CrossMark

ORIGINAL RESEARCH ARTICLE

Discordance Between Physician- and Patient-Reported Disease Severity in Adults with Atopic Dermatitis: A US Cross-Sectional Survey

Wenhui Wei[1,2,3] · Peter Anderson[2] · Abhijit Gadkari[3] · Stuart Blackburn[2] ·
Rachel Moon[2] · James Piercy[2] · Shashank Shinde[1] · Jorge Gomez[3] ·
Eric Ghorayeb[1]

Abstract

Background There is limited understanding of severity rating of atopic dermatitis in clinical practice.

Objectives To evaluate the agreement between physician- and patient-rated severity of atopic dermatitis.

Methods Data were collected from the 2014 Adelphi US Atopic Dermatitis Disease Specific Programme, a cross-sectional survey of physicians and their patients with a history of moderate-to-severe atopic dermatitis; patients voluntarily completed a questionnaire. Current disease severity (mild/moderate/severe), based on personal judgment, was rated independently by patients and their physicians. The weighted kappa statistic identified level of agreement between physicians and patients. Bivariate analyses characterized agreement; multi-nomial logistic regression identified factors associated with discordance.

Results Overall, 678 patients were included (369 [54.4%] were women, 525 [77.4%] were White, mean age was 39.3 years). Agreement was moderate (weighted kappa = 0.52): compared with physician ratings, patient-rated severity was higher in 76 patients (11.2%), lower in 137 patients (20.2%), and matched in 465 patients (68.6%). There were no differences in the rates of agreement between physician and patient ratings based on physician

specialty ($p = 0.6781$), objective severity measures [Eczema Area and Severity Index score ($p = 0.5308$), percent body surface area affected ($p = 0.9872$), and current systemic immunosuppressant use ($p = 0.9197$)]. Multivariate analysis showed patients with a worse quality of life (Dermatology Life Quality Index) were more likely to rate a higher severity (relative risk ratio 1.04, 95% confidence interval 1.00–1.08; $p = 0.0460$). Physicians were more likely to rate a higher severity with a greater physician-reported sleep disturbance (relative risk ratio 1.71, 95% confidence interval 1.01–2.89; $p = 0.0440$).

Conclusions Almost one-third of patients rated atopic dermatitis severity differently from their physicians, supporting the importance of the patient perspective in the severity assessment of atopic dermatitis and the need for greater communication between patients and physicians.

✉ Wenhui Wei
  wenhui.wei@regeneron.com

1   Sanofi, Bridgewater, NJ, USA

2   Adelphi Real World, Bollington, UK

3   Present Address: Regeneron Pharmaceuticals, Inc., 777 Old Saw Mill River Road, Tarrytown, NY 10591, USA

Key Points

There is limited information on how the severity of atopic dermatitis is rated by patients and their physicians in the real world.

This study shows a discrepancy between how physicians and patients rate the severity of atopic dermatitis; patients focus more on skin-related quality-of-life outcomes and physicians focus more on sleep disturbance.

Better communication between patients and physicians should be encouraged to ensure the management of atopic dermatitis is directed towards the needs of the patient.

1 Introduction

Atopic dermatitis (AD), a systemic inflammatory disease with autoimmune components, is a severe form of eczema pathologically characterized by skin barrier disruption mediated by a type 2 helper T-cell immune response [1]. Atopic dermatitis is part of the atopic march that may result from underlying biologic mechanisms [2], and while AD can develop in adults, onset is generally during early childhood and may extend into adult life in approximately half of the cases [3]; it has been estimated that the lifetime prevalence of AD in adults is 2–10% [4, 5]. In the absence of a cure, AD requires treatment over the lifetime of the patient.

In addition to the hallmark clinical symptoms of eczema and intense pruritus, AD is associated with a substantial and multi-faceted patient-reported burden that encompasses sleep disturbances, anxiety/depression, reduced function/productivity, and a lower quality of life (QoL) [6–14]. The presence of these effects of AD highlights the importance of understanding disease severity and impact from the patient's perspective and the need to ensure that disease management is tailored to the needs of the individual patient.

Although AD presents in moderate and severe forms in 20–37% and 10–34% of patients, respectively [15, 16], there is limited understanding on how physicians assess the severity of AD in routine clinical practice and its relationship to how patients self-report their disease severity. This challenge of appropriately and consistently assessing disease severity is exacerbated by the availability of more than 20 measures that can be used to assess severity, which are mainly used in clinical trials [17, 18], and the lack of a consensus on how to define the severity of AD. While cut-off values for defining severity have been suggested for multi-component clinical scales such as the Eczema Area and Severity Index (EASI) [19] and Scoring Atopic Dermatitis (SCORAD) [20], these scales are predominantly based on specific clinical signs, and the use of these measures may not necessarily be appropriate for daily clinical practice, nor may they adequately capture factors that patients use to define severity. Simpler global impression scales may be more readily interpretable, but the wide variability in content and use in the absence of standardization limits their value in defining severity and making comparisons across populations and studies [21].

There is a need to better understand how AD severity is rated in clinical practice, and the objective of this study was to evaluate the level of agreement between physician- and patient-rated AD severity, and to identify factors that may be associated with discordance between these severity ratings. The study also evaluated physician awareness of clinical and patient-reported outcome measures available to assess disease severity and the impact of AD on QoL.

