Current Medical Research and Opinion

ISSN: 0300-7995 (Print) 1473-4877 (Online) Journal homepage: https://www.tandfonline.com/loi/icmo20

Evaluating patient-perceived control of atopic dermatitis: design, validation, and scoring of the Atopic Dermatitis Control Tool (ADCT)

David M. Pariser, Eric L. Simpson, Abhijit Gadkari, Thomas Bieber, David J. Margolis, Michelle Brown, Lauren Nelson, Puneet Mahajan, Matthew Reaney, Isabelle Guillemin, Usha G. Mallya & Laurent Eckert

To cite this article: David M. Pariser, Eric L. Simpson, Abhijit Gadkari, Thomas Bieber, David J. Margolis, Michelle Brown, Lauren Nelson, Puneet Mahajan, Matthew Reaney, Isabelle Guillemin, Usha G. Mallya & Laurent Eckert (2019): Evaluating patient-perceived control of atopic dermatitis: design, validation, and scoring of the Atopic Dermatitis Control Tool (ADCT), Current Medical Research and Opinion, DOI: 10.1080/03007995.2019.1699516

To link to this article: https://doi.org/10.1080/03007995.2019.1699516

Evaluating patient-perceived control of atopic dermatitis: design, validation, and scoring of the Atopic Dermatitis Control Tool (ADCT)

David M. Pariser[1], Eric L. Simpson[2], Abhijit Gadkari[3], Thomas Bieber[4], David J. Margolis[5], Michelle Brown[6], Lauren Nelson[6], Puneet Mahajan[7], Matthew Reaney[8], Isabelle Guillemin[9], Usha G. Mallya[7], Laurent Eckert[9]

[1]Eastern Virginia Medical School, Norfolk, VA, USA
[2]Department of Dermatology, Oregon Health & Science University, Portland, OR, USA
[3]Regeneron Pharmaceuticals, Inc., Tarrytown, NY, USA
[4]Department of Dermatology and Allergy, University Hospital Bonn, Bonn, Germany
[5]Department of Dermatology, University of Pennsylvania Medical Center, Philadelphia, PA, USA
[6]RTI Health Solutions, Research Triangle Park, NC, USA
[7]Sanofi, Bridgewater, NJ, USA
[8]Sanofi, Guildford, UK
[9]Sanofi, Chilly Mazarin, France

Correspondence: Laurent Eckert, Sanofi, Chilly Mazarin, France, email: Laurent.Eckert@sanofi.com

Transparency

Declaration of funding

This study was funded by Sanofi and Regeneron Pharmaceuticals, Inc.

Declaration of financial/other relationships

DMP has been a consultant to Abbott Laboratories; Amgen; Asana; Biosciences, LLC; Atacama Therapeutics; Bickel Biotechnology; Biofrontera AG; Celgene Corporation; Dermira; Dermavant Sciences; DUSA Pharmaceuticals, Inc.; Eli Lilly and Company; LEO Pharma; US, Merck & Co., Inc; Novartis Pharmaceuticals Corp.; Novo Nordisk A/S; Ortho Dermatologics; Peplin, Inc.; Pfizer, Inc.; Photocure ASA; Promius Pharmaceuticals; Regeneron; Sanofi; Stiefel, a GSK company; TDM SurgiTech, Inc.; TheraVida; and Valeant Pharmaceuticals International. ELS has received grants/research support from Amgen, Celgene, Chugai, Galderma, and Regeneron Pharmaceuticals, Inc. and is a consultant for Anacor, Asubio, Celgene, Galderma, Genentech, Medicis, and Merck. TB was lecturer and/or consultant for Regeneron, Sanofi, GSK, Celgene, Abbvie, AnaptysBio, Novartis, Abbvie, Asana Biosciences, LEO, Galapagos, Galderma, Kymab, Glenmark, Astellas, Daiichi-Sankyo, Lilly, Pfizer, Dermavant, and Allmiral. DJM has received research grants from Valeant and the National Institutes of Health for studies on atopic dermatitis and consulting monies from Valeant, Regeneron, Sanofi, Astellas, and Anacor with respect to atopic dermatitis. MB and LN are employees of RTI Health Solutions, which received research funding for the current study. PM, MR, IG, UGM, and LE are employees and shareholders of Sanofi. AG is an employee and shareholder of Regeneron. Peer reviewers on this manuscript have received an honorarium from CMRO for their review work but have no other relevant financial relationships to disclose.

Author contributions

Conceptualization: LE, AG, PM, UGM; technical input: MB, LN, MR, IG; data analysis: MB, LN; medical input: DMP, ELS, TB, DJM; writing of drafts: all authors.

Acknowledgements

Medical writing support provided by Gauri Saal, MA Economics and Akua Adu-Boahene, MD, MPH Prime, Knutsford, Cheshire, UK, funded by Sanofi and Regeneron Pharmaceuticals, Inc.

Abstract

Objectives: The Atopic Dermatitis Control Tool (ADCT) was designed to evaluate patient-perceived AD control and facilitate patient–physician discussion on long-term disease control.

Methods: The study was performed in adult patients with AD. Development of the ADCT followed FDA guidelines on patient-reported outcome measures (PROMs). Qualitative research, including targeted literature review, interviews with clinical experts and combined concept elicitation/cognitive debriefing with patients with AD, was conducted to provide a list of comprehensive concepts capturing AD control per physician and patient perspectives. Quantitative methods assessed psychometric properties of the instrument and defined the threshold for AD control.

Results: The resulting pilot six-item ADCT, reflecting key concepts related to AD control, had 7-day recall and assessed symptoms and impacts on patients' everyday lives by severity and/or frequency. The ADCT showed good content validity (well understood by adult patients with AD), and quick completion time (<2 minutes). Psychometric analysis indicated no floor/ceiling effects for response distributions, particularly strong ($r \geq 0.80$) inter-item correlations for the six ADCT items, robust construct validity ($r > 0.50$), and item-level discriminating ability ($p < 0.03$); this supported the derivation of a total score based on responses to all items. ADCT total score showed evidence of strong internal consistency reliability (Cronbach's alpha >0.80). A score ≥7 points was identified as an optimum threshold to identify patients whose AD is 'not in control'.

