Measurement properties of the Patient-Reported Outcomes Measurement Information System Itch Questionnaire item banks in adults with atopic dermatitis

Jonathan I. Silverberg, MD, PhD, MPH,[a,b,c,d,e] Jin-Shei Lai, PhD,[b] Paras P. Vakharia, MD, PharmD,[a] Kevin Patel, MD,[a] Vivek Singam, MD,[a] Rishi Chopra, MD, MS,[a] Ryan Sacotte, MD,[a] Neha Patel, BS,[a] Supriya Rastogi, MD,[a] Robert Kantor, MD,[a] Derek Y. Hsu, MD,[a] and David Cella, PhD[b,c,f]
Chicago, Illinois and Washington, DC

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS) Itch Questionnaire (PIQ) was recently developed.

Objective: To validate PIQ short forms in adults with AD.

Methods: Self-administered questionnaires and skin examinations were performed in 239 adults with atopic dermatitis (AD) in a dermatology practice setting.

Results: PIQ items had good content validity. PIQ item bank T-scores strongly correlated with each other, moderately correlated with numeric and verbal rating scales for worst or average itch and with itch frequency, moderately to strongly correlated with patient-oriented eczema measure, and weakly to moderately correlated with the Eczema Area and Severity Index and Objective-Scoring AD (Spearman correlations, $P < .0001$). There were significant and stepwise increases of T-scores for all item banks with increasing patient-reported global severity (Wilcoxon rank sum test, $P < .0001$). However, there was limited ability to discriminate between the lowest or highest 2 levels of AD or itch severity. Item banks showed good internal consistency (Cronbach α, 0.91-0.95). No differential item functioning was identified by age, sex, race/ethnicity, or educational level. There were floor effects for total scores, particularly in almost clear/mild AD or itch.

Limitations: Single-center study.

Conclusions: PIQ item bank short forms showed good content and construct validity and are feasible for potential use in clinical trials and practice. (J Am Acad Dermatol 2020;82:1174-80.)

Key Words: atopic dermatitis; burden; eczema; itch; patient-reported outcomes; quality of life; severity; pruritus.

From the Department of Dermatology,[a] Department of Medical Social Sciences,[b] and Department of Preventive Medicine, Feinberg School of Medicine at Northwestern University, Chicago[c]; Northwestern Medicine Multidisciplinary Eczema Center, Chicago[d]; Department of Dermatology, The George Washington University School of Medicine and Health Sciences, Washington, DC[e]; and Department Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago.[f]

Funding sources: Supported by the Agency for Healthcare Research and Quality grant K12 HS023011, the Dermatology Foundation, and a research grant from Galderma.

Conflicts of interest: None disclosed.

IRB approval status: Reviewed and approved by the IRBs of Northwestern University.

Accepted for publication November 22, 2019.

Reprints not available from the authors.

Correspondence to: Jonathan I. Silverberg, MD, PhD, MPH, The George Washington University School of Medicine and Health Sciences, Department of Dermatology, Ste 2B-425 2150 Pennsylvania Ave, Washington, DC 20037. E-mail: JonathanISilverberg@gmail.com.

Published online December 5, 2019.

0190-9622/$36.00

https://doi.org/10.1016/j.jaad.2019.11.057

Itch is the most common[1-3] and burdensome symptom in atopic dermatitis (AD).[4] There are different approaches for measuring itch, each with pros and cons. A systematic review found considerable overlap between the burden of chronic itch[5] and AD in particular.[6] Furthermore, a study found that the ItchyQOL and 5-Dimensions of Itch (5D-Itch) scale had reasonable measurement properties for assessing the burden of AD.[7] This suggests that measures of itch-related quality of life (QOL) impact are conceptually appropriate for assessing the burden of AD. However, several deficiencies were observed with existing itch-related QOL scales. Patient-reported outcome (PRO) measures such as the ItchyQOL and 5D-Itch may not capture the full extent of the patient burden from itch[8] and suffer from limited or undocumented content, structural, and/or cross-cultural validity in AD.[7,9]

CAPSULE SUMMARY

- This study showed that Patient-Reported Outcomes Measurement Information System (PROMIS) Itch Questionnaire (PIQ) had good content; concurrent, convergent, and discriminant validity; and feasibility, with floor effects observed in almost clear/mild disease.
- PIQ item bank short forms all appear to have sufficient validity and feasibility to be used as assessments of burden in adults with atopic dermatitis in clinical practice.