2 Methods

2.1 Data Source and Populations

The Adelphi AD Disease Specific Programme (DSP) is a cross-sectional real-world survey fielded between the fourth quarter of 2014 and the first quarter of 2015, which captures data from the perspective of physicians and their consulting patients [22]. The collection of patient data is compliant with the Health Insurance Portability and Accountability Act of 1996 and the Health Information Technology for Economic and Clinical Health Act of 2009.

Screening and recruitment of doctors reflect nationally representative samples subject to meeting the DSP inclusion criteria, which for the current study were primary care physicians (PCPs; including internal medicine specialists), dermatologists, or allergists/immunologists who qualified between 1978 and 2011; were practicing in the USA with active involvement in the pharmaceutical management of patients with AD; and saw a minimum number of patients with moderate-to-severe AD per month (five or more for PCPs, six or more for dermatologists, and 15 or more for allergists/immunologists); and severity was assessed by the physician.

For inclusion in the current AD DSP analysis, patients were required to be adults aged 18 years or older with, in the physician's subjective opinion, a confirmed diagnosis or history of moderate or severe AD and not currently enrolled in a clinical trial. Patients who the physicians subjectively considered to have mild AD at the time of consultation could be included, provided that at some point previously in the course of their disease they were considered to have moderate or severe AD.

2.2 Outcomes

2.2.1 Physician-Reported

Physicians were requested to complete a detailed patient record form (PRF) for the next five prospective patients eligible for inclusion. Each specialist physician was requested to provide two additional prospective PRFs for patients currently receiving systemic immunosuppressant (IM) therapy. The PRFs captured characteristics of the AD and its treatment including the use of IMs, as well as information on the interference of itch and sleep disturbance on the patient's daily activities in response to the question "Based on your discussions with the patient or

perceptions, during the last week how much interference has each of the following aspects of the patient's condition caused to their activities of daily living (excluding work)?" Physicians indicated interference using a Likert scale from 0 (none at all) to 6 (complete interference), with a cut-off value ≥4 considered substantial interference. Current severity of AD was rated by physicians based on the question "What is your overall assessment of the severity of atopic dermatitis symptoms in this patient currently based on your own definitions of mild, moderate and severe?"

Physicians also evaluated the individual components of the EASI [23], including body regions, area score, and severity scores, which were then used to calculate the EASI score for each patient. The EASI score provides an overall assessment of disease severity. In addition, physicians were asked to report their awareness and frequency of use of a range of clinical and patient-reported outcome measures that may be used for assessing AD severity and disease activity including: EASI; SCORAD (a clinical measure of AD severity) [20]; Patient-Oriented Eczema Measure (POEM; evaluates the presence and impact of AD signs and symptoms) [24]; Dermatology Life Quality Index (DLQI; evaluates the impact of skin conditions on health-related QoL) [25]; peak pruritus numerical rating scale (assesses severity of the worst itch in the past 24 h) [26]; and Pruritus 5D (assesses five dimensions of background itch: degree, duration, direction, disability, and distribution) [27].

2.2.2 Patient-Reported

Patients were invited to complete a patient self-completion (PSC) form. Those who accepted completed the PSC independently of the physician and returned it in a sealed envelope. The PSC consisted of questions eliciting information on demographics and details relating to their AD, and patients self-rated their current AD severity in response to the question "Generally how bad was your atopic dermatitis in the past 24 hours?" with possible responses of mild, moderate, and severe. The PSC also included validated measures that assessed QoL and impairment of daily work productivity. Quality of life was assessed using the DLQI (score range 0–30 with higher scores indicating a greater impact on QoL) [25] and the generic five-dimension Euro-Qol (EQ-5D-3L) [28], with health utility scores estimated using the US value set [29]. The Work Productivity and Activity Impairment Questionnaire for Specific Health Problems (WPAI:SHP) [30] was used to evaluate the effects of AD on productivity. The WPAI:SHP assesses the impact of a specific disease on work productivity and non-work activity during the previous 7 days, and in the current analysis, the overall work impairment item was used to evaluate the effects of AD among patients who were employed.

2.3 Statistical Analyses

Descriptive statistics were used to characterize the number and proportions of patients whose ratings matched or disagreed with physician ratings. Physician and patient ratings reflect perceptions of 'current' AD severity, and thus could be rated as mild despite the exclusion of patients with a history of only having mild AD. A weighted kappa coefficient was derived to determine the magnitude of the inter-rater agreement between the patient and physician ratings [31], and a concordance variable was based on the observed categories of patient assessment of severity higher (vs. physician), physician assessment of severity higher (vs. patient), and matched (reference group).

Bivariate analyses identified variables that showed significant differences in patterns of agreement. For categorical data, Pearson's Chi-square tests were performed to test for significant differences between subgroups. For continuous data not considered to be normally distributed or ordinal data, Kruskal–Wallis tests were performed. All statistical tests performed were two-sided using a 5% significance level. Multinomial logistic regression was used to identify factors independently associated with a higher patient or physician assessment of severity using 'matched' as the base outcome, expressed as relative risk ratios (RRR) with their 95% confidence intervals (CIs). The choice of variables for inclusion in the model was based on a number of factors guided by a combination of bivariate analysis results in conjunction with disease knowledge to aid in the understanding of the likely correlation between variables. Standard errors were adjusted in the model to allow for the intra-group correlation (or clustering) of patients within physicians, using the Huber and White sandwich estimator of variances [32]. Where missing data were observed in a specific variable, that patient was excluded from analysis pertaining to that specific variable with no imputation of any missing data. Multi-variate analysis was performed on the set of patients that had complete data on the set of variables being analyzed. Goodness of fit for the logistic regression analysis was assessed using methods described by Cox and Snell and Cragg-Uhler/Nagelkerke as an indicator of effect sizes [33, 34]. All analyses were performed using Stata Statistical Software: Release 14 (StataCorp LP, College Station, TX, USA).

