

Article type : Original Article: Epidemiology and Genetics

Editor : Thomas Bieber

Title: Validation of patient-reported global severity of atopic dermatitis in adults.

Running Header: Severity of atopic dermatitis.

**Paras P. Vakharia, PharmD¹, Rishi Chopra, MS¹, Ryan Sacotte, BS¹, Neha Patel, BS¹,
Supriya Immaneni, BS¹, Takeshia White, CNA¹, Robert Kantor, MD¹, Derek Y. Hsu,
MD¹, Jonathan I. Silverberg, MD, PhD, MPH^{2,3}**

1. Department of Dermatology, Feinberg School of Medicine at Northwestern University, Chicago, IL USA
2. Departments of Dermatology, Preventive Medicine and Medical Social Sciences, Feinberg School of Medicine at Northwestern University, Chicago, IL USA
3. Northwestern Medicine Multidisciplinary Eczema Center, Chicago, IL USA

Key Words: atopic dermatitis; eczema; severity

Abbreviations used: AD=atopic dermatitis; DLQI = Dermatology Life Quality Index; EASI = Eczema Area and Severity Index; NRS = Numeric Rating Scale; POEM = Patient Oriented Eczema Measure; oSCORAD = objective Scoring Atopic Dermatitis; SCORAD = Scoring Atopic Dermatitis

Corresponding author:

Jonathan I. Silverberg, MD, PhD, MPH
Northwestern University Feinberg School of Medicine
Department of Dermatology
Suite 1600
676 N. St Clair St.
Chicago, IL 60611
Telephone: (312) 503-4958
Fax: (312) 695-0664
JonathanISilverberg@Gmail.com

Acknowledgements:

JI Silverberg had full access to all the data in the study and takes responsibility for the integrity of the data and accuracy of the data analysis.

Study concept and design: JI Silverberg

Acquisition of Data: JI Silverberg, P Vakharia, R Chopra, R Sacotte, N Patel, S Immaneni, T White, R Kantor, D Hsu

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/all.13309

This article is protected by copyright. All rights reserved.

Analysis and interpretation of data: JI Silverberg, P Vakharia, R Chopra, R Sacotte, N Patel, S Immaneni, T White, R Kantor, D Hsu
Drafting of the manuscript: JI Silverberg
Critical revision of the manuscript for important intellectual content: JI Silverberg, P Vakharia, R Chopra, R Sacotte, N Patel, S Immaneni, T White, R Kantor, D Hsu
Statistical analysis: JI Silverberg
Administrative technical or material support: None
Study supervision: None
Financial disclosures: None
Funding Support: This publication was made possible with support from the Agency for Healthcare Research and Quality (AHRQ), grant number K12 HS023011, and the Dermatology Foundation.

Conflicts of interest: None
Funding/Sponsor was involved? No
Design and conduct of the study: Yes__ No_X_
Collection, management, analysis and interpretation of data: Yes__ No_X_
Preparation, review, or approval of the manuscript: Yes__ No_X_
Decision to submit the manuscript for publication: Yes__ No_X_

Abstract

Background: Atopic dermatitis (AD) is associated with a heterogeneous presentation and clinical course. There is a lack of simple and validated severity assessments that are feasible for clinical practice and epidemiological research.

Objectives: We sought to validate patient-reported global AD severity in adults.

Methods: We performed a prospective dermatology practice-based study using questionnaires and evaluation by a dermatologist (n=265).

Results: At baseline and follow-up, patient-reported global AD severity significantly correlated with oSCORAD (Spearman rho=0.56 and 0.49), SCORAD (0.64 and 0.56), EASI (0.56 and 0.50), BSA (0.52 and 0.45), NRS-itch (0.60 and 0.53), POEM (0.50 and 0.48), and DLQI (0.50 and 0.49) ($P<0.0001$ for all). Patient-reported moderate and severe AD vs. mild AD were associated with significantly higher oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM and DLQI ($P<0.0001$ for all). There was moderate concordance between patient-reported AD severity (mild, moderate and severe) and previously developed severity strata for oSCORAD (kappa=0.39), SCORAD (kappa=0.47), EASI (kappa=0.37), NRS-itch (kappa=0.49), POEM (kappa=0.37), and DLQI (kappa=0.40). Among patients with severe disease at baseline, those who reported mild or moderate disease on follow-up had significantly greater absolute reductions of oSCORAD (-23.4/-9.7/-1.8), SCORAD (-33.0/-13.2/-2.3), EASI (-17.1/-9.8/-3.2), BSA (-46%/-15%/-4%), NRS-itch (-5/-2/0), POEM (-5/-2/0), and DLQI (-8/-6/-1) than those who continued to report severe disease (Kruskal-Wallis, $P\leq 0.0003$ for all).

