patient-Oriented Scoring Atopic Dermatitis vs Eczema Area and Severity Index and other measures of atopic dermatitis

A validation study

Jonathan I. Silverberg, MD, PhD, MPH^{*}; Donald Lei, MS[†]; Muhammad Yousaf, BS[‡]; Sherief R. Janmohamed, MD, PhD[†]; Paras P. Vakharia, MD, PharmD[‡]; Rishi Chopra, MD, MS[§]; Rajeev Chavda, MD^{||}; Sylvie Gabriel, MD^{||}; Kevin R. Patel, MD[#]; Vivek Singam, MD^{**}; Robert Kantor, MD[§]; Derek Y. Hsu, MD[†]

^{*} Department of Dermatology, The George Washington University School of Medicine and Health Sciences, Washington, District of Columbia

[†] Department of Dermatology, Feinberg School of Medicine at Northwestern University, Chicago, Illinois

[‡] Department of Dermatology, University of Texas Southwestern Medical Center, Dallas, Texas

[§] State University of New York Downstate Medical Center, Brooklyn, New York

^{||} Galderma SA | Rx Strategy & Innovation Group, La Tour-de-Peluz, Switzerland

[#] Department of Internal Medicine, Massachusetts General Hospital, Boston, Massachusetts

^{**} Department of Internal Medicine, Weiss Memorial Hospital, Chicago, Illinois

ARTICLE INFO

Article history:

Received for publication February 12, 2020.

Received in revised form March 5, 2020.

Accepted for publication March 9, 2020.

ABSTRACT

Background: Little is known about the measurement properties of Patient-Oriented Scoring Atopic Dermatitis (PO-SCORAD) in adults with atopic dermatitis (AD). Even less is known about how PO-SCORAD performs compared with the Patient-Oriented Eczema Measure (POEM).

Objective: To examine the measurement properties of PO-SCORAD and compare them with those of POEM.

Methods: A prospective dermatology practice-based study of 291 patients with AD (age range, 18–72 years).

Results: PO-SCORAD and POEM were moderately correlated with each other (Spearman $\rho = 0.56$) and had weak-moderate correlations with the Numeric Rating Scale (NRS) worst itch and average itch, Dermatology Life Quality Index (DLQI), ItchyQOL, Patient-Reported Outcomes Measurement Information System (PROMIS) Sleep Disturbance (SD) and Sleep-Related Impairment (SRI), Patient Health Questionnaire-9 (PHQ-9), and Eczema Area and Severity Index (EASI) ($P < .001$). POEM had significantly stronger correlations with DLQI, ItchyQOL, and EASI than did PO-SCORAD. PO-SCORAD and POEM had fair discriminant validity. Changes from baseline in PO-SCORAD and POEM were moderately correlated with each other; were weakly to strongly correlated with NRS worst itch and average itch, DLQI, ItchyQOL, PROMIS SD, PROMIS SRI, PHQ-9, and EASI; and had good test-retest reliability. There was no differential item functioning of items or floor or ceiling effects for PO-SCORAD or POEM. The thresholds for meaningful change for PO-SCORAD and POEM were -15.5 and -5.0 , respectively. Median completion times for PO-SCORAD and POEM were 3 minutes and 1 minute, respectively. **Conclusion:** PO-SCORAD and POEM had good construct and cross-cultural validity, reliability, and responsiveness in adults with AD and were feasible for use in clinical trials and practice. However, POEM had better measurement properties than PO-SCORAD.

© 2020 American College of Allergy, Asthma & Immunology. Published by Elsevier Inc. All rights reserved.

Reprints: Jonathan I. Silverberg, MD, PhD, MPH, Department of Dermatology, The George Washington University School of Medicine and Health Sciences, 2150 Pennsylvania Ave NW, Suite 2B-425, Washington, DC 20037; E-mail: JonathanSilverberg@gmail.com.

Disclosures: Drs Chavda and Gabriel are employees of Galderma. Dr Silverberg has

received honoraria from Galderma. The other authors have nothing to disclose.

Funding Sources: This publication was made possible with support from the Agency for Healthcare Research and Quality (AHRQ, grant number K12 HS023011), the Dermatology Foundation, and an unrestricted research grant from Galderma.

