

Article type : Original Article

Title: Severity strata for five patient-reported outcomes in adults with atopic dermatitis

Running Header: PRO strata in atopic dermatitis.

P.P. Vakharia¹, R. Chopra¹, R. Sacotte¹, N. Patel¹, S. Immaneni¹, T. White², R. Kantor¹, D.Y. Hsu¹, E.L. Simpson³, J.I. Silverberg^{2,4}

1. Departments of Dermatology, Preventive Medicine and Medical Social Sciences, Feinberg School of Medicine at Northwestern University, Chicago, IL USA
2. Northwestern Medicine Multidisciplinary Eczema Center, Chicago, IL USA
3. Department of Dermatology, Oregon Health & Science University, Portland, OR USA
4. Department of Dermatology, Feinberg School of Medicine at Northwestern University, Chicago, IL USA

Key Words: atopic dermatitis; eczema; severity; patient-reported outcomes

Abbreviations used: AD=atopic dermatitis; POEM=Patient Oriented Eczema Measure; NRS=Numeric Rating Scale; DLQI=Dermatology Life Quality Index

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.

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/bjd.16078

This article is protected by copyright. All rights reserved.

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

Analysis and interpretation of data: JI Silverberg, P Vakharia, R Chopra, R Sacotte, N Patel, S Immaneni, T White, R Kantor, D Hsu, EL Simpson

Drafting of the manuscript: JI Silverberg, EL Simpson

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, EL Simpson

Statistical analysis: JI Silverberg

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

What's already known about this topic?

- Severity strata were previously developed for POEM, DLQI and total ItchyQOL, but lack confirmation in atopic dermatitis.
- Mean ItchyQOL, NRS-itch, and 5-D itch scales currently do not have accepted severity strata.

What does this study add?

- The present study confirmed that previously developed severity strata for POEM and DLQI perform well in adults with AD.
- We also developed novel severity strata for NRS-itch, mean ItchyQoL and 5-D itch scales.

Abstract

Background: Multiple patient-reported outcomes have been used to assess the burden of atopic dermatitis (AD). Some are disease specific, e.g. the Patient Oriented Eczema Measure (POEM). While others pertain to itch, e.g. Numerical Rating Scale (NRS)-itch, ItchyQoL and 5-D itch, or dermatologic disease in general, e.g. Dermatology Life Quality Index (DLQI). Development of severity strata is essential for proper interpretability of these assessments. We sought to confirm previously developed strata for POEM, DLQI and raw ItchyQoL, and develop strata for the NRS-itch, mean ItchyQOL and 5-D itch scale for use in adults with AD.

Methods: Self-administered questionnaires were completed by 210 adults with AD in a dermatology practice setting. Strata were selected using an anchoring approach based on patient-reported disease severity.

Results: We confirmed the existing strata for POEM (mild=0-7, moderate=8-16, severe=17-28) ($\kappa=0.440$), DLQI (mild=0-5, moderate=6-10, severe=11-30) ($\kappa=0.398$) and NRS-itch (mild=0-3, moderate=4-6, severe=7-10) ($\kappa=0.499$). However, the preferred band for *raw*

ItchyQoL was mild=22-58, moderate=59-74 and severe=75-110 (kappa=0.379) and *mean* ItchyQOL was mild=1-2.9, moderate=3.0-3.9, severe=4.0-5.0 (kappa=0.374). The preferred band for 5-D itch scale was mild=0-11, moderate=12-17 and severe=18-25 (kappa=0.331).

Conclusions: Existing strata for POEM and DLQI perform well in adult AD. Previously reported strata for VAS-itch performed best for NRS-itch. We identified banding for the raw ItchyQoL for our AD population that varies slightly from the banding published for a more heterogeneous population. Finally, we proposed strata for mean ItchyQOL and 5-D itch scale in adult AD.

Introduction

Atopic dermatitis (AD) is a chronic inflammatory skin disease that is associated with a significant patient-burden. AD is associated with pruritus, sleep disturbance^{1,2}, impaired quality of life (QoL), including limitations of function and activities of daily living, poor school and work performance³⁻⁷, and impaired mental health^{8,9}. Some aspects of AD can be reliably assessed using objective assessments, accounting for the extent and severity of skin lesions¹⁰. However, many of the symptoms and detrimental aspects of the disease are subjective and may not be adequately reflected by investigator-derived assessments of disease severity.

Patient-reported outcomes (PRO) are becoming increasingly incorporated into clinical practice and trials in dermatology¹¹. PRO are particularly useful in the assessment of the symptoms of AD. The symptoms of AD are multidimensional and not comprehensively assessed by any single PRO instrument. A variety of PRO have been used in AD, including those that measure intensity and frequency of symptoms, patient-reported signs, QoL impairment and disease control. Recently, Harmonizing Outcome Measures in Eczema (HOME), an international consensus group aimed at standardizing outcomes in clinical trials of AD, selected the Patient-Oriented Eczema Measure (POEM) as the preferred PRO of AD symptoms in clinical trials of AD¹². However, POEM does not assess the intensity of pruritus, or quality of life (QoL)

impairment secondary to AD. No QOL assessment met sufficient rigor for selection in AD, though DLQI and QoLIAD were potential candidates in adults¹³.

Itch intensity is commonly assessed in clinical research and practice for AD. A recent systematic review by the HOME group found that the visual analog scale (VAS-itch) was the most commonly used PRO for itch, followed by the numerical rating scale (NRS-itch)¹⁴.

However, another study showed there was less missing data for NRS¹⁵ compared to VAS, and that patients preferred NRS over VAS¹⁶.

Multiple PRO have been used to assess QoL impairment in AD, including Dermatology Life Quality Index (DLQI)¹⁷, ItchyQOL^{18,19} and 5-D itch scales²⁰. The former assesses QoL across dermatologic disease, whereas the latter two assess itch-related QoL impairment. Each of these PRO has been used in clinical research for AD and has a potential role for the assessment of AD in clinical practice. The HOME group deemed that no QOL instruments met their criteria for recommendation to be included in AD RCT,^{13,21} but that the CADIS in children²¹ and DLQI and QoLIAD in adults¹³ have the potential to be recommended based on further validation studies.

One challenge of using these PRO is interpretability, i.e. how to interpret a given PRO score, when each PRO has a different numbers of questions, response options, and scoring systems. Severity strata or bands have been developed for some of these instruments to improve interpretability of results and help determine clinical meaningfulness^{15,22-24}. Additional studies are warranted to confirm that previously proposed severity strata are reproducible in other cohorts. Further, previously reported strata for DLQI, VAS-itch and raw ItchyQOL were determined in broader cohorts of dermatologic and pruritic disease, respectively^{15,22,23,25}. It is unclear how well strata developed in more heterogeneous cohorts perform in an AD-specific

population. Similar strata identified in a cohort of AD patients would suggest that these strata perform well in AD. Moreover, several instruments are still lacking interpretable severity strata altogether, e.g. 5-D itch scale²⁰ and *mean* ItchyQOL scores²⁵. Further, while severity strata for VAS-itch have been developed, no severity strata have been formally proposed for NRS-itch. There are a number of advantages of applying severity strata to outcome measures, which have recently been discussed in a recent study of AD²⁶. We sought to confirm whether previously reported strata for POEM and DLQI are applicable to adults with AD. We also sought to develop interpretable severity strata for NRS-itch, mean ItchyQOL and 5-D itch scale.