Conclusions: Patient-reported AD severity appears to be sufficiently valid for assessing AD severity in the clinical and epidemiological setting.

Introduction

Atopic dermatitis (AD) is a chronic, inflammatory skin disease, associated with a heterogeneous presentation and clinical course. There is no standardized definition of AD severity. Rather, AD severity encompasses the converging effects of the intensity and extent of lesions, chronicity and flares, intensity of itch, pain, sleep disturbance and mental health symptoms, and the impact activities of daily living and quality of life (QOL). Hence, no single objective assessment or patient-reported outcome (PRO) fully captures all of these aspects of disease severity. Moreover, there is a lack of simple and validated severity assessments that are feasible for clinical practice and epidemiological research. One approach worth consideration is the use of a single-question assessment for patient-reported global severity of AD. This approach is short, intuitive and feasible, though its validity has never been established.

Much of our knowledge of the epidemiology of AD in the United States comes from a number of US population-based studies, such as the National Survey of Children's Health (NSCH) (1-12). These studies typically employ questionnaires with either self- or caregiver-report of various disorders and their severity. Given the scale of and number of questions used in these studies, it is important that disease defining or classifying approaches be as efficient as possible. A single question for classifying the self- or caregiver-reported severity of AD would certainly be efficient. However, the validity of this approach has not been established. We sought to determine the performance characteristics/or validity of a single question in determining the severity of adult AD.

Methods

Study design

We performed a prospective, dermatology practice-based, observational study to determine the validity of patient-reported AD severity. Self-administered questionnaires were completed by adult (≥ 18 years) patients of the eczema clinic prior to their encounter. Patients were then evaluated with a medical history and total body skin examination by a dermatologist (JS). Subjects were enrolled between September, 2014 and January, 2017. The study was approved by the institutional review boards of Northwestern University and informed consent was waived.

Assessments for AD

The diagnosis of AD was confirmed using the Hanifin and Rajka diagnostic criteria(13). The objective assessments of AD were the Eczema Area and Severity Index (EASI) [range: 0-72] (14), objective-Scoring AD (oSCORAD) [range: 0-83], SCORAD [range: 0-103] (15), and body surface area (BSA) of involved skin [range: 0-100%]. Patients completed a questionnaire that included the question “Would you describe your atopic dermatitis or eczema as mild, moderate, or severe)?”, as well as PRO related to AD, including numeric rating scale (NRS) for itch [range: 0-10], Patient-Oriented Eczema Measure (POEM) [7 questions; range: 0-28] (16) and Dermatology Life Quality Index (DLQI) [10 questions; range: 0-30] (17).

Data processing and statistical methods

All data analyses and statistical processes were performed using SAS version 9.4 (SAS Institute). Statistical significance was determined based on two-sided P-values <0.05 .

Baseline AD severity scores were not normally distributed. Therefore, correlations between baseline patient-reported global AD severity and other assessments were performed using non-parametric Spearman correlations. Mann-Whitney U tests were used to compare oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM, and DLQI scores between patient-reported mild, moderate and severe AD. Patient-reported AD severity was also compared with previously reported severity strata combined into mild, moderate and severe for oSCORAD (0-14.9=mild, 15.0-40.9=moderate and 41-83=severe) (18), SCORAD (0-25.9=mild, 26.0-50.9=moderate, 51-103=severe) (19), EASI (0-7.0=mild, 7.1-21.0=moderate, 21.1-72=severe) (20), NRS-itch (0-3=mild, 4-6=moderate, 7-10=severe) (21), POEM (0-7=mild, 8-16=moderate, 17-28=severe) (16), and DLQI (0-5=mild, 6-10=moderate, 11-30=severe) (22). Spearman correlations and Fisher's exact test were performed and kappa coefficients were determined for comparisons of patient-reported AD severity and the abovementioned severity strata.

