Supplementary material has been published as submitted. It has not been copyedited, or typeset by Acta Dermato-Venereologica

**Table SI(A). Validity Measure and Methods**

| Reference | Known group | Construct | Convergent: correlation with..... | Concurrent | Divergent | Content | Criterion | Specificity | Face | Predictive | Type_notes | Other_measures | Mann-Whitney U-test | Spearman's rho | Pearson's correlation coefficient | Student's t-test | PCA | EFA | CFA | Interviews | Not stated | Method other | Other | Validity results |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| An 2010(57) | | + | | | | | | | | | | | | | | | | | | | | | | |
| Badia 1999(34) | + | + | + | | | + | | | | | | | | | | + | | | | | | | Content validity was assessed by calculating the proportion of 'Not relevant' answers. | Content validity: Total percentage of 'not relevant' answers was 5%, and no one subject had over 10% of 'not relevant' answers. Known group: Differences between eczema and psoriasis patients were not statistically significant. Construct validity: DLQI differences between patients and the general population were different (P< 0·001). |
| Baranzoni 2007(38) | | | | | | | | | | | | Feasibility tested by calculating number of missing items | | | | | | | | | | | | |
| Barbieri 2021(177) | + | + | + | | | | | | | | Construct validity: To examine the construct validity of the DLQI-R, both the DLQI and DLQI-R were calculated and their correlation with POEM, PO-SCORAD, and SF-12 scores was assessed. Known-groups validity of the DLQI and DLQI-R were assessed by comparing DLQI and DLQI-R scores across the severity categories for the POEM and PO-SCORAD | | | + | | | | | | | | | + | Known group Kruskal-Wallis H-test and effect size: POEM: DLQI, p<0.001, effect size: 0.28; DLQI-R, p<0.001, effect size 0.27, relative efficiency: 0.96. PO-SCORAD: DLQI, p<0.001, effect size: 0.28; DLQI-R, p<0.001, effect size 0.28, relative efficiency: 0.98. | Construct validity: The DLQI-R scoring modification performed similarly to the traditional DLQI score with respect to correlation with POEM and PO-SCORAD scores. Known group: Consistent with prior studies of the DLQI, more severe disease as assessed by POEM and PO-SCORAD was associated with higher DLQI scores indicating larger impact on health-related quality of life |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Beamer 2019(75) | | + | + | Concurrent: participant ranked responses on the DLQI and narrative responses on the radiation skin changes form ranged from 71% to 98%. Content:12 expert radiation oncology nurses expert opinions: adiation oncology nurses did not recommend additions to or deletion of any DLQI items | | | | |
| Bouland 2016(223) | + | | | Subgroups defined by Autoimmune Bullous Skin Disorder Intensity Score (ABSIS) and Pemphigus Disease Area Index (PDAI) | | | | DLQI score (Kruskal-Wallis nonparametric test) by subgroups moderate vs significant vs extensive (defined by 15- and 45-point cutoff values): both PADI and ABSIS p<0.001 |
| da Silva 2022(224) | + | | | | Effect size by partial eta squared (η2 p) calculated from the sum of squares of the effect in relation to the sum of squares of the effect and the sum of squares of the error associated with the effect, considering ηp 2 ≥ 0.01, ηp 2 ≥ 0.06 and ηp 2 ≥ 0.14 as small, medium, and large effects, respectively. | | Two- way univariate analyses of covariance (ANCOVA) including presence of intertriginous psoriasis, biological treatment and PASI as covariates. | Subgroup analysis by none/mild (NRS numerical rating scale ≤ 3) or medium/severe (NRS ≥ 4) pruritus F=21.46 p<0.001. Subgroup analysis by anogenital involvement F=0.36 ηp2 = 0.00. Subgroup analysis by pruritus anogenita interaction p<0.01, F=0.71 ηp2=0.01 |
| Ferraz 2006(36) | + | + | | DLQI scores (N=71) active mean =8.1 (SD 6.2) and inactive cutaneous lesions mean=3.5 (2.2) t-test p=0.0006; active and inactive cutaneous lesions 8.0 (6.6) and 4.7 (3.6) (p=0.01) | | + | | Correlation: There were highly statistically significant correlations between DLQI and SF-36 component scores, as well as between DLQI and SF-36 components. |
| Finlay 1994(1) | + | | | Scores of the patients with atopic eczema, pruritus, psoriasis, viral warts, acne were all strongly significantly higher (P<0.0001) than for the control population | There was no significant difference (P-0-67) in the D1.,QJ scores of patients with skin disease between men and women | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gabes 2021(225) | | | + | | Groups were composed of 8-12 patients or parents of patients, clinicians, methodologists and pharmaceutical industry delegates, at least one of these stakeholders was present in each group | | | | | | | |
| Gergely 2020(226) | + | + | | + | Known group effect size (ES, eta2), relative efficiency (RE), no CI given | Multivariate linear regression of HRQoL outcomes R2 (F-test) =0.275 (P < 0.001) | + | + | + | | Skindex-16. Emotions subscale: P-value < 0.001, ES: 0.090, RE: 0.555. Functioning subscale: P-value < 0.001, ES: 0.134, RE: 0.819. Symptoms subscale: P-value < 0.001, ES: 0.146, RE: 0.894. (c) EQ-5D-5L index and EQ VAS. EQ-5D-5L: P-value < 0.001, ES: 0.116, RE: 0.709. EQ VAS: P-value < 0.001, ES: 0.111, RE: 0.683. | Known group: Skindex-16 (emotions 0.555, functioning 0.819, symptoms 0.894), EQ-5D-5L (0.709) and EQ VAS (0.683) lagged behind the DLQI in differentiating between severity groups. Convergent validity: DLQI, DLQI-R, Skindex-16 total score and EQ-5D-5L index score had strong Spearman's correlations with each other (range of DQLI/EQ5D-5L rs= 0.697,DLQI/Skindex-16 total score rs=0.859,DLQI/Skindex-16 symptoms rs=-0.750,Skindex-16 emotions rs=-0.725,Skindex-16 functioning rs=0.847; HS-PGA correlated moderately with DLQI (rs = 0.418) and DLQI-R (rs = 0.433). All P < 0.05 |
| He 2013(61) | + | + | | | | | | + | | | | Spearman's: excellent correlation between the DLQI and the PDI (r =0.78, P < 0.001). Four subscales (RP, BP, SF and RE) of the SF-36 showed moderate to good correlations with the DLQI, whereas the other subscales of SF-36 and the PASI score indicated a fair correlation with DLQI (all p<0.001). Known-groups comparison, found a difference in DLQI scores among patients with different demographics and severity of psoriasis. |
| Henok 2008(55) | | | + | | | | + | | | | | Compared DLQI in new and treated patients. DLQI scores for new |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | patients ranged from 4 to 21 (median 13) and those for treated patients ranged from 1 to 18 (median 3). difference between the two groups was statistically significant (P <0.001, |
| Herédi 2014(125) | + | | | In 11 categories (incl. Clinical type of psoriasis, Localisation of psoriasis and Medical history), patients with more severe disease (responded 'Yes') reported significantly worse quality of life than the control group (Mann-Whitney U test, all p<0.05), Cohen's d effect sizes 0.24 to 0.41. | + | | | | Each of the evaluated tools (incl. DLQI) was able to discriminate between groups regarding the severity of psoriasis. |
| Hunt 2018(72) | + | | | | | | + | Kruskal-Wallis one-way analysis of variance with Wilcoxon signed-rank tests as a post-test for the subdomains | For the DLQI, Groups A&B had significantly higher (lower QOL) scores than Group C for multiple subdomain scores, including symptoms & feelings (A vs C, p ¼ 0·0004; B vs C, p = 0·001), work & school (A vs C, p = 0·003; B vs C, p = 0·006), and the total DLQI score (A vs C, p = 0·0009; B vs C, p = 0·0025). |
| Ilgen 2005(50) | + | | | Global Acne Grading System (GAGS): 0=none, 1-18=mild, 19-30=moderate, 31-38=severe, >39 very severe | + | | | | no statistically significant correlation between the DLQI scores and mild, moderate or severe acne GAGS score (p=0.575) |
| Janse 2017(147) | + | | | Patients with active disease had significantly higher DLQI scores and lower FSFI scores than patients without active disease (DLQI 12.5 +/- 7.5 vs. 4.8 +/- 4.7, P < 0.001; FSFI 21.6 +/- 9.6 vs. 27.9 +/- 8.5, P =0.009). | + | | | | Pearson's Correlation DLQI with Female Sexual Function Index (FSFI)r = -0.20, P =0.003. After removing question nine the correlation was still significant (r = □019, P =0003), indicating that poor sexual health associates with poor QoL. |
| Jesmin 2021(44) | | + | + | Convergent validity with SF-36; Criterion validity was by Pearson Correlation between PASI Score and DLQI | | + | | | Convergent validity: SF-36 Role Physical r=-0.69, Role Emotional r=-0.70, Mental Health r=-0.81, Social Functioning r=-0.79, Vitality r=-0.76 , Pain r= -0.45, p<0.001. Criterion validity: |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Person's r=0.56 signifying a moderate relationship |
| Jobanputra 2000(35) | | | + | | Logistic regression using anchor: Dermatologists assessment of clinical severity | | | + | | Severity Mild N=140 odds ratio=1.00; Moderate N=150 odds ratio=0.98, 95% CI=0.45-2.13 p=0.952; Moderately severe N=103 odds ratio=3.45, 95% CI=1.66-7.15 p=0.001; Severe N=80 odds ratio=5.52, 95% CI=2.61-11.73 p<0.0001 |
| Kage 2022(227) | + | | | AE patients with/without suicidal thoughts, plans, or suicide during the last 12 months | + | | | | | With suicidal thoughts DLQI 17 (IQR 9.5-22.5) vs without 8.0 (IQR 3-14) p=0.004 and their QoL (DLQI) was significantly decreased (p = 0.004) |
| Kent 1999(91) | + | | | | | | | | One-way ANOVA | DLQI scores: no stigma 3.1+/- 3.6, felt stigma 6.7 +/-5.3, enacted stigma 6.7+/-5.6, F=40.5, p<0.001 |
| Khoudri 2013(41) | | + | | | | | + | + | | |
| Kirby 2017(70) | | | | From mediation analysis, resilience score measured by Brief Resilient Coping Scale (BCRS)was significantly associated with depressive symptoms score (regression coefficient a = −0.21; P < .001), and depressive symptoms score (c =0.637; P < .001) was significantly associated with lower HRQOL (c′=0.644; P < .001). However, both the direct association (b = 0.033; P = .86) and the indirect association (a × b = 0.007; P = 0.87) of resilience on HRQOL (DLQI) were not significant. | | | | | |