Methods

Study design

We performed a prospective dermatology practice-based study of adults (≥ 18 years) with AD as defined by the Hanifin & Rajka diagnostic criteria²⁷. Self-administered questionnaires were completed by patients of the eczema clinic prior to their encounter. At each encounter, questionnaires included a validated question assessing self-reported severity of AD (“Would you describe your atopic dermatitis or eczema as mild, moderate, or severe?”)²⁸, the Patient Oriented Eczema Measure (POEM) (7 questions, range: 0-28), Dermatology Life Quality Index (DLQI) (10 questions, range: 0-30)¹⁷, Numerical Rating Scale for average itch (NRS-itch) (1 question, range: 0-10)¹⁵, Itchy Quality of Life (ItchyQoL) (22 questions; range for raw scores: 22-110; range for mean scores: 1.0-5.0)¹⁸, including symptoms (6 questions; range for raw scores: 6-30; range for mean scores: 1.0-5.0), physical functioning (7 questions; range for raw scores: 7-35; range for mean scores: 1.0-5.0) and emotion (9 questions; range for raw scores: 9-45; range for mean scores: 1.0-5.0) subscores, and 5-dimensions (5-D) itch scale (5 domains, range: 5-25)²⁰.

Surveys were administered between January, 2014 and June, 2016. The study was approved by the institutional review boards of Northwestern University and informed consent was waived.

Data Processing and Statistical Methods

An anchor-based approach was used as recently described²⁶. Briefly, the thresholds for severity strata were determined by comparing the assessments with a global assessment. POEM, DLQI, NRS-itch, raw and mean ItchyQOL and 5-D itch scores were stratified against the self-reported AD severity, which was used as the severity anchor. Missing or invalid responses occurred in <1% of subjects for any of the assessments. A detailed description of the methods can be found in the Supplemental Methods.

Results

Patient characteristics

Overall, 210 adults (ages 18-93 years) with 728 encounters were included in the study, including 131 self-reported women (62.7%) and 128 Caucasian/white (61.2%) (Table 1). The mean $\pm$ std. dev. age at enrollment was 39.8 ± 17.8 years. All subjects met the Hanifin and Rajka criteria for AD. Forty-eight subjects (22.9%) reported mild, 64 (30.5%) moderate and 98 (46.7%) severe disease upon initial encounter. POEM, DLQI, ItchyQoL and 5-D itch scale were all significantly correlated with each other, with Spearman rho statistics ranging from 0.36-0.73 ($P < 0.0001$).

Self-reported AD severity (the anchor) was most strongly correlated to NRS-itch (Spearman correlation; $\rho = 0.61$), followed by POEM ($\rho = 0.57$), DLQI ($\rho = 0.51$), raw total ItchyQoL ($\rho = 0.51$), mean total ItchyQOL ($\rho = 0.52$), and 5-D itch scale ($\rho = 0.28$) ($P \leq 0.0005$).

for all). In addition, self-reported AD severity correlated most with the raw and mean symptoms subscores of ItchyQOL ($\rho=0.55$ and 0.55), followed by functioning limitations ($\rho=0.50$ and 0.50) and emotion ($\rho=0.40$ and 0.41) subscores, respectively.

Strata

The distribution, mean, median and mode of self-reported AD severity for each POEM, DLQI, NRS-itch, raw and mean ItchyQoL total score and symptoms, functioning limitation and emotion subscores, and 5-D itch are presented in Supplemental Tables 1-12. An illustration of the anchor-based approach is presented in the Supplemental Results.

Based on assessment of mean, median and mode POEM values, lower thresholds of 7 or 8 were identified for moderate and 17 and 19 were identified for severe disease. In combination, four different strata were tested. Multiple strata were selected for testing for DLQI ($n=16$), NRS-itch ($n=3$), raw total ItchyQOL ($n=9$), mean ItchyQOL ($n=6$), and 5-D itch scale ($n=4$). Optimal severity strata for each scale are presented in Table 2.

For POEM, the first band was 0-6=mild, 7-16=moderate and 17-28=severe and had the highest k coefficient (0.443) (Supplemental Table 13). However, the second band of 0-7=mild, 8-16=moderate and 17-28=severe had an almost identical k coefficient (0.439), and was similar to the stratum previously reported for POEM²².

For DLQI, the previously reported band²³ of 0-5=mild, 6-10=moderate and 11-30=severe had the highest k coefficient (0.398). For NRS-itch, the previously reported band (for VAS-itch¹⁵) of 0-3=mild, 4-6=moderate and 7-10=severe had the highest k coefficient for NRS-itch (0.499). For 5-D itch, the band of 0-11=mild, 12-17=moderate and 18-25=severe had the highest k coefficient (0.331).

For raw total ItchyQOL, the stratum of 22-58=mild, 59-74=moderate and 75-110=severe had the highest k coefficient (0.379). The previously reported raw ItchyQoL strata²⁵ had a lower k coefficient (0.332). For mean ItchyQOL, the stratum with the highest k coefficient was 1-2.6=mild, 2.7-4.2=moderate and 4.3-5.0=severe. Since these thresholds are uneven and potentially difficult to remember, we also tested the stratum 1-2.9=mild, 3.0-3.9=moderate and 4.0-5.0=severe to see if there was similar concordance. In fact, this band had the highest k coefficient (0.374).(Supplemental Table 13).

Robustness of strata

Eighteen patients (3.5%) had a self-reported AD severity score >1 point outside of that predicted by the final POEM strata. There were 92 (17.7%) of patients whose severity score was 1 point higher and 116 (22.3%) with 1 point lower than the final POEM strata predicted. Fifty-nine patients (9.3%) had an AD severity score >1 point outside of that predicted by the final DLQI strata. There were 117 (18.5%) of patients whose severity score was 1 point higher and 100 (15.8%) with 1 point lower than the final DLQI strata predicted. Forty-nine patients (7.6%) had an AD severity score >1 point outside of that predicted by the final NRS-itch strata. There were 61 (9.4%) of patients whose severity score was 1 point higher and 132 (20.4%) with 1 point lower than the final NRS-itch strata predicted. Forty-three patients (8.6%) and 39 (7.8%) had AD severity scores >1 point outside of that predicted by the final total and mean ItchyQOL strata, respectively. There were 82 (16.4%) and 119 (23.8%) of patients whose severity score was 1 point higher, and 100 (20.0%) and 58 (11.6%) with 1 point lower than the final total and mean ItchyQOL strata predicted, respectively. Patients falling outside the possible strata had a similar distribution of sex, age, race/ethnicity as the overall study cohort.