3 Results

3.1 Physician Demographic Characteristics and Awareness of Outcomes Measures

The physician population consisted of 102 PCPs (including 52 internal medicine specialists), 75 dermatologists, and 25

allergists/immunologists. Overall, 120 (59.4%) were male, and they represented all geographic regions [East, $n = 69$ (34.2%); South, $n = 28$ (13.9%); Midwest, $n = 58$ (28.7%); West, $n = 47$ (23.3%)]; 93 (46.0%) qualified between 1978 and 1990, 65 (32.2%) between 1991 and 2000, and 44 (21.8%) between 2001 and 2011. Dermatologists and allergists generally reported greater awareness relative to PCPs across all the clinical and patient-reported outcomes measures (Table 1). However, even among dermatologists and allergists/immunologists, the awareness of key measures of itch, such as the pruritus numerical rating scale, Pruritus 5D, and POEM, was low (Table 1). Additionally, there was low use of all measures, including the pruritus numerical rating scale and the DLQI to assess the impact of AD on the patient's daily life, both of which were commonly used by <5% of physicians across specialties (Table 1). While EASI was the most frequently used measure, with 69 physicians overall (34.2%) reporting use, it was commonly used by only 34 physicians (16.8%).

3.2 Patient Demographic and Disease Characteristics

The physicians provided PRFs for 1196 patients, 678 (56.7%) of whom completed a corresponding PSC, provided a rating of their current AD severity, and were included in the analysis. Of these, 369 were female (54.4%) and 525 were White (77.4%), with a mean age of

Table 1 Awareness of outcomes measures used for the assessment of atopic dermatitis

Measure	Number (%) of physicians			
	Overall ($N = 202$)	PCPs ($n = 102$)	Dermatologists ($n = 75$)	Allergists/immunologists ($n = 25$)
Eczema Area and Severity Index (EASI)				
Not aware	58 (28.7)	43 (42.2)	13 (17.3)	2 (8.0)
Aware, never used	75 (37.1)	31 (30.4)	34 (45.3)	10 (40.0)
Aware, rarely used	35 (17.3)	17 (16.7)	12 (16.0)	6 (24.0)
Aware, commonly used	34 (16.8)	11 (10.8)	16 (21.3)	7 (28.0)
SCORing Atopic Dermatitis (SCORAD)				
Not aware	83 (41.1)	57 (55.9)	23 (30.7)	3 (12.0)
Aware, never used	65 (32.2)	26 (25.5)	28 (37.3)	11 (44.0)
Aware, rarely used	26 (12.9)	13 (12.7)	9 (12.0)	4 (16.0)
Aware, commonly used	28 (13.9)	6 (5.9)	15 (20.0)	7 (28.0)
Dermatology Life Quality Index (DLQI)				
Not aware	101 (50.0)	72 (70.6)	18 (24.0)	11 (44.0)
Aware, never used	58 (28.7)	23 (22.5)	28 (37.3)	7 (28.0)
Aware, rarely used	38 (18.8)	6 (5.9)	26 (34.7)	6 (24.0)
Aware, commonly used	5 (2.5)	1 (1.0)	3 (4.0)	1 (4.0)
Patient-Oriented Eczema Measure (POEM)				
Not aware	103 (51.0)	53 (52.0)	39 (52.0)	11 (44.0)
Aware, never used	57 (28.2)	31 (30.4)	22 (29.3)	4 (16.0)
Aware, rarely used	33 (16.3)	13 (12.7)	12 (16.0)	8 (32.0)
Aware, commonly used	9 (4.5)	5 (4.9)	2 (2.7)	2 (8.0)
Pruritus numerical rating scale				
Not aware	129 (63.9)	71 (69.6)	48 (64.0)	10 (40.0)
Aware, never used	51 (25.2)	22 (21.6)	19 (25.3)	10 (40.0)
Aware, rarely used	19 (9.4)	8 (7.8)	7 (9.3)	4 (16.0)
Aware, commonly used	3 (1.5)	1 (1.0)	1 (1.3)	1 (4.0)
Pruritus 5D				
Not aware	149 (73.8)	81 (79.4)	55 (73.3)	13 (52.0)
Aware, never used	37 (18.3)	17 (16.7)	16 (21.3)	4 (16.0)
Aware, rarely used	16 (7.9)	4 (3.9)	4 (5.3)	8 (32.0)
Aware, commonly used	0 (0)	0 (0)	0 (0)	0 (0)

PCPs primary care physicians

39.3 years, and the majority (315/492 with data available; 64.0%) had been diagnosed with AD as adults (Table 2). Approximately half of the patients ($n = 312/629$ with data available; 49.6%) were currently flaring (Table 2), and the mean body surface area affected was 15.2% (Electronic Supplementary Material (ESM) Table 1).

