Psychometric Evaluation of Three Patient-Reported Outcome Questionnaires Assessing the Symptoms and Impacts of Atopic Dermatitis in Adults and Adolescents

Jonathan I. Silverberg[1], Eric L. Simpson[2], Leighann Litcher-Kelly[3], Jeffrey McDonald[3], Brian M. Calimlim[4], Yael A. Leshem[5]

[1] Department of Dermatology, The George Washington University School of Medicine and Health Sciences, Washington DC, USA;
[2] Department of Dermatology, Oregon Health & Science University, Portland, OR, USA;
[3] Patient-Centered Outcomes, Adelphi Values LLC, Boston, MA, USA;
[4] AbbVie Inc., North Chicago, IL, USA;
[5] Division of Dermatology, Rabin Medical Center, Petah Tikva, Israel;
Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

Submission: Revolutionizing Atopic Dermatitis (Virtual Conference; 13-14 December 2020)
Word count: [700/700] (including title and headings)

Objectives: There is a need for patient-reported outcome (PRO) questionnaires specific to atopic dermatitis (AD) that are fit for the purpose of evaluating treatment efficacy in regulated clinical trials. The Atopic Dermatitis Symptom Scale (ADerm-SS) and the Atopic Dermatitis Impact Scale (ADerm-IS) were developed according to Food and Drug Administration guidance and best measurement practices to measure the symptoms and impacts of AD, respectively; content validity has previously been established. The ADerm-SS is an 11-item questionnaire (24-hour recall; 0-10 numerical rating scale [NRS] for each item). To minimize participant burden and avoid repeated measurement of concepts potentially assessed by other instruments in clinical trials, a 7-item total symptom score (ADerm-SS TSS-7) was developed by summing items that assess concepts not measured by clinician-reported questionnaires; specifically, daily itch and skin pain, skin cracking and pain associated with cracking, and dry and flaking skin (range 0-70). The ADerm-IS uses 10 items (0-10 NRS) to calculate three domains scores: ADerm-IS Sleep sums three daily items assessing sleep impact (24-hour recall; range 0-30); ADerm-IS Daily Activities sums four items measuring limitations of household, physical, and social activities, and difficulty concentrating (7-day recall, range 0-40); ADerm-IS Emotional State sums three items measuring self-consciousness, embarrassment, and sadness (7-day recall, range 0-30). The Worst Pruritus NRS measures itch severity at its worst (single-item; 24-hour recall; range 0-10), scored by taking a weekly average of daily assessments. For all three PROs, higher scores represent worse symptoms or impact. This study evaluated the psychometric properties of the ADerm-SS TSS-7, ADerm-IS, and Worst Pruritus NRS scores.

Methods: Data from a global, randomized, double-blind, placebo-controlled multi-center clinical trial of adolescent and adult patients with moderate-to-severe AD were used to evaluate the dimensionality, reliability, and validity of scores, and calculate potential

responder definitions for minimally important change for each questionnaire. Confirmatory factor analyses (CFAs) evaluated *a priori* factor structures for the ADerm-SS TSS-7 and ADerm-IS (based on prior exploratory factor analyses) and utilized the bifactor method to support unidimensionality. Model fit was determined using multiple indices including the Comparative Fit Index (CFI >0.95) for CFA and the proportion of variance explained by the general factor for unidimensionality (ω_h>0.70). Test-retest reliability, internal consistency reliability, and convergent validity were assessed by intraclass correlation coefficient (ICC), Cronbach's Alpha (Cr-α), and correlation coefficient (r), respectively. Anchor-based analyses using patient-reported global assessments of severity and change were used to evaluate change scores for potential responder definitions. Analyses were conducted for both adolescent and adult sub-groups.

Results: Adolescents (age 12-17; n=113; 45.1% male; 74.3% white) and adults (age 18-75; n=769, 38.4% male, 71.0% white) were included in these analyses. Distribution of scores on the items of the questionnaires showed no floor or ceiling effects. The CFAs supported three domain scores of the ADerm-IS (Sleep [Items 1-3], Daily Activities [Items 4-7], Emotional State [Items 8-10]), for both adolescents and adults (CFI: 0.97-0.98). Although poor model fit and large values in the residual correlation matrices were observed in the CFAs for the ADerm-SS TSS-7, bifactor modeling consistently supported unidimensionality (ω_h: $0.78 - 0.91$). The scores for the ADerm-SS TSS-7 and ADerm-IS domains also demonstrated internal consistency reliability (Cr-α > 0.89), and all three questionnaires had adequate test-retest reliability in adults (ICC > 0.60). To evaluate convergent validity, correlations between scores on conceptually related questionnaires were calculated: ADerm-SS TSS-7 and Worst Pruritus NRS scores were strongly correlated with other patient-reported assessments of symptom severity, including the Patient Oriented Eczema Measure (POEM; range r=0.70-0.80); ADerm-IS domain scores were strongly correlated with the Dermatology Life Quality Index (DLQI; range r=0.63-0.78). Estimates of minimally important within-person change ranged as follows: 19-29 points for ADerm-SS TSS-7; 8-13 points for ADerm-IS Sleep; 10-16 points for ADerm-IS Daily Activities; 8-12 points for ADerm-IS Emotional State; and 3-4 points for Worst Pruritus NRS.

Conclusions: Results demonstrate the reliability, convergent validity, and meaning of change for the ADerm-SS TSS-7, ADerm-IS, and Worst Pruritus NRS scores, and support their use to assess the symptoms and impacts experienced by adults and adolescents with moderate-to-severe AD. The potential responder definitions for minimally important within-person change may be used to evaluate treatment effects in future clinical trials.

Disclosures:

Jonathan I. Silverberg received honoraria for advisory board, speaker, and consultant services from AbbVie, Asana, Bluefin, Boehringer-Ingelheim, Celgene, Dermavant, Dermira, Eli Lilly, Galderma, GlaxoSmithKline, Glenmark, Incyte, Kiniksa, Leo Pharma, Novartis, Pfizer, Regeneron, Sanofi, and research grants for investigator services from Galderma.

Eric L. Simpson reports grants, personal fees, and non-financial support from Eli Lilly; grants and personal fees from Anacor Pharma, GlaxoSmithKline, Regeneron Pharmaceuticals, Sanofi Genzyme, Pfizer, Leo Pharma, Eli Lilly, and Valeant Pharmaceuticals; personal fees from AbbVie, Celgene Corporation, Dermira, Galderma, Genentech, Leo Pharma, Menlo Therapeutics; and grants from MedImmune, Novartis, Roivant Sciences, Tioga Pharmaceuticals, and Vanda Pharmaceuticals.
Leighann Litcher-Kelly and **Jeffrey McDonald** are employed by Adelphi Values LLC, which received payment from AbbVie Inc. to support the research activities presented in this publication.
Brian M. Calimlim is an employee of AbbVie and may own AbbVie stock or stock options.
Yael Leshem has received honoraria or fees as a consultant from AbbVie, Sanofi, and Genentech and as an advisory board member from Sanofi and Regeneron Pharmaceuticals, Pfizer, and Dexcel Pharma, an independent research grant from Abbvie, and has, without personal compensation, provided investigator services for Eli Lilly, Pfizer, and AbbVie.

Funding Statement: AbbVie Inc., funded this study and participated in the study design; study research; collection, analysis and interpretation of data; and writing, reviewing and approving of this publication. All authors had access to the data, and participated in the development, review, and approval, and in the decision to submit this publication. No honoraria or payments were made for authorship.