Discussion

The present study determined potential severity strata for POEM and DLQI in adults with AD, which were consistent with previously reported strata^{22,23}. In addition, we found potential severity strata for NRS-itch that were very similar to previously reported strata for VAS-itch¹⁵. Severity strata for DLQI and VAS-itch were previously determined in broad cohorts, including multiple disorders, and found similar potential strata^{15,23}. The potential strata identified in this study were categorized as mild, moderate and severe. This is a notable difference from the previously reported strata, which also included categories for clear or almost clear, and very severe. This was an a priori decision based on preliminary feedback from participant stakeholders that these categories were less meaningful to patients²⁶. Another noteworthy difference of this study is that the anchoring question used in the study was worded differently than those used in prior studies. Despite these subtle differences, the optimal severity strata observed in the present study for POEM, DLQI and NRS-itch align almost perfectly with those previously reported. Together, the present study confirms that these strata are indeed optimal for adults with AD.

In addition, we found potential severity strata for the raw ItchyQOL score. This banding scheme performed better in our cohort of adult AD than strata determined using anchors from the Global Itch Severity Questionnaire (GISQ) (scale of 0-10 for itch severity) in 54 US veterans with chronic pruritus of unspecified etiology²⁵. Both banding schemes may be valid in different contexts, but the strata determined in the present study appear to be optimal for the interpretation of ItchyQOL scores in adults with AD. Further, we developed novel severity strata for the mean ItchyQOL score, which are both intuitive and applicable to the original ItchyQOL scoring approach. We also developed strata for the raw and mean symptom, functioning disturbance and

emotion subscores of ItchyQOL. Finally, we developed potential severity strata for the 5-D itch scale.

An interesting observation of this study is that the banded scores for NRS-itch had the greatest concordance with patient-reported AD severity. This suggests that patients' report of AD severity is most strongly influenced by the intensity of their itch. Alternatively, more severe AD could be associated with more intense itch, potentially because it is often more widespread. The stronger correlation of patient-reported AD severity with NRS-itch than POEM is intriguing. The HOME consensus group has recently selected POEM as the preferred PRO to assess AD symptoms. The results of the present study confirm that POEM correlates fairly well with overall disease severity. However, they also suggest that NRS-itch correlates with AD severity as well or perhaps even better than POEM in some populations. NRS-itch, as a single question instrument, is shorter, simpler, and more feasible than virtually any of the currently available PRO for AD symptoms. While the NRS has been well validated for rating pruritus^{16,29}, it is possible that a multi-item instrument, e.g. POEM, may be more reliable than NRS-itch. POEM has been well validated and shown to demonstrate construct validity, convergent validity, divergent validity, internal consistency, sensitivity to change, and test-retest reliability^{30,31}. POEM was chosen by the HOME group at the fourth international consensus meeting as the preferred core instrument to assess patient-reported symptoms in future AD trials^{32,33}. We concur with this recommendation, and suggest that NRS-itch is an additional assessment to consider for the symptoms in AD.

The concordance of self-reported AD severity with POEM, NRS-itch, DLQI, ItchyQOL and 5-D itch scale were all modest at best. This underscores the heterogeneity of AD and its multidimensional impact on patients, including disturbances of function, activities of daily

living, sleep, and emotional well-being. Each of these instruments assesses different aspects of these disturbances, with only partial overlap. Thus, each instrument has its merits and no one instrument can replace the others. Selection of an instrument in clinical practice or trials will depend on feasibility and specific domains to be probed.

This study has several strengths, including large sample size, and good representation across gender, race/ethnicity and AD severity. However, there are some limitations of this study. The study cohort was recruited from a single academic center, which may limit generalizability. However, the remarkable consistency between potential strata for POEM, DLQI and NRS-itch identified in our cohort and previous cohorts is reassuring that this is not a major issue. This suggests that the results from our cohort are generalizable to the broader population of AD patients. Despite their common use in clinical trials and research of AD, there are limited data available for ItchyQOL, NRS-itch, and 5-D itch in AD patients. We studied the interpretability of these PRO because of their commonality and potential utility in AD. Nevertheless, future studies are needed to elaborate on the measurement properties of these assessments in AD.

In conclusion, the present study confirmed that previously determined severity strata for POEM and DLQI are optimal in adults with AD. We also found that previously reported strata for VAS-itch were optimal for NRS-itch. Moreover, we found a different banding scheme for raw ItchyQOL to be optimal in adult AD, and developed novel strata for mean ItchyQOL and 5-D itch scale. We recommend these strata be used in clinical practice and/or research studies of adult AD. These strata can be used to improve interpretation of clinical and research data using these instruments.

Acknowledgments

Northwestern Medicine Enterprise Data Warehouse (NMEDW) was supported, in part, by the Northwestern University Clinical and Translational Science Institute, funded, in part, by Grant Number 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 NIH. The Clinical and Translational Science Award (CTSA) is a registered trademark of DHHS.

References

- 1 Silverberg JI, Garg NK, Paller AS *et al.* Sleep disturbances in adults with eczema are associated with impaired overall health: a US population-based study. *The Journal of investigative dermatology* 2015; **135**: 56-66.
- 2 Yu SH, Attarian H, Zee P *et al.* Burden of Sleep and Fatigue in US Adults With Atopic Dermatitis. *Dermatitis : contact, atopic, occupational, drug : official journal of the American Contact Dermatitis Society, North American Contact Dermatitis Group* 2016; **27**: 50-8.
- 3 Coghi S, Bortoletto MC, Sampaio SA *et al.* Quality of life is severely compromised in adult patients with atopic dermatitis in Brazil, especially due to mental components. *Clinics* 2007; **62**: 235-42.
- 4 Fivenson D, Arnold RJ, Kaniecki DJ *et al.* The effect of atopic dermatitis on total burden of illness and quality of life on adults and children in a large managed care organization. *Journal of managed care pharmacy : JMCP* 2002; **8**: 333-42.

- 5 Haeck IM, ten Berge O, van Velsen SG *et al.* Moderate correlation between quality of
life and disease activity in adult patients with atopic dermatitis. *Journal of the European*
Academy of Dermatology and Venereology : JEADV 2012; **26**: 236-41.
- 6 Kim DH, Li K, Seo SJ *et al.* Quality of life and disease severity are correlated in patients
with atopic dermatitis. *Journal of Korean medical science* 2012; **27**: 1327-32.
- 7 Monti F, Agostini F, Gobbi F *et al.* Quality of life measures in Italian children with
atopic dermatitis and their families. *Italian journal of pediatrics* 2011; **37**: 59.
- 8 Yu SH, Silverberg JI. Association between Atopic Dermatitis and Depression in US
Adults. *The Journal of investigative dermatology* 2015; **135**: 3183-6.
- 9 Yaghmaie P, Koudelka CW, Simpson EL. Mental health comorbidity in patients with
atopic dermatitis. *The Journal of allergy and clinical immunology* 2013; **131**: 428-33.
- 10 Schmitt J, Langan S, Deckert S *et al.* Assessment of clinical signs of atopic dermatitis: a
systematic review and recommendation. *The Journal of allergy and clinical immunology*
2013; **132**: 1337-47.
- 11 Copley-Merriman C, Zelt S, Clark M *et al.* Impact of Measuring Patient-Reported
Outcomes in Dermatology Drug Development. *Patient* 2016.
- 12 Spuls PI, Gerbens LA, Simpson E *et al.* POEM a core instrument to measure symptoms
in clinical trials: a HOME statement. *The British journal of dermatology* 2016.
- 13 Heintz D, Prinsen CA, Deckert S *et al.* Measurement properties of adult quality-of-life
measurement instruments for eczema: a systematic review. *Allergy* 2016; **71**: 358-70.
- 14 Gerbens LA, Chalmers JR, Rogers NK *et al.* Reporting of symptoms in randomized
controlled trials of atopic eczema treatments: a systematic review. *The British journal of*
dermatology 2016; **175**: 678-86.