Conclusions: No single validated instrument has been available to holistically evaluate patient-perceived AD control. The newly developed ADCT displays good-to-excellent content validity, construct validity, internal consistency, reliability, and discriminating ability.

Keywords: Atopic dermatitis; patient-reported outcomes; long-term disease control; psychometric validation

Introduction

Atopic dermatitis (AD) is a systemic skin condition [1,2] characterized by chronic inflammation and intense pruritus [3-5], which is associated with substantial patient burden, especially in moderate-to-severe disease [6-10]. Lack of long-term disease control has been identified as a critical factor of disease burden compared with controlled AD [6], and persistent itch, sleep disturbance, functional impairment, depression, anxiety, and reduced health-related quality of life (HRQoL) and work productivity/activity participation have been identified as hallmarks [6-10]. Therefore, achieving adequate disease control, a core outcome domain for AD trials as identified by the global Harmonising Outcome Measures for Eczema (HOME) initiative [7,8], continues to be an important focus of clinical research [9]. Assessment of control is a critical input for guiding overall AD management in the clinical setting.

Currently, a limited number of AD instruments are available to evaluate patients during the clinical visit. These include the SCORing Atopic Dermatitis (SCORAD) index [10], the Eczema Area and Severity Index (EASI) [11], and the Investigator's Global Assessment (IGA) [12]. However, as these are clinician-reported outcome measures of disease activity and symptoms [13], their correlations against patient assessments of HRQoL are weak [14]. Beyond these, additional instruments that are available perform less adequately [15]. Conversely, although patient-reported outcomes measures (PROMs) are available, including the Peak Pruritus Numerical Rating Scale (Peak Pruritus NRS) [16,17] for itch, the Patient-Oriented Eczema Measure (POEM) [18] for overall AD symptoms, and the Dermatology Life Quality Index (DLQI) for HRQoL, none serve to holistically capture the multidimensional concept of patient-perceived disease control or are adapted for use in routine clinical practice. Therefore, a brief and straightforward PROM to comprehensively capture AD control, as perceived by the patient, is highly warranted.

This report describes the development and validation of the Atopic Dermatitis Control Tool (ADCT), a new instrument designed to assess patient-perceived disease control and foster patient–physician communication around disease control. It is anticipated that a brief measure such as this, targeted to adult patients with AD managed in primary or specialist care, reflecting their personal experience and requiring minimal time to complete, would have high utility in routine clinical practice as well as in non-clinical settings.

Methods

Development and validation of the ADCT followed the US Food and Drug Administration PRO Guidance [19] and was conducted in three stages of research: (Stage 1) targeted literature review and interviews with clinical experts; (Stage 2) concept elicitation and cognitive debriefing interviews with patients with AD; (Stage 3) scoring and validation of the pilot ADCT. This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki 2008 and reviewed by RTI International's Institutional Review Board. Informed consent was obtained prior to patient's participation (Stages 2 and 3).

Stage 1: Targeted literature review and interviews with clinical experts

The targeted literature review aimed to identify existing measures for use as a disease control/communication tool, or key concepts related to disease control for inclusion in the new ADCT. or. Peer-reviewed AD literature was searched in PubMed on October 16, 2015, limited to English-language studies published within 10 years of the search date. Search terms focused on AD, PROMs, clinician-reported signs and outcomes, symptom control and severity, and HRQoL, among other topics (Table S1).

A series of 60-minute guided telephone interviews were then conducted between January and February 2016 with a three-country sample of dermatologists who have extensive clinical and research experience in AD (US [n = 3], Germany [n = 1], and the United Kingdom [n = 1]). The clinician's experience in managing and communicating with adult patients with AD was first explored and followed by targeted questions to identify concepts defining disease control and encouraging dermatologist–patient interaction.

Based on a framework of the concepts identified from the literature and expert interviews, and including the American Academy of Dermatology (AAD)-issued guidelines to address important clinical questions pertaining to the care and management of AD [20], the pilot tool (ADCT v0.1) was developed. The tool was designed to be brief, easily understood, generalizable across the AD severity spectrum, and available for electronic administration during a patient's clinic visit and non-clinical settings.

Stage 2: Concept elicitation and cognitive debriefing interviews with participants with AD
The conceptual framework for the ADCT, as developed based on the published literature and feedback from expert clinicians, was further assessed based on feedback from patients with AD. As the ADCT aimed to support clinicians in collecting information on AD control as reported by patients themselves, it was deemed more appropriate to develop the items based on the literature and clinician input before engaging patients in concept elicitation.

In-person interviews were conducted with 16 adult (aged 18–75 years) participants self-reporting a clinical diagnosis of AD. The patient sample size was based on the study objective and research question and predictions about the number of study participants needed to comprehensively cover the concept (support and definition) of "AD control."

Using a recruitment screener, medical recruiters from a qualitative research facility (L&E Research, Raleigh, NC) screened and recruited interview participants. Potential participants identified from a database of individuals who had expressed interest in participating in qualitative research and provided health-related information. Recruiters then contacted these individuals via e-mail and/or telephone to determine their interest in, and eligibility for, the study.

Two rounds of interviews comprising eight patients each were conducted, allowing for testing of the changes that could have been implemented following the first round of interviews. Each one-to-one interview, following informed consent, was led by an interviewer with extensive qualitative research experience and followed a purpose-made interview guide.

Interviews were structured into two parts. In the first part, concept elicitation was aimed at identifying concepts relevant to patients' self-perception of AD control. Participants were asked to describe their experiences with AD, including the symptoms experienced and associated levels of bother, and their current treatments. The concept of 'control' as it related to the patient's AD was also explored.