| Koszoru 2022(196) | + | | | EASI Degree of Severity: Clear or Mild (0.0-5.9), Moderate (6-22.9), Severe (23-72) | | | | + | Effect size computed using the H statistic obtained in the Kruskal-Wallis test.55 Effect size values were interpreted as follows: small, 0.01 or greater; moderate, 0.06 or | N=218 Kruskal-Wallis test. DLQI score: clear or mild 8.04 (+/-6.80), moderate 13.24 (+/-7.86), severe 19.46 (+/-7.45); effect size=0.197; p<0.001 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | greater; and large, 0.14 or greater | |
| Koszoru 2023(205) | + | | | | EQ-5D-3L and EQ-5D-5L with DLQI grouped by Hombu banding | | | Kruskal-Wallis test and Effect size (ESs ≥ 0.01 as small, ≥ 0.06 as moderate and ≥ 0.14 as large) | EQ-5D-3L p< 0.001 Es= 0.489; EQ-5D-5L p< 0.001 Es= 0.108 |
| Lei 2016(228) | + | | | | Acne scar grading: macular, mild, moderate/severe | | | | Kruskal-Wallis test: macular N=11 median DLQI= 2.0 (IQR 7), mild median N=18 median DLQI=4.5 (IQR 6), moderate/severe N=21 median DLQI=7.0 (IQR 11); chi-sq=5.00, p<0.082 |
| Lilly 2013(62) | | + | + | + | + | Demonstrated by examining concurrent associations between self-reported severity (no measure of vitiligo severity is widely accepted) and the VitiQoL total score (r = 0.51, P<.0001). Construct validity was demonstrated using known groups comparisons and convergent associations with patients with unexposed vitiligo patches. VitiQoL behavior subscale score was significantly higher in patients with exposed vitiligo (3.50) as compared with those with nonexposed vitiligo (1.17) (P= 0.013). | | | |
| Long 2022(229) | + | | | | Hamilton Depression Scale (HAMD, 24 HAMD score ≥18 was defined as accompanied by depressive symptoms | + | | | Psoriasis with depression DLQI score 11(10-17), Psoriasis without depression DLQI score 4(2-7), F=-6.152, p <0.001 |
| Loo 2003(230) | | + | | | Score differences between two versions | + | | | 122 (63.9%) patients scored the same for text and illustrated versions. Score difference was 1 in 35 patients (18.3%), 2 in 21 (11.0%), 3 in 8 (4.2%), 4 in four (2.0%) and 5 in 1 patient (0.5%). 157 (82.2%) patients either scored the same or had a score difference of only 1. |
| Madarasingha 2011(40) | | | + | | SF36 Pearson's correlation (N=125): Role - physical r=-0.639, Vitality r=-0.499, Mental health r=-0.579, | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Social functioning r=-0.645, Role-emotional r=0.541, all p<0.01 | | | | | |
| Martínez-Ortega 2019(231) | + | + | With genital location (N=41) and without genital location (N=29) | Multiple linear regression of DLQI global score N=70 | + | | | Mean DLQI (SD) With genital location 31.5 (12.8), without genital location, 39.1 (13.4), p=0.03. Multiple linear regression DLQI with HADS depression score partial r=0.56, t(exp)=5.42, p<0.001. |
| Mazharinia 2007(54) | + | + | Burn Index | | + | | | Known group Burn index <15% DLQI=13.8 +/-5.2, >=15 DLQI= 19.7+/-4.9, Mann-Whitney U-test p=0.045. Convergent validity: There was strong association between the DA (Daily Activity Q3 and Q4), PR (Personal Relationship Q8 and Q9) scale, and total DLQI score with educational level (P = 0.023, P< 0.001, P< 0.01, respectively). |
| Mazzotti 2003(232) | + | | Different levels of clinical change as measured by variations in SAPASI scores | | | + | | Change in DLQI scores by SAPSI outcome: Improved (N=290) -3.05 ± 6.23; Unchanged (N=45) -1.37 ± 4.37 p=0.07 vs improved; Worsened (N=24) +2.08 ± 5.49 p<0.001 vs improved |
| Mazzotti 2005(51) | | + | | Psoriasis Disability Index (PDI) , Skindex-29 | | + | | Convergent validity was provided by high correlation (r=0.64) between DLQI and Psoriasis Disability Index (PDI) and Skindex-29 functioning, emotions and symptoms scales (0.72, 0.64 and 0.56 respectively) |
| McKenzie 2020(233) | + | | Grouped by Hurley stage | | | + | | Mean DLQI score (range) by Hurley Stage I 11.3 (5-19), Hurley Stage II 13.9 (8.8-21), Hurley Stage III 20.2 (13-27) ; stage I and II P<0.001, and stage II and III P=0.001 |
| Meneguin 2021(45) | | + | Spearman's: Skindex-16 Total r=0.75; Sk-16 symptoms r=0.57; k-16 emotions r=0.66; k-16 functionality r=0.70 | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nagpal 2019(220) | + | | | | white versus non-white | + | | | | | Multivariable Linear Regression Predicting Dermatologic Life Quality Index | Known group DLQI scores median (IQR): white 3.0 (1.8-7.0), non-white 8.0 (4.0-12.0; p<0.001. Multivariable Linear Regression beta = -2.261 (white vs non-white), SE=0.907, t=-2.492 p=0.014, 95% CI (-4.058, -0.464). illness perception (IP) is significantly associated with DLQI, with a 0.192 point increase in DLQI for every one point increase in IP (B=0.192, 95% CI=0.101-0.282, P<0.001). |
| Nahidi 2022(197) | + | | | | 1. Psoriasis patients versus healthy controls. 2. Psoriasis severity mild or moderate to severe | + | | | | | | DLQI scores: patients 11.75 ± 2.75, healthy controls 3.27 ± 3.57, p=0.03. DLQI scores severity mild (10.75 ± 6.8) or moderate to severe (16.4 ± 7.6), p=0.03 |
| Narang 2019(164) | + | | | | General Health Questionnaire (GHQ-12) | | | | | | | GHQ-12 < 12 DLQI =9.50 (4.0-13.50), GHQ-12 >=12 DLQI = 13.0 (8.0-19.0) chi-sq p<0.09 |
| Ozturkcan 2006(52) | | + | + | + | | + | + | + | | | | Convergent validity of DLQI with SF-36 Scale; all eight subscales (except "general health") of the SF-36 with relatively greater correlation ( r > 0.5) with pain and social relationship subscales of SF-36. Discriminant validity Mann-Whitney U-test : The overall score was found to be significantly worse in hospitalized patients (inpatients) compared with outpatients, in women compared with men, and in patients having visible dermatological illnesses (eczema, contact dermatitis and acne) compared with those who have other dermatological diagnoses, which shows the discriminative ability of the scale |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Park 2021(234) | + | + | + | | Psoriasis Life Stress Inventory (PLSI) Pearson's r = 0.74, p < 0.001; Center for Epidemiologic Studies-Depression scale (CES-D) (r = 0.70, p < 0.001 | Known group: Patients DLQI scores were were significantly different (all p<0.001) for PASI scores (mild <5, moderate 5-15, severe >=15, F = 13.09, p < 0.001), PLSI scores (low <10, high >10, t = 6.17, p < 0.001), and CES-D scores (normal <16, mild <16 or moderate depression >25, F = 32.00, p < 0.001) | | | | | |
| Patel 2019(33) | | | + | | | No patients reported that the items assessed in DLQI, ItchyQoL or 5-D itch were conceptually irrelevant to AD. | + | | + | In multivariable log-linear models controlling for EASI, NRS-itch and POEM scores DLQI was significantly associated with female sex (P=0.003), but inversely associated with age (P=0.008). | Spearman's correlation: SCORAD r = 0.55 p<0.001; mean ItchyQoL r = 0.79 p<0.001 |
| Patro 2019(165) | + | | | | Statistically significant difference in the DLQI among the different sexes, distribution pattern, duration of disease, socio-economic class (Kuppuswamy's revised socio-economic status (SES) scale), and educational status | | | | | mean ± SD DLQI according to the sex 10.67 ± 5.63 (males), 13.48 ± 4.28 (females) (P = 0.0034); duration of lesions 10.90 ± 5 (≤6 months), 14 ± 4.8 (>6 months); (P = 0.0003), distribution of lesions BSA 10.06 ± 5.34 (≤10% BSA), 12.60 ± 5.01 (>10% BSA) (P = 0.0016), SES 17.45 ± 2.73 (High SES), 10.76 ± 3.64 (Medium SES), 7.79 ± 4.88 (Low SES) (P < 0.0001), and educational status 17.24 ± 3.00 (High ES), 11.64 ± 4.11 (Medium ES), and 5.28 ± 2.00 (Low ES) (P < 0.0001). |
| Paudel 2020(81) | | | + | + | The face to face semi-structured interview using a | | | | | Concluded that the DLQI fails to adequately |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | a topic guide followed completion of the questionnaires encouraged participants through open questions, to critically evaluate the structure and content of these two questionnaires. Content validity considerqd Questionnaire content, recall period, ambiguous and repetitive phrases, missing items, sensitive and irrelevant question items. | | | capture the emotional and mental aspects of the patient's QoL. Indeed, the DLQI does not capture relevant items such as sleep and swimming but includes questions regarding relationships with friends and relatives which might have little relevance to patients with mild to moderate atopic dermatitis. Close to 40% of patients provided at least one 'not relevant' response, particularly women, elderly patients and those with low educational background |
| Rencz 2018(235) | + | | | + | Descriptive results of 'not relevant' responses: Of the 428 patients, 166 (38.8%) gave at least one NRR on the DLQI. Of these, there were 84 patients (19.6%) with one NRR, 49 (11.5%) with two NRRs, 22 (5.1%) with three NRRs, seven (1.6%) with four NRRs, one (0.2%) with five NRRs, two (0.5%) with six NRRs, none with seven NRRs and one (0.2%) with eight NRRs. Patients with DLQI scores 0-1, 2-5, 6-10 and 11-20 were 28%, 38%, 52% and 53% likely to have at least one NRR, respectively. The proportion of NRRs in patients with a DLQI ≥ 21 was 13%, on average. | 38.8% of psoriasis patients provided at least one NRR. Furthermore, more patients with DLQI scores of 6 to 20 had at least one NRR than those who did not. By eliminating DLQI items that were answered NRRs in the calculation of the total score and then converting these raw scores to scores on a 0 to 30 scale, the mean total DLQI score of the 166 patients with NRRs in our sample would increase from 7.23 to 8.94 ($P < 0.001$) |