Introduction

Atopic dermatitis (AD) is a heterogeneous disorder associated with variable signs,¹ lesional distribution,^{1,2} and symptoms (eg, itch, pain,³ sleep disturbance,^{4–8} depression, anxiety, and inattentiveness^{5,7,9,10}), all of which contribute toward perception of disease severity and health-related quality of life.^{11,12} Standardized patient-reported outcomes (PROs) are important tools for measuring AD severity in clinical practice and trials.

Multiple standardized PROs for assessing AD severity exist, including Patient-Oriented Scoring AD (PO-SCORAD) and Patient-Oriented Eczema Measure (POEM), with varying measurement constructs, validity, and reliability. POEM assesses the frequency of 7 different patient-reported signs and symptoms of AD.¹³ POEM was previously found to have good content and construct validity but only fair internal consistency and cross-cultural validity.^{13–15} PO-SCORAD is a PRO analog of SCORAD^{16,17} and assesses the severity of 6 signs of AD, lesional extent, and intensity of itch and sleep loss.¹⁶ PO-SCORAD correlates well with SCORAD,¹⁶ has good convergent and discriminant validity, and has fair cross-cultural validity.^{15,18–20} However, only a few studies have examined the validity of PO-SCORAD, and little is known about its responsiveness, reliability, and meaningful change.

Recently, the Harmonizing Outcome Measures in Eczema international consensus group recommended POEM as the preferred assessment for AD symptoms in clinical trials based on the available data to support their validity and reliability.²¹ The Harmonizing Outcome Measures in Eczema group is now deliberating which assessments should be recommended for use in clinical practice. In this study, we sought to determine the construct and cross-cultural validity, internal consistency, floor or ceiling effects, responsiveness, reliability, meaningful change, and feasibility of PO-SCORAD in adults with AD. In addition, we sought to compare the measurement properties of PO-SCORAD with those of POEM.

Methods

Study Design

A prospective, dermatology practice–based study of adults (≥ 18 years), male or female, with AD as defined by the Hanifin and Rajka diagnostic criteria²² was performed. Exclusion criteria were as follows: (1) those without a definite diagnosis of AD and (2) being unwilling or unable to complete the assessments. Patients were serially recruited from the eczema clinic at Northwestern Medicine, and virtually all patients ($>99\%$) who were invited agreed to participate. Patients received standard-of-care follow-up and treatment, including emollients, prescription topical, systemic therapy, and/or phototherapy, where appropriate. Surveys were administered between June 2017 and September 2019. The study was approved by the institutional review board of Northwestern University. Informed consent was obtained by patients electronically.

Outcome Measures

Patients completed questionnaires on self-reported severity of AD (“Would you describe your atopic dermatitis or eczema as: clear, almost clear, mild, moderate, or severe?”)¹¹; Patient-Reported Outcomes Measurement Information System (PROMIS) itch severity, including Numeric Rating Scale (NRS) (whole number scale using a single question; range of responses during a 7-day recall period, 0–10) and Verbal Rating Scale (Likert-type scale using a single question; 5 responses during a 7-day recall period: had no itch, mild, moderate, severe, and very severe) for worst itch and average itch²³; POEM (7 questions; range, 0–28)¹³; PO-SCORAD (6 self-reported signs [erythema, excoriation, swelling, oozing/crusting, lichenification, and dryness], visual analog scale for itch and

sleep loss in the past 3 days; range, 0–103)¹⁶; DLQI (10 questions; range, 0–30)^{24,25} and ItchyQoL (22 questions; range for mean scores, 1.0–5.0)²⁵ to measure health-related quality of life; Patient-Health Questionnaire-9 (PHQ-9) to measure mental health symptoms (9 questions; range, 0–27)²⁶; and PROMIS sleep disturbance (SD) and sleep-related impairment (SRI) (each with 8 items; range of raw scores, 0–40) to assess sleep symptoms.^{27–29}

Patients were assessed with full-body skin examination by a dermatologist (J.S.). The Eczema Area and Severity Index (EASI; 4 signs [erythema, excoriation, swelling, lichenification] on 4 body sites; range, 0–72)^{30–33} was assessed.

Statistical Analysis

Summary statistics were estimated for baseline population characteristics. Concurrent validity of scores for PO-SCORAD and POEM were established using Spearman correlations with each other. Convergent validity was established using Spearman correlations of PO-SCORAD and POEM with other PROs and EASI. Correlation coefficients scores of 0.70 or higher were considered strong, scores of 0.50 to 0.69 were considered moderate, and scores of 0.30 to 0.49 were considered weak.