3.3 Severity Ratings

3.3.1 Level of Agreement

As shown in Table 3, of the 678 patients, 237 (35.0%) rated their AD as mild, 343 (50.6%) as moderate, and 98 (14.5%) as severe. In contrast, the physicians rated AD as mild in 175 patients (25.8%), as moderate in 404 (59.6%), and as severe in 99 (14.6%). The concordance variable showed that while 465 patient ratings (68.6%) matched the physician ratings, where discordance is present, patients tended to rate their disease severity lower than their treating physicians; in 137 cases (20.2%), the physician rated severity higher than the patients, and in 76 cases (11.2%), the patient ratings of severity were higher than the physician ratings. These ratings resulted in an overall severity discordance in 213 cases (31.4%), with a weighted kappa of 0.52 (bootstrapped 95% CI 0.518–0.525), indicating a moderate level of agreement.

3.3.2 Bivariate Analysis

While race (White vs. other) was identified as being a significant demographic factor for discordance in AD severity ratings ($p = 0.0099$), neither specific ethnicity nor other demographic variables showed significance (Table 2). Among the clinical characteristics, neither time since diagnosis nor age at diagnosis was associated with patterns of agreement, although AD on the head and neck showed a significant association, as did the baseline symptom (day-to-day) of permanent scarring (Table 2). Although the baseline symptom (day-to-day) of papules was not deemed significant, the p value of 0.0500 was indicative of a possible relationship (ESM Table 1).

Patterns of agreement between physician and patient ratings showed no differences based on physician specialty ($p = 0.6781$; Table 2) or objective AD severity rating assessed by EASI score ($p = 0.5308$; ESM Table 1), percent body surface area affected ($p = 0.9872$; ESM Table 1), or current use of IMs ($p = 0.9197$; Table 2). In contrast, agreement patterns were significant for increased interference of AD on sleep (Fig. 1a; ESM Table 2), with less matching of severity ratings when there was sleep interference. Additionally, QoL was significantly lower, assessed using the DLQI (Fig. 1b; ESM Table 3; higher scores) and the EQ-5D-3L (Fig. 1c; ESM Table 3; lower scores), among patients who rated their disease severity higher than their physician's rating. Similarly, among employed patients, overall work impairment on the WPAI:SHP was significantly greater for patients who gave higher ratings of their AD severity than their physician (Fig. 1d; ESM Table 3).

3.3.3 Multivariate Analysis

Several variables that either trended toward significance or were identified as significant in the bivariate analysis were not significant in the regression analysis, including the baseline symptoms (day-to-day) of papules and permanent scarring, and AD on head/neck body regions (Fig. 2). However, two variables were demonstrated to be significantly associated with severity rating discordance; patient-reported QoL assessed using the DLQI, and sleep disturbance (Fig. 2). Patients were more likely to rate higher AD severity than their physician if they also reported worse QoL on the DLQI (RRR 1.04, 95% CI 1.00–1.08; $p = 0.046$) (Fig. 2a). Correspondingly, physicians were less likely to rate AD as more severe than the patient rating if the patient reported worse QoL (RRR 0.94, 95% CI 0.90–0.99; $p = 0.017$) (Fig. 2b). Conversely, physicians were more likely to rate higher AD severity than the patient when there was greater physician-reported sleep disturbance (RRR 1.71, 95% CI 1.01–2.89; $p = 0.044$) (Fig. 2b). Goodness-of-fit estimates using the methods of Cox and Snell R^2 (0.049) and Cragg-Uhler/Nagelkerke R^2 (0.061) were indicative of small effect sizes.

4 Discussion

Results from this study show a discordance between patient- and physician-reported AD severity, with an inter-rater agreement that was only of moderate magnitude (weighted kappa = 0.52); almost one-third of patients rated severity of their AD differently from how their physicians rated severity. These results are consistent with Torrelo et al. [35] who also found only moderate agreement between patient and physician perceptions of AD severity and a kappa coefficient that was identical to that of the current study. Such discordance between patient and physician perceptions of disease is common in the clinical setting and has been previously reported in other chronic conditions including osteoarthritis [36], rheumatoid arthritis [37], painful diabetic peripheral neuropathy [38], and psoriasis [39].

The use of bivariate analysis in this study was important not only for initially identifying significant variables that contribute to the observed disparity, but also for showing which variables were not significant. The patterns of

Table 2 Demographic and clinical characteristics of the study population, overall and by severity rating agreement between patients[a] and their treating physicians[b]