| Safikhani 2013(236) | + | + | Open-ended and cognitive debriefing interviews | | | Overall, the results of the interview process indicated that all 10 items |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | followed a semi-structured schedule. | | | | | | | of the DLQI were clear and easy to understand (Figure 3) and relevant for measuring psoriasis-related symptoms and the impact of symptoms. The response scales were well understood and clear to 90% of the study participants, and 95% understood and provided responses consistent with the definitions in the DLQI. |
| Sahin 2022(237) | + | | | | | Good and poor sleepers by PSQI, Pittsburgh Sleep Quality Index (cut-off = 5) | | + | | | | | Mean DLQI (IQR): Good sleepers N=106 2.0 (1.0-7.0), poor sleepers N=152 7.0 (3.0-12.8), p<0.001 |
| Schwartzman 2021(46) | + | + | + | + | + | Concurrent validity Speaman's correlations: PGH-P4 T score r=-0.40, PGH-M4 T score r=-0.36, PGH-P2 T score r=-0.39, PGH-M2 T score r=-0.30, mEQ-5D r=-0.43, all p<0.001 | | | + | | | + | AUC of ROC: DLQI scores showed multilevel (significant and stepwise increases) area under the curve, indicating poor known-groups validity in predicting self-reported global atopic dermatitis severity overall. DLQI scores were unable to differentiate between the lowest 3 levels of atopic dermatitis severity. However, DLQI was slightly better at distinguishing between different levels of self-reported disease than PGH. | Convergent validity: PO-SCORAD r=0.50, PHQ9 r=0.45, PROMIS Sleep disturbance r=0.41, PROMIS Sleep related impairment r=0.38, Objective SCORAD r=0.55, SOCRAD r=0.48, all p<0.001. Known group: None of the PGH or DLQI scores were able to differentiate between the lowest 3 levels of atopic dermatitis severity. However, DLQI was slightly better at distinguishing between different levels of self-reported disease |
| Shikiar 2006(53) | + | | + | | | Change in DLQI among responder groups (defined as PASI improvement ≥ 75%; non-responder is defined as PASI improvement <50%): | One way ANOVA four categories of responders: responders, defined as those with PASI improvements ≥ 75%; "partial responders," those with PASI improvement 50- | + | + | | | | Mean Change Score for Responders (n = 66) -12.17 (SD 6.78) versus Mean Change Score for Non-Responders (n = 53) -1.77 (SD 5.52), difference -10.39, t-value 90., p<0.001, effect size = 0.40 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | 74%, inclusively; "near responders," those with PASI improvement 25-49%, inclusively; and non-responders, with <PASI25. Improvement <25% (n = 31) - 0.16 (SD 5.41), improvement 25-49% (n = 22) - 4.05 (SD 4.95) , Improvement 50-74% (n = 21) - 6.95 (SD 5.71), Improvement ≥75% (n = 66) - 12.17 (SD 6,78), overall F=30.4, all p<0.05 | | |
| Shimizu 2018(73) | | + | Correlation between DLQI and WAA-QoL (Women's Androgenetic Alopecia Quality of Life Questionnaire) | | | | rho= 0.81 (p <0.01). |
| Silpa-archa 2020(172) | + | | PHQ-9 (Depression; PHQ-9 ≥9) | | | + | PHQ-9 <9 none/mild (n=90) mean DLQI +/- SD = 6.29±5.22; PHQ-9 ≥9 moderate to severe (n=14) mean DLQI +/- SD = 15.00±5.8, p<0.001 |
| Silverberg 2020(88) | + | | Baseline numeric rating scales (NRS) and verbal rating scales (VRS) | | + | | Baseline DLQI correlations: NRS worse 0.51, NRS average 0.53, VRS worse 0.46. NRS average 0.48. frequency of itch 0.53, all p<0.001 |
| Sojevic Timotijevic 2013(238) | | + | Correlations between DLQI and EQ5D and Psoriasis Disability Index (PDI) | | + | | Strong significant correlations were found between usual activities, dimension of EQ-5D and DLQI total score (r = 0.60), daily activities (r = 0.60) and work or school (r = 0.64). All other correlations between EQ-5D and DLQI were moderate to weak (0.27 - 0.46). Mainly strong and moderate significant correlations ranging 0.26-0.84 were seen between DLQI and PDI instruments. |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Solak 2022(239) | + | | 1. Psoriasis (N=70) and control groups (N=70) 2. Patients with (N=13) and without (N=57) restless leg syndrome by International RLS Rating Scale | | + | | | DLQI score: psoriasis group 6 (3-10), control group 1 (0-3.3), p<0.001. Wit RLS 16 (6.5-20), without RLS 5 (2-7.8, p<0.001 |
| Storck 2018(77) | | + | Concordance correlation between paper and electronic methods | | | | Correlation method not stated | paper vs electronic r=0.84; electronic 1 vs 2 r=0.98 |
| Sung 2015(134) | + | | ANOVA by Pemphigus Disease Area Index (PDAI) score total and mucosal | | | | | PDAI total: ≤15 N=43 DLQI=5.95, >15 N=23 DLQI=18.09, p<0.0001. PDAI mucosa: ≤5 N=47 DLQI=7.85, >5 N=19 DLQI=15.95, p<0.0004. |
| Szabo 2022(32) | + | + | global question (GQ), 'How much does your dermatological condition affect your life?' on a 5-point scale (no effect, small effect, moderate effect, very large effect and extremely large effect on life)(Hongbo, 2005) (206) | | + | | | Skindex-16 subscale and total scores exhibited a strong Spearman's correlation both with DLQI and DLQI-R scores (range of rs = 0.664 to 0.751). World Health Organization-5 Well-Being Index (WHO-5) cores showed weak negative correlations with DLQI (ps= -0.315, p<0.05). LQI was able to better discriminate between known groups of patients based on overall HRQoL impairment (GQ rating): ANOVA F=118.7, p<0.001 |
| Takahashi 2006(37) | + | + | Weakest correlation was between DLQI-J scores and role-physical scores. As might be expected, the respondents apparently did not attribute effects of acne on role functioning to the purely physical aspects of the acne. | Known group: Subjects were divided into two "severity" groups by physicians' evaluations: severe, mean DLQI-J score of moderate-or-severe group was significantly higher than that of mild group (N=182, p<0.01); Acne symptoms were converted to a global score (7 to 35 points) and subjects divided into three groups (7-717, 18-22, 23-35). Subjects with | | + | | Correlation with SF36: Role-physical (-0.33), Vitality (-0.42), Mental health -0.48), Social functioning (-0.49), Role-emotional (-0.49), DLQI-J scores were found to be correlated with clinical severity. This was true for both physician-reported severity N=182 (mild vs moderate/severe groups p<0.01) and patient-reported severity (N=193, p<0.001) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | more severe acne symptoms had higher DLQI-J scores (N=193, p<0.001) | | | | |
| Talamonti 2021(191) | + | | + | t-test mild-to-moderate AD versus moderate-to-severe AD (EASI score < 16 or ≥16), age (<40 or ≥40 years) and disease duration (<15 and ≥15 years. | | | | EASI < 16 DLQI = 12.3 ± 5.5 (3-24); EASI ≥ 16 DLQI=14.4 ± 7.4 (0-30); p=0.0879. No statistically significant differences in mean DLQI score among patients stratified by age (<40 or ≥40 years) and disease duration (<15 and ≥15 years) |
| Tan 2022(30) | + | | | Self-assessment of Clinical Acne-Related Scars (SCARS) | | + | Chi-squared independence test with Yates correction and by Fisher's exact test.: No differences in DLQI scores across Fitzpatrick (skin type) scale grades (I to VI) p=0.223, or by gender (p=0.120) or age (p=0.116) | Chi-squared independence test with Yates correction and by Fisher's exact test.: SCARS 0-2, clear/almost clear scarring DLQI 5.00 (SEM 0.30); 3-6, mild scarring 6.34 (0.22); 7-10, moderate scarring; 11-20, severe/very severe scarring. P values for the comparison of scores across acne scarring severity groups for adjusted analyses including age, sex, residential background and country of residence, acne scar severity grade, type of skin, education, and employment status are as follows: Q1 = 0.036; Q2 = 0.001; Q3 = 0.001; Q4 = 0.051; Q5 = 0.001; Q6 = 0.087; Q7 = 0.002; Q8 = 0.017; Q9 = 0.033; Q10 = 0.020. |
| Thomas 2014(65) | | + | + | Discriminant Validity WHODAS 2.0, DLQI, and LFSQQ | | + | All three tools demonstrated lower HRQoL in LF subjects as compared to the control group. Although all domains of the WHODAS 2.0, DLQI, and LFSQQ discriminated well between LF subjects and the control group, no global tool score | Construct validity: Strongest correlation between the WHODAS 2.0 and DLQI was noted between the two total scores (r =0.748, p<0.001). Total LFSQQ score highly correlated with DLQI total score (r =-0.808, p<0.001). |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | was able to discriminate between stage II and stage III lymphedema subjects. LF stage discrimination was only noted with the DLQI symptoms subscale (p = 0.045) | |
| Wachholz 2014(66) | | + | Quick Inventory of Depressive Symptomatology-Self Report - QIDS-SR16 | | + | + | | | | Correlation between total scores (Pearson, p=0.013) and the categories of the DLQI and QIDS-SR16 (Mann-Whitney, p<0.001) was significant |
| Wallenhammar 2004(240) | + | | DLQI score distribution between genders | Regression analysis: For testing age and gender effects, and possible interaction, a logistic regression analysis was used. It revealed neither significant main effects of age and gender nor any interaction effects on DLQI dichotomized total score. | | | | + | Wilcoxon's rank sum test | There was no statistically significant difference in distribution between genders (Wilcoxon's rank sum test), females 7.3 +/- 6.3 (range 0-25), and males 7.5 +/- 5.3 (range 0-27). |
| Yang 2022(241) | + | | Between Different Severity Levels Investigator's Global Assessment (IGA) | | + | | | | | Almost Clear (n=68), Mild (n=159), Moderate (n=148), Severe (n=77), Very Severe (n=17), p=0.022 |
| Yazici 2004(98) | + | | The patients at risk for anxiety according to HAD-A had significantly higher scores on AQOL and DLQI compared to those who were not at risk. | | | + | | | | HADS Anxiety subscale <=10 N=45 DLQI=4.8 ± 4, >10 N=16 DLQ=8.6 ± 5.5, t = −2.976 df = 59 P = 0.004; HADS Depression subscale <=7 N=43 DLQI=5.1 ± 3.6, >7 N=18 DLQI=7.3 ± 6.6, t = −1.680 df = 59 P = 0.098 |
| Ye 2022(203) | + | | Urticaria Activity Score over 7 days (UAS7): well-controlled (≤ 6) N=111, mild (7-15) N=156 , moderate (16-27) N=142, severe (28-42) N=91 | | | | | + | | DLQI Mean score: UAS7 well controlled 3.1, mild 8.5, moderate 12.7, severe 17.6, p<0.001 (ANOVA) |