Criterion validity was determined using the Kruskal-Wallis test. There is no gold standard assessment for AD severity. Thus, we compared PO-SCORAD and POEM with self-reported global AD severity to examine criterion validity. We hypothesized that there would be significant and stepwise differences of itch assessments with each level of AD severity. Discriminant construct validity of PO-SCORAD and POEM was established using logistic regression models with self-reported global AD severity as the ordinal dependent variables.³⁴ The independent variables were PO-SCORAD and POEM. Discriminant validity was determined using area under the curve (AUC). AUC scores of 0.90 or greater were considered excellent, 0.80 to 0.89 were considered good, 0.70 to 0.79 were considered fair, less than 0.70 were considered poor, and less than 0.60 were considered failure.³⁵ Ordinal logistic regression was used because the data met the proportional odds assumption (score test, $P < .01$).

Responsiveness of scores was determined using Pearson correlations between change from baseline and follow-up visit for PO-SCORAD and POEM with change of each other and other PROs (NRS average itch, NRS worst itch, DLQI, ItchyQoL, PROMIS SD and SRI, and PHQ-9) and EASI. We hypothesized that the changes of PO-SCORAD and POEM would have moderate correlations with each other and weaker correlations with changes in other PROs of AD severity.

Floor or ceiling effects arise when an assessment has a lower or upper limit to the values it can reliably measure. Floor or ceiling effects of total scores were considered present if 15% of responses fell in the lowest or highest scores.^{36,37}

An anchor-based approach was used to determine thresholds for meaningful clinically important difference (MCID) of PO-SCORAD and POEM at follow-up from baseline. Thresholds were determined by comparing the change of PO-SCORAD and POEM with 1-point improvement of patient-reported global AD severity.

Test-retest reliability was assessed by intraclass correlation coefficient (ICC) and 95% CIs, using a mixed-effects model for absolute agreement among patients with no change of patient-reported global AD severity between the baseline and follow-up visits. ICCs less than 0.50 were considered poor, 0.50 to 0.74 were considered moderate, 0.75 to 0.89 were considered good, and 0.90 or higher were considered excellent.³⁸

The above statistical analyses were performed in SAS software, version 9.4.3 (SAS Institute, Cary, North Carolina). Missing values were encountered in 0.1% or less of respondents for all analyzed variables. Complete case analysis was performed (ie, missing values were excluded). A 2-sided $P < .05$ was considered statistically

significant. No formal power calculation was performed. However, a sample size of more than 100 participants per analysis has been recommended as sufficient for validation studies.³⁹

Results

Patient Characteristics

Overall, 291 adults (age range, 18–72 years) were included in the study, with a mean (SD) follow-up visit duration of 0.3 (0.4) years (maximum, 1.1 years). The patient cohort included 144 females (51.6%) and 132 self-reported white individuals (47.3%), with a mean (SD) age at enrollment of 39.9 (14.9) years. Baseline characteristics of AD severity are presented in Table 1.

Concurrent Validity

Regarding concurrent validity, PO-SCORAD was moderately correlated with POEM overall ($\rho = 0.56$, $P < .001$) and in those with moderate-severe self-reported AD severity ($\rho = 0.53$, $P < .001$) but only weakly correlated with POEM in those with clear-mild AD ($P = .39$, $P = .001$) (Table 2). PO-SCORAD vs POEM was associated with a lower proportion of severe scores (13.1% vs 25.2%) but a higher proportion of moderate scores (40.6% vs 31.0%). There was only modest concordance between preestablished severity strata for POEM and PO-SCORAD (weighted $\kappa = 0.41$).

Convergent Validity

Regarding convergent validity, PO-SCORAD was moderately correlated with DLQI, ItchyQOL, PROMIS SRI, and PHQ-9 and weakly correlated with NRS worst itch and average itch, PROMIS SD, and EASI (Spearman correlations, $P < .001$ for all) (Table 2). POEM was strongly correlated with DLQI and ItchyQOL, moderately correlated with EASI, and weakly correlated with NRS worst itch and average itch, PROMIS SD and SRI, and PHQ-9 ($P < .001$ for all).