In the second part, cognitive debriefing assessed how participants interpreted and selected an answer for each of the draft items of the ADCT v0.1, with additional questions posed to gather information about participants' understanding and perceptions of the concepts addressed by the items and the various response scales. In addition, patients were asked about any concepts/items they found redundant, missing, or not relevant, based on their own experience with the disease.

Feedback from the first round of patient interviews allowed refinement of ADCT v0.1. ADCT v0.2 was then tested in interviews (following the same structure as the first round, i.e. combined concept elicitation and cognitive debriefing) with a separate set of patients (n = 8) and further refined, resulting in the final content-valid version of the ADCT (ADCT v1).

All qualitative data from clinician and patient interviews were examined using Applied Thematic Analysis [21].

Stage 3: Scoring and validation of the pilot ADCT

Scoring and validation of the pilot ADCT (ADCT v1) were conducted on data collected from a cross-sectional, web-based survey of 270 adults (aged 18–75 years) in the United States self-reporting a clinical diagnosis of AD or eczema. Participants were identified and recruited through Survey Sampling International's (SSI's) online consumer panel; a prescription AD treatment (not including oral antihistamines or oral antibiotics) within the preceding 6 months was a requisite. Recruitment quotas were established based on participant self-reported AD severity to ensure a balanced sample across levels (i.e., clear/mild, moderate, and severe) and variability in ADCT scores for the evaluation of construct validity [22,23]. Participants completed the ADCT v1, demographics, AD treatment, and clinical information forms, and PROMs including the Peak Pruritus NRS [16], POEM [18], DLQI [24], Skindex-16 [25], and the AD Control Patient Global Assessment (PGA) measuring patients' overall impression of their AD control over the past week. Timestamps for each survey participant, based on a select number of web-survey screens, items, or from start to finish of the entire survey, were used to estimate average completion time of the ADCT.

Psychometric measurement properties of the ADCT v1, including internal consistency reliability, and construct validity were evaluated on the survey data.

Item-level psychometric evaluation

Item-level response distributions indicated the degree of endorsement among all survey participants. ADCT item-level floor or ceiling effects were defined as ≥40% (two times the expected value of a 5-point scale at 20%) of patients responding in the lowest or highest response category, respectively. Inter-item (polychoric) correlations indicated the strength of the associations between each pair of ADCT items and between each item-to-total score. Item-level construct validity (i.e., convergent and discriminant validity) was evaluated based on correlations between ADCT item scores and other PROMs included in the study (approximately 12 per item).

Pearson correlations were computed among continuous variables and Spearman correlations for ordinal variables. The magnitude and direction of the resulting correlation coefficients were compared against 21 *a priori* hypotheses for symptom severity (item 1) and to seven *a priori* hypotheses for nights with trouble falling or staying asleep (item 4) (Table S2), following Cohen's guideline for interpreting correlation coefficients [26]. Exploratory chi-squared tests assessed differences in ADCT item score patterns between participants classified into subgroups based on POEM severity groups, self-reported severity (screening), AD control PGA, and by well-controlled weeks (WCWs) [27-29].

ADCT scoring and total score psychometric evaluation

Based on the item-level internal consistency, construct validity and exploratory factor analysis results, a scoring algorithm was developed for the ADCT v1. A total score was created by summing the six items. The following measurement properties were assessed for the total score based on responses to all the ADCT items:

1) Construct validity: via correlation between ADCT total score and scores on other PROMs, using Cohen's guidelines for interpreting correlation and coefficients [26].

2) Discriminating ability of the ADCT total scores (known-group validity): via analyses of variance (ANOVAs); *a priori* hypotheses were used to examine mean differences in ADCT scores between participants classified into subgroups based on the other PROMs of the study (Table S2).

3) Exploratory factor analysis: via polychoric correlation matrix and weighted least-squares estimation.

ADCT thresholds for AD control

Logistic regressions were performed to predict lack of control using the AD control PGA as a 'reference standard' for 'global assessment of disease control' outcome (i.e., the dependent variable) from the total ADCT score (i.e., independent variable). For the AD control PGA, study participants who reported their AD was 'not at all controlled,' 'a little controlled,' or 'moderately controlled' were classified as 'not in control'. Study participants who reported their AD was 'mostly controlled' or 'completely controlled' were classified as 'in control'. The total ADCT score that best optimized sensitivity and specificity was selected as the ADCT threshold classifying 'in control' and 'not in control'. Area under the curve (AUC) and receiver–operator curves (ROCs) indicated the optimum total ADCT threshold. Finally, in keeping with the auxiliary goal of promoting patient–physician communication on disease control, response patterns on individual items were also explored as scoring algorithms that would maximize sensitivity while maintaining acceptable specificity.

Results

Stage 1

Results of the targeted literature review and the interviews conducted with experts during the study are detailed in Table S3. In brief, concepts of AD identified in the literature included: objective clinical signs; subjective symptoms; disease extent; flares, relapses, or exacerbations or chronicity of disease; triggers; treatment and disease management; perceived bother of AD; and impact on sleep and HRQoL. No single patient- or clinician-reported outcome instruments identified from the review comprehensively measured all the control concepts identified from the literature review. Therefore, it was deemed necessary to develop a new tool that would fill this gap and that patients could easily complete, score, and discuss with their physicians.

The interviews with clinical experts corroborated the findings from the literature review that the notion of AD control spans such a wide range of concepts.

Development of ADCT v0.1

Based on Stage 1 findings, the ADCT v0.1 comprised 11 draft items, had a 7-day recall period, and covered six concepts defining AD control: symptom severity (one item), itch (three items), bother (one item), sleep impact (two items), impact on daily activities (two items), and impact on mood and emotions (two items). No items related to flares were included in the ADCT v0.1 given the overlapping and varying definitions of both flares and AD control, and the tendency to reference symptom exacerbations in defining an AD flare, which was already being captured. Draft items also included alternative versions of the questions and response options, including those based on severity and frequency.