- 15 Reich A, Heisig M, Phan NQ *et al.* Visual analogue scale: evaluation of the instrument
for the assessment of pruritus. *Acta dermato-venereologica* 2012; **92**: 497-501.
- 16 Reich A HJ, Ramus M, Ständer S, Szepietowski J. New data on the validation of VAS
and NRS in pruritus assessment: minimal clinically important difference and itch
frequency measurement. *Acta dermato-venereologica* 2011; **91**: 636.
- 17 Basra MK, Fenech R, Gatt RM *et al.* The Dermatology Life Quality Index 1994-2007: a
comprehensive review of validation data and clinical results. *The British journal of*
dermatology 2008; **159**: 997-1035.
- 18 Desai NS, Poindexter GB, Monthrope YM *et al.* A pilot quality-of-life instrument for
pruritus. *Journal of the American Academy of Dermatology* 2008; **59**: 234-44.
- 19 ItchyQol: A Pruritus-Specific Quality of Life Instrument. In, Vol. 2017. Atlanta: Emory
University. 2009.
- 20 Elman S, Hynan LS, Gabriel V *et al.* The 5-D itch scale: a new measure of pruritus. *The*
British journal of dermatology 2010; **162**: 587-93.
- 21 Heintz D, Prinsen CA, Sach T *et al.* Measurement properties of quality-of-life
measurement instruments for infants, children and adolescents with eczema: a systematic
review. *The British journal of dermatology* 2016.
- 22 Charman CR, Venn AJ, Ravenscroft JC *et al.* Translating Patient-Oriented Eczema
Measure (POEM) scores into clinical practice by suggesting severity strata derived using
anchor-based methods. *The British journal of dermatology* 2013; **169**: 1326-32.
- 23 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**: 333-42, x.

- 24 Hongbo Y, Thomas CL, Harrison MA *et al.* Translating the science of quality of life into
practice: What do dermatology life quality index scores mean? *The Journal of*
investigative dermatology 2005; **125**: 659-64.
- 25 Love EM, Marrazzo GA, Kini S *et al.* ItchyQoL bands: pilot clinical interpretation of
scores. *Acta dermato-venereologica* 2015; **95**: 114-5.
- 26 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. *The British journal of dermatology* 2017.
- 27 Hanifin J, Rajka G. Diagnostic features of atopic eczema. *Acta dermato-venereologica*
1980; **92**: 44-7.
- 28 Vakharia PP, Chopra R, Sacotte R *et al.* Validation of patient-reported global severity of
atopic dermatitis in adults. *Allergy* 2017.
- 29 Phan NQ, Blome C, Fritz F *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 dermato-venereologica* 2012; **92**:
502-7.
- 30 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. *Archives of dermatology* 2004; **140**: 1513-9.
- 31 Schmitt J, Langan S, Williams HC. What are the best outcome measurements for atopic
eczema? A systematic review. *J Allergy Clin Immunol* 2007; **120**: 1389-98.

- 32 Chalmers JR, Simpson E, Apfelbacher CJ *et al.* Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *The British journal of dermatology* 2016; **175**: 69-79.
- 33 Spuls PI, Gerbens LA, Simpson E *et al.* POEM a core instrument to measure symptoms in clinical trials: a HOME statement. *The British journal of dermatology* 2016.

Table 1. Subject characteristics.	
Variable	Value
Age – mean $\pm$ std. dev.	39.8 $\pm$ 17.8
Female sex – freq (%)	131 (62.7%)
Race/ethnicity – freq (%)	
Caucasian/white	128 (61.2%)
African-American/black	28 (13.4%)
Hispanic	9 (4.3%)
Asian	42 (20.1%)
Multiracial/other	2 (1.0%)

Table 2. Concordance of final proposed severity strata with self-reported AD severity.

Scale	Self-reported AD severity			κ-coefficient
	Mild	Moderate	Severe	
POEM	0-7	8-16	17-28	0.440
DLQI	0-5	6-10	11-30	0.398
NRS-itch	0-3	4-6	7-10	0.499
Raw ItchyQOL				
Total	22-58	59-74	75-110	0.379
Symptoms	6-18	19-21	22-30	0.433
Functioning	6-18	19-27	28-35	0.413
Emotion	7-26	27-33	34-45	0.327
Mean ItchyQOL				
Total	1-2.9	3.0-3.9	4.0-5.0	0.374
Symptoms	1-2.9	3.0-3.9	4.0-5.0	0.402
Functioning	1-2.6	2.7-3.9	4.0-5.0	0.413
Emotion	1-2.9	3.0-3.9	4.0-5.0	0.314
5-D itch	0-11	12-17	18-25	0.331

Supplemental Table 1. Distribution of self-reported AD severity across different POEM scores.

POEM score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
0	22	20	2	0	1.1	1	1
1	21	21	0	0	1.0	1	1
2	32	26	5	1	1.2	1	1
3	25	18	5	0	1.2	1	1
4	21	15	5	0	1.3	1	1
5	23	15	6	2	1.4	1	1
6	34	22	8	3	1.4	1	1
7	12	5	6	1	1.7	2	2
8	35	13	15	6	1.8	2	2
9	45	21	7	14	1.8	1.5	1
10	27	15	9	3	1.6	1	1
11	22	7	10	3	1.8	2	2
12	31	8	11	8	2.0	2	2
13	22	11	8	3	1.6	1.5	1
14	13	5	5	3	1.8	2	1
15	18	4	9	5	2.1	2	2
16	31	4	14	10	2.2	2	2
17	16	3	3	7	2.3	3	3
18	15	3	6	6	2.2	2	2
19	9	0	3	5	2.6	3	3
20	12	2	4	5	2.3	2	3
21	9	1	0	7	2.8	3	3
22	15	0	3	11	2.8	3	3
23	8	0	2	5	2.7	3	3
24	12	1	3	8	2.6	3	3
25	5	0	1	3	2.8	3	3
26	2	0	2	0	2.0	2	2
27	3	0	1	2	2.7	3	3
28	7	1	0	6	2.7	3	3

Supplemental Table 2. Distribution of self-reported AD severity across different DLQI scores.