Compared with POEM, PO-SCORAD had significantly weaker correlations with DLQI ($P < .001$), ItchyQOL ($P = .02$), and EASI ($P = .03$) but no significant differences of correlations with NRS worst itch or average itch, PROMIS SD and SRI, and PHQ-9. In particular, POEM had stronger correlations with DLQI and ItchyQOL than did POEM in those with self-reported clear-mild AD or those of nonwhite race, but there were no significant differences of

correlations in those with self-reported moderate-severe AD or those of white race. POEM had stronger correlations with EASI than PO-SCORAD in those with clear-mild and moderate-severe AD and those of nonwhite race.

Discriminant Validity

There were significant and stepwise increases of PO-SCORAD and POEM at each level of severity for self-reported global severity (Wilcoxon rank sum test, $P < .001$). However, both PO-SCORAD and POEM had poor discriminant validity in predicting self-reported global AD severity (Table 3). There was only a minimal increase in AUC in models that included both PO-SCORAD and POEM.

Floor or Ceiling Effects

The proportions of patients with lowest and highest values for PO-SCORAD (0.3% or 0.4%) and POEM (10.1% or 2.2%) were below 15%, indicating there were no floor or ceiling effects overall.

Differential Item Functioning

No items from PO-SCORAD or POEM were found to have differential item functioning.

Feasibility

Assessments were completed by more than 99% of respondents. The mean (SD) time to completion of PO-SCORAD was 3.2 (3.0) minutes, with a median completion time of 3 minutes (range, <1–25 minutes). The mean (SD) time to completion of POEM was 1.3 (2.2) minutes, with a median completion time of 1 minute (range, <1–25 minutes).

There were no significant differences of completion times for PO-SCORAD or POEM by age (Wilcoxon rank sum test; $P = .68$ and $.75$, respectively), sex ($P = .48$ and $.22$, respectively), race/ethnicity ($P = .38$ and $.15$, respectively), level of education ($P = .48$ and $.08$, respectively), and patient-reported global AD severity ($P = .68$ and $.19$, respectively).

Responsiveness

Changes from baseline in PO-SCORAD and POEM were moderately correlated with each other ($P = .52$). However, changes from baseline in NRS worst itch and average itch, DLQI, ItchyQOL, PROMIS SD and SRI, PHQ-9, and EASI scores all had stronger correlations with changes in POEM compared with changes in PO-SCORAD (Table 2).

Meaningful Change

Overall, the thresholds for MCID for PO-SCORAD and POEM based on an anchor of a 1-point improvement or higher of patient-reported global severity were -15.5 and -5.0 , respectively. In particular, the thresholds were similar for PO-SCORAD (14.6/16.8) but differed for POEM (3.7/6.1) by baseline patient-reported global AD severity (clear-mild or moderate-severe).

Reliability

A total of 189 cases were included in the analysis for reliability. The ICC for PO-SCORAD was 0.77 (95% CI, 0.68–0.84) overall and 0.86 (95% CI, 0.78–0.92) in whites and 0.58 (95% CI, 0.34–0.75) in nonwhites. The ICC for POEM was 0.86 (95% CI, 0.79–0.91) overall and 0.90 (95% CI, 0.83–0.95) in whites and 0.79 (95% CI, 0.63–0.89) in nonwhites.

Table 1
Characteristics of the Study Patients

Characteristic	Finding (N = 291)
Study duration, mean (SD), y	0.3 (0.4)
Age, mean (SD), y	41.8 (14.1)
Female sex	152 (52.2)
Race/ethnicity	
White	140 (48.1)
African American/Black	30 (10.3)
Hispanic	54 (18.6)
Asian	67 (23.0)
Level of education	
High school or less	22 (7.6)
More than high school	269 (92.4)
NRS score for worst itch, median (range)	4.5 (0–10)
NRS score for average itch, median (range)	3 (0–10)
Patient-reported global AD severity	
Clear	18 (6.6)
Almost clear	43 (15.7)
Mild	80 (29.2)
Moderate	82 (29.9)
Severe	51 (18.6)
POEM score, median (range)	9 (0–28)
PO-SCORAD score, median (range)	29.8 (0.3–89.0)
DLQI score, median (range)	5 (0–28)
EASI score, median (range) (n = 141)	5.8 (0.0–49.2)

Abbreviations: AD, atopic dermatitis; DLQI, Dermatology Life Quality Index; EASI, Eczema Area and Severity Index; NRS, Numeric Rating Scale; POEM, Patient Oriented Eczema Measure; PO-SCORAD, Patient-Oriented SCORing Atopic Dermatitis.