We recently developed the Patient-Reported Outcomes Measurement Information System (PROMIS) Itch Questionnaire (PIQ), a novel suite of PRO measures for itch.[10] The PIQ includes 4 item banks that assess different aspects of the burden of chronic itch. Bank 1 measures itch interference; bank 2, mood and sleep; bank 3, clothing and physical activity; and bank 4, scratching behavior. The PIQ item banks had good concurrent, convergent, and discriminant validity with itch intensity, good internal consistency, and no floor or ceiling effects among US adults with chronic itch. However, little is known about the measurement properties of PIQ item banks across different pruritic disorders.

The PIQ was developed by using the PROMIS methodology to improve on the limitations of previous instruments. We hypothesized that the PIQ item banks are valid for assessing QOL impact in AD. In this study, we sought to determine the content validity, construct validity, internal consistency, floor or ceiling effects, differential item functioning (DIF), and feasibility of PIQ item banks for assessing the burden of itch in adult AD.

METHODS

Content validity

Content validity was assessed with probing cognitive interviews in 12 adults with chronic itch. To reduce respondent burden and fatigue, items were divided so that each item was reviewed by 3 different patients. Multiple questions were asked pertaining to the interpretation, difficulty, and relevance of the items.

Validation study design

A prospective, dermatology practice–based study of adults (≥18 years old) was performed with AD as defined by the Hanifin-Rajka diagnostic criteria.[6] Exclusion criteria included those without a definite diagnosis of AD and those unwilling or unable to complete assessments. Virtually all (>99%) patients who were invited agreed to participate. Patients received standard-of-care follow-up and treatment, including emollients, prescription topical or systemic therapy, and/or phototherapy, where appropriate.

Self-administered questionnaires were completed by patients of the eczema clinic at an academic medical center before they were seen by a clinician. Questionnaires included the PIQ 8-item short forms (SFs) for item banks 1 (itch interference), 2 (mood and sleep), and 3 (clothing and physical activity); the 5-item SF for bank 4 (scratching behavior) (Supplemental Methods; available at https://data.mendeley.com/datasets/publish-confirmation/bnj8znmdwp/1); self-reported severity of AD (*Would you describe your atopic dermatitis or eczema as clear, almost clear, mild, moderate, or severe?*)[11]; Numeric Rating Scale (NRS) (1 question; range, 0-10) and Visual Rating Scale (VRS) (1 question; range, *had no itch* to *very severe*) for worst itch and average itch; frequency of itch; and Patient-Oriented Eczema Measure (POEM) (7 questions; range, 0-28).

Patients were assessed with full-body skin examination by a dermatologist (JS). The Eczema Area and Severity Index (4 signs [erythema, excoriation, swelling, lichenification] on 4 body sites; range, 0-72)[12] and objective component of Scoring AD (oSCORAD) (6 signs [erythema, excoriation, swelling, oozing/crusting, lichenification, dryness] on 8 body sites, no symptoms; range, 0-83)[13] were the clinically reported outcomes (ClinROs) examined. Surveys were administered between June 2017 and February 2019. The study was approved by the institutional review boards of Northwestern

Abbreviations used:

5D-Itch:	5 Dimensions of Itch
AD:	atopic dermatitis
ClinRO:	clinically reported outcome
DIF:	differential item functioning
DLQI:	Dermatology Life Quality Index
EASI:	Eczema Area and Severity Index
NRS:	Numeric Rating Scale
oSCORAD:	objective component of Scoring Atopic Dermatitis
PIQ:	Patient-Reported Outcomes Measurement Information System Itch Questionnaire
POEM:	Patient Oriented Eczema Measure
PRO:	patient-reported outcomes
QOL:	quality of life
SCORAD:	Scoring Atopic Dermatitis
SD:	standard deviation
SF:	short form
VRS:	Visual Rating Scale

University, and informed consent was obtained electronically.