Variable[c]	Overall (N = 678)	By severity agreement			
		Matched (n = 465)	Patient assessed higher (n = 76)	Physician assessed higher (n = 137)	p value
Age, years, mean ± SD	39.3 ± 15.5	39.0 ± 15.4	40.6 ± 15.4	39.6 ± 16.1	0.6678
Sex, n (%)					0.7909
Male	309 (45.6)	209 (67.6)	34 (11.0)	66 (21.4)	
Female	369 (54.4)	256 (69.4)	42 (11.4)	71 (19.2)	
Race, n (%)					0.0099[d]
White	525 (77.4)	372 (70.9)	49 (9.3)	104 (19.8)	
Other	153 (22.6)	93 (60.8)	27 (17.6)	33 (21.6)	
Employment status, n (%)					0.3004
Unemployed	121 (18.2)	76 (62.8)	16 (13.2)	29 (24.0)	
Employed	544 (81.8)	381 (70.0)	58 (10.7)	105 (19.3)	
Missing	13	8	2	3	
Body mass index, n (%)					0.8572
Normal (18.5–24.9 kg/m^2)	259 (38.4)	179 (69.1)	29 (11.2)	51 (19.7)	
Overweight (25.0–29.9 kg/m^2)	295 (43.7)	205 (69.5)	30 (10.2)	60 (20.3)	
Obese (≥30 kg/m^2)	121 (17.9)	79 (65.3)	16 (13.2)	26 (21.5)	
Missing	3	2	1	0	
Time since diagnosis, years, mean ± SD	12.0 ± 12.2	12.2 ± 12.3	12.3 ± 11.7	11.3 ± 12.4	0.5486
Missing	186	134	17	35	
Age at diagnosis (derived), n (%)					0.8969
<12 years	99 (20.1)	70 (70.7)	10 (10.1)	19 (19.2)	
12–17 years	78 (15.9)	49 (62.8)	10 (12.8)	19 (24.4)	
≥18 years	315 (64.0)	212 (67.3)	39 (12.4)	64 (20.3)	
Missing	186	134	17	35	
Charlson Comorbidity Index, mean ± SD	0.2 ± 0.7	0.1 ± 0.4	0.2 ± 0.9	0.3 ± 1.0	0.3063
Current immunosuppressant use, n (%)	174 (25.7)	118 (67.8)	19 (10.9)	37 (21.3)	0.9197
Current disease progression, n (%)					0.2328
Improving/stable	470 (73.0)	328 (69.8)	58 (12.3)	84 (17.9)	
Deteriorating	109 (16.9)	79 (72.5)	6 (5.5)	24 (22.0)	
Changeable	65 (10.1)	40 (61.5)	9 (13.8)	16 (24.6)	
Missing	34	18	3	13	
Current flare status, n (%)					0.9176
Flaring	312 (49.6)	218 (69.9)	37 (11.9)	57 (18.3)	
Missing	49	26	4	19	
Head and neck affected by AD, n (%)	281 (41.4)	181 (64.4)	41 (14.6)	59 (21.0)	0.0437
Baseline symptoms (day-to-day), n (%)[e]					
Permanent scarring	159 (25.3)	96 (60.4)	22 (13.8)	41 (25.8)	0.0093
Missing	49	26	4	19	
Physician primary specialty, n (%)					0.6781
PCP	319 (47.1)	215 (67.4)	35 (11.0)	69 (21.6)	
Dermatologist	279 (41.2)	198 (71.0)	29 (10.4)	52 (18.6)	

Table 2 continued

Variable[c]	Overall (N = 678)	By severity agreement			
		Matched (n = 465)	Patient assessed higher (n = 76)	Physician assessed higher (n = 137)	p value
Allergist/immunologist	80 (11.8)	52 (65.0)	12 (15.0)	16 (20.0)	

AD atopic dermatitis, PCP primary care physicians, SD standard deviation

[a] Patients rated their disease severity on a patient self-completion form in answer to the question "Generally how bad was your atopic dermatitis in the past 24 hours?"

[b] Physicians prospectively rated disease severity on the patient record form during the patient visit based on the following question "What is your overall assessment of the severity of atopic dermatitis symptoms in this patient currently based on your own definitions of mild, moderate and severe?"

[c] Patient data available for all variables unless indicated as missing

[d] p value for non-Hispanic White vs. all others; p = 0.1272 when comparing all ethnicity groups

[e] Only those symptoms present at baseline (day-to-day) that were significant are shown in Table 2. Additional baseline information is provided in the Electronic Supplementary Material

Table 3 Physician- vs. patient-reported severity of atopic dermatitis

Physician-reported severity	Patient-reported severity, n			Row total, n (%)
	Mild	Moderate	Severe	
Mild	134	39	2	175 (25.8)
Moderate	99	270	35	404 (59.6)
Severe	4	34	61	99 (14.6)
Column total, n (%)	237 (35.0)	343 (50.6)	98 (14.5)	678 (100)

agreement did not appear to be dependent on physician specialty, suggesting that physicians may have a similar misunderstanding of the patient's perspective regardless of their specialty. Additionally, three key objective clinical measures also did not appear to influence pattern of agreement, i.e., the extent of AD as measured by body surface area, use of IMs, and EASI score, which itself has been proposed as a measure of disease severity [19]. In contrast, variables that evaluated the impact on the patient demonstrated significance, including sleep disturbance, QoL using both a dermatology-specific measure and a generic measure, and work productivity.

In the multi-variate model, day-to-day (baseline) symptoms of papules and permanent scarring were not significant, although a previous study using multiple regression analysis reported that the specific clinical signs of excoriations, erythema, and edema/papulation were independent predictors of patient-rated disease severity [40]. However, in that study, these signs accounted only for 25% of the variation in patient-reported severity, with the patient-reported severity primarily based on the bothersomeness of the condition rather than overall disease activity. Additionally, in the present study, excoriations, erythema, and edema/papulation were captured as part of the EASI score rather than as individual symptoms. Of the

variables that did demonstrate significance in bivariate analysis and were included in the multi-variate model, the only two that retained significance were patient-reported QoL assessed using a dermatology-specific measure (DLQI), and sleep disturbance, which is both a frequent complaint of patients with AD and a component of their disease perception [11]. The finding that these two variables retained significance suggests that QoL-related issues contribute to patient perceptions of their overall disease severity and may likely be key drivers of discordance between the patient and physician perspectives. The importance of QoL and sleep is not surprising given the impact that AD, and the associated pruritus in particular, has on these outcomes [12, 14, 41, 42], and that sleep disturbance has also been reported to be a significant predictor of QoL in patients with AD [42].