Table 2
Spearman Correlations of PO-SCORAD and POEM Scores With Itch Severity, Quality of Life, Sleep Disturbance and Sleep-Related Impact at Baseline, EASI Scores, and Change in Scores From Baseline at Follow-Up

Baseline AD severity	POEM	NRS itch		DLQI	ItchyQOL	PHQ-9	PROMIS sleep		EASI
		Worst	Average				Disturbance	Sleep-related impairment	
Overall									
PO-SCORAD	0.53	0.45	0.46	0.52	0.62	0.51	0.45	0.52	0.31
POEM	1.00	0.45	0.50	0.59	0.71	0.58	0.49	0.49	0.52
P value		.93	.59	<.001	.02	.25	.51	.65	.03
Clear, almost clear, and mild self-reported AD severity									
PO-SCORAD	0.39	0.34	0.34	0.35	0.48	0.32	0.35	0.41	0.40
POEM	1.00	0.35	0.45	0.70	0.73	0.56	0.46	0.45	0.60
P value		.93	.59	<.001	.001	.25	.51	.65	.01
Moderate-severe self-reported AD severity									
PO-SCORAD	0.53	0.36	0.36	0.59	0.66	0.56	0.42	0.50	0.03
POEM	1.00	0.38	0.34	0.56	0.58	0.45	0.39	0.38	0.28
P value		.84	.86	.73	.30	.23	.78	.25	.01
White race									
PO-SCORAD	0.51	0.57	0.58	0.57	0.68	0.50	0.41	0.52	0.43
POEM	1.00	0.42	0.48	0.65	0.65	0.44	0.46	0.41	0.48
P value		.84	.86	.73	.30	.23	.78	.25	.66
Nonwhite race									
PO-SCORAD	0.57	0.34	0.35	0.48	0.60	0.52	0.48	0.49	0.17
POEM	1.00	0.48	0.51	0.72	0.74	0.68	0.50	0.56	0.54
P value		.20	.11	.001	.02	.01	.80	.40	.02
Change in scores									
PO-SCORAD	0.53	0.23	0.26	0.38	0.61	0.25	0.39	0.18	0.13
POEM	1.00	0.48	0.48	0.67	0.70	0.49	0.57	0.42	0.47
P value		<.001	<.001	<.001	.005	<.001	<.001	<.001	.007

Abbreviations: AD, atopic dermatitis; DLQI, Dermatology Life Quality Index; EASI, Eczema Area and Severity Index; NRS, Numeric Rating Scale; PHQ-9, Patient Health Questionnaire 9; POEM, Patient Oriented Eczema Measure; PO-SCORAD, Patient-Oriented SCORing Atopic Dermatitis; PROMIS, Patient-Reported Outcomes Measurement Information System.

Discussion

This study found that PO-SCORAD and POEM had overall good concurrent validity, convergent validity, discriminant validity, cross-cultural validity, feasibility, reliability, and fair responsiveness, with no observed floor or ceiling effects. PO-SCORAD and POEM had similar discriminant validity. However, POEM correlated better with quality-of-life measures than PO-SCORAD overall and particularly in those with milder AD and of nonwhite race (convergent validity). Subsequently, an adaptation of PO-SCORAD was developed for use in black skin,⁴⁰ which likely should improve the performance of PO-SCORAD in adults of nonwhite race. POEM also had better responsiveness than

PO-SCORAD, with improvement of itch, AD severity, quality of life, and mental health. POEM also had better reliability than PO-SCORAD overall and particularly in nonwhites. POEM appeared to be more feasible than PO-SCORAD, with shorter completion times and simpler scoring. Together, PO-SCORAD and POEM have sufficient validity, reliability, and feasibility to be used as assessments of AD severity. However, POEM had better measurement properties than PO-SCORAD.