Stage 2

The 16 patients who took part in the concept elicitation interviews in each of rounds 1 and 2 were of mean age 39.4 (range 22–55) and 46.0 (range 28–61) years, respectively. In both rounds, half of the participants (n = 4, respectively) were female (detailed results, including demographic and clinical characteristics of interview participants, are presented in Table S4). The interviews generally confirmed the results from the literature review and expert interviews, with no new concepts emerging from the concept elicitation part of the round 1 interviews. In addition, patient feedback supported the subjective nature of the overall construct of AD control and each of the concepts that comprise it — being an experience unique to each patient, known, and reportable only by the patient.

On cognitive debriefing, participants confirmed easy recall of their symptoms of eczema and impact over the past week, given that their symptoms tended to change every few days or from week to week. Participants did not find any concepts missing, nor that any concepts were redundant. In parallel, patients indicated that five alternative versions of items were less clear or easy to answer than the original ones. The pilot ADCT v0.1 was then refined, with five items removed and other items refined to improve clarity (Table S5), to create ADCT v0.2 comprising six items, defining AD control by perceptions of symptom severity, itch and bother, and impacts on sleep, daily activities, and emotional function.

No new concepts emerged from the round 2 interviews, suggesting the saturation principle had been achieved, thus the good comprehensiveness of the research question. Data saturation characterizes the point at which no new relevant information is gained from the conduct of additional interviews [30] and is an indicator that the interview sample size yielded results that are stable and trustworthy [31,32]. Hence, no changes were made to the tool after round 2 (Table S5), as participants did not report any concepts/items to be missing, irrelevant or redundant, nor did they raise issues regarding clarity of the wording. The final version of the tool (ADCT v1; https://patient-questionnaires.sanofi.com/questionnaires/adct-atopic-dermatitis-control-communication-tool) consisted of six concepts/items assessing severity or frequency of these over a 7-day recall period (Table 1).

Stage 3

The mean age of the patients participating in the quantitative evaluation sample was 45.5 (SD 16.3) years; 77.8% were female and 73.3% were White/Caucasian (Table 2); approximately 19% were educated to only high school or less. Consistent with the recruitment quotas defined for the study, 33.3% were included in each of the three AD severity categories: clear/mild, moderate, and severe. Descriptive statistics for the other study PROMs further confirmed a uniform distribution of AD levels (Table 2). Fifty per cent of the sample reported that their AD was less than ‘mostly controlled,’ with PGA responses ranging from ‘moderately controlled’ to ‘not at all controlled’. The average survey completion time was <2 minutes.

Item-level psychometric evaluation

The following patterns were observed across the distributions of responses: the highest mean scores were observed for ‘bother,’ there were no problematic floor or ceiling effects observed, just over one-third of participants reported no impact of AD on ‘sleep,’ and nearly one-third of participants experienced no impact on daily activities or mood/emotions (Table S6).

Inter-item polychoric correlations (Table S7) informed ADCT scoring and indicated particularly strong correlations between symptom severity (item 1)/itching (item 2) and bother (item 3) (r = 0.93 and 0.85, respectively), and between daily activities (item 5) and mood/emotions (item 6) (r = 0.85). There was evidence of strong internal consistency and reliability, as measured by Cronbach's coefficient alpha (0.94) and strong item-to-total correlations.

Spearman correlations between ADCT items and other PROMs supported convergent and divergent validity; pairs of measures assessing similar constructs tended to be more strongly associated than measures assessing less related constructs. While the directions of all correlation coefficients were positive as anticipated, the magnitudes were considerably stronger than the hypothesized moderate or small correlations. Correlations ranged from 0.22 (between POEM/Dry or rough and ADCT/Night with trouble falling asleep or staying asleep) to 0.86 (between POEM/Night sleep disturbance and ADCT/ Night with trouble falling asleep or staying asleep) (Table S8).

ADCT item-level descriptive statistics by POEM severity group, patient-reported AD severity levels, PGA levels, and WCWs indicated that participants who were classified as more severe reported worse (higher) mean and median ADCT item scores than participants classified as better (lower) (Table S8). Thus, the item-level discriminant validity was confirmed. Altogether, these properties confirmed good construct validity of the ADCT and suggested there was no need for item reduction.

ADCT scoring

The strong inter-item correlations (ranging from 0.71 to 0.93), large Cronbach's alpha (≥0.80), and the robust item-level construct and discriminant validity results provide strong evidence of reliability (internal consistency) and validity, thus supporting the derivation of an ADCT total score based on the responses from all six items. The factor analysis also supported a single total score. The eigenvalue for a one-factor solution was 4.98 and all items significantly loaded on this factor, with all factor loadings above 0.80 ($p < 0.05$). An alternate two-factor solution produced a lower eigenvalue (<0.40), a high correlation between the factors (0.90) and some cross-loading. The ADCT total score mean (11.1) and median (11.0) fell approximately in the middle of the response scale (range of 0 to 24 points), and was statistically significantly higher in the PGA 'not in control' subgroup (13.0) than in the PGA 'in control' subgroup (3.0) (Table S9).

Correlations between the ADCT total and the POEM, Pruritus NRS, DLQI, Skindex-16, and PGA scores were, in general, strong (ranging from 0.76/0.79 with the Pruritus NRS, to 0.82 with the POEM and DLQI) (Table 3), thus confirming construct validity for the total score. The ADCT total score also differentiated between degrees of AD disease severity and control as assessed by other PROMs, thus confirming the discriminant ability of the total score (Table 4).