Most patients were assessed at ≥ 1 follow-up visit. To determine the responsiveness of patient-reported AD severity, we determined the proportion of patients with improvement of patient-reported severity. Mann-Whitney U tests were used to compare the absolute differences of oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM, and DLQI scores at follow-up between patients with vs. without improvement of their moderate and/or severe AD at baseline. Chi square tests were used to compare the proportion of patients who achieved a minimal clinically important difference (MCID) of oSCORAD (8.2 points) (23), SCORAD (8.7 points) (23), EASI (6.6 points) (23), NRS-itch (3 points) (24), POEM (3.4

points) (23) and DLQI (4 points) (25) scores at follow-up between patients with vs. without improvement of their moderate and/or severe AD at baseline.

Missing values occurred for some variables in >5% of subjects. Several approaches were therefore used to address missing values. First, complete data analysis was performed, i.e. subjects with missing data were excluded. Second, sensitivity analyses using multiple imputations (n=5 imputations) were performed for correlations and concordance of patient-reported global AD severity with other outcomes (continuous) and severity strata thereof (ordinal), respectively.

Results

Patient characteristics

Overall, 265 adults (ages 18-97 years) were included in the study, of which 192 were evaluated at follow-up. The cohort included 162 females (61.1%) and 170 whites (64.2%). The mean $\pm$ std. dev. age at enrollment was 42.5 ± 17.5 years and age of patient-reported AD onset was 21.2 ± 22.7 years. The mean $\pm$ std. dev. BSA of involvement was $33.2\% \pm 31.8\%$, NRS-itch was 6.3 ± 2.9 , oSCORAD was 29.6 ± 16.6 , SCORAD was 40.0 ± 19.7 , EASI was 13.7 ± 14.5 , POEM was 11.5 ± 7.4 , and DLQI was 10.7 ± 7.6 .

Patient-reported global AD severity and other AD assessments

In complete case analyses, at baseline, patient-reported global AD severity significantly correlated with oSCORAD (Spearman rho = 0.56), SCORAD (0.64), EASI (0.56), BSA (0.52), NRS-itch (0.60), POEM (0.50), and DLQI (0.50) ($P < 0.0001$ for all). At least one follow-up evaluation was available for 192 (75.0%) patients, with a variable follow-up schedule ranging from 2-26 visits (median: 4 visits) over 0.04-2.5 years (median: 0.6 years). Similarly at follow-up, patient-reported AD severity significantly correlated with

oSCORAD (0.49), SCORAD (0.56), EASI (0.50), BSA (0.45), NRS-itch (0.53), POEM (0.48), and DLQI (0.49) ($P < 0.0001$ for all).

In sensitivity analyses using multiple imputations, patient-reported global AD severity similarly correlated with oSCORAD (Spearman $\rho = 0.54$), SCORAD (0.59), EASI (0.52), BSA (0.48), NRS-itch (0.54), POEM (0.49), and DLQI (0.52) ($P < 0.0001$ for all).

Patient-reported moderate and severe AD vs. mild AD were associated with significantly higher oSCORAD (median [IQR, range]: 25.3 [17.3, 69.4] and 34.5 [22.7, 82.8] vs. 14.9 [12.8, 47.7]), SCORAD (32.7 [19.8, 81.3] and 48.4 [24.8, 89.8] vs. 20.6 [16.5, 58.2]), EASI (5.7 [9.0, 54.3] and 14.8 [19.3, 66.7] vs. 2.0 [3.4, 27.6]), BSA (15.5 [25.5, 100.0] and 39.0 [57.5, 100.0] vs. 4.5 [11.9, 100.0]), NRS-itch (5 [3, 10] and 8 [3, 10] vs. 3 [4, 10]), POEM (11 [10, 28] and 17 [11.5, 26] vs. 5 [8, 28]) and DLQI (7 [7, 28] and 13 [13, 30] vs. 3 [5, 24]) (Kruskal-Wallis, $P < 0.0001$ for all) (Figure 1).

The correlations between patient-reported global AD severity and the abovementioned objective assessments and PRO were similar in both men and women, whites and non-whites, ages <40 and ≥ 40 years ($P < 0.0001$ for all; data not shown).

Patient-reported global AD severity and other AD severity strata

There was moderate concordance between patient-reported AD severity (mild, moderate and severe) and previously developed severity strata for oSCORAD (kappa [95% CI]: 0.39 [0.33-0.44]), SCORAD (0.47 [0.42-0.52]), EASI (0.37 [0.32-0.43]), NRS-itch (0.49 [0.44-0.54]), POEM (0.37 [0.24-0.35]), and DLQI (0.40 [0.35-0.46]) (Fisher exact test, $P < 0.001$ for all) (Figure 2).