| | | | | | |
|---|---|---|---|---|---|
| Yi 2022 (242) | + | | | + | Genital involvement N=115 mean 8.8 (range +/-6.9); No genital Involvement N=138 mean 6.5 (range ±6.6); p=0.006. They also reported impaired sexual functioning based on question 9 in DLQI (Question 9 mean score of 0.8 vs 0.5, P = 0.046). |

**Table SI(B). Factor structure and dimensionality studies**

| Reference | Country | DLQI completed | Disease | Method used | Other | Results | Comments | COSMIN |
|---|---|---|---|---|---|---|---|---|
| Aghaei 2004 (49) | Iran | 70 | Vitiligo | FA | Construct validity was checked by factor analysis (EFA?, SPSS 11.0) | The Persian version is a two-dimensional measure including social and psychological parameters | | |
| Wallenhammar 2004 (240) | Niger | 100 | Eczema/ Hand eczema | EFA, CFA | | EFA: two underlying dimensions with loadings 0.63 to 0.83 and 0.58 - 0.90). CFA estimated correlation between DLQI '"mental" and SF-36 MCS was -0.25, and between DLQI "physical" and SF-36 PCS -0.53. | Each question (except DLQI 1) was dichotomised. In CFA, two-factor, analysis model, there was good agreement between the data and the model | 8 |
| Mazzotti 2005 (51) | Italy | 900 | Psoriasis | EFA, CFA | Exploratory FA followed by confirmatory FA (using two parameter model) | Four-factor solution (61% variance); a 2nd order factor supports unidimensionality | Two-parameter latent trait model applied to test unidimensionality, but could not be established at item level | |
| Ozturkcan 2006 (52) | Niger | 79 | Eczema-contact dermatitis, acne, psoriasis, urticaria, tinea, alopecia areata | EFA, PCA | | PCA (Varimax rotation) showed two factors (factor1 Symptoms and Feelings only, loading 0.7000 - 0.835, factor 2, all others, loading 0.438 - 0.808) | | 8, 10 |
| Takahashi 2006 (37) | Japan | 197 | Acne | PCA, FA | | Eigenvalue of the 1st component was 4.26, and 2nd was 1.02, with 42.6% of variance explained by the 1st factor. Loading on 1st factor was 0.397 to 0.863 | PCA revealed that the correlation between the "sexual difficulties" item and the total score was weak. Even though the question was limited to "because of acne" | |
| Mazharinia 2007 (54) | Iran | 109 | Burns | FA | | Factor loadings (varimax rotated) of 3-factor solution: F1 Physical Q1=0.480, Q3=0.608, Q7=0.580, Q10=0.286; F2 Psychological Q2=0.664, Q4=0.641, | FA determined the Persian version of DLQI questionnaire in burned patients is a three dimensional | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | Q6=0.340, Q8=0.556; F3 Sexual Q9=0.908 | instrument (with physical, psychological, and sexual domains) | |
| Nijsten 2007 (243) | Belgium | 450 | Psoriasis | Rasch IRT | | Overall, the DLQI significantly misfitted the Rasch model but had excellent person separation index (0.88). After stratification by country, DLQI misfitted the Rasch model in half the countries | | 9 |
| Schmitt 2007 (244) | Niger | 265 | Psoriasis | CFA | SEM correlation matrix was analysed. Model fit was analysed by CFA. | The range of estimated DLQI factor correlations were 0.556 to 0.848, which is consistent with discriminant validity. Model fit was also adequate (RMSEA 0.067; CFI 0.945; NNFI 0.980) | | 1 |
| Henok 2008 (55) | Ethiopia | 74 | Podoconiosis | PCA, FA | | FA suggested that item 4 (clothes choice) and item 10 (treatment) together accounted for 64% of variance in DLQI score in this setting | | 8 |
| An 2010(57) | China | 128 | Leprosy | FA | | Factor loadings (Varimax rotated) of three-factor solution given | Q1 (symptoms) and Q2 (feeling self-conscious) together accounted for 50.95% of the variance in DLQI | |
| Liu 2012 (42) | China | 131 | Urticaria | EFA, CFA, PCA, FA, Eigen | CFA item loading for 2 factors ranged from 0.56 to 0.76 (after standardization). Model fit, RMSEA < 0.05. CFI > 0.9 and TLI > 0.95 | PCA revealed 2 initial eigenvalues >1 with a cumulative contribution of 56.88%. Items 2 to 6 had large load on 1 factor and items 7 to 10 had large load on the other | Bartlett's test of sphericity showed v2 = 368.26 (P < 0.001); the Kaiser-Meyer- Olkin measure of sampling adequacy had a value of 0.88, which indicated that the dataset could be analysed by the factor analysis | 2 |
| Twiss 2012(59) | Niger | 292 | Psoriasis and Atopic dermatitis | PCA | Item locations and logit coverage showed items were bunched around middle of the logit scale, with few | The DLQI misfit the Rasch model (item-trait interaction, P<0.010). Confirms that the DLQI does not form a unidimensional measure of health- | PCM fit statistics for the DLQI showed Items 2 (P<0.005), 5 (P<0.013), and 7 (P<0.013) misfitted to the model. Item 5 also | 10 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | items covering the mild end. This indicates the scale does not work well with individuals who have mild disease | related quality of life (HRQL) for combined psoriasis and AD samples. The PSI indicated that the DLQI had adequate internal reliability | had a low-fit residual (<-2.5), suggesting it was redundant. Item response thresholds were disordered for items 4, 6, 7, 8, and 9, showing that the response formats for these items did not work logically | |
| An 2013 (60) | China | 395 | Neurodermatitis or psoriasis vulgaris | FA, Eigen | | Scree plot showed a sharp drop in eigenvalues from the first to the second component score | Q1 (symptoms) accounted for 50.80% of the variance in DLQI score. All loadings of the DLQI items> 0.40 | |
| He 2013 (61) | China | 851 | Psoriasis | EFA, CFA, Eigen | | EFA identified 1-factor structure accounting for 55.9% total DLQI variance. CFA indicated a good fit to original one-factor | | 1 |
| Khoudri 2013 (41) | Morocco | 244 | Psoriasis | CFA, PCA, Eigen | | CFA: GFI = 0.96, RMSEA = 0.090, and CFI = 0.93; Principal component analysis with varimax rotation loaded two factors explaining 44% and 11% of variance, loadings on factor 1 0.633-0.837 and factor 2 0.473-0.686 | | 2 |
| Lilly 2013(62) | Niger | 90 | Vitiligo | EFA | | EFA oblique rotation yielded 3 factors | | 8 |
| Ofenloch 2014 (245) | Germany | 1038 | Eczema/ Hand eczema | | | Four items (items 3, 5, 7 and 8) showed significant misfit because of fit-residuals lying outside the ± 2.5 range. Overall, the DLQI had a person separation index (pSI) of 0.82, indicating good reliability | | 9 |
| Qi 2015 (67) | China | 698 | Alopecia | FA | | Two factors were extracted from the factor solution accounting for 61.45% of the variance in DLQI score | | 9, 10 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Yale 2016 (246) | China | 149 | Neurodermatitis (lichen simplex chronicus, LSC) | PCA | Only item 9 (Infit MnSq values > 1.3) did not demonstrate acceptable goodness-of-fit to the Rasch model | DLQI did not measure a single underlying construct. However, dimension 1 accounted for 50.8 % of the variance in DLQI score in this setting | |