Overall, the results of this study are consistent with those of previous studies that found good validity and reliability for PO-SCORAD and POEM in adults.¹⁵ A US population-based study of 602 adults with AD also found that PO-SCORAD and POEM had good convergent validity for itch severity, quality-of-life impact and mental health symptoms, and criterion validity for AD severity.¹⁵ However, in contrast to the results of this study, PO-SCORAD was previously found to correlate better with quality of life^{15,41} and mental health symptoms¹⁵ than POEM. In addition, differential item functioning was previously identified for multiple items in PO-SCORAD and POEM, indicating limited cross-cultural validity,¹⁵ but was not identified for any items in PO-SCORAD or POEM in this study. We found differences of reliability for PO-SCORAD between whites and nonwhites. Finally, we did not compare PO-SCORAD with SCORAD. However, the objective components of PO-SCORAD and SCORAD were previously found to be moderately correlated in 75 children with AD from a single outpatient clinic¹⁸ and 185 adults and 286 children with AD in hospitals from 9 European countries.¹⁶

Both PO-SCORAD and POEM were completed in a mean of 3 minutes or less, which is acceptable or adequate for use in clinical practice (<3 or <3-5 minutes).¹⁴ However, POEM was generally more time-efficient than PO-SCORAD. POEM can be administered via a paper form or electronically, with a simple scoring system by summing the responses across all 7 items. PO-SCORAD can also be administered via a paper form or electronically, but it has a more complicated scoring system that may be challenging for clinicians to

Table 3
AUCs for PO-SCORAD and/or POEM as Predictors of Self-Reported Global AD Severity at Baseline

AD severity	AUC		
	PO-SCORAD	POEM	PO-SCORAD and POEM
Multilevel	0.67	0.66	0.69
Almost clear vs clear	0.50	0.50	0.50
Mild vs almost clear	0.53	0.53	0.55
Mild vs clear	0.59	0.53	0.60
Moderate vs mild	0.62	0.67	0.64
Moderate vs almost clear	0.74	0.72	0.76
Moderate vs clear	0.84	0.80	0.86
Severe vs moderate	0.53	0.52	0.53
Severe vs mild	0.70	0.70	0.71
Severe vs almost clear	0.79	0.74	0.82
Severe vs clear	0.86	0.83	0.89

Abbreviations: AD, atopic dermatitis; AUC, area under the receiver operating characteristic curve; POEM, Patient Oriented Eczema Measure; PO-SCORAD, Patient-Oriented SCORing Atopic Dermatitis.

perform in the office. PO-SCORAD has an online app available, which can simplify administration. However, this may be impossible to integrate into existing electronic health records and research data capture software packages. Although both may be integrated into day-to-day practice, POEM appears to have greater feasibility.

POEM demonstrated good responsiveness and reliability in this study, which is generally in line with previous studies.^{13,14,42} Although PO-SCORAD demonstrated only fair responsiveness to changes of other PROs and EASI in this study, it had good responsiveness to change of SCORAD in a multicenter study of children and adults.¹⁶ PO-SCORAD also had good reliability; reliability was not previously assessed for PO-SCORAD.⁴² Five- and 15.5-point reductions of POEM and PO-SCORAD were the thresholds for meaningful improvement based on a 1-point improvement of patient-reported global AD severity. The MCID of 5 is slightly higher than the MCID of 3.4 points previously identified for POEM using a pooled cohort from 3 clinical trials.⁴³ Furthermore, we found a higher MCID for POEM in those with moderate-severe vs clear-mild AD at baseline. It is logical that patients with more severe AD at baseline should have greater improvements of POEM scores over time. However, these data raise questions about the rationale of using a single MCID for patients with mild-severe AD.

This study has several strengths, including clinical confirmation of the AD diagnosis; good representation across sex, race/ethnicity, and AD severity; evaluation of feasibility; and testing of multiple PROs and EASI when examining the psychometric properties. However, there are some limitations. Patients were recruited from a single academic center, which may limit generalizability. We did not assess content validity. However, content validity was previously assessed and found to be good.⁴² Measurement properties were not tested in children or adolescents. We used the original PO-SCORAD tool that was developed with standardized images for white skin. We did not use the adaptation of PO-SCORAD for use in black skin.⁴⁰ Future studies are needed to address these points.

In conclusion, this study demonstrates that PO-SCORAD and POEM have good construct and cross-cultural validity, responsiveness, reliability, and feasibility in the assessment of adult AD in clinical practice, with no overall floor or ceiling effects. However, POEM had better measurement properties than PO-SCORAD. These instruments provide important information about AD severity that can guide therapeutic decision making.