The associations between patient-reported AD severity and the abovementioned objective assessments and PRO was significant in both men and women, whites and non-whites, ages <40 and ≥ 40 years ($P \leq 0.03$ for all; data not shown).

In sensitivity analyses using multiple imputations, there was similar concordance between patient-reported global AD severity and severity strata for oSCORAD (mean kappa: 0.37), SCORAD (0.43), EASI (0.35), NRS-itch (0.48), POEM (0.35), and DLQI (mild=0-5, moderate=6-10, severe=11-30) (0.39).

Improvement of patient-reported global AD severity and other AD assessments

Among patients with severe disease at baseline, those who reported mild or moderate disease on follow-up had significantly greater reductions of oSCORAD (median [IQR]: -23.4 [29.9] vs. -9.7 [25.8] vs. -1.8 [18.5]; Kruskal-Wallis, $P<0.0001$), SCORAD (-33.0 [36.4] vs. -13.2 [26.7] vs. -2.3 [24.6]; $P<0.0001$), EASI (-17.1 [32.1] vs. -9.8 [23.5] vs. -3.2 [23.0]; $P<0.0001$), BSA (-46% [62.8%] vs. -15% [42.5%] vs. -4.0% [48.1%]; $P<0.0001$), NRS-itch (-5 [6] vs. -2 [4] vs. 0 [3]; $P<0.0001$), POEM (-5 [10] vs. -2 [9] vs. 0 [4]; $P\leq 0.0003$), and DLQI (-8 [11] vs. -6 [10] vs. -1 [6]; $P<0.0001$) than those who continued to report severe disease (Figure 3).

Among patients with moderate disease at baseline, those who reported mild disease on follow-up had significantly greater reductions of oSCORAD (-14.4 [16.4] vs. -7.5 [13.5]), SCORAD (-17.2 [23.3] vs. -7.8 [19.2]), EASI (-8.2 [11.2] vs. -3.0 [8.6]), BSA (-15% [24.5%] vs. -7% [19.5%]), NRS-itch (-2 [4] vs. -1 [2]), POEM (-1 [7] vs. 0 [4]), and DLQI (-3.5 [8.5] vs. 0 [5]) than those who continued to report moderate disease ($P<0.0001$ for all) (Figure 3).

Similarly, among patients with severe disease at baseline, those who reported mild or moderate disease on follow-up significantly higher proportions achieved MCID [95% CI] of oSCORAD (77.0% [70.2-83.8%] vs. 51.2% [40.4-62.1%] vs. 14.6% [8.9-22.2%]), SCORAD (79.7% [73.2-86.2%] vs. 58.5% [47.8-69.2%] vs. 15.2% [9.5-21.0%]), EASI (74.3% [67.3-81.4%] vs. 58.5% [47.8-69.2%] vs. 13.8% [8.3-19.3%]), NRS-itch (62.8% [55.0-70.7%] vs.

46.3% [35.5-57.2%] vs. 6.6% [2.6-10.5%]), POEM (62.8% [54.9-70.7%] vs. 33.3% [23.6-43.1%] vs. 9.1% [4.7-13.6%]), and DLQI (74.5% [67.6-81.4%] vs. 64.8% [55.0-74.7%] vs. 15.0% [9.8-20.2%]) than those who continued to report severe disease (Chi-square, $P < 0.0001$ for all) (Figure 4). Further, among patients with moderate disease at baseline, those who reported mild disease on follow-up significantly higher proportions achieved MCID [95% CI] of oSCORAD (68.4% [56.2-80.6%] vs. 18.7% [10.6-26.8%]), SCORAD (68.4% [56.2-80.6%] vs. 19.8% [11.5-28.1%]), EASI (58.9% [45.9-72.0%] vs. 16.9% [9.0-24.7%]), NRS-itch (38.6% [25.8-51.4%] vs. 9.9% [3.7-16.1%]), POEM (62.8% [54.9-70.7%] vs. 9.1% [4.7-13.6%]), and DLQI (50.0% [37.2-62.8%] vs. 12.0% [6.0-17.9%]) than those who continued to report moderate disease (Chi-square, $P < 0.0001$ for all).