DLQI score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
0	55	48	6	1	1.1	1	1
1	50	36	11	3	1.3	1	1
2	60	41	16	3	1.4	1	1
3	51	36	12	3	1.4	1	1
4	40	18	14	8	1.8	2	1
5	52	26	18	8	1.7	1.5	1
6	46	19	13	14	1.9	2	1
7	30	12	11	7	1.8	2	1
8	27	5	16	6	2.0	2	2
9	26	9	10	7	1.9	2	2
10	27	9	12	6	1.9	2	2
11	21	8	4	9	2.0	2	3
12	20	4	10	6	2.1	2	2
13	18	6	5	7	2.1	2	3
14	11	3	2	6	2.3	3	3
15	11	3	2	6	2.3	3	3
16	13	3	5	5	2.2	2	2
17	12	2	6	4	2.2	2	2
18	11	0	5	6	2.5	3	3
19	9	1	4	4	2.3	2	2
20	13	2	0	11	2.7	3	3
21	4	0	0	4	3.0	3	3
22	5	0	0	5	3.0	3	3
23	5	0	1	4	2.8	3	3
24	7	0	1	6	2.9	3	3
25	3	0	1	2	2.7	3	3
26	1	0	0	1	3.0	3	3
27	1	0	0	1	3.0	3	3
28	1	1	0	0	1.0	1	1
29	1	0	0	1	3.0	3	3
30	1	0	0	1	3.0	3	3

Supplemental Table 3. Distribution of self-reported AD severity across different NRS-itch scores.

NRS-itch	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
0	37	35	1	1	1.1	1	1
1	53	48	4	1	1.1	1	1
2	70	56	12	2	1.2	1	1
3	90	65	19	6	1.3	1	1
4	71	31	33	7	1.7	2	2
5	71	31	33	7	1.7	2	2
6	51	15	25	11	1.9	2	2
7	52	16	21	15	2.0	2	2
8	63	12	22	29	2.3	2	3
9	32	9	7	16	2.2	2.5	3
10	57	2	5	50	2.8	3	3

Supplemental Table 4. Distribution of self-reported AD severity across different raw ItchyQOL total scores.

Raw total ItchyQOL score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
22	8	6	1	1	1.4	1.0	1.0
23	1	1	0	0	1.0	1.0	1.0
24	2	2	0	0	1.0	1.0	1.0
25	1	1	0	0	1.0	1.0	1.0
27	4	4	0	0	1.0	1.0	1.0
28	4	4	0	0	1.0	1.0	1.0
29	3	2	1	0	1.3	1.0	1.0
30	4	4	0	0	1.0	1.0	1.0
31	3	3	0	0	1.0	1.0	1.0
32	2	2	0	0	1.0	1.0	1.0
33	5	4	1	0	1.2	1.0	1.0
35	2	2	0	0	1.0	1.0	1.0
36	4	2	1	1	1.8	1.5	1.0
37	2	1	1	0	1.5	1.5	.
38	10	9	1	0	1.1	1.0	1.0
39	7	7	0	0	1.0	1.0	1.0
40	7	6	1	0	1.1	1.0	1.0
41	8	7	1	0	1.1	1.0	1.0
42	3	3	0	0	1.0	1.0	1.0
43	4	2	2	0	1.5	1.5	1.0
44	2	2	0	0	1.0	1.0	1.0
45	5	5	0	0	1.0	1.0	1.0
46	6	4	2	0	1.3	1.0	1.0
47	9	5	4	0	1.4	1.0	1.0
48	4	3	1	0	1.3	1.0	1.0
49	5	3	2	0	1.4	1.0	1.0
50	3	2	1	0	1.3	1.0	1.0
51	4	2	0	2	2.0	2.0	1.0
52	12	8	3	1	1.4	1.0	1.0
53	7	5	2	0	1.3	1.0	1.0
54	9	6	0	3	1.7	1.0	1.0
55	11	5	5	1	1.6	2.0	1.0
56	10	5	4	1	1.6	1.5	1.0
57	10	5	4	1	1.6	1.5	1.0
58	10	8	2	0	1.2	1.0	1.0
59	4	0	3	1	2.3	2.0	2.0
60	11	6	5	0	1.5	1.0	1.0
61	11	6	2	3	1.7	1.0	1.0
62	11	5	5	1	1.6	2.0	1.0
63	12	6	4	2	1.7	1.5	1.0
64	8	4	2	2	1.8	1.5	1.0

65	7	4	2	1	1.6	1.0	1.0
66	6	4	1	1	1.5	1.0	1.0
67	9	3	5	1	1.8	2.0	2.0
68	11	3	3	5	2.2	2.0	3.0
69	7	5	2	0	1.3	1.0	1.0
70	6	2	2	2	2.0	2.0	1.0
71	10	3	4	3	2.0	2.0	2.0
72	10	4	3	3	1.9	2.0	1.0
73	7	1	4	2	2.1	2.0	2.0
74	8	2	3	3	2.1	2.0	2.0
75	8	1	3	4	2.4	2.5	3.0
76	2	0	0	2	3.0	3.0	3.0
77	8	2	6	0	1.8	2.0	2.0
78	8	2	3	3	2.1	2.0	2.0
79	11	3	4	4	2.1	2.0	2.0
80	6	2	2	2	2.0	2.0	1.0
81	3	0	2	1	2.3	2.0	2.0
82	2	1	0	1	2.0	2.0	.
83	4	1	1	2	2.3	2.5	3.0
84	9	2	4	3	2.1	2.0	2.0
85	11	0	8	3	2.3	2.0	2.0
86	7	2	1	4	2.3	3.0	3.0
87	3	0	2	1	2.3	2.0	2.0
88	4	1	1	2	2.3	2.5	3.0
89	3	1	1	1	2.0	2.0	.
90	8	4	1	3	1.9	1.5	1.0
91	6	1	3	2	2.2	2.0	2.0
92	5	0	2	3	2.6	3.0	3.0
93	6	0	4	2	2.3	2.0	2.0
94	5	2	1	2	2.0	2.0	1.0
95	3	0	0	3	3.0	3.0	3.0
96	1	0	0	1	3.0	3.0	3.0
97	1	0	0	1	3.0	3.0	3.0
98	4	3	1	0	1.3	1.0	1.0
99	3	1	1	1	2.0	2.0	.
100	2	0	2	0	2.0	2.0	2.0
101	1	0	0	1	3.0	3.0	3.0
102	4	1	1	2	2.3	2.5	3.0
103	3	0	1	2	2.7	3.0	3.0
104	3	1	0	2	2.3	3.0	3.0
105	2	0	1	1	2.5	2.5	.
106	5	1	0	4	2.6	3.0	3.0
107	2	0	1	1	2.5	2.5	.
108	2	0	0	2	3.0	3.0	3.0

109	2	0	1	1	2.5	2.5	.
110	4	1	0	3	2.5	3.0	3.0

Supplemental Table 5. Distribution of self-reported AD severity across different mean ItchyQOL scores.