| He 2018 (247) | China | 9845 | Any skin disease | CFA | | 7.8% underfitted and 7.3% of persons overfitted Person and item fit. Local independence: PCA of the residuals identified no substantive residual latent dimensions (eigenvalue ≤ 1.5), supporting unidimensionality. PSI of 2.30 (r = 0.84) suggested that DLQI had adequate internal reliability, and was able to distinguish between the 3 sample subgroups in HRQL impairments | The likelihood ratio test ($\chi 2$ ($\Delta$32524) all supported a PCM over a RSM, and therefore a PCM was chosen. 3244 = 2937; $p < 0.0001$), AIC ($\Delta$9425), and BIC |
| Stull 2018 (248) | Multiple | 675 | Psoriasis | CFA | Structural equation modelling (SEM) using latent variables and manifest variables of change in DLQI total score from baseline to week 16 and from baseline to week 52 | GOF statistics for the DLQI models were all excellent, indicating that the relationships hypothesized were representative, were much better for the model where DLQISF was antecedent to the DLQI Total (revised) | |
| Xiao 2018 (74) | China | 465 | Arsenic-related skin lesions and symptoms | EFA, CFA | | EFA identified two factors with eigenvalue > 1.0 (50.3% of the variance). Factor 1 loadings 0.62 to 0.78, factor 2 0.66 to 0.72. A bi-dimensional (two-factors from EFA) CFA model with local dependency pathways between the residuals had good fit with p = 0.06, CFI = 0.95, TLI = 0.92, and RMSEA = 0.031 | 1,2 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Beamer 2019 (75) | Niger | 40 | Radio-dermatitis | PCA | | 8 of the rotated DLQI items (except work and study), loaded exclusively on 1 of 3 components that together explained 87% of total variance, with rotated factor loadings of 0.45 to 0.98 | PCA SPSS, oblimin (orthogonal) rotation | |
| Jorge 2020 (80) | Brazil | 1286 | 14 dermatoses (see Suppl. Data for the full list) | IRT, FA, PA, Eigen | Ordinal IRT using Samejima's GRM produced the best fit. Items q6 and q7 exhibited unsatisfactory fit ($p < 0.01$). All items demonstrated good discrimination ($a > 0.8$). Ordering was adequate by IRT analysis. Factor analysis eigenvalues and scree plot showed only one factor. | 1 factor found. Samejima's graded response model produced the best adjustment (AIC = 22.157; CFI = 0.98; RMSEA = 0.05; X2 = 318.9; p = 0.22) | DLQI dimensionality was assessed by Horn's parallel analysis method, using a random matrix (sphericity calculated after a Monte Carlo simulation method with 99% reliability). Factor analysis eigenvalues and scree plot showed only one factor with a random spherical matrix. Horn's parallel analysis (n = 1286; Kaiser-Meyer-Olkin test = 0.92; Bartlett's statistic = 6219.9, df= 45; $p < 0.01$) | 1,2, 5 |
| Paudel 2020 (81) | Nepal | 149 | Urticaria | PCA, FA, Eigen | | PCA method revealed two factors with eigenvalues > 1 and cumulative contribution of 65.91%. Items 1 to 3 had large loading on factor (component) 1, and items 4 to 10 loaded on factor 2. Items 4 and 5 had higher load on factor 2, though their load on factor 1 was more than 0.5 and suggested to have a shared loading on both factors | Bartlett's test of sphericity showed $\chi 2 = 781.19$ ($p< 0.001$); the Kaiser-Meyer-Olkin measure of sampling adequacy had a value of 0.89, which indicated that this dataset could be analysed for factor analysis. The cutoff value for item loading (α coefficient) was set at 0.4, with Varimax rotation and Kaiser normalization | 8 |
| Jesmin 2021 (249) | Banglades h | 80 | Psoriasis | EFA, FA | | Extracted communalities revealed all the items had values >0.50. PCA with Varimax Rotation of DLQI Bangla ranged from 0.41 to 0.83. The highest score was 0.83 for the item-5. The lowest score was 0.41 for item-3. There was only one component extracted | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Rencz 2021 (250) | Hungary | 425 | Psoriasis | PCA, Eigen | No items misfitted when scoring NRRs as missing. With zero-scoring, outfit and infit MNSQ statistics ranged from 0.564 to 1.212 | PCA (orthogonal varimax rotation) on standardized residuals of the Rasch model gave one factor (explaining 60.9% of variance). All the eigenvalues of the residuals (range 0.160-1.699) of the latent trait were < 2, and the correlations between the items' standardized residuals (range \|0.001\| - \|0.282\|) were below 0.3 supporting the unidimensional construct | PCA on the residuals of the Rasch model revealed one factor explaining 60.9% of the variance in DLQI. All the eigenvalues of the residuals (range 0.160-1.699) of the latent trait were < 2, and the correlations between the items' standardized residuals (range \|0.001\| - \|0.282\|) were below 0.3 supporting the unidimensional construct of the DLQI | 2,6, 7 |
| Tosun 2022 (251) | Niger | 390 | Any skin disease | FA | Structural model standardized path coefficients and analysis | Relationships between factor loadings and latent variables were significant ($p<0.05$) | | |

Note: Data was extracted from referenced publications.

Abbreviations. SEM: structural equation modelling, EFA: exploratory factor analysis, CFA: confirmatory factor analysis, FA: factor analysis, PCA: principal component analysis, PA: parallel analysis, correl coeff: correlation coefficient, Eigen: eigenvalues, RMSEA: Root Mean Square Error of Approximation, GOF: goodness-of-fit, PSI: Person Separation Index, PCM: partial credit model for item response theory, GRM: graded response model.

COSMIN analysis: Structural validity CTT: [1]CFA: CFI or TLI or comparable measure >0.95 OR RMSEA < 0.06 OR SRMR < 0.08.
Structural validity IRT "+" [2]IRT/Rasch: No violation of unidimensionality[a]: CFI or TLI or comparable measure >0.95 OR RMSEA < 0.06 OR SRMR < 0.08; [3]no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37; [4]no violation of monotonicity: adequate looking graphs OR item scalability >0.30; [5]adequate model fit: [6]IRT: $\chi 2$ >0.01, [7]Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z standardized values > -2 and < 2; [8]CTT "?" : Not all information for '+' reported; [9]IRT/Rasch "?" : Model fit not reported; [10]Structural validity "-" Criteria for '+' not met.
"+" = sufficient"; "?" = indeterminate. Studies with no COSMIN entry contained no data that could be analysed by COSMIN criteria.
[a]Unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) patient reported outcome measure. The criteria are based on Prinsen et al.(17)