References

- Yew YW, Thyssen JP, Silverberg JI. A systematic review and meta-analysis of the regional and age-related differences in atopic dermatitis clinical characteristics. *J Am Acad Dermatol*. 2019;80(2):390–401.
- Silverberg JI, Margolis DJ, Boguniewicz M, et al. Distribution of atopic dermatitis lesions in United States adults. *J Eur Acad Dermatol Venereol*. 2019;33(7):1341–1348.
- Vakharia PP, Chopra R, Sacotte R, et al. Burden of skin pain in atopic dermatitis. *Ann Allergy Asthma Immunol*. 2017;119(6):548–552.e3.
- Fishbein AB, Vitaterna O, Haugh IM, et al. Nocturnal eczema: review of sleep and circadian rhythms in children with atopic dermatitis and future research directions. *J Allergy Clin Immunol*. 2015;136(5):1170–1177.
- Silverberg JI. Selected comorbidities of atopic dermatitis: atopy, neuropsychiatric, and musculoskeletal disorders. *Clin Dermatol*. 2017;35(4):360–366.
- Silverberg JI, Garg NK, Paller AS, Fishbein AB, Zee PC. Sleep disturbances in adults with eczema are associated with impaired overall health: a US population-based study. *J Invest Dermatol*. 2015;135(1):56–66.
- Silverberg JI, Simpson EL. Association between severe eczema in children and multiple comorbid conditions and increased healthcare utilization. *Pediatr Allergy Immunol*. 2013;24(5):476–486.
- Yu SH, Attarian H, Zee P, Silverberg JI. Burden of sleep and fatigue in US adults with atopic dermatitis. *Dermatitis*. 2016;27(2):50–58.
- Patel KR, Immaneni S, Singam V, Rastogi S, Silverberg JI. Association between atopic dermatitis, depression and suicidal ideation: a systematic review and meta-analysis. *J Am Acad Dermatol*. 2019;80(2):402–410.
- Yu SH, Silverberg JI. Association between atopic dermatitis and depression in US adults. *J Invest Dermatol*. 2015;135(12):3183–3186.
- Vakharia PP, Chopra R, Sacotte R, et al. Validation of patient-reported global severity of atopic dermatitis in adults. *Allergy*. 2018;73(2):451–458.
- Silverberg JI, Chiesa Fuxench ZC, Gelfand JM, et al. Content and construct validity, predictors, and distribution of self-reported atopic dermatitis severity in US adults. *Ann Allergy Asthma Immunol*. 2018;121(6):729–734.e4.
- Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol*. 2004;140(12):1513–1519.
- Schmitt J, Langan S, Williams HC, European Dermato-Epidemiology Network. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol*. 2007;120(6):1389–1398.
- Silverberg JI, Margolis DJ, Boguniewicz M, et al. Validation of five patient-reported outcomes for atopic dermatitis severity in adults. *Br J Dermatol*. 2020;182(1):104–111.
- Stalder JF, Barbarot S, Wollenberg A, et al. Patient-oriented SCORAD (PO-SCORAD): a new self-assessment scale in atopic dermatitis validated in Europe. *Allergy*. 2011;66(8):1114–1121.
- Vourc'h-Jourdain M, Barbarot S, Taieb A, et al. Patient-oriented SCORAD: a self-assessment score in atopic dermatitis: a preliminary feasibility study. *Dermatology*. 2009;218(3):246–251.
- van Oosterhout M, Janmohamed SR, Spierings M, Hiddinga J, de Waard-van der Spek FB, Oranje AP. Correlation between Objective SCORAD and three-item severity score used by physicians and Objective PO-SCORAD used by parents/patients in children with atopic dermatitis. *Dermatology*. 2015;230(2):105–112.
- Boccardi D, D'Auria E, Turati F, et al. Disease severity and quality of life in children with atopic dermatitis: PO-SCORAD in clinical practice. *Minerva Pediatr*. 2017;69(5):373–380.
- Pustisek N, Vurnek Zivkovic M, Situm M. Quality of life in families with children with atopic dermatitis. *Pediatr Dermatol*. 2016;33(1):28–32.
- Spuls PI, Gerbens LAA, Simpson E, et al. Patient-oriented eczema measure (POEM), a core instrument to measure symptoms in clinical trials: a Harmonising Outcome Measures for Eczema (HOME) statement. *Br J Dermatol*. 2017;176(4):979–984.