ADCT thresholds

Logistic regression indicated a threshold of ≥7 points on the ADCT total score for optimal sensitivity/specificity values (0.87/0.86), AUC (0.91), and ROC (Figure S1). The response pattern with the highest sensitivity/specificity (0.96/0.68) to identify a patient as 'not in control' was: at least 'Moderate/Moderately' on items 1, 3, 5 or 6, or '3 to 4 days' on item 2, or at least '1 or 2 nights' on item 4 (sleep) (Table 5 and Table S10). An alternative four-item pattern produced comparable sensitivity, specificity, and AUC (Table 5). Cohen's kappa between the ADCT total score threshold (≥7 points) and the primary ADCT six-item response pattern was 0.71, indicating acceptable agreement between the two schemes (Table S11).

Discussion

The six-item ADCT is a new tool developed and validated to capture, in clinical and non-clinical settings, patient-perceived AD control. The ADCT allows for the patients' self-assessment of AD control based on the patient's own perceptions of AD symptoms and impacts on their life and function (e.g., itch, sleep, daily activities, mood, and emotions), and is supported by an underlying conceptual framework derived from the literature and feedback from expert clinicians and patients. Whereas other AD PROMs developed to date do not allow for a comprehensive, standardized, and rapid way of collecting data on AD control, the structure and content of the ADCT will be of great value and help to clinical practitioners and researchers. The tool's utility is underscored by the findings of our literature review, which highlighted the lack of qualitative studies addressing patient perceptions and definitions of control.

Our intention was to develop the ADCT with sufficient rigor and following standardized methodology, involving both experts and patients, with the aim that it will be well adapted and endorsed by patients, dermatological associations, and clinical experts. The brevity of the tool and the simplicity of the items, thus resulting in a quick completion, are also likely to contribute to its good acceptability. To optimize its use, we designed the ADCT to be brief, easily understood, and self-administered via paper, online, or handheld device; easily scored and interpreted by patients; and easily completed at home or during a physician office visit. Content validity of the ADCT was confirmed through qualitative interviews with participants diagnosed with AD; psychometric measurement properties of the ADCT were evaluated and confirmed good construct validity (item-level response distributions, ADCT item-level floor or ceiling effects, inter-item correlations and item-level correlations) and good discriminant validity and reliability (internal consistency). Finally, the optimal scoring threshold was determined to identify patients whose AD would be considered 'not in control'.

This study also validated the ADCT against a patient-reported global assessment (PGA) of disease control, likely producing the most accurate and relevant data for establishing the validity of a new PRO because it is based on the patients' own experiences of the symptoms and/or functioning limitations [33]. In addition, a PGA of disease severity was used to measure patients' overall impression of their current AD severity. Both PGA scales were designed for the study specifically. Global assessment scales, whether based on patient report or clinician report, including those previously tested or validated or not (e.g., developed *de novo* for a specific study), are commonly used and accepted for the purpose of sample stratification in survey studies, instrument validation (to establish construct validity), and the development of responder definitions or meaningful change in clinical trials [34]. It is also to be noted that the PGA of disease control complemented the POEM and DLQI to provide reassurance about the quality of the findings. However, as none of the scales have been developed for the purpose of assessing AD control, it was necessary to include a global measure; in addition the results were then discussed by evaluating the totality of the evidence. Nonetheless, without formal validation of these scales, the findings must be interpreted with some caution.

As previous studies have shown a lack of concordance between patient and physician assessment of AD disease control [35], validation studies for the ADCT measure (in press [36] and ongoing), further assess the classification of 'in control' versus 'not in control,' based on the patient-reported ADCT against a physician global assessment of disease control informed by a routine clinical evaluation. Concordance indicated by these studies further highlight the need to incorporate PROMs, such as the ADCT, in clinical practice and non-clinical settings for the discussion of disease control with the patient. Ongoing research will also indicate test–retest reliability and longitudinal properties of the

ADCT, such as responsiveness to change [36]. Results from this work are also expected to support potential use of the tool in clinical trials for the treatment of AD [37].

In a patient-centric clinical model, the physician should, and often does, rely on the patient's own assessment of the degree to which their disease is controlled on current medications [38]. Indeed, improved self-management has been formally identified by some healthcare systems as a critical factor in optimizing outcomes for patients with chronic conditions [39], and the use of PROMs to enhance the patient–physician encounter, ongoing relationship, and associated quality of care and health outcomes is well established in numerous healthcare fields, and in chronic and acute health conditions [40-43]. When assessing patients with AD, the AAD recommends that physicians inquire about itch, sleep, impact on daily activity, and persistence of disease. Yet, many patients with AD infrequently visit their dermatologist and, as such, the opportunity for such conversations is limited. In addition, a large proportion of patients have been reported to rate the severity of their AD differently from their physicians [35]. However, the lack of, and need for AD PROMs for clinical use has been noted by the AAD [20] and HOME global initiative [37]. We anticipate that the ADCT will facilitate meaningful patient–physician dialogue about disease control, enabling improvements in clinical monitoring and the identification of patients requiring treatment escalation.

A few limitations should be considered regarding the methodology of the present study and the potential application of the ADCT to clinical practice. Firstly, patients were not recruited from dermatology clinics nor from patient associations (those associations contacted preferred to be involved only after the instrument was validated) and examined to secure the diagnosis and disease severity. Rather, the diagnosis of AD for the qualitative and quantitative research relied only on self-report and were not confirmed by a physician/dermatologist. However, samples of convenience (involving patient reporting of conditions [given the condition is known]) are well accepted for qualitative research for the purpose of instrument development and obtaining patient feedback on such, including the acceptance by regulatory bodies [19]. In addition, we have since performed studies in which patients have been recruited with physicians confirmed diagnosis [36] and it appears that our initial research design does not affect the scales qualities and relevance to patients.