Mean ItchyQOL score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
1.0	9	7	1	1	1.3	1.0	1.0
1.1	3	3	0	0	1.0	1.0	1.0
1.2	4	4	0	0	1.0	1.0	1.0
1.3	7	6	1	0	1.1	1.0	1.0
1.4	7	7	0	0	1.0	1.0	1.0
1.5	7	6	1	0	1.1	1.0	1.0
1.6	6	4	1	1	1.5	1.0	1.0
1.7	12	10	2	0	1.2	1.0	1.0
1.8	14	13	1	0	1.1	1.0	1.0
1.9	11	10	1	0	1.1	1.0	1.0
2.0	11	9	2	0	1.2	1.0	1.0
2.1	15	9	6	0	1.4	1.0	1.0
2.2	9	6	3	0	1.3	1.0	1.0
2.3	7	4	1	2	1.7	1.0	1.0
2.4	19	13	5	1	1.4	1.0	1.0
2.5	30	16	9	5	1.6	1.0	1.0
2.6	20	13	6	1	1.4	1.0	1.0
2.7	15	6	8	1	1.7	2.0	2.0
2.8	22	11	7	4	1.7	1.5	1.0
2.9	20	10	6	4	1.7	1.5	1.0
3.0	22	11	8	3	1.6	1.5	1.0
3.1	18	8	5	5	1.8	2.0	1.0
3.2	16	5	6	5	2.0	2.0	2.0
3.3	17	5	7	5	2.0	2.0	2.0
3.4	16	3	6	7	2.3	2.0	3.0
3.5	18	4	9	5	2.1	2.0	2.0
3.6	17	5	6	6	2.1	2.0	2.0
3.7	5	1	2	2	2.2	2.0	2.0
3.8	13	3	5	5	2.2	2.0	2.0
3.9	18	2	9	7	2.3	2.0	2.0
4.0	10	2	4	4	2.2	2.0	2.0
4.1	14	5	4	5	2.0	2.0	1.0
4.2	11	0	6	5	2.5	2.0	2.0
4.3	8	2	1	5	2.4	3.0	3.0
4.4	2	0	0	2	3.0	3.0	3.0
4.5	9	4	4	1	1.7	2.0	1.0
4.6	5	1	1	3	2.4	3.0	3.0

4.7	6	1	1	4	2.5	3.0	3.0
4.8	7	1	1	5	2.6	3.0	3.0
4.9	4	0	1	3	2.8	3.0	3.0
5.0	6	1	1	4	2.5	3.0	3.0

Supplemental Table 6. Distribution of self-reported AD severity across different raw ItchyQOL symptom scores.

Raw ItchyQOL symptom score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
6	14	12	1	1	1.2	1.0	1.0
7	3	3	0	0	1.0	1.0	1.0
8	7	7	0	0	1.0	1.0	1.0
9	14	14	0	0	1.0	1.0	1.0
10	24	20	4	0	1.2	1.0	1.0
11	14	10	4	0	1.3	1.0	1.0
12	15	14	0	1	1.1	1.0	1.0
13	18	10	7	1	1.5	1.0	1.0
14	31	20	10	1	1.4	1.0	1.0
15	16	10	5	1	1.4	1.0	1.0
16	32	20	9	3	1.5	1.0	1.0
17	28	15	10	3	1.6	1.0	1.0
18	37	19	11	7	1.7	1.0	1.0
19	38	13	17	8	1.9	2.0	2.0
20	32	14	11	7	1.8	2.0	1.0
21	26	7	12	7	2.0	2.0	2.0
22	25	5	5	15	2.4	3.0	3.0
23	21	3	12	6	2.1	2.0	2.0
24	28	7	7	14	2.3	2.5	3.0
25	14	2	6	6	2.3	2.0	2.0
26	10	3	3	4	2.1	2.0	3.0
27	11	1	3	7	2.5	3.0	3.0
28	8	0	5	3	2.4	2.0	2.0
29	8	0	2	6	2.8	3.0	3.0
30	16	2	4	10	2.5	3.0	3.0

Supplemental Table 7. Distribution of self-reported AD severity across different mean ItchyQOL symptom scores.

Mean ItchyQOL symptom score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
1	14	12	1	1	1.2	1.0	1.0
1.2	3	3	0	0	1.0	1.0	1.0
1.3	7	7	0	0	1.0	1.0	1.0
1.5	14	14	0	0	1.0	1.0	1.0
1.7	24	20	4	0	1.2	1.0	1.0
1.8	14	10	4	0	1.3	1.0	1.0
2	15	14	0	1	1.1	1.0	1.0
2.2	18	10	7	1	1.5	1.0	1.0
2.3	31	20	10	1	1.4	1.0	1.0
2.5	16	10	5	1	1.4	1.0	1.0
2.7	32	20	9	3	1.5	1.0	1.0
2.8	28	15	10	3	1.6	1.0	1.0
3	37	19	11	7	1.7	1.0	1.0
3.2	38	13	17	8	1.9	2.0	2.0
3.3	32	14	11	7	1.8	2.0	1.0
3.5	26	7	12	7	2.0	2.0	2.0
3.7	25	5	5	15	2.4	3.0	3.0
3.8	21	3	12	6	2.1	2.0	2.0
4	28	7	7	14	2.3	2.5	3.0
4.2	14	2	6	6	2.3	2.0	2.0
4.3	10	3	3	4	2.1	2.0	3.0
4.5	11	1	3	7	2.5	3.0	3.0
4.7	8	0	5	3	2.4	2.0	2.0
4.8	8	0	2	6	2.8	3.0	3.0
5	16	2	4	10	2.5	3.0	3.0

Supplemental Table 8. Distribution of self-reported AD severity across different raw ItchyQOL functional scores.

Raw ItchyQOL functional score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
7	11	9	1	1	1.3	1.0	1.0
8	7	6	1	0	1.1	1.0	1.0
9	4	3	1	0	1.3	1.0	1.0
10	12	12	0	0	1.0	1.0	1.0
11	16	15	1	0	1.1	1.0	1.0
12	11	7	2	2	1.5	1.0	1.0
13	17	12	5	0	1.3	1.0	1.0
14	16	12	3	1	1.3	1.0	1.0
15	27	21	4	2	1.3	1.0	1.0
16	29	23	6	0	1.2	1.0	1.0
17	22	12	6	4	1.6	1.0	1.0
18	21	14	5	2	1.4	1.0	1.0
19	33	12	11	10	1.9	2.0	1.0
20	29	9	14	6	1.9	2.0	2.0
21	32	15	13	4	1.7	2.0	1.0
22	32	13	13	6	1.8	2.0	1.0
23	24	6	12	6	2.0	2.0	2.0
24	22	5	12	5	2.0	2.0	2.0
25	21	3	10	8	2.2	2.0	2.0
26	11	2	5	4	2.2	2.0	2.0
27	13	3	5	5	2.2	2.0	2.0
28	14	6	1	7	2.1	2.5	3.0
29	16	3	5	8	2.3	2.5	3.0
30	8	1	2	5	2.5	3.0	3.0
31	7	2	3	2	2.0	2.0	2.0
32	12	1	3	8	2.6	3.0	3.0
33	8	3	2	3	2.0	2.0	1.0
34	5	0	1	4	2.8	3.0	3.0
35	10	1	1	8	2.7	3.0	3.0

Supplemental Table 9. Distribution of self-reported AD severity across different mean ItchyQOL functional scores.