- Hanifin JM, Rajka G. Diagnostic features of atopic eczema. *Acta Derm Venereol*. 1980;92:44–47.
- Silverberg JI, Lai J, Patel K, et al. Measurement properties of the PROMIS Itch Questionnaire – itch severity assessments in adults with atopic dermatitis. *Br J Dermatol*. Published online February 27, 2020. <https://doi.org/10.1111/bjd.18978>
- Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994–2007: a comprehensive review of validation data and clinical results. *Br J Dermatol*. 2008;159(5):997–1035.
- Desai NS, Poindexter GB, Monthroppe YM, Bendeck SE, Swerlick RA, Chen SC. A pilot quality-of-life instrument for pruritus. *J Am Acad Dermatol*. 2008;59(2):234–244.
- Gilbody S, Richards D, Brealey S, Hewitt C. Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *J Gen Intern Med*. 2007;22(11):1596–1602.
- Silverberg JI, Lai JS, Vakharia PP, et al. Measurement properties of the Patient-Reported Outcomes Measurement Information System Itch Questionnaire Item Banks in adults with atopic dermatitis [e-pub ahead of print]. *J Am Acad Dermatol*. <https://doi.org/10.1016/j.jaad.2019.11.057>, accessed December 5, 2019.
- Silverberg JI, Lai JS, Kantor RW, et al. Development, validation, and interpretation of the PROMIS Itch Questionnaire: a patient-reported outcome measure for the quality of life impact of itch [e-pub ahead of print]. *J Invest Dermatol*. <https://doi.org/10.1016/j.jid.2019.08.452>, accessed October 16, 2019.
- Lei D, Yousaf M, Janmohamed SR, et al. Validation of patient-reported outcomes information system sleep disturbance and sleep-related impairment in adults with atopic dermatitis [e-pub ahead of print]. *Br J Dermatol*. <https://doi.org/10.1111/bjd.18920>, accessed February 4, 2020.
- Hanifin JM, Thurston M, Omoto M, Cherill R, Tofte SJ, Graeber M. The Eczema Area and Severity Index (EASI): assessment of reliability in atopic dermatitis. EASI Evaluator Group. *Exp Dermatol*. 2001;10(1):11–18.
- Chopra R, Silverberg JI. Assessing the severity of atopic dermatitis in clinical trials and practice. *Clin Dermatol*. 2018;36(5):606–615.
- Chopra R, Vakharia PP, Sacotte R, et al. Relationship between EASI and SCORAD severity assessments for atopic dermatitis. *J Allergy Clin Immunol*. 2017;140(6):1708–1710.e1.
- Chopra R, Vakharia PP, Sacotte R, et al. Severity strata for Eczema Area and Severity Index (EASI), modified EASI, Scoring Atopic Dermatitis (SCORAD), Objective SCORAD, Atopic Dermatitis Severity Index and body surface area in adolescents and adults with atopic dermatitis. *Br J Dermatol*. 2017;177(5):1316–1321.
- Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Mach Learn*. 2001;45(2):171–186.
- Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8(4):283–298.

36. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007; 60(1):34–42.
37. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open*. 2015; 5(7): e007765.
38. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
39. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012; 21(4):651–657.
40. Faye O, Meledie N'Djong AP, Diadie S, et al. Validation of the Patient-Oriented SCORing for Atopic Dermatitis tool for black skin. *J Eur Acad Dermatol Venereol*. 2020;34(4):795–799.
41. Coutanceau C, Stalder JF. Analysis of correlations between Patient-Oriented SCORAD (PO-SCORAD) and other assessment scores of atopic dermatitis severity and quality of life. *Dermatology*. 2014;229(3):248–255.
42. Gerbens LA, Prinsen CA, Chalmers JR, et al. Evaluation of the measurement properties of symptom measurement instruments for atopic eczema: a systematic review. *Allergy*. 2017;72(1):146–163.
43. Schram ME, Spuls PI, Leeftang MM, Lindeboom R, Bos JD, Schmitt J. EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy*. 2012;67(1):99–106.