Furthermore, disease severity was also only determined by PROMs. Secondly, patient samples for the qualitative and quantitative research, although generally representative of the US target population, were limited by practical constraints; therefore, validation in different regions and globally is warranted before wider use. Additional research is also required to assess any practical or logistical challenges for implementing the tool in daily clinical practice and to confirm the utility of the tool in shaping clinical decision-making. Finally, this version of the ADCT was developed and validated only among adult patients with AD. Subsequent research will test the content validity and psychometric properties of the tool in pediatric/adolescent AD populations. While cultural adaptation and linguistic validation, including cognitive debriefing, are needed to develop other language versions of this tool, several translations have already been conducted according to the scientific standards and are available by request to the corresponding author.

The high correlation of the ADCT with other AD PROMs could question the need to create a new instrument. Yet, the concept of 'AD control' has never been the intent of any existing questionnaire, making the work of the present study defining concepts pertaining to AD control entirely novel. None of the existing instruments fully capture the concepts of the ADCT. Some instruments will focus on symptoms but lack aspects of mood and emotions, whereas instruments that measure HRQoL will not comprehensively cover AD symptoms. These shortcomings in the existing

instruments are underscored by the fact that AD control varies with both AD symptoms and HRQoL, as indicated by the measurement properties of the ADCT.

The present study has also demonstrated that AD control can be measured efficiently and validly through a relatively small number of items, rather than a battery of measures which may cause potential burden to patients in completing them. Not only does the ADCT evaluate AD control in a comprehensive and standardized approach but the tool can easily be completed at home or during a clinical visit given its features.

In addition to an ADCT total score that can be used to inform clinical decision-making, we have also established a threshold score to capture AD control; the present study indicated that a total score ≥7 points to a lack of AD control at a given timepoint, whereas a separate longitudinal study has shown that the meaningful within-person change threshold of 5 points indicates improved disease control for a patient over time [36]. Given these favorable features of the ADCT, the tool has been recommended by HOME as one of a set of core outcomes instruments to measure long-term control [37].

In sum, this study has shown that the current version of the ADCT is deemed to be valid and reliable for assessing patient-perceived AD control, with a total score of ≥7 points corresponding to a lack of AD control. The tool may be valuable in facilitating patient–physician communication on disease control in clinical and non-clinical settings in the US. Finally, the use of the ADCT is free for all potential users, including patients, clinicians, and researchers (https://patient-questionnaires.sanofi.com/questionnaires/adct-atopic-dermatitis-control-communication-tool).

References

[1] Brunner PM, Silverberg JI, Guttman-Yassky E, et al. Increasing comorbidities suggest that atopic dermatitis is a systemic disorder. J Invest Dermatol. 2017;137:18-25.

[2] Thijs JL, Strickland I, Bruijnzeel-Koomen C, et al. Serum biomarker profiles suggest that atopic dermatitis is a systemic disease. J Allergy Clin Immunol. 2018;141:1523-1526.

[3] Bieber T. Atopic dermatitis. Ann Dermatol. 2010;22:125-137.

[4] Boguniewicz M, Leung DY. Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. Immunol Rev. 2011;242:233-246.

[5] Guttman-Yassky E, Nograles KE, Krueger JG. Contrasting pathogenesis of atopic dermatitis and psoriasis--part I: clinical and pathologic concepts. J Allergy Clin Immunol. 2011;127:1110-1118.

[6] Simpson EL, Guttman-Yassky E, Margolis DJ, et al. Association of inadequately controlled disease and disease severity with patient-reported disease burden in adults with atopic dermatitis. JAMA Dermatol. 2018;154:903-912.

[7] Chalmers JR, Thomas KS, Apfelbacher C, et al. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). Br J Dermatol. 2018;178:e332-e341.

[8] Schmitt J, Spuls P, Boers M, et al. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. Allergy. 2012;67:1111-1117.

[9] Leung DY, Guttman-Yassky E. Assessing the current treatment of atopic dermatitis: Unmet needs. J Allergy Clin Immunol. 2017;139:S47-S48.

[10] European Task Force on Atopic Dermatitis. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. Dermatology. 1993;186:23–31.

[11] Tofte S, Graeber M, Cherill R, et al. Eczema area and severity index (EASI): A new tool to evaluate atopic dermatitis. J Eur Acad Dermatol Venereol. 1998;11:S197.

[12] Siegfried E, Korman N, Molina C, et al. Safety and efficacy of early intervention with pimecrolimus cream 1% combined with corticosteroids for major flares in infants and children with atopic dermatitis. J Dermatolog Treat. 2006;17:143-150.

[13] Rehal B, Armstrong AW. Health outcome measures in atopic dermatitis: a systematic review of trends in disease severity and quality-of-life instruments 1985-2010. PloS One. 2011;6:e17520.

[14] Simpson E, Beck LA, Wu R, et al. Correlation between clinical and patient-reported outcomes in patients with moderate-to-severe atopic dermatitis in a phase 2 clinical trial of dupilumab. Poster presented at the 23rd World Congress of Dermatology; Vancouver (BC); 8-13 June 2015.

[15] Schmitt J, Langan S, Williams HC. What are the best outcome measurements for atopic eczema? A systematic review. J Allergy Clin Immunol. 2007;120:1389-1398.

[16] Yosipovitch G, Eckert L, Chen Z, et al. P480 Correlations of itch with quality of life and signs of atopic dermatitis across dupilumab trials. Ann Allergy Asthma Immunol. 2017;119:S95.

[17] Yosipovitch G, Reaney M, Mastey V, et al. Peak Pruritus Numerical Rating Scale: psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. 2019;181:761-769.

[18] 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;140:1513-1519.

[19] U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims 2009 [cited 2018 September 14].