**Table SI(C). Known group validity analysis using the DLQI**

| References | Country | DLQI completed | Disease | Methods | Results | COSMIN |
|---|---|---|---|---|---|---|
| Badia 1999(34) | Spain | 246 | Eczema and psoriasis | Student's t-test, other | Content validity: Total percentage of 'not relevant' answers was 5%, and no one subject had over 10% of 'not relevant' answers. Known group: Differences between eczema and psoriasis patients were not statistically significant, but differences between patients and the general population (N=100) were (P< 0·001). Construct validity: DLQI between patients and general population were different (P< 0·001) | DR=VG SM=A |
| Kent 1999(91) | United Kingdom | 614 | Vitiligo | one way ANOVA | Responses to open-ended question placed into three mutually exclusive categories corresponding to relevant concepts of enacted stigma, felt stigma, and no stigma. Independent judges categorised 20 randomly selected descriptions of each type. Kappa for the categories were: no stigma, k = 0.90; felt stigma, k = 1 .0; and enacted stigma, k =0.90, all p < 0.01. DLQI scores by group: no stigma 3.1+/- 3.6, felt stigma 6.7 +/-5.3, enacted stigma 6.7+/-5.6, F=40.5, p<0.001 | DR=VG SM=VG |
| Mazzotti 2003(232) | Italy | 359 | Psoriasis | ANOVA | Change in DLQI scores by Self-administered PASI (SAPSI) outcome: Improved (decrease in SAPASI score >= 2 points or more at follow-up, N=290) -3.05 ± 6.23; Unchanged (decrease or increase <2 points, N=45) -1.37 ± 4.37 p=0.07 vs improved; Worsened (increase > 2 points, N=24) +2.08 ± 5.49 p<0.001 vs improved | DR=VG SM=VG |
| Wallenhammar 2004(240) | Sweden | 100 | Hand eczema | Wilcoxon's rank sum test | There was no statistically significant difference in distribution between genders (), females 7.3 +/- 6.3 (range 0-25), and males 7.5 +/- 5.3 (range 0-27) | DR=VG SM=VG |
| Yazici 2004(98) | Türkiye | 61 | Acne | Student's t-test | Hospital Anxiety and Depression Scale (HADS) Anxiety subscale <=10 N=45 DLQI=4.8 ± 4, >10 N=16 DLQI=8.6 ± 5.5, t= −2.976 df = 59 P = 0.004; HADS Depression subscale <=7 N=43 DLQI=5.1 ± 3.6, >7 N=18 DLQI=7.3 ± 6.6, t = −1.680 df = 59 P = 0.098 | DR=VG SM=VG |
| Ilgen 2005(50) | Türkiye | 108 | Acne | Kruskal-Wallis | No statistically significant correlation between the DLQI scores and Global Acne Grading System (GAGS) score (0=none, 1–18=mild, 19–30=moderate, 31–38=severe, >39 very severe) (p=0.575) | DR=VG SM=VG |
| Ferraz 2006(36) | Brazil | 115 | Multiple for reliability incl. onychomycosis, psoriasis, contact dermatitis, solar keratosis, viral | Student-t test | The mean (SD) DLQI score in lupus erythematosus (LE) patients with active cutaneous lesions was 8.1 (6.2) while in LE patients with inactive lesions it was 3.5 (2.2), highly statistically significant (p=0.0006). When the 71 LE patients were classified as presenting alopecia or not the respective DLQI mean (SD) scores were 8.0 (6.6) and 4.7 (3.6) (p=0.01). | DR=VG SM=VG |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | warts, vitiligo. Lupus Erthematous for validity. | | | |
| Shikiar 2006(53) | United States | 147 | Psoriasis | Spearman's, Student's t-test | Mean Change Score for Responders (a patient with >75% improvement in PASI, n = 66) -12.17 (SD 6.78) versus Mean Change Score for Non-Responders (a patient with a PASI improvement <50%, n = 53) -1.77 (SD 5.52), difference -10.39, t-value 90., p<0.001, effect size = 0.40 | DR=VG SM=VG |
| Takahashi 2006(37) | Japan | 197 | Acne | One way ANOVA | DLQI-J scores were found to be correlated with clinical severity. This was true for both physician-reported severity N=182 (mild vs moderate/severe groups p<0.01) and patient-reported severity (N=193, p<0.001). No differences were found among the three age groups (teens, 20s, and all others) by one-way ANOVA (p = 0.25). | DR=VG SM=VG |
| Mazharinia 2007(54) | Iran | 109 | Burns | Mann-Whitney U-test | Known group Burn Index (total body surface areas of burned skin %TBSA) <15% DLQI=13.8 +/-5.2, >=15 DLQI= 19.7+/-4.9, Mann-Whitney U-test p=0.045. Convergent validity: There was strong association between the DA (Daily Activity Q3 and Q4), PR (Personal Relationship Q8 and Q9) scale, and total DLQI score with educational level (P = 0.023, P< 0.001, P< 0.01, respectively). | DR=VG SM=VG |
| He 2013(61) | China | 851 | Psoriasis | t-test or one-way ANOVA | Known-groups comparison found a difference in DLQI scores among patients with different demographics and severity of psoriasis; age p=0.036, gender p=0.730, region groups p=0.002, chronic/non- chronic disease p= 0.457, duration p= <0.001, PASI (0–8, n = 409; 8–12, n = 148), >12, n = 294) p= <0.001 | DR=VG SM=VG |
| Herédi 2014(125) | Hungary | 200 | Psoriasis | Mann-Whitney U-test | Each of the evaluated tools (EQ-5D, EQ-5D=VAS, PASI and DLQI) was able to discriminate between 11 categories, including clinical types and localisation, regarding the severity of psoriasis (p values and effect sizes given) | DR=VG SM=VG |
| Sung 2015(134) | Korea South | 66 | Pemphigus | ANOVA | Pemphigus Disease Area Index (PDAI) total: ≤15 N=43 DLQI=5.95, >15 N=23 DLQI=18.09, p<0.0001. PDAI mucosa: ≤5 N=47 DLQI=7.85, >5 N=19 DLQI=15.95, p<0.0004. General Health Questionnaire (GHQ-12): negative (GHQ＜4) DLQI= 5.11, positive (GHQ≥4) DLQI= 15.90, p<0.0001. Also known group by age, duration of disease, co-morbidities, treatment, disease state active or in remission | DR=VG SM=VG |

| | | | | | | |
|---|---|---|---|---|---|---|
| Bouland 2016(223) | France | 96 | Pemphigus | Kruskal-Wallis | DLQI score by three pemphigus activity subgroups (moderate vs significant vs extensive defined by (first and third quartiles): 15- and 45-point cutoff values for Pemphigus Disease Area Index (PDSI) and 17 and 53 for the Autoimmune Bullous Skin Disorder Intensity Score (ABSIS) score): PDAI p=0.02 and ABSIS p=0.03 | DR=VG SM=VG |
| Lei 2016(228) | Brunei | 50 | Acne | Kruskal-Wallis Chi-squared, Mann-Whitney | Median DLQI vs Goodman and Barron's post-acne scar grading: Macular N=11 median = 2.0 (IQR 7), mild median N=18 median DLQI=4.5 (IQR 6), moderate/severe N=21 median DLQI=7.0 (IQR 11); Kruskal-Wallis $\chi^2$=5.00, p<0.082. There was no significant difference between DLQI for male and female patients (Mann-Whitney test, p=0.132) | DR=VG SM=VG |
| Janse 2017(147) | Netherlands | 300 | Hidradenitis suppurativa | Student's t-test | Patients with active disease had significantly higher DLQI scores and lower Female Sexual Function Index (FSFI) scores than patients without active disease (DLQI 12.5 +/- 7.5 vs. 4.8 +/- 4.7, P < 0.001; FSFI 21.6 +/- 9.6 vs. 27.9 +/- 8.5, P =0.009) | DR=VG SM=VG |
| Hunt 2018(72) | Vietnam | 102 | Leprosy | Kruskal–Wallis or chi2 test (categorical variables) | For the DLQI, Groups with leprosy (Group A) or cured of leprosy (Group B) had significantly higher (lower QOL) DLQI scores than controls (Group C) for multiple subdomain scores, including symptoms & feelings (A vs C, p = 0·0004; B vs C, p = 0·001), work & school (A vs C, p = 0·003; B vs C, p = 0·006), and the total DLQI score (A vs C, p = 0·0009; B vs C, p = 0·0025) | DR=VG SM=VG |
| Martínez-Ortega 2019(231) | Spain | 70 | Psoriasis | Mann-Whitney U-test | Mean DLQI (SD) with genital location (n=42) 13.9 (7.2), without genital location (n=29) 10.0 (8.1), p=0.01. With articular location (n=42) 12.6 (8.5), Without articular location (n=28) 10.1 (6.9), p=0.24 | DR=VG SM=VG |
| Nagpal 2019(220) | United States | 132 | Acne, psoriasis, eczema | Mann-Whitney U-test, Kruskall-Wallis H test | Known group DLQI scores median (IQR): white 3.0 (1.8-7.0), non-white 8.0 (4.0-12.0); p<0.001 (Mann-Whitney U test). Multivariable Linear Regression beta = -2.261 (white vs non-white), SE=0.907, t=-2.492 p=0.014, 95% CI (-4.058, -0.464). illness perception (IP) is significantly associated with DLQI, with a 0.192 point increase in DLQI for every one point increase in IP (B=0.192, 95% CI=0.101-0.282, P<0.001). There was a significantly different overall (Kruskall-Wallis H test) DLQI score (mean,SD) between diagnosis groups (acne 4.0 (1.0-9.5), psoriasis 4.0 (2.2-13.0), eczema 7.0 (4.0-11.0), $\chi$2 (2)=7.927, P=0.019 | DR=VG SM=VG |
| Narang 2019(164) | India | 179 | Superficial cutaneous dermatophytosis | Chi-squared test, Kruskal-Wallis test | General Health Questionnaire (GHQ-12) < 12 DLQI Mean (IQR) =9.50 (4.0-13.50), GHQ-12 >=12 DLQI = 13.0 (8.0-19.0) $\chi$2 p<0.019. Median DLQI values for chronic and recurrent cases were 13.0 (IQR = 8.0-19.50) and 11.5 (IQR = 8.0-19.0), respectively; differences among the three groups were not found to be significant (Kruskal-Wallis test) | DR=VG SM=VG |