Because QoL does not necessarily correlate with disease activity assessed using standard clinical measures [43], its identification as a factor contributing to the discordance between physician and patient perceptions suggests a need for the incorporation of QoL evaluation when assessing patients with AD. Such QoL assessment is clinically relevant considering that awareness and use of clinical and patient-reported AD measures among physicians was generally low. To our knowledge, this is the first study to
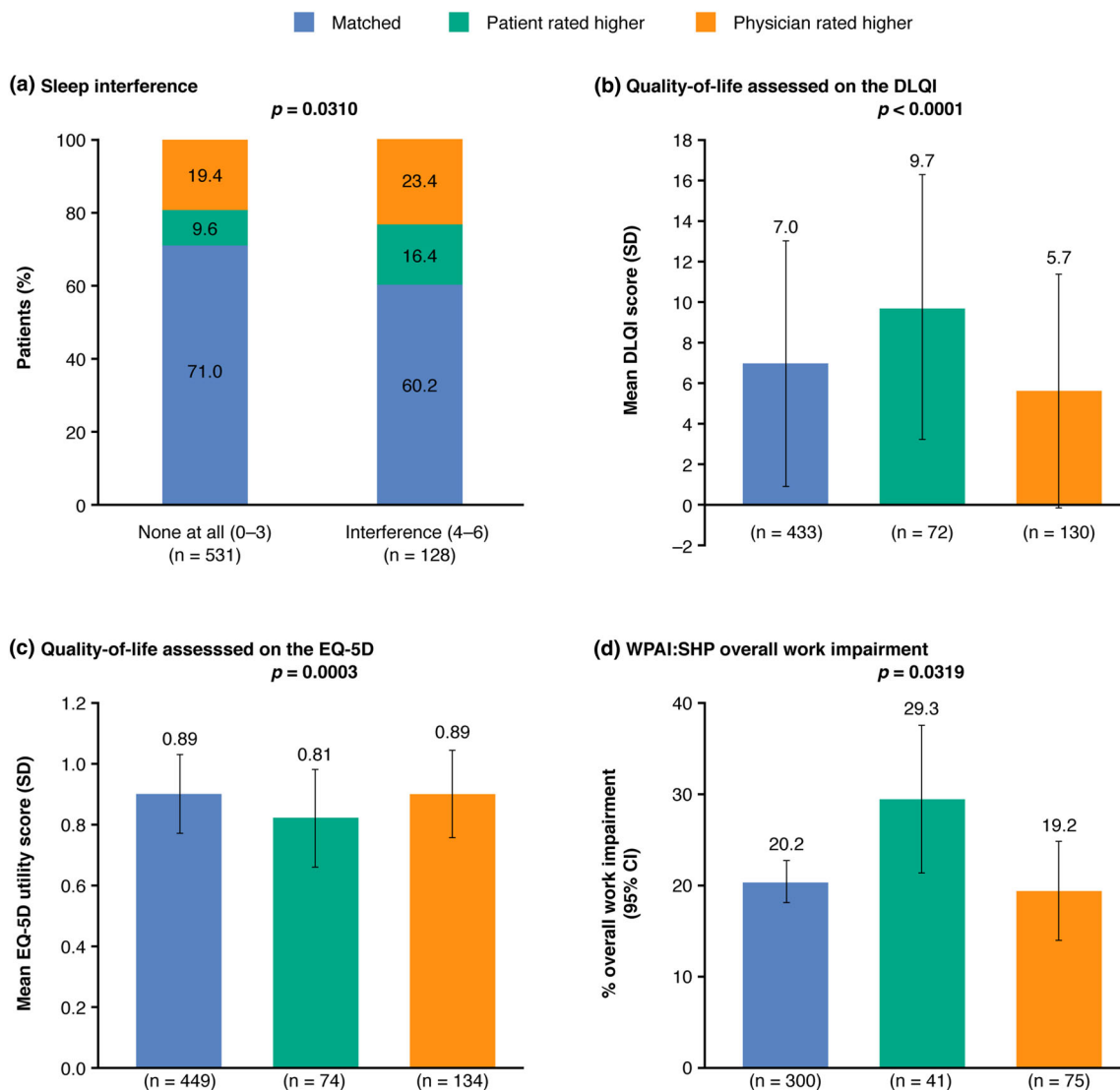
Fig. 1 Agreement patterns based on subjective measures. Sleep disturbance interference in (**a**) was rated by the clinician on a scale of 0–6. Dermatology Life Quality Index (DLQI), 5-dimension EuroQol (EQ-5D), and Work Productivity and Activity Impairment Questionnaire for Specific Health Problems (WPAI:SHP) were patient-reported outcomes. *CI* confidence interval, *SD* standard deviation

evaluate the extent of awareness and clinical use of these measures among physicians. The low awareness even among dermatologists for measures that evaluated pruritus was surprising, especially given the importance of itch to the burden of AD including daily effects on sleep disturbance and functional impairment [10, 14]. Such low awareness of AD measures also indicates that physician estimation of severity may not necessarily be based on quantitative measures of disease activity, further suggesting that clinical assessment may be less than optimal. Importantly, these results also indicate that physicians may not be having meaningful conversations with patients about QoL or the impact of AD on the daily life of the patient. The patient impact likely represents an important indicator of a need for treatment because patients may be more

concerned about what they can do and how they are perceived than by an estimated value on a severity scale.