[20] Eichenfield LF, Tom WL, Chamlin SL, et al. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. J Am Acad Dermatol. 2014;70:338-351.
[21] Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3:77-101.
[22] Frost MH, Reeve BB, Liepa AM, et al. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health. 2007;10 Suppl 2:S94-S105.
[23] 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.
[24] Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)-a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19:210-216.
[25] Chren MM, Lasek RJ, Sahay AP, et al. Measurement properties of Skindex-16: a brief quality-of-life measure for patients with skin diseases. J Cutan Med Surg. 2001;5:105-110.
[26] Cohen J. A power primer. Psych Bull. 1992;112:155-159.
[27] Langan SM, Schmitt J, Williams HC, et al. How are eczema 'flares' defined? A systematic review and recommendation for future studies. Br J Dermatol. 2014;170:548-556.
[28] Langan SM, Stuart B, Bradshaw L, et al. Measuring long-term disease control in patients with atopic dermatitis: A validation study of well-controlled weeks. J Allergy Clin Immunol. 2017;140:1580-1586.
[29] Howells LM, Chalmers JR, Cowdell F, et al. 'When it goes back to my normal I suppose': a qualitative study using online focus groups to explore perceptions of 'control' among people with eczema and parents of children with eczema in the UK. BMJ Open. 2017;15:11.
[30] Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity-establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 2-Assessing Respondent Understanding. Value Health. 2011;14:978-988.
[31] Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qual Res. 2008;8:137-152.
[32] Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field Methods. 2006;18:59-82.
[33] Coon CD, Cook KF. Moving from significance to real-world meaning: methods for interpreting change in clinical outcome assessment scores. Qual Life Res. 2018;27:33-40.
[34] Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. 2nd edition ed. New York: Oxford University Press; 1995.
[35] Wei W, Anderson P, Gadkari A, et al. Discordance between physician- and patient-reported disease severity in adults with atopic dermatitis: A US cross-sectional survey. Am J Clin Dermatol. 2017;18:825-835.
[36] Simpson E, Eckert L, Gadkari A, et al. Validation of the Atopic Dermatitis Control Tool (ADCT©) using a longitudinal survey of biologic-treated patients with atopic dermatitis. BMC Dermatol. 2019;19:15.
[37] Harmonising Outcome Measures for Eczema (HOME). Core Outcome Set (COS) and core outcome instruments (for clinical trials). 2019 [cited 2019 November 8]. Available from: http://www.homeforeczema.org/index.aspx.
[38] Ridd MJ, King AJL, Le Roux E, et al. Systematic review of self-management interventions for people with eczema. Br J Dermatol. 2017;177:719-734.
[39] Department of Health. A mandate from the Government to NHS England: April 2014 to March 2015. 2015 [cited 2017 June 29]. Available from: https://www.gov.uk/government/publications/nhs-mandate-2014-to-2015

[40] Olde Rikkert MGM, van der Wees PJ, Schoon Y, et al. Using patient reported outcomes measures to promote integrated care. Int J Integr Care. 2018;18:8.
[41] Small M, Vickers A, Anderson P, et al. The patient-physician partnership in asthma: real-world observations associated with clinical and patient-reported outcomes. Adv Ther. 2010;27:591-599.
[42] Hamdy FC, Donovan JL. Patient-reported outcomes following treatment for localized prostate cancer: Helping decision making for patients and their physicians. JAMA. 2017;317:1121-1123.
[43] Calvert M, Kyte D, Price G, et al. Maximising the impact of patient reported outcome assessment for patients and society. BMJ. 2019;364:k5267.

Table 1. Final pilot ADCT (ADCTv1) – concepts and response scales.

ADCTv1: Six concepts in the overarching measurement of 'AD control'	Anchors of the 5-point response scales
1. Overall severity of symptoms	None ⟺ Very severe
2. Frequency of intense episodes of itching	Not at all ⟺ Every day
3. Intensity of bother	Not at all ⟺ Extremely
4. Frequency of sleep impact	No nights ⟺ Every night
5. Intensity of daily activities impact	Not at all ⟺ Extremely
6. Intensity of mood or emotions impact	Not at all ⟺ Extremely

Table 2. Demographic and clinical characteristics of stage 3-survey participants.

Characteristic	Total (N = 270)
Age (years), mean (SD)	45.53 (16.3)
Median, min to max	45.5, 18 to 75
Sex, n (%)	
Male	60 (22.2)
Female	210 (77.8)
Race/ethnicity, n (%)	
White	198 (73.3)
Black/African American	31 (11.5)
Asian	20 (7.4)
Hispanic or Latino	9 (3.3)
Other	12 (4.4)
Highest education level, n (%)	
Less than high school	2 (0.7)
High school diploma or equivalent (GED)	48 (17.8)
Some college (or Associate's degree)	88 (32.6)
College degree (Bachelor's degree)	77 (28.5)
Some graduate school	14 (5.2)
Professional or advanced degree (e.g., master's, doctoral)	41 (15.2)
Prefer not to answer	0 (0.0)
How would you rate your eczema? n (%)	
Clear	12 (4.4)
Mild	78 (28.9)
Moderate	90 (33.3)
Severe	90 (33.3)
Age of first AD symptoms, n (%)	

Younger than 15	113 (41.9)
15–29 years	67 (24.8)
30–49 years	46 (17.0)
50 and older	37 (13.7)
I don’t remember	7 (2.6)
Age diagnosed with AD, n (%)	
Younger than 15	92 (34.1)
15–29 years	72 (26.7)
30–49 years	53 (19.6)
50 and older	49 (18.1)
I don’t remember	4 (1.5)

AD, atopic dermatitis; SD, standard deviation.

Table 3. ADCT total construct validity correlations (N = 270).

PRO Measure Item/Domain	Correlation Coefficient (r) on ADCT Total Score
POEM	
Total score	0.82*
Pruritus NRS	
Peak	0.76*
Average	0.79*
DLQI[a]	
Total	0.82*
Symptoms and feelings	0.83*
Daily activities	0.76*
Leisure	0.72*
Work and school	0.59*
Personal relationships	0.57*
Treatment	0.64*
Skindex-16	
Symptoms	0.83*
Emotions	0.79*
Functioning	0.80*
Work or do what you enjoy	0.79*
AD Control PGA	0.65*

AD, atopic dermatitis; ADCT, Atopic Dermatitis Control Tool; DLQI, Dermatology Life Quality Index; NRS, numeric rating scale; PGA, Patient Global Assessment; POEM, Patient-Oriented Eczema Measure.