| | | | | | | |
|---|---|---|---|---|---|---|
| Patro 2019(165) | India | 294 | Superficial dermatophytic infection | Method not given | Mean ± SD DLQI according to the sex 10.67 ± 5.63 (males), 13.48 ± 4.28 (females) (P = 0.0034); duration of lesions 10.90 ± 5 (≤6 months), 14 ± 4.8 (>6 months); (P = 0.0003), distribution of lesions BSA 10.06 ± 5.34 (≤10% BSA), 12.60 ± 5.01 (>10% BSA) (P = 0.0016), Kuppuswamy's revised socio-economic status (SES): 17.45 ± 2.73 (High SES), 10.76 ± 3.64 (Medium SES), 7.79 ± 4.88 (Low SES) (P < 0.0001), and educational status 17.24 ± 3.00 (High ES), 11.64 ± 4.11 (Medium ES), and 5.28 ± 2.00 (Low ES) (P < 0.0001) | DR=VG SM=D |
| Gergely 2020(226) | Hungary | 200 | Hidradenitis Suppuratvia | Mann-Whitney U-test, Kruskal–Wallis H-test | Known group relative efficiency of the HRQoL measures with reference to the DLQI varied noticeably. Skindex-16 (emotions 0.555, functioning 0.819, symptoms 0.894), EQ-5D-5L (0.709) and EQ VAS (0.683) lagged behind the DLQI in differentiating between severity groups. Physicians' Global Assessment of HS severity (HS-PGA): Clear-minimal 4.8, Mild 8.6, Moderate 11.0, Severe 14.0, Very severe 17.3; P < 0.001 (Kruskal–Wallis), ES=0.163 | DR=VG SM=VG |
| McKenzie 2020(233) | United States | 145 | Hidradenitis Suppuratvia | Student's t-test | Mean DLQI score (Q1-Q3) by Hurley Stage I 11.3 (5-19), Hurley Stage II 13.9 (8.8-21), Hurley Stage III 20.2 (13-27); stage I and II P<0.001, and stage II and III P=0.001 | DR=VG SM=VG |
| Silpa-archa 2020(172) | Thailand | 104 | Vitiligo | t-test, chi-squared test, ANOVA | Patient Health Questionnaire-9 (PHQ-9) <9 none/mild (n=90) mean DLQI +/- SD = 6.29±5.22; PHQ-9 ≥9 moderate to severe (n=14) mean DLQI +/- SD = 15.00±5.8, p<0.001 | DR=VG SM=VG |
| Silverberg 2020(88) | United States | 410 | Atopic dermatitis | Spearman | Baseline DLQI correlations: numeric rating scales (NRS) worse 0.51, NRS average 0.53, VRS worse 0.46. NRS average 0.48. Frequency of itch 0.53. All p<0.001 | DR=VG SM=D |
| Barbieri 2021(177) | United States | 764 | Atopic dermatitis | Used post-stratification sample weights to account for survey design | Known-groups validity of the DLQI and DLQI-R were assessed by comparing DLQI and DLQI-R scores across the severity categories for POEM (mild = 0–7, moderate = 8–19, and severe = 20–28) and PO-SCORAD (mild = 1–27, moderate = 28–56, severe = 57–104). Consistent with prior studies of the DLQI, more severe disease as assessed by POEM and PO-SCORAD was associated with higher DLQI scores indicating larger impact on QoL | DR=VG SM=A |
| Park 2021(234) | Korea South | 118 | Psoriasis | t-test, ANOVA; Scheffé post-hoc analysis. | Patients DLQI scores were significantly different (all p<0.001) for PASI scores (mild <5, moderate 5-15, severe >=15, F = 13.09, p < 0.001), Psoriasis Life Stress Inventory (PLSI) scores (low <10, high >10, t = 6.17, p < 0.001), Center for Epidemiologic Studies-Depression scale (CES-D) scores (normal <16, mild <16 or moderate depression >25, F = 32.00, p < 0.001) | DR=VG SM=A |

| | | | | | | |
|---|---|---|---|---|---|---|
| Schwartzman 2021(46) | United States | 994 | Atopic dermatitis | Area under receiver operating characteristic curve (ROC) | None of the PROMIS Global Health (PGH) or DLQI scores were able to differentiate between the lowest 3 levels of atopic dermatitis severity. However, DLQI was slightly better at distinguishing between different levels of self-reported disease. Mapped EuroQol-5D health utility score (mEQ-5D), PGH-P4, PGH-M4, and PGH-M2 T scores and DLQI scores showed similar multilevel area under the curve, indicating poor known-groups validity in predicting self-reported global atopic dermatitis severity overall | DR=VG SM=VG |
| Talamonti 2021(191) | Italy | 174 | Atopic dermatitis | $\chi 2$ Fisher exact test | EASI < 16 DLQI = 12.3 ± 5.5 (3-24); EASI ≥ 16 DLQI=14.4 ± 7.4 (0-30); p=0.0879. No statistically significant differences in mean DLQI score among patients stratified by age (<40 or ≥40 years) and disease duration (<15 and ≥15 years) | DR=VG SM=VG |
| da Silva 2022(224) | Germany | 107 | Psoriasis | ANOVA - post hoc Bonferroni correction | Subgroup analysis by none/mild (NRS numerical rating scale ≤ 3) or medium/severe (NRS ≥ 4) pruritus F=21.46 p<0.001. Subgroup analysis by anogenital involvement F=0.36 $\eta p2 = 0.00$. Subgroup analysis by pruritus anogenita interaction p<0.01, F=0.71 $\eta p2=0.01$ | DR=VG SM=VG |
| Kage 2022(227) | Germany | 83 | Atopic eczema | Mann-Whitney U-test | With suicidal thoughts, plans, or suicide DLQI 17 (IQR 9.5-22.5) vs without 8.0 (IQR 3-14) p=0.004 and their QoL (DLQI) was significantly decreased (p = 0.004) | DR=A SM=VG |
| Koszoru 2022(196) | Hungary | 218 | Atopic dermatitis | Kruskal-Wallis U-test | DLQI score (SD) for EASI groups: clear or mild (EASI 0.0-5.9) 8.04 +/-6.80, moderate (EASI 6-22.9) 13.24 +/-7.86, severe (EASI 23-72) 19.46 +/-7.45; effect size=0.197; p<0.001 | DR=VG SM=VG |
| Long 2022(229) | China | 90 | Psoriasis | Mann-Whitney U-test | Hamilton Depression Scale (HAMD, 24 HAMD score ≥18 was defined as accompanied by depressive symptoms Psoriasis with depression DLQI score 11(10-17), Psoriasis without depression DLQI score 4(2-7), F=-6.152, p <0.001 | DR=VG SM=VG |
| Nahidi 2022(197) | Iran | 80 | Psoriasis | Mann-Whitney U-test | DLQI scores: patients 11.75 ± 2.75, healthy controls 3.27 ± 3.57, p=0.03. DLQI scores severity mild (PASI < 10) 10.75 ± 6.8 or moderate to severe (PASI >=10) 16.4 ± 7.6, p=0.03 | DR=VG SM=VG |
| Sahin 2022(237) | Germany | 258 | Psoriasis | Mann-Whitney U-test | Median DLQI (IQR): Good sleepers (global Pittsburgh Sleep Quality Index (PSQI) score ≤ 5) N=106 2.0 (1.0-7.0), poor sleepers (global PSQI score > 5) N=152 7.0 (3.0-12.8), p<0.001 | DR=VG SM=VG |
| Solak 2022(239) | Türkiye | 140 | Psoriasis | Mann-Whitney U-test | Median (IRQ) DLQI score: psoriasis group 6 (3-10), control group 1 (0-3.3), p<0.001. With restless legs syndrome (RLS) 16 (6.5-20), without RLS 5 (2-7.8, p<0.001 | DR=VG SM=VG |

| | | | | | | |
|---|---|---|---|---|---|---|
| Szabo 2022(32) | Hungary | 618 | Warts, Eczema, Onychomycosis, Acne, Psoriasis, Tinea pedis, Basal cell carcinoma, Rosacea, Urticaria, Herpes zoster | ANOVA | DLQI was able to better discriminate between known groups of patients based on overall HRQoL impairment (global question GQ rating 'How much does your dermatological condition affect your life?' on a 5-point scale): DLQI N, Mean (SD); No effect N=212 0.9 (1,7), Small effect N=163 2.8 (2.9), Moderate effect N=175 5.3 (4.5), Very large effect N=52 9.4 (6.33), Extremely large effect N=16 17.0 (9.4), ANOVA F=118.7, p<0.001 | DR=A SM=VG |
| Tan 2022(30) | Multiple | 723 | Acne | Chi-squared, Fisher's exact test | Chi-squared independence test with Yates correction and by Fisher's exact test: Self-assessment of Clinical Acne-Related Scars (SCARS): 3-6 mild scarring DLQI mean 5.00 (SEM 0.30); 7-10 moderate scarring 6.34 (0.22); 11-20, severe/very severe scarring 8.12 (0.39), p= 0.014 | DR=A SM=VG |
| Yang 2022(241) | China | 469 | Rosacea | Kruskal–Wallis test | Investigator's Global Assessment (IGA): Almost Clear (n=68), Mild (n=159), Moderate (n=148), Severe (n=77), Very Severe (n=17), p=0.022. Correlation DLQI with IGA r=0.104 p=0.024. Clinician's Erythema Assessment (CEA): Clear (n=23), Almost clear (n=56), Mild (n=220), Moderate (n=140), Severe (n=30), p=0.271. Correlation DLQI with CEA r=0.052 p=0.257 | DR=VG SM=VG |
| Ye 2022(203) | Korea South | 500 | Urticaria | ANOVA | DLQI Mean score: Urticaria Activity Score over 7 days (UAS7) ≤ 6 well controlled 3.1, 7–15 mild 8.5, 16–27 moderate 12.7, 28–42 severe 17.6, p<0.001. Pearson correlation DLQI with UAS7 r=0.677, p<0.001. | DR=VG SM=VG |
| Yi 2022(242) | Malaysia | 262 | Psoriasis | Mann-Whitney U-test | Current/history of genital involvement N=115 mean 8.8 (range +/- 6.9); Ne genital Involvement N=138 mean 6.5 (range ± 6.6); p=0.006. They also reported impaired sexual functioning based on question 9 in DLQI (Question 9 mean score of 0.8 (range +/- 0.8) vs 0.5 (range +/- 0.9), P = 0.046). | DR=VG SM=VG |
| Koszoru 2023(205) | Hungary | 218 | Atopic dermatitis | Kruskal–Wallis test | EQ-5D-3L and EQ-5D-5L with DLQI grouped by Hongbo banding (206) Kruskal-Wallis test and Effect size (ESs ≥ 0.01 as small, ≥ 0.06 as moderate and ≥ 0.14 as large). EQ-5D-3L p< 0.001 Es= 0.489; EQ-5D-5L p< 0.001 Es= 0.108 | DR=VG SM=VG |