To put the DSP physician-perceived severity findings into context, information regarding overall physician workload (from the same physicians that completed the DSP), in terms of the total numbers of AD patients seen over a 2-week period split into mild, moderate, and severe, was used to produce an estimate of the overall distribution of AD. This indicated proportions of patients with mild, moderate, and severe AD of 53.9, 42.0, and 4.1%, respectively. In addition, using information regarding total numbers of patients managed over a 12-month period by each physician included in the DSP, and taking into account the total number of physicians of each type in the USA, we estimated that the proportion of patients
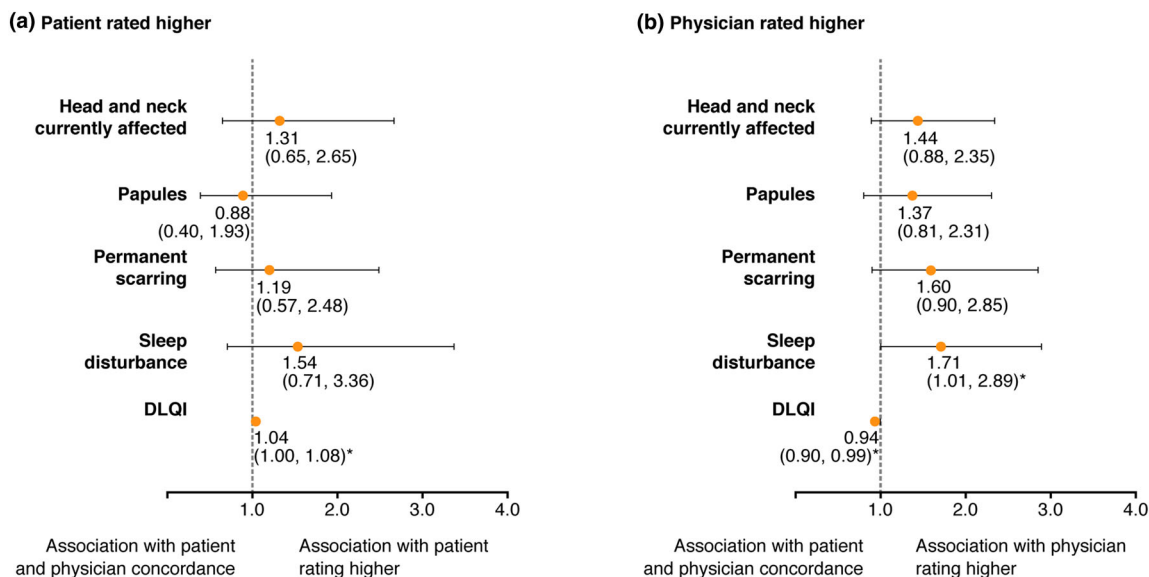
Fig. 2 Relative risk ratios for higher severity assessment and their 95% confidence intervals. Cox and Snell's R^2: 0.049; Cragg-Uhler/Nagelkerke R^2: 0.061. *DLQI* Dermatology Life Quality Index, $*p < 0.05$

treated by PCPs, dermatologists, and allergists/immunologists was 66, 28, and 5% of patients with AD, respectively. Results indicate that when AD severity is projected to the total population, severe AD represents a smaller proportion of the overall AD population than the sample included in this analysis.

4.1 Strengths and Limitations

The strengths of this study are that the patients had a clinically confirmed diagnosis of AD, and that these patients reflect the consulting population from real-world clinical practice, providing insight into how patients who met the inclusion criteria and their consulting physician rate AD severity. However, interpretation and generalizability of these results should also take into account the study limitations, including that this analysis was limited only to patients with a history of moderate-to-severe AD. In addition, disease severity rating by the patient was based on personal judgment, rather than an objective measure such as POEM. Although this may limit the generalizability of the findings to other populations, this was considered important to determine patient perceptions in this real-world setting. Another limitation is that data acquisition was reliant on the accuracy of the physician's report. There is also the potential for selection bias because those who agreed to participate may have characteristics different from those who did not agree. Last, the DSP is a cross-sectional study, and thus relationships should be considered associative rather than causal.

5 Conclusions

A discordance between patient and physician rating of AD severity was observed in almost one third of patients in this study, with QoL-related outcomes appearing to be the primary driver of this discordance. Patients tended to rate their disease of lower severity than their treating physicians, highlighting a need for greater communication between patients with AD and their treating physicians and confirming the importance of including the patient perspective when making management decisions. In addition, there is a general lack of awareness and use of clinical and patient-reported measures for assessing the severity and impact of AD, suggesting a need for greater education of physicians regarding the availability and use of such measures for the assessment of AD.