Statistical analysis

Summary statistics were estimated for baseline population characteristics. Concurrent validity of T-scores for SF-8 and SF-4 for banks 1 through 3 and SF-5 for bank 4 was established by using Spearman correlations with each other. Convergent validity was established by using Spearman correlations of T-scores with other PROs and ClinROs. Correlation coefficients scores of 0.70 or greater were considered strong, 0.50 to 0.69 were moderate, and 0.30 to 0.49 were weak. We hypothesized that there would be moderate to strong positive correlations between PIQ scores and itch severity.

Criterion validity was determined with the Kruskal-Wallis test. Discriminant construct validities of item bank T-scores were established by using logistic regression models with either self-reported global AD severity, VRS-worst itch, or VRS-average itch as the ordinal dependent variables.[14] The independent variables were item bank T-scores. Discriminant validity was determined by using the area under the curve. Area under the curve scores of 0.90 or greater were considered excellent; 0.80 to 0.89, good; 0.70 to 0.79, fair; less than 0.70, poor; and less than 0.60, fail.[15] Ordinal logistic regression was used because the data met the proportional odds assumption (score test, $P < .01$).

DIF was analyzed by age (<50 vs ≥50 years), education (high school graduate or less, more than high school), sex, and race/ethnicity (white, nonwhite). DIF occurs when respondents from different groups have differing probabilities of success on an item, after overall ability based on their total scores is controlled for. If DIF is present, item responses may be related to both the underlying trait being measured and other factors. DIF was tested by using ordinal logistic regression with items flagged based on $P < .01$ and $R^2 > 0.02$.

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.[16,17] Feasibility was examined by survey completion rates and time to completion. The statistical analyses were performed in SAS, version 9.4.3 (SAS Institute, Cary, NC). Missing values were encountered in 0.1% of respondents or less for all analyzed variables. Complete case analysis was performed; that is, missing values were excluded. A 2-sided P value of .05 was considered statistically significant.

RESULTS

Patient characteristics

Content validity was assessed in 12 adults with chronic itch (mean ± standard deviation [SD] age, 61.8 ± 10.3 years; 50% female; 92% white), including 7 (58.3%) with AD, 1 each (8.3%) with nummular eczema, idiopathic pruritus, and cutaneous T-cell lymphoma, and 2 (16.7%) with lamellar ichthyosis.

Construct validity and feasibility were assessed in 239 adults (ages 18.3-97.9 years) with AD, which included 149 women (62.3%) and 139 self-reported white individuals (57.9%); the mean age ± SD at enrollment was 46.8 ± 18.2 years. Baseline characteristics of AD severity are presented in Table I.

Content validity

All items except 1 were properly interpreted by patients (Supplemental Table I; available at https://data.mendeley.com/datasets/publish-confirmation/bnj8znmdwp/1); … *I was more sedentary* was interpreted as refraining from physical activity and was revised to … *I sat around more than usual*. Virtually all reported that the questions were clear and that they were confident or very confident answering the questions. Three items were revised for improved clarity or translatability. Some patients found that some items were not relevant to their experience with itch but appreciated how they could be relevant to others' experiences with itch. All patients reported that the item response options made sense, were easy to respond to, and provided enough but not too many response options.

