Table SV. Summary of findings table

	Summary or pooled results	Overall rating	Quality of evidence
ACSD			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	Principal component analysis (PCA) found 6 factors: social anxiety/avoidance, itch-scratch cycle, helplessness, anxious-depressive mood, impact of QoL, and inadequate active coping. (?)	Indeterminate	High
Internal consistency			
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
CDLQI			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity			
Internal consistency			
Reliability	Spearman correlation=0.86 ($p<0.0001$) (?)	Indeterminate	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	Scores for children with skin disease were higher than score for children other diseases and health children ($p<0.0001$); Scores for eczema, psoriasis and acne were higher than for moles and naevi ($p<0.002$) (+)	Sufficient	Low (due to risk of bias)
Measurement invariance			
Responsiveness			
CDLQI-C			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity			
Internal consistency	Cronbach's alpha=0.83 (+), but this insufficient without information on structural validity	Indeterminate	Low (due to risk of bias and imprecision)
Reliability	Correlation of test-retest results is strong (cs ¼ 0.958, $p<0.01$) but results reported do not correspond to COSMIN statistics	Indeterminate	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	CDLQI scores were positively correlated with physician-rated disease severity, such correlation being strong (cs ¼ 0.850, $p<0.01$)(+)	Sufficient	Very low (due to risk of bias and imprecision)
Measurement invariance			
Responsiveness			
CQI-CSD			
Content validity	Relevance (±), comprehensiveness (+), comprehensibility (+)	Inconsistent	Very low (due to risk of bias)
Structural validity	7 scales: information about the care process, healthcare provided by physicians, healthcare provided by nurses, cooperation of healthcare providers, information provision by healthcare providers, patient participation and safety. The remaining 23 items did not fit into any of these scales statistically and/or by content.	Indeterminate	Very low (due to risk of bias)
Internal consistency	Cronbach's alpha=0.74–0.92 (+), but this is insufficient as it was calculated for the total scale not subscales	Insufficient	Very low (due to risk of bias)
Reliability			
Measurement error			
Construct validity	Tests performed were not relevant to a priori hypotheses	Indeterminate	Very low (due to risk of bias)
Measurement invariance			
Responsiveness			
DIS			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity			
Internal consistency			
Reliability			
Measurement error			
Construct validity	Group differences found, but no hypothesis reported.	Indeterminate	Low (due to risk of bias)
Measurement invariance			
Responsiveness			
DLQI			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	Unidimensional scale	Indeterminate	High (based on only studies of very good or adequate methodological quality)
Internal consistency	Cronbach's alpha=0.89–0.92	Sufficient (based on studies that reported COSMIN relevant statistics)	High
Reliability	Spearman's r=0.99 ($p<0.0001$)	Indeterminate	Low (due to risk of bias)
Measurement error			
Construct validity	Scores were significantly correlated with the Short Form (36) Health Survey (SF-36), EuroQOL 5D (EQ-5D), and patient- and clinician-rated severity. [6–8] [7–9] [7–9] [6–8] [5–7] [5–7] [4–6] Results 84% consistent with a priori hypotheses.	Sufficient	High
Measurement invariance			
Responsiveness	Treatment success was a significant predictor of improvement in DLQI score and significant differences between treatment responders and partial responders were found. Results 79% consistent with a priori hypotheses	Sufficient	High

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
DLQI-B			
Content validity			
Structural validity			
Internal consistency			
Reliability	Pearson's correlation coefficient=0.96 ($p < 0.001$)	Sufficient	Very low (due to risk of bias and imprecision
Measurement error			
Construct validity	Results 100% consistent with a priori hypotheses	Sufficient	Moderate (due to imprecision)
Measurement invariance			
Responsiveness			
DLQI-C			
Content validity			
Structural validity	Unidimensional	Sufficient	High (based on only studies of very good or adequate methodological quality)
Internal consistency	Cronbach's alpha=0.82–0.91	Sufficient	High
Reliability			
Measurement error			
Construct validity	Known-groups comparison showed that the DLQI discriminated well between patients who differed in their age, geographical region, duration of psoriasis and PASI score, but not discriminated between subgroups based on gender and presence of any other chronic condition. Results showed excel-lent 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 the DLQI. The correlation coefficients between DLQI and SF-36 were in negative values, because a higher DLQI score indicated greater impairment in QoL, whereas a higher score on SF-36 indicated better health or performance.	Sufficient	High
Measurement invariance	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.	Indeterminate	Very low (due to risk of bias)
Responsiveness			
DLQI-D			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha=0.88	Sufficient	Very low (due to risk of bias)
Reliability	Test-retest coefficient=0.93 ($p < 0.01$)	Insufficient	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	Significant rank order correlations were found between DLQI scores and physician's ratings of severity of skin symptoms for