Mean ItchyQOL functional score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
1	11	9	1	1	1.3	1.0	1.0
1.1	7	6	1	0	1.1	1.0	1.0
1.3	4	3	1	0	1.3	1.0	1.0
1.4	12	12	0	0	1.0	1.0	1.0
1.6	16	15	1	0	1.1	1.0	1.0
1.7	11	7	2	2	1.5	1.0	1.0
1.9	17	12	5	0	1.3	1.0	1.0
2	16	12	3	1	1.3	1.0	1.0
2.1	27	21	4	2	1.3	1.0	1.0
2.3	29	23	6	0	1.2	1.0	1.0
2.4	22	12	6	4	1.6	1.0	1.0
2.6	21	14	5	2	1.4	1.0	1.0
2.7	33	12	11	10	1.9	2.0	1.0
2.9	29	9	14	6	1.9	2.0	2.0
3	32	15	13	4	1.7	2.0	1.0
3.1	32	13	13	6	1.8	2.0	1.0
3.3	24	6	12	6	2.0	2.0	2.0
3.4	22	5	12	5	2.0	2.0	2.0
3.6	21	3	10	8	2.2	2.0	2.0
3.7	11	2	5	4	2.2	2.0	2.0
3.9	13	3	5	5	2.2	2.0	2.0
4	14	6	1	7	2.1	2.5	3.0
4.1	16	3	5	8	2.3	2.5	3.0
4.3	8	1	2	5	2.5	3.0	3.0
4.4	7	2	3	2	2.0	2.0	2.0
4.6	12	1	3	8	2.6	3.0	3.0
4.7	8	3	2	3	2.0	2.0	1.0
4.9	5	0	1	4	2.8	3.0	3.0
5	10	1	1	8	2.7	3.0	3.0

Supplemental Table 10. Distribution of self-reported AD severity across different raw ItchyQOL emotional scores.							
Raw ItchyQOL emotional score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
9	24	21	2	1	1.2	1.0	1.0
10	11	10	1	0	1.1	1.0	1.0
11	15	11	4	0	1.3	1.0	1.0
12	11	8	2	1	1.4	1.0	1.0
13	7	5	1	1	1.4	1.0	1.0
14	14	9	4	1	1.4	1.0	1.0
15	16	10	5	1	1.4	1.0	1.0
16	17	13	3	1	1.3	1.0	1.0
17	10	5	4	1	1.6	1.5	1.0
18	12	6	5	1	1.6	1.5	1.0
19	14	5	6	3	1.9	2.0	2.0
20	13	8	4	1	1.5	1.0	1.0
21	16	8	2	6	1.9	1.5	1.0
22	15	9	4	2	1.5	1.0	1.0
23	19	12	4	3	1.5	1.0	1.0
24	14	5	6	3	1.9	2.0	2.0
25	21	13	7	1	1.4	1.0	1.0
26	22	14	4	4	1.5	1.0	1.0
27	22	10	8	4	1.7	2.0	1.0
28	17	5	7	5	2.0	2.0	2.0
29	6	2	3	1	1.8	2.0	2.0
30	13	5	2	6	2.1	2.0	3.0
31	15	4	7	4	2.0	2.0	2.0
32	13	0	8	5	2.4	2.0	2.0
33	10	2	6	2	2.0	2.0	2.0
34	11	2	1	8	2.5	3.0	3.0
35	19	6	9	4	1.9	2.0	2.0
36	16	3	7	6	2.2	2.0	2.0
37	10	3	2	5	2.2	2.5	3.0
38	11	5	3	3	1.8	2.0	1.0
39	4	1	1	2	2.3	2.5	3.0
40	5	0	3	2	2.4	2.0	2.0
41	8	2	3	3	2.1	2.0	2.0
42	7	2	2	3	2.1	2.0	3.0
43	4	0	2	2	2.5	2.5	2.0
44	4	2	2	0	1.5	1.5	1.0
45	24	5	4	15	2.4	3.0	3.0

Supplemental Table 11. Distribution of self-reported AD severity across different mean ItchyQOL emotional scores.							
Mean ItchyQOL emotional score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
1.0	24	21	2	1	1.2	1.0	1.0
1.1	11	10	1	0	1.1	1.0	1.0
1.2	15	11	4	0	1.3	1.0	1.0
1.3	11	8	2	1	1.4	1.0	1.0
1.4	7	5	1	1	1.4	1.0	1.0
1.6	14	9	4	1	1.4	1.0	1.0
1.7	16	10	5	1	1.4	1.0	1.0
1.8	17	13	3	1	1.3	1.0	1.0
1.9	11	6	4	1	1.5	1.0	1.0
2.0	12	6	5	1	1.6	1.5	1.0
2.1	15	6	6	3	1.8	2.0	1.0
2.2	13	8	4	1	1.5	1.0	1.0
2.3	16	8	2	6	1.9	1.5	1.0
2.4	16	10	4	2	1.5	1.0	1.0
2.6	19	12	4	3	1.5	1.0	1.0
2.7	14	5	6	3	1.9	2.0	2.0
2.8	21	13	7	1	1.4	1.0	1.0
2.9	23	15	4	4	1.5	1.0	1.0
3.0	22	10	8	4	1.7	2.0	1.0
3.1	17	5	7	5	2.0	2.0	2.0
3.2	6	2	3	1	1.8	2.0	2.0
3.3	13	5	2	6	2.1	2.0	3.0
3.4	15	4	7	4	2.0	2.0	2.0
3.5	1	0	1	0	2.0	2.0	2.0
3.6	13	0	8	5	2.4	2.0	2.0
3.7	10	2	6	2	2.0	2.0	2.0
3.8	12	3	1	8	2.4	3.0	3.0
3.9	19	6	9	4	1.9	2.0	2.0
4.0	17	4	7	6	2.1	2.0	2.0
4.1	10	3	2	5	2.2	2.5	3.0
4.2	11	5	3	3	1.8	2.0	1.0
4.3	4	1	1	2	2.3	2.5	3.0
4.4	5	0	3	2	2.4	2.0	2.0
4.6	8	2	3	3	2.1	2.0	2.0
4.7	7	2	2	3	2.1	2.0	3.0
4.8	4	0	2	2	2.5	2.5	2.0
4.9	5	3	2	0	1.4	1.0	1.0
5.0	25	5	4	16	2.4	3.0	3.0

Supplemental Table 12. Distribution of self-reported AD severity across different 5-D itch scale scores.