Concurrent validity

All item-bank T-scores based on 4-item and 8-item SFs had moderate to strong correlations with each

Table I. Participant characteristics for the construct validity cohort

Variable	Value
Demographics (n = 239)	
Age, y	
Mean ± SD	46.8 ± 18.2
Range	18.3-97.9
Female sex, n (%)	149 (62.3)
Race/ethnicity, n (%)	
White	139 (57.9)
African American/black	33 (13.8)
Hispanic	19 (7.9)
Asian	39 (16.3)
Multiracial/other	10 (4.2)
Level of education, n (%)	
High school or less	21 (8.8)
Greater than high school	218 (91.2)
Patient-reported outcomes (n = 239)	
PROMIS Itch Questionnaire T-score, median (min, max)	
Item bank 1: itch interference SF-4	39.5 (35.8-64.2)
Item bank 1: itch interference SF-8	42.9 (35.2-75.6)
Item bank 2: mood and sleep SF-4	41.7 (30.2-68.5)
Item bank 2: mood and sleep SF-8	43.9 (30.8-74.8)
Item bank 3: physical activity and clothing SF-4	40.7 (32.4-69.9)
Item bank 3: physical activity and clothing SF-8	42.5 (33.9-75.3)
Item bank 4: scratching behavior	48.0 (32.6-72.8)
NRS worst itch, median (min, max)	6 (0-10)
NRS average itch, median (min, max)	4 (0-10)
POEM, median (min, max)	14 (0-28)
Patient-reported global AD severity, n (%)	
Clear	6 (2.5)
Almost clear	25 (10.5)
Mild	68 (28.5)
Moderate	85 (35.6)
Severe	55 (23.0)
Clinically reported outcomes (n = 151), median (min, max)	
EASI, median	8.0 (0-59.4)
Objective component of SCORAD	30.4 (0-68.6)

AD, Atopic dermatitis; *EASI*, Eczema Area and Severity Index; *NRS*, Numeric Rating Scale; *POEM*, Patient-Oriented Eczema Measure; *PROMIS*, Patient-Reported Outcomes Measurement Information System; *SCORAD*, Scoring Atopic Dermatitis; *SD*, standard deviation.

other (Spearman correlations, $P < .001$ for all) (Fig 1). In particular, scores from item banks 1 and 2 had the strongest correlation.

Convergent validity

The T-scores for item banks 1, 2, 3, and 4 in general had moderate correlations with all assessments of itch severity (NRS and VRS for worst or average itch, frequency of itch, and SCORAD-itch), moderate to strong correlations with AD symptoms (POEM), and weak to moderate correlations with ClinROs (Eczema Area Severity Index [EASI] and oSCORAD) ($P < .0001$ for all). Numerically, itch severity and frequency had the strongest correlation with T-scores from item bank 1 and the lowest correlation with item bank 4. ClinROs (EASI and oSCORAD) had similar correlations with T-scores from all item banks.

Discriminative validity

There were significant and stepwise increases of T-scores for all item banks at each level of patient-reported global severity (Wilcoxon rank sum test, $P < .0001$ for all); however, there were no major differences in item bank 1 T-scores between those with almost clear or mild self-reported global AD severity (Supplemental Fig 1, *A-D*; available at https://data.mendeley.com/datasets/publish-confirmation/bnj8znmdwp/1). In addition, all item bank T-scores had significant and stepwise increases with each level of worsening worst (Supplemental Fig 1, *E-H*) or average VRS-itch (Supplemental Fig 1, *I-L*).

The area under the curve for item bank T-scores were fair for distinguishing between AD severity and worst and average VRS-itch levels (Supplemental Table II; available at https://data.mendeley.com/datasets/publish-confirmation/bnj8znmdwp/1). None of the item banks was able to discriminate between the lowest 2 or highest 2 levels of AD and itch severity.

Internal consistency

Within each factor, correlations among the included items were positive (range of Pearson *r* for bank 1, 0.43-0.81; bank 2, 0.46-0.94; bank 3, 0.56-0.86; and bank 4, 0.57-0.80) and statistically significant ($P < .0001$ for all). The Cronbach α values were 0.93, 0.94, 0.95, and 0.91, and the item total correlations ranged from 0.65 to 0.84, 0.71 to 0.85, 0.69 to 0.90, and 0.73 to 0.88 for factors 1, 2, 3, and 4, respectively.