References

1. Boguniewicz M, Leung DY. Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. Immunol Rev. 2011;242(1):233–46.
2. Dharmage SC, Lowe AJ, Matheson MC, et al. Atopic dermatitis and the atopic march revisited. Allergy. 2014;69(1):17–27.
3. Napolitano M, Megna M, Patruno C, et al. Adult atopic dermatitis: a review. G Ital Dermatol Venereol. 2016;151(4):403–11.
4. Bieber T. Atopic dermatitis. N Engl J Med. 2008;358(14):1483–94.
5. Silverberg JI, Greenland P. Eczema and cardiovascular risk factors in 2 US adult population studies. J Allergy Clin Immunol. 2015;135(3):721–8.
6. 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. J Manag Care Pharm. 2002;8(5):333–42.
7. Barbeau M, Lalonde H. Burden of atopic dermatitis in Canada. Int J Dermatol. 2006;45(1):31–6.
8. Chamlin SL. The psychosocial burden of childhood atopic dermatitis. Dermatol Ther. 2006;19(2):104–7.
9. Dalgard FJ, Gieler U, Tomas-Aragones L, et al. The psychological burden of skin diseases: a cross-sectional multicenter study among dermatological out-patients in 13 European countries. J Invest Dermatol. 2015;135(4):984–91.
10. Bender BG, Leung SB, Leung DY. Actigraphy assessment of sleep disturbance in patients with atopic dermatitis: an objective life quality measure. J Allergy Clin Immunol. 2003;111(3):598–602.
11. Wittkowski A, Richards HL, Griffiths CE, Main CJ. Illness perception in individuals with atopic dermatitis. Psychol Health Med. 2007;12(4):433–44.
12. Yano C, Saeki H, Ishiji T, et al. Impact of disease severity on sleep quality in Japanese patients with atopic dermatitis. J Dermatol Sci. 2013;72(2):195–7.
13. Yano C, Saeki H, Ishiji T, et al. Impact of disease severity on work productivity and activity impairment in Japanese patients with atopic dermatitis. J Dermatol. 2013;40(9):736–9.
14. Simpson EL, Bieber T, Eckert L, et al. Patient burden of moderate-to-severe atopic dermatitis: insights from a phase 2b clinical trial of dupilumab in adults. J Am Acad Dermatol. 2016;74(3):491–8.
15. Kim JP, Chao LX, Simpson EL, Silverberg JI. Persistence of atopic dermatitis (AD): a systematic review and meta-analysis. J Am Acad Dermatol. 2016;75(4):681–7.
16. The World Allergy Organization. The World Allergy Organization white book on allergy: 2013 update. Available from: http://www.worldallergy.org/UserFiles/file/WhiteBook2-2013-v8.pdf. Accessed 13 Apr 2017.
17. 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(4):e17520.
18. Schmitt J, Spuls PI, Thomas KS, et al. The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. J Allergy Clin Immunol. 2014;134(4):800–7.
19. Leshem YA, Hajar T, Hanifin JM, Simpson EL. What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis: an interpretability study. Br J Dermatol. 2015;172(5):1353–7.
20. Kunz B, Oranje AP, Labreze L, et al. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. Dermatology. 1997;195(1):10–9.
21. Futamura M, Leshem YA, Thomas KS, et al. A systematic review of Investigator Global Assessment (IGA) in atopic dermatitis (AD) trials: many options, no standards. J Am Acad Dermatol. 2016;74(2):288–94.
22. Anderson P, Benford M, Harris N, et al. Real-world physician and patient behaviour across countries: disease-specific programmes: a means to understand. Curr Med Res Opin. 2008;24(11):3063–72.
23. Hanifin JM, Thurston M, Omoto M, et al. The Eczema Area and Severity Index (EASI): assessment of reliability in atopic dermatitis: EASI Evaluator Group. Exp Dermatol. 2001;10(1):11–8.
24. Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. Arch Dermatol. 2004;140(12):1513–9.
25. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI): a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19(3):210–6.
26. 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 Derm Venereol. 2012;92(5):502–7.
27. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. Br J Dermatol. 2010;162(3):587–93.
28. The EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199–208.
29. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care. 2005;43(3):203–20.
30. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. Pharmacoeconomics. 1993;4:353–65.
31. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988.
32. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Berkeley: University of California Press; 1967. p. 221–33.
33. Cox DR, Snell EJ. The analysis of binary data. 2nd ed. London: Chapman and Hall; 1989.
34. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika. 1991;78(3):691–2.
35. Torrelo A, Ortiz J, Alomar A, et al. Atopic dermatitis: impact on quality of life and patients' attitudes toward its management. Eur J Dermatol. 2012;22(1):97–105.
36. Bushmakin AG, Cappelleri JC, Taylor-Stokes G, et al. Relationship between patient-reported disease severity and other

clinical outcomes in osteoarthritis: a European perspective. J Med Econ. 2011;14(4):381–9.

37. Barton JL, Imboden J, Graf J, et al. Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis. Arthritis Care Res (Hoboken). 2010;62(6):857–64.

38. Taylor-Stokes G, Pike J, Sadosky A, et al. Association of patient-rated severity with other outcomes in patients with painful diabetic peripheral neuropathy. Diabetes Metab Syndr Obes. 2011;4:401–8.

39. Paul C, Bushmakin AG, Cappelleri JC, et al. Do patients and physicians agree in their assessment of the severity of psoriasis? Insights from tofacitinib phase 3 clinical trials. J Dermatol Clin Res. 2015;3(3):1048.

40. Charman CR, Venn AJ, Williams H. Measuring atopic eczema severity visually: which variables are most important to patients? Arch Dermatol. 2005;141(9):1146–51.

41. Bender BG, Ballard R, Canono B, et al. Disease severity, scratching, and sleep quality in patients with atopic dermatitis. J Am Acad Dermatol. 2008;58(3):415–20.

42. Beikert FC, Langenbruch AK, Radtke MA, et al. Willingness to pay and quality of life in patients with atopic dermatitis. Arch Dermatol Res. 2014;306(3):279–86.

43. 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.