Note: Spearman correlations were calculated.

*$P < 0.01$.

[a]The sample size is 270 for correlations between the ADCT total and DLQI items with the exception that DLQI Item 7a ('[if no] Problem at work or studying') has n = 170 due to the skip pattern.

ADCT total is the sum of items 1 through 6 (original scale).

Table 4. ADCT total discriminating ability.

	ADCT total		
	N	**Mean (SD)**	***p* value**
POEM severity subgroups			
None or mild	62	4.18 (2.93)	
Moderate	99	9.71 (4.82)	
Severe or very severe	109	16.34 (4.77)	
ANOVA		**Mean difference (95% CI)**	
None or mild vs. moderate		–5.53 (–6.87, –4.19)	<0.0001
None or mild vs. severe or very severe		–12.16 (–13.48, –10.84)	<0.0001
Moderate vs. severe or very severe		–6.63 (–7.94, –5.32)	<0.0001
Patient-reported AD severity (item S9)			
Clear	12	2.08 (1.93)	
Mild	78	5.91 (3.75)	
Moderate	90	11.02 (4.78)	
Severe	90	16.92 (4.89)	
ANOVA		**Mean difference (95% CI)**	
Clear vs. mild		–3.83 (–6.03, –1.63)	0.0009
Clear vs. moderate		–8.94 (–11.72, –6.16)	<0.0001
Clear vs. severe		–14.84 (–17.68, –12.00)	<0.0001
Mild vs. moderate		–5.11 (–6.44, –3.79)	<0.0001
Mild vs. severe		–11.01 (–12.36, –9.67)	<0.0001
Moderate vs. severe		–5.90 (–7.32, –4.48)	<0.0001

AD control PGA			
Not at all controlled	36	17.97 (4.52)	
A little controlled	94	13.02 (4.97)	
Moderately controlled	77	11.19 (5.56)	
Mostly controlled	54	4.78 (4.21)	
Completely controlled	9	1.11 (1.17)	
ANOVA		**Mean difference (95% CI)**	
Not at all controlled vs. a little controlled		4.95 (3.07, 6.83)	<0.0001
Not at all controlled vs. moderately controlled		6.78 (4.68, 8.88)	<0.0001
Not at all controlled vs. mostly controlled		13.19 (11.34, 15.05)	<0.0001
Not at all controlled vs. completely controlled		16.86 (13.77, 19.95)	<0.0001
A little controlled vs. moderately controlled		1.83 (0.24, 3.42)	0.0247
A little controlled vs. mostly controlled		8.24 (6.65, 9.83)	<0.0001
A little controlled vs. completely controlled		11.91 (8.60, 15.22)	<0.0001
Moderately controlled vs. mostly controlled		6.42 (4.64, 8.19)	<0.0001
Moderately controlled vs. completely controlled		10.08 (6.37, 13.80)	<0.0001
Mostly controlled vs. completely controlled		3.67 (0.83, 6.51)	0.0122
WCW			
Yes	13	2.69 (1.38)	
No	179	12.44 (6.00)	
ANOVA		**Mean difference (95% CI)**	
Yes vs. no		–9.74 (–13.04, –6.45)	<0.0001

AD, atopic dermatitis; ADCT, Atopic Dermatitis Control Tool; ANOVA; analysis of variance; PGA, Patient Global Assessment; POEM, Patient-Oriented Eczema Measure; WCWs, well-controlled weeks.

Table 5. ADCT thresholds based on response patterns: predicting PGA 'not in control' subgroup.

	Predicting PGA 'Not in Control'[a]						
ADCT Response 'Not in Control'	**True Positive**	**True Negative**	**False Positive**	**False Negative**	**Sensitivity**	**Specificity**	**AUC**
Six-item pattern (primary) Item 1 = Moderate **OR** Item 2 = 3 to 4 days **OR** Item 3 = Moderately **OR** Item 4 = 1 or 2 nights **OR** Item 5 = Moderate **OR** Item 6 = Moderate	199	43	20	8	0.96	0.68	0.82
Four-item pattern (alternate) Item 1 = Moderate **OR** Item 2 = 3 to 4 days **OR** Item 3 = Moderately **OR** Item 4 = 1 or 2 nights	199	43	20	8	0.96	0.68	0.82

ADCT, Atopic Dermatitis Control Tool; AUC, area under the curve; PGA, Patient Global Assessment.

[a]Results are from a logistic model predicting the probability of PGA not in control from each response pattern, where not in control included response categories 'not at all controlled,' 'a little controlled,' and 'moderately controlled.'

ADCT cut-off values were selected based on the maximum sensitivity and specificity.

The sensitivity is defined as the probability that the ADCT threshold will classify the participant as positive when the participant was not in control based on the PGA (i.e., true positive rate).

The specificity is defined as the probability that the ADCT threshold will classify the participant as negative when the participant was in control based on the PGA (i.e., true negative rate).

Supporting information

Table S1. PubMed literature search strategy

Table S2. *a priori* hypotheses for ADCT item-level construct validity

Table S3. Stage 1 results

Table S4. Demographic and clinical characteristics of qualitative interview participants

Table S5. Item-tracking matrix

Table S6. ADCT item-level response distributions

Table S7. ADCT inter-item correlations (N = 270)

Table S8. Construct validity correlations (N = 270)

Table S9. ADCT total scores according to level of AD control defined based on PGA

Table S10. Primary pattern using all six items producing the highest sensitivity (0.96) and acceptable level of specificity (0.68)

Table S11. Agreement between the ADCT total threshold and the ADCT 6-item response pattern

Figure S1. ADCT total score ROC