Discussion

The present study demonstrated the validity of patient-reported AD severity, using a simple scoring system of mild, moderate and severe. Patient-reported AD severity correlated well with the objective AD assessments oSCORAD and EASI. EASI and SCORAD have been extensively validated and used frequently to define AD severity in outcome measures (26) and inclusion criteria (27). The Harmonizing Outcome Measures in Eczema (HOME) group found that only oSCORAD and EASI were adequately validated and assessed the four essential signs of AD.(28) However, EASI was preferred over oSCORAD as for the assessment of AD signs in clinical trials.(29) Patient-reported AD severity also correlated well with multiple validated PRO, including NRS-itch, POEM, and DLQI. POEM was recently selected as the preferred assessment for AD symptoms in clinical trials (30). Together, patient-reported global AD severity appears to be a simple, feasible and valid tool for assessing AD severity in clinical trials and epidemiological research of AD.

Accepted Article

Patient-reported AD severity correlated only moderately with all measures, suggesting that no single instrument comprehensively characterizes the patient-burden of AD. This is further illustrated by the discordance observed between severity strata for different outcome measures and patient-reported global AD severity. Some patients reported having severe disease owing to severe itch, but had only mild or moderate disease as judged by EASI or oSCORAD. Thus, patient-reported global AD severity may be an important PRO for assessing the overall burden of disease that accounts for both lesional severity and extent, symptom burden and QOL impact. Patient-reported AD severity had a stronger correlation with NRS-itch than the other validated PRO - POEM and DLQI. Together, these findings suggest that AD severity is determined by multiple aspects of disease, but patients' dominant concerns regarding AD severity are related to the burden of itch.

A decrease of patient-reported AD severity was associated with significant improvements of and a greater proportion of patients achieving a MCID for the abovementioned assessments. The results suggest that this patient-reported global assessment is a valid approach for assessing AD severity in the clinical and epidemiological setting. A similar question was used for caregiver-report of pediatric AD severity in the 2007-2008 National Survey of Children's Health. In that study, 67% of children with AD were reported to have mild, 26% moderate and 7% severe AD (10, 12). The demonstrated validity for patient-reported AD severity in adults suggests that caregiver-reported AD severity may similarly be valid. Nevertheless, future studies are explicitly needed to validate caregiver-report in pediatric cohorts. Another important consideration is that patient-reported AD severity is responsive to treatment. Thus, it is more appropriate for estimating the point-prevalences of disease severity, and may underestimate the true prevalence of moderate-severe disease in a longer assessment period, e.g. 1-year or lifetime prevalence. This is particularly relevant to AD, which can have a highly variable or episodic course. Thus, the

proportion of children with at least some period of moderate-severe AD may be even higher than reported in NSCH.

Interestingly, patient-reported AD severity had a stronger correlation with NRS-itch than POEM, suggesting itch intensity is a major driver of disease severity. The HOME group selected POEM as the preferred PRO to assess AD symptoms in clinical trials (30, 31). POEM demonstrates construct validity, convergent validity, divergent validity, internal consistency, sensitivity to change, and test-retest reliability (32, 33). NRS-itch, as a single question instrument, is validated for rating pruritus (34, 35), shorter, simpler, and more feasible than most PRO for AD symptoms. We agree with the recommendation of the HOME group, and suggest that NRS-itch be used in addition to POEM to assess AD symptoms.

This study has several strengths, including being prospective, well-powered, and comparison with validated objective assessments and PRO. However, the study has limitations. Patient-reported global AD severity was classified into three categories: mild, moderate and severe. This was an a priori decision based on preliminary feedback from participant stakeholders that the categories of clear, almost clear and very severe were less meaningful to patients, particularly those with chronic AD. Thus, the mild stratum encompasses both the clear/almost clear and mild categories, whereas the severe stratum encompasses both the severe and very severe categories previously presented. Multiple patient stakeholders raised concerns that even when their skin appeared clear or to have no apparent eczema, that they still experienced itch, sensory disturbances, e.g. allodynia, and fear of unpredictable flares. The significant correlations of patient-reported global AD severity with other existing scores suggests that it may be a clinically useful outcome measure. Ultimately, future prospective studies would be needed to determine how well it performs in clinical practice and therapeutic decision making. Finally, the study was performed in an urban academic dermatologic setting and may not be generalizable to the

entire US population. Future, larger-scale multicenter studies are needed to confirm the results of this study.

In conclusion, a single question examining patient-reported AD or eczema severity may be sufficiently valid for the assessment of AD severity in adults in the clinical and epidemiological setting. Future studies are underway to demonstrate the validity of caregiver-reported severity in childhood AD.