Floor or ceiling effects

Overall, the proportions of patients with the lowest T-scores for item banks 1, 2, 3, and 4 (39.8%, 25.1%, 33.3%, and 18.3%, respectively) were greater than 15%, indicating there were floor effects (Supplemental Fig 2, *A-D*; available at https://data.mendeley.com/datasets/publish-confirmation/bnj8znmdwp/1). The proportion of patients with the

PIQ Item bank		PIQ Item bank				NRS-itch		VRS-itch		POEM	Frequency of itch (POEM)	EASI	Objective SCORAD	SCORAD-itch
		Bank-1	Bank-2	Bank-3	Bank-4	Worst	Average	Worst	Average					
SF-4	Bank-1	1.00	0.75	0.71	0.68	0.45	0.52	0.45	0.43	0.50	0.43	0.40	0.40	0.46
	Bank-2		1.00	0.76	0.76	0.51	0.57	0.50	0.48	0.58	0.54	0.46	0.48	0.54
	Bank-3			1.00	0.69	0.49	0.53	0.48	0.46	0.50	0.53	0.49	0.47	0.50
SF-8	Bank-1	1.00	0.82	0.79	0.72	0.52	0.60	0.52	0.49	0.54	0.47	0.47	0.45	0.51
	Bank-2		1.00	0.77	0.79	0.54	0.60	0.53	0.50	0.60	0.55	0.48	0.49	0.55
	Bank-3			1.00	0.70	0.49	0.54	0.47	0.46	0.52	0.44	0.50	0.48	0.52
	Bank-4				1.00	0.49	0.57	0.44	0.47	0.64	0.57	0.45	0.46	0.44

* $P < .0001$ for all

Fig 1. Spearman correlations between assessments of PROMIS Itch Questionnaire item bank T-scores, and severity of itch and atopic dermatitis in adult patients with atopic dermatitis.* *PIQ*, Patient-Reported Outcomes Measurement Information System Itch Questionnaire; *PROMIS*, Patient-Reported Outcomes Measurement Information System; *SF*, short form.

lowest T-scores for item banks 1, 2, 3, and 4 was lower in patients with moderate to severe AD and moderate to very severe worst or average itch (Supplemental Fig 2, *E-G*). However, there were no ceiling effects for item banks 1, 2, 3, or 4 (0.4%, 1.7%, 1.3%, and 2.5%, respectively).

Differential item functioning

Several items showed statistically significant ($P < .01$) uniform or nonuniform DIF by age (question 10, $R_U^2 = 0.05$; question 11, $R_U^2 = 0.03$; question 12, $R_U^2 = 0.03$; question 28, $R_N^2 = 0.03$), sex (question 23, $R_U^2 = 0.03$), race (question 4, $R_U^2 = 0.06$, question 17, $R_U^2 = 0.03$), or education level (question 1, $R_N^2 = 0.05$; question 19, $R_N^2 = 0.03$). However, the magnitudes were very low, and no items met the a priori exclusion criteria.

Feasibility

One respondent did not complete all items in item banks 1 and 2. However, all items in item banks 3 and 4 were completed by all participants. The mean ± SD time to completion of all item banks was 1.8 ± 2.1 minutes, with a median (min, max) completion time of 2 (<1 to 11) minutes. There were no significant differences of completion time by age (Wilcoxon rank sum test, $P = .85$), sex ($P = .21$), race/ethnicity ($P = .17$), level of education ($P = .39$), patient-reported global AD severity ($P = .28$), or POEM scores ($P = .56$).

Only 30 (12.5%) respondents gave the same response for all items for bank 1, all of which were for the responses of *never*; 74.0% of these reported having no or only mild itch. No patients gave the same responses for all items for banks 2, 3, or 4.