Note: Data was extracted from referenced publications. df = degrees of freedom
COSMIN: Box 9b. Comparison between subgroups (discriminative or known-groups validity).
*Design requirements* DR - 5 Was an adequate description provided of important characteristics of the subgroups? VG= very good, A = adequate, D= doubtful
*Statistical methods* - SM 6 Were design and statistical methods adequate for the hypotheses to be tested? VG= very good, A = adequate, D= doubtful, IN=inadequate
Lidwine et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) user manual. V1.0 February 2018. www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf

**Table SI(D). Studies assessing the differential item functioning (DIF) of the DLQI**

| References | Country | DLQI completed | Disease | Results | COSMIN |
|---|---|---|---|---|---|
| Nijsten 2007(243) | Belgium | 450 | Psoriasis | No DIF was seen across gender and age. All items had uniform DIF and four items also had nonuniform DIF across cultures. Rasch model only, no PCM | + |
| Twiss 2012(59) | United Kingdom | 292 | Psoriasis and Atopic dermatitis | Uniform DIF by disease was found for items 1,3,5,7 for age item 2 and gender items 2,4,6,7. Non-uniform DIF was found for disease item3 and gender item6 (all $p<0.05$). | - |
| Ofenloch 2014(245) | Germany | 1038 | Eczema/Hand eczema | Seven items showed uniform differential item functioning (DIF) according to gender or age group, 2 items showed non-uniform DIF (items 7 and 10) between centres or gender and 2 items had disordered thresholds. | - |
| Yale 2016(246) | China | 149 | Neurodermatitis (lichen simplex chronicus, LSC) | There was no uniform differential item functioning. DLQI items functioned similarly in relation to participant sex and location. However, items 2 and 4 functioned differently by age group, and items 4 and 8 functioned differently by severity of illness. Item 4 functioned differently in relation to educational level. When the significance level was adjusted for the number of comparisons ($p < 0.01$) none of the DIFs were significant in the 10-item version. | + |
| He 2018(247) | China | 9845 | Any skin disease | In total, DIF was observed in 4 of 10 items, and was associated with the hospital's geographical location for item 7 and with the disease for items 1, 2 and 5. The visual inspection suggested that all DIF except for that of item 5 related to disease were non-uniform. On the other hand, the ordinal logistic regression classified all DIF as uniform. No DIF was observed for sex, age, or diagnosed disease severity | + |
| Xiao 2018(74) | China | 465 | Arsenic-related skin lesions and symptoms | This indicated that DLQI did not have measurement invariance across these subgroups. | - |
| Patel 2019(33) | United States | 340 | Atopic dermatitis | Uniform and nonuniform differential item functioning by age, sex and/or race/ethnicity was found for multiple items in DLQI. DLQI scores were significantly lower in older patients (Wilcoxon rank sum test, $P<0.0001$). | + |
| Jorge 2020(80) | Brazil | 1286 | 14 dermatoses. (see Suppl. Data for the full list) | Analysis of variance according to sex, age group (< 30, 30-60 and > 60 years old) and type of disease (symptomatic or psychosocial) disclosed non-uniform behavior for several items according to sex, age and disease type after multivariate adjustment. | + |

| | | | | | |
|---|---|---|---|---|---|
| Rencz 2021(250) | Hungary | 425 | Psoriasis | With zero-scoring NRRs, six and three items showed DIF for gender and age, respectively. A uniform DIF was found in the majority of instances. With missing-scoring NRRs, four and three items indicated DIF for gender and age, respectively. | |
| Schwartzman 2021(46) | United States | 994 | Atopic dermatitis | No items from PGH or DLQI were found to suffer from differential item functioning by sex, age, race, or level of education, indicating good cross-cultural validity. | - |
| Tan 2022(30) | Multiple | 723 | Acne | No significant differences were observed in DLQI scores across countries (P = 0.308) | |

Note: Data was extracted from referenced publications.

COSMIN: Cross-cultural validity\measurement invariance bias: "+" No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R2 < 0.02); "**?**" No multiple group factor analysis OR DIF analysis performed; "**-**" Important differences between group factors OR DIF was found.

**Table SI(€). Translations and cross-cultural adaptations**

| References | Country | Forward translation | Back translation | Cognitive debriefing | Face validity by dermatology experts | Field testing | Focus group | Pilot test for content validity | Reliability assessment | Validity assessment | Validation | Other | Results | Original language | Final language |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Jobanputra 2000(35) | South Africa | Y | Y | Y | | | | Y | Y | | | Piloted on a mixed sample of patients (n= 140) during the first week of the study , test-retest with N=65. Main study n=607 plus 53 controls. | The adapted and translated DLQI was valid and reliable. In this multicultural setting, social class and language group, but not gender, influenced the impact of skin disease on overall QoL. Xhosa speakers were apparently less affected than other patients. | English | Afrikaans and Xhosa |
| Zachariae 2000(47) | Denmark | Y | Y | | | Y | | | | | | Hospitalized patients (n=200), outpatients (n=100) | Internal consistency and test-retest reliability were comparable the original English version. The Danish translation of the DLQI showed satisfactory reliability and preliminary results indicate that this version is a valid measure, which can be used in both research and clinical settings. | English | Danish |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Loo 2003(230) | United Kingdom | | | | Y | | N=191. Time to complete different versions. Median time to complete text-only version was 124 s (mean ±SD 126 ± 65, n=27). vs. illustrated version 88 s (mean ±SD 101 ± 52, n-25) (P= 0.08, Mann-Whitney U-test) | Showed satisfactory reliability and the preliminary results indicate that this version is a valid measure, which can be used in both research and clinical settings. | English | Illustrated |
| Aghaei 2004(49) | Iran | Y | Y | | | | N=70. The reliability and internal consistency of the questionnaire were assessed by Cronbach's alpha coefficient and Spearman's correlation, respectively. Validity was performed using convergent validity. | The Persian version of the DLQI questionnaire has a good structural characteristic and is a reliable and valid instrument that can be used for measuring the effects of vitiligo on quality of life. | | Farsi |
| Ferraz 2006(36) | Brazil | Y | Y | | Y | | Y | Reliability (N=44), validity (N=71). | As none of the items were determined to lack cultural equivalence, so none were replaced. Very few questions have very slightly modifications. Results suggest the Brazilian–Portuguese version is a reliable and valid outcome measure to be used in LE clinical studies. | English | Portuguese Brazil |
| Ozturkcan 2006(52) | Türkiye | Y | Y | Y | | Y | | N=69. The main problem was the translation of the word “partner”, which was overcome by all individual | It was found that the Turkish version of the DLQI was an acceptable index for dermatologists | English | Turkish |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | patients and the dermatology specialists' consensus. | and dermatology patients and, moreover, to be valid and reliable in a cross-sectional level. | | |
| Takahashi 2006(37) | Japan | Y | Y | Y | | | Y | | N=197. Participants reported no difficulties in answering the DLQI-J items | No problems were found with regard to content validity. Responses were found to be reproducible and stable. The DLQI-J provides valid and reliable data despite having only a small number of items. | English | Japanese |
| Nijsten 2007(243) | Belgium | | | Y | | | Y | Y | N=450. Rasch model of 450 psoriasis patients in Belgium, Germany, Ireland, Italy The Netherlands, UK, USA. Unidimensionality, DIF. No IRT partial credit (PCM) or graded response (GRM) models | In addition to suboptimal psychometric properties, the majority of the items of the DLQI and Skindex dysfunctioned between psoriasis patients from different cultural backgrounds. All DLQI items showed DIF across the 7 countries. | N/A | N/A |
| Henok 2008(55) | Ethiopia | Y | Y | | | | | | N=74. Expert review of back translation. The DLQI was quick and simple to use, taking on average 4 min to administer. | The Amharic DLQI appears feasible, reliable and valid among patients with podoconiosis in southern Ethiopia. | English | Amharic Ethiopia |
| Madarasingha 2011(40) | Sri Lanka | Y | Y | | | | | | N=200. Back-translated version was sent to original authors and modifications were done with discussion of the two parties. | Sinhala version took 3-5 minutes to complete on average and had no confusing, embarrassing or difficult to answer questions. The DLQI | English | Sinhala Sri Lanka |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | (Sinhala) version is a simple, acceptable and reliable tool to measure the effect of dermatological diseases on quality of life. | | |
| Khoudri 2013(41) | Morocco | Y | Y | N = 176 | The Arabic version for Morocco of the DLQI is reliable and valid. | English | Arabic Morocco |
| Jesmin 2021(44) | Bangladesh | Y | Y | N-80. Face validity, content validity, convergent validity, and criterion validity were found acceptable. | Cronbach's alpha = 0.86 and significant test-retest reliability (ICC= 0.97). Adapted Bangla version of DLQI appears to be an acceptable, reliable, and valid instrument for measuring QOL in Bangla speaking patients with psoriasis. | English | Bangla Bangladesh |

Note: Data was extracted from referenced publications. Some studies started from the point where forward and backward translation had already been carried out.

**Table SI(F). DLQI scores of study datasets and healthy controls**

| | DLQI study datasets | DLQI healthy controls |
|---|---|---|
| N | 153 | 6 |
| Average score | 8.9 | 0.9 |
| Minimum score | 1.0 | 0.3 |
| Maximum score | 18.8 | 2.1 |
| Standard deviation | 3.4 | 0.8 |