References

1. Garg N, Silverberg JI. Association between eczema and increased fracture and bone or joint injury in adults: a US population-based study. *JAMA Dermatol* 2015;**151**(1):33-41.
2. Garg NK, Silverberg JI. Eczema is associated with osteoporosis and fractures in adults: a US population-based study. *J Allergy Clin Immunol* 2015;**135**(4):1085-1087 e1082.
3. Shaw TE, Currie GP, Koudelka CW, Simpson EL. Eczema prevalence in the United States: data from the 2003 National Survey of Children's Health. *J Invest Dermatol* 2011;**131**(1):67-73.
4. Silverberg JI. Healthcare Utilization, Patient Costs, and Access to Care in US Adults With Eczema: A Population-Based Study. *JAMA Dermatol* 2015.
5. 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.
6. Silverberg JI, Greenland P. Eczema and cardiovascular risk factors in 2 US adult population studies. *J Allergy Clin Immunol* 2015;**135**(3):721-728 e726.
7. Silverberg JI, Hanifin J, Simpson EL. Climatic factors are associated with childhood eczema prevalence in the United States. *J Invest Dermatol* 2013;**133**(7):1752-1759.
8. Silverberg JI, Hanifin JM. Adult eczema prevalence and associations with asthma and other health and demographic factors: a US population-based study. *J Allergy Clin Immunol* 2013;**132**(5):1132-1138.
9. Silverberg JI, Lee-Wong M, Silverberg NB. Complementary and alternative medicines and childhood eczema: a US population-based study. *Dermatitis* 2014;**25**(5):246-254.
10. 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.
11. Silverberg JI, Simpson EL. Association between obesity and eczema prevalence, severity and poorer health in US adolescents. *Dermatitis* 2014;**25**(4):172-181.
12. Silverberg JI, Simpson EL. Associations of childhood eczema severity: a US population-based study. *Dermatitis* 2014;**25**(3):107-114.
13. Hanifin J, Rajka G. Diagnostic features of atopic dermatitis. *Acta Derm Venereol (Stockh)* 1980;**92**(Suppl):44-47.
14. 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.

15. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. *Dermatology* 1993;**186**(1):23-31.
16. Charman CR, Venn AJ, Ravenscroft JC, Williams HC. Translating Patient-Oriented Eczema Measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. *Br J Dermatol* 2013;**169**(6):1326-1332.
17. 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.
18. Kunz B, Oranje AP, Labreze L, Stalder JF, Ring J, Taieb A. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. *Dermatology* 1997;**195**(1):10-19.
19. Wollenberg A, Oranje A, Deleuran M, Simon D, Szalai Z, Kunz B, et al. ETFAD/EADV Eczema task force 2015 position paper on diagnosis and treatment of atopic dermatitis in adult and paediatric patients. *J Eur Acad Dermatol Venereol* 2016;**30**(5):729-747.
20. Leshem YA, Hajar T, Hanifin JM, Simpson EL. What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis: an interpretability study. *Br J Dermatol* 2015;**172**(5):1353-1357.
21. Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. *Acta Derm Venereol* 2012;**92**(5):497-501.
22. Rogers A, DeLong LK, Chen SC. Clinical meaning in skin-specific quality of life instruments: a comparison of the Dermatology Life Quality Index and Skindex banding systems. *Dermatol Clin* 2012;**30**(2):333-342, x.
23. 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.
24. Reich A, Riepe C, Anastasiadou Z, Medrek K, Augustin M, Szepletowski JC, et al. Itch Assessment with Visual Analogue Scale and Numerical Rating Scale: Determination of Minimal Clinically Important Difference in Chronic Itch. *Acta Derm Venereol* 2016;**96**(7):978-980.
25. Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology* 2015;**230**(1):27-33.
26. Schmitt J, Spuls PI, Thomas KS, Simpson E, Furue M, Deckert S, et al. The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. *J Allergy Clin Immunol* 2014;**134**(4):800-807.
27. Chopra RV, P; Simpson, E; Paller, A; Silverberg, JI. Severity assessments used for inclusion criteria and baseline severity evaluation in atopic dermatitis clinical trials: a systematic review. *JEADV* 2017.
28. Schmitt J, Langan S, Deckert S, Svensson A, von Kobyletzki L, Thomas K, et al. Assessment of clinical signs of atopic dermatitis: a systematic review and recommendation. *J Allergy Clin Immunol* 2013;**132**(6):1337-1347.
29. Chalmers JR, Schmitt J, Apfelbacher C, Dohil M, Eichenfield LF, Simpson EL, et al. Report from the third international consensus meeting to harmonise core outcome measures for atopic eczema/dermatitis clinical trials (HOME). *Br J Dermatol* 2014;**171**(6):1318-1325.
30. Spuls PI, Gerbens LA, Simpson E, Apfelbacher CJ, Chalmers JR, Thomas KS, et al. POEM a core instrument to measure symptoms in clinical trials: a HOME statement. *Br J Dermatol* 2016.