DISCUSSION

This study showed that PIQ item bank SFs had good content, concurrent, convergent, discriminative, and cross-cultural validity; internal consistency; and feasibility. All item banks showed floor effects, particularly in patients with mild AD and/or itch, but there were no ceiling effects. Thus, the SFs may perform best in patients with moderate to severe AD and/or itch. In contrast, the Dermatology Life Quality Index (DLQI), ItchyQOL, and 5D-Itch were not found to have floor or ceiling effects in adults with AD.[9] Item bank T-scores correlated well with multiple PROs and ClinROs but had the strongest correlations with NRS for average itch. Similar results were found for the DLQI, ItchyQOL, and 5D-Itch.[9] Overall, PIQ item bank SFs showed fair discriminative ability. In particular, PIQ SFs showed good to excellent ability to discriminate 2- or 3-point differences of AD and itch severity but poor ability to discriminate between the lowest or highest 2 severity groups. Although these properties were not quantitatively assessed previously, the DLQI, ItchyQOL, and 5D-Itch appear to have similar limitations.[9] The 8-item SFs had slightly better measurement properties than the 4-item SFs for banks 1 through 3. These results are consistent with previous studies showing that PIQ item banks had good convergent and discriminant validity with itch intensity and internal consistency and had no significant floor or ceiling effects in adults with chronic itch.[10] Taken together, PIQ item bank SFs appear to be valid and feasible for assessing QOL impact secondary to itch in adults with AD and other causes of chronic itch.

There are different approaches to assessing QOL impairment in AD, including AD- and symptom-specific assessments. The DLQI[18] assesses QOL across dermatologic disease in general and was recently selected by the Harmonizing Outcome Measures in Eczema group as the preferred QOL assessment in clinical trials of adult AD. The ItchyQOL[19,20] and 5D-Itch scales[21] assess QOL-impact related to itch. Both were recently found to be valid for assessing QOL impact in AD.[7] Taken together, it appears that assessing QOL impact secondary to itch is a valid approach to assessing the burden of AD. Moreover, the PIQ has several advantages over existing QOL measures. First, the PIQ was developed by using rigorous PROMIS

methodology in accordance with the guidance from the US Food and Drug Administration.[22] One particular aspect is the use of a patient-centric approach to have content validity for a PRO. The PIQ was developed based on a conceptual model that incorporated patient-centric qualitative interviews with 33 patients[8]; the 5D-Itch was not. Second, the PIQ can be assessed with computerized adaptive testing, which reduces the number of questions asked and improves efficiency; the DLQI, 5D-Itch, and ItchyQOL cannot. Third, uniform and nonuniform DIFs were previously observed for the DLQI, ItchyQOL, and 5D-Itch in adults with AD,[7,23] but not with the PIQ. This suggests that the PIQ has better cross-cultural validity. However, we did not directly compare the measurement properties of the PIQ with other established QOL assessments. Future studies are needed to determine how the measurement properties of the PIQ compare with those of other QOL assessments in AD and chronic itch in general.

The PIQ was time efficient, was easy to interpret for patients, and may be integrated into day-to-day practice. All 29 questions for the 8-item SFs for item banks 1 through 3 and the 5-item SF for item bank 4 were completed in approximately 2 minutes, which is acceptable or adequate for use in clinical practice (<3 or <3-5 minutes, respectively).[24] SFs for PIQ item banks can be easily be implemented in the clinical practice setting with either paper-based or electronic administration. Computerized adaptive testing can be administered in the PROMIS Assessment Center or by using an application programming interface for REDCap (Vanderbilt University, Nashville, TN) and other research platforms. PIQ SFs can be used in conjunction with other assessments of itch (eg, NRS or VRS itch) or AD-specific measures (eg, POEM and EASI).

This study has several strengths, including examination of content validity and good representation across sex, race/ethnicity, and AD severity; testing of multiple itch assessments; and use of multiple PROs and ClinROs when examining the psychometric properties. There are some limitations. Patients were recruited from a single academic center, which may limit generalizability. We did not assess test-retest reliability. Future studies are needed to address these points.