5-D itch score	Freq.	Self-reported severity			Mean	Median	Mode
		Mild	Moderate	Severe			
5	245	117	74	54	1.7	2	1
6	13	13	0	0	1.0	1	1
7	10	10	0	0	1.0	1	1
8	37	32	4	1	1.2	1	1
9	46	32	11	3	1.4	1	1
10	53	43	8	2	1.2	1	1
11	45	27	13	5	1.5	1	1
12	25	11	8	6	1.8	2	1
13	42	15	19	8	1.8	2	2
14	50	19	20	11	1.8	2	2
15	23	5	13	5	2.0	2	2
16	45	11	19	15	2.1	2	2
17	21	4	9	8	2.2	2	2
18	16	1	3	12	2.7	3	3
19	27	2	7	18	2.6	3	3
20	10	2	4	4	2.2	2	2
21	7	0	1	6	2.9	3	3
22	9	0	0	9	3	3	3
23	4	1	1	2	2.3	3	3

Supplemental Table 13. Concordance of proposed severity strata with self-reported AD severity.			
Self-reported AD severity			© coefficient of agreement
Mild	Moderate	Severe	
Possible POEM bandings			
0-6	7-16	17-28	0.443
0-7	8-16	17-28	0.439
0-6	7-18	19-28	0.429
0-7	8-18	19-28	0.425
Possible DLQI bandings			
0-4	5-13	14-30	0.387
0-4	5-14	15-30	0.378
0-4	5-17	18-30	0.368
0-4	5-18	19-30	0.358
0-4	5-19	20-30	0.354
0-7	8-13	14-30	0.372
0-7	8-14	15-30	0.363
0-7	8-17	18-30	0.352
0-7	8-18	19-30	0.343
0-7	8-19	20-30	0.339
0-8	9-13	14-30	0.342
0-8	9-14	15-30	0.334
0-8	9-17	18-30	0.321
0-8	9-18	19-30	0.312
0-8	9-19	20-30	0.308
0-5	6-10	11-30	0.398
Possible NRS-itch bandings			
0-3	4-6	7-10	0.499
0-3	4-8	9-10	0.489
0-3	4-9	10	0.471
Possible raw ItchyQOL bandings			
22-50	51-74	75-110	0.338
22-50	51-75	76-110	0.332
22-50	51-82	83-110	0.330
22-50	51-85	86-110	0.328
22-58	59-74	75-110	0.379
22-58	59-75	76-110	0.372
22-58	59-82	83-110	0.373
22-58	59-85	86-110	0.373
22-50	51-80	81-110	0.332
Possible raw ItchyQOL symptoms bandings			
6-18	19-21	22-30	0.433
6-18	19-23	24-30	0.408
6-18	19-26	27-30	0.374
6-19	20-21	22-30	0.408

6-19	20-23	24-30	0.381
6-19	20-26	27-30	0.346
Possible raw ItchyQOL functioning limitation bandings			
6-18	19-27	28-35	0.413
6-18	19-28	29-35	0.401
6-18	19-29	30-35	0.387
6-19	20-27	28-35	0.393
6-19	20-28	29-35	0.381
6-19	20-29	30-35	0.367
Possible raw ItchyQOL emotion bandings			
7-18	19-29	30-45	0.288
7-18	19-33	34-45	0.281
7-23	24-29	30-45	0.310
7-23	24-33	34-45	0.305
7-26	27-29	30-45	0.331
7-26	27-33	34-45	0.327
Possible mean ItchyQOL bandings			
1-2.6	2.7-4.2	4.3-5.0	0.338
1-2.6	2.7-4.3	4.4-5.0	0.323
1-3.0	3.1-4.2	4.3-5.0	0.336
1-3.0	3.1-4.3	4.4-5.0	0.322
1-3.1	3.2-4.2	4.3-5.0	0.326
1-3.1	3.2-4.3	4.4-5.0	0.330
1-2.9	3.0-3.9	4.0-5.0	0.374
Possible mean ItchyQOL symptoms bandings			
1-3.1	3.2-3.6	3.7-5.0	0.377
1-3.1	3.2-3.9	4.0-5.0	0.381
1-3.2	3.3-3.6	3.7-5.0	0.377
1-3.2	3.3-3.9	4.0-5.0	0.381
1-3.4	3.5-3.6	3.7-5.0	0.370
1-3.4	3.5-3.9	4.0-5.0	0.374
1-2.9	3.0-3.9	4.0-5.0	0.402
Possible mean ItchyQOL functioning limitation bandings			
1-2.6	2.7-3.9	4.0-5.0	0.413
1-2.6	2.7-4.0	4.1-5.0	0.401
1-2.8	2.9-3.9	4.0-5.0	0.368
1-2.8	2.9-4.0	4.1-5.0	0.356
1-2.9	3.0-3.9	4.0-5.0	0.368
Possible mean ItchyQOL emotion bandings			
1-2.0	2.1-3.7	3.8-5.0	0.258
1-2.0	2.1-4.0	4.1-5.0	0.260
1-2.6	2.7-3.7	3.8-5.0	0.283
1-2.6	2.7-4.0	4.1-5.0	0.286
1-2.9	3.0-3.7	3.8-5.0	0.305
1-2.9	3.0-4.0	4.1-5.0	0.309

1-3.0	3.1-3.7	3.8-5.0	0.302
1-3.0	3.1-4.0	4.1-5.0	0.307
1-2.9	3.0-3.9	4.0-5.0	0.314
Possible 5-D itch scale bandings			
0-11	12-17	18-25	0.331
0-11	12-18	19-25	0.306
0-12	13-17	18-25	0.327
0-12	13-18	19-25	0.303

Supplemental Methods

All data analyses and statistical processes were performed using SAS version 9.4 (SAS Institute). Spearman rank correlation was performed to compare self-reported AD severity with POEM, DLQI, NRS-itch, raw and mean ItchyQOL and 5-D itch scores. The frequencies, mean, median and mode of self-reported mild, moderate and severe disease were determined for each integer value of the POEM, DLQI, NRS-itch, raw ItchyQoL and its symptom, physical functioning and emotion subscores and 5-D itch scales and single-decimal precision values of the mean ItchyQOL and subscores. Different values for each assessment that were associated with a one-grade increase in mean, median and/or mode self-reported AD severity were considered possible threshold values for severity strata. Potential strata were then constructed using combinations of these different threshold values. Weighted κ coefficients were determined for concordance between the potential bandings and self-reported AD severity. In addition, we tested any bandings that were previously reported for the abovementioned scales. Subgroup analyses were performed to determine how many patients had self-reported severity scores that were outside the final severity strata for each measure. Complete case analysis was performed, i.e. subjects with missing values were excluded.

Supplemental Results

Illustration of the anchor-based approach

To illustrate the anchor-based approach, selection of potential thresholds for POEM based on self-reported moderate and severe AD will be elaborated. A median and mode severity score of 2 (moderate) first occurred at a POEM score of 7, with a mean severity score of 1.7. Most POEM scores between 7 and 16 had median and mode severity scores of 2. Therefore, a POEM score of 7 was selected as a proposed threshold for moderate AD. Since not all POEM scores above 7 had median and mode severity of ≥ 2 , we also tested a POEM score of 8 as a potential threshold. A median and mode severity score of 3 (severe) first occurred at a POEM score of 17, with consistent median and/or mode severity scores of 3 occurring at a POEM score of ≥ 19 . Therefore, a POEM score of 17 was considered as a threshold for severe AD. However, a POEM score of 18 was associated with median and mode severity scores of 2. A POEM of 19 was the next score to be associated with a median and mode severity of 3. Therefore, a POEM of 19 was also considered as a potential threshold for severe AD.