31. Chalmers JR, Simpson E, Apfelbacher CJ, Thomas KS, von Kobyletzki L, Schmitt J, et al. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol* 2016;**175**(1):69-79.
32. 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.
33. Schmitt J, Langan S, Williams HC. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol* 2007;**120**(6):1389-1398.
34. Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. *Acta Derm Venereol* 2012;**92**(5):502-507.
35. Reich A HJ, Ramus M, Ständer S, Szepletowski J. New data on the validation of VAS and NRS in pruritus assessment: minimal clinically important difference and itch frequency measurement. *Acta Derm Venereol* 2011;**91**:636.

Figure Legends

Figure 1. Relationship between patient-reported severity and other outcome measures in atopic dermatitis. Atopic dermatitis severity was divided into mild, moderate and severe. Box-whisker plots and overlaid jitter plots of (A) oSCORAD, (B) SCORAD, (C) EASI, (D) BSA, (E) NRS-itch, (F) POEM, and (G) DLQI are presented for each level of AD severity.

Figure 2. Relationship between patient-reported atopic dermatitis severity and established severity strata for oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM and DLQI. Proportion (95% CI) of subjects with patient-reported mild, moderate or severe AD who also had mild, moderate or severe disease as judged by established severity strata for oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM and DLQI.

Figure 3. Relationship between improvement of patient-reported atopic dermatitis severity and other outcome measures. Percent reduction of oSCORAD, SCORAD, EASI, BSA, NRS-itch, POEM and DLQI are plotted for different patient-reported atopic dermatitis severity levels, stratified by baseline atopic dermatitis severity.

Figure 4. Relationship between improvement of patient-reported atopic dermatitis severity and achieving the minimal clinically important difference (MCID) of other outcome measures. Proportion of patients achieving a MCID of oSCORAD, SCORAD, EASI, NRS-itch, POEM and DLQI are plotted for different patient-reported atopic dermatitis severity levels, stratified by baseline atopic dermatitis severity.

Table 1. Demographics of the study cohort (n=265).

Variable	Value
Gender – freq (%)	
Male	103 (38.9%)
Female	162 (61.1%)
Race/ethnicity – freq (%)	
White, non-Hispanic	170 (64.2%)
Black	34 (12.8%)
Hispanic	12 (4.5%)
Asian	45 (17.0%)
Multiracial/Other	4 (1.5%)
Age at enrollment (yr) – mean $\pm$ std. dev	42.5 $\pm$ 17.5
Age of AD onset (yr) – freq (%)	
Childhood (<18)	164 (62.8%)
Adulthood ($\geq$ 18)	97 (37.2%)
Patient-reported global AD severity – freq (%)	
Mild	72 (27.2%)
Moderate	76 (28.7%)
Severe	117 (44.2%)
BSA – mean $\pm$ std. dev	33.2% $\pm$ 31.8%
NRS-itch – mean $\pm$ std. dev	6.3 $\pm$ 2.9
oSCORAD – mean $\pm$ std. dev	29.6 $\pm$ 16.6
SCORAD – mean $\pm$ std. dev	40.0 $\pm$ 19.7
EASI – mean $\pm$ std. dev	13.7 $\pm$ 14.5
POEM – mean $\pm$ std. dev	11.5 $\pm$ 7.4
DLQI – mean $\pm$ std. dev	10.7 $\pm$ 7.6

* Missing values were encountered in 0 (0.0%) for gender, race/ethnicity and patient-reported global AD severity, 4 (1.5%) for age of AD onset, 33 (12.5%) for BSA, 34 (10.6%) for NRS-itch, 34 (12.8%) for oSCORAD and SCORAD, 39 (14.7%) for EASI, 21 (7.9%) for POEM, and 11 (4.2%) for DLQI.