In conclusion, PIQ SFs have good content, construct, discriminant and cross-cultural validity; internal consistency; and feasibility to assess AD in clinical practice, particularly moderate to severe AD. These instruments may be incorporated into the assessment of patients with AD and chronic itch. They provide important information about the burden of itch that can guide therapeutic decision making.

The Northwestern Medicine Enterprise Data Warehouse was supported, in part, by the Northwestern University Clinical and Translational Science Institute, funded, in part, by grant UL1TR000150 from the National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The Clinical and Translational Science Award is a registered trademark of the US Department of Health and Human Services.

REFERENCES

1. Wittkowski A, Richards HL, Griffiths CE, Main CJ. Illness perception in individuals with atopic dermatitis. *Psychol Health Med*. 2007;12:433-444.
2. Dawn A, Papoiu AD, Chan YH, Rapp SR, Rassette N, Yosipovitch G. Itch characteristics in atopic dermatitis: results of a web-based questionnaire. *Br J Dermatol*. 2009;160:642-644.
3. Simpson EL, Bieber T, Eckert L, et al. Patient burden of moderate to severe atopic dermatitis (AD): insights from a phase 2b clinical trial of dupilumab in adults. *J Am Acad Dermatol*. 2016;74:491-498.
4. Silverberg JI, Gelfand JM, Margolis DJ, et al. Patient burden and quality of life in atopic dermatitis in US adults: a population-based cross-sectional study. *Ann Allergy Asthma Immunol*. 2018;121:340-347.
5. Kantor R, Dalal P, Cella D, Silverberg JI. Research letter: impact of pruritus on quality of life—a systematic review. *J Am Acad Dermatol*. 2016;75:885-886.e4.
6. Hanifin J, Rajka G. Diagnostic features of atopic eczema. *Acta Dermatovenereolo*. 1980;92:44-47.
7. Patel KR, Singam V, Vakharia PP, et al. Measurement properties of three assessments of burden used in atopic dermatitis in adults. *Br J Dermatol*. 2019;180:1083-1089.
8. Silverberg JI, Kantor RW, Dalal P, et al. A comprehensive conceptual model of the experience of chronic itch in adults. *Am J Clin Dermatol*. 2018;19:759-769.
9. Schoch D, Sommer R, Augustin M, Stander S, Blome C. Patient-reported outcome measures in pruritus: a systematic review of measurement properties. *J Invest Dermatol*. 2017;137:2069-2077.
10. 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. *J Invest Dermatol*, in press. 2019. https://doi.org/10.1016/j.jid.2019.08.452.
11. Vakharia PP, Chopra R, Sacotte R, et al. Validation of patient-reported global severity of atopic dermatitis in adults. *Allergy*. 2018;73:451-458.
12. 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;10:11-18.
13. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology*. 1993;186:23-31.
14. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning*. 2001;45:171-186.

15. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:283-298.
16. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42.
17. 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:e007765.
18. 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;159:997-1035.
19. Desai NS, Poindexter GB, Monthrope YM, Bendeck SE, Swerlick RA, Chen SC. A pilot quality-of-life instrument for pruritus. *J Am Acad Dermatol*. 2008;59:234-244.
20. ItchyQol. *A Pruritus-Specific Quality of Life Instrument*. Atlanta, GA: Emory University; 2009.
21. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D Itch Scale: a new measure of pruritus. *Br J Dermatol*. 2010;162:587-593.
22. FDA Center for Drug Evaluation and Research, FDA Center for Biologics Evaluation and Research, FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.
23. Silverberg JI, Gelfand JM, Margolis DJ, et al. Validation and interpretation of Short Form 12 and comparison with Dermatology Life Quality Index in atopic dermatitis in adults. *J Invest Dermatol*. 2019;139:2090-2097.e3.
24. Schmitt J, Langan S, Williams HC, European Dermato-Epidemiology N. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol*. 2007;120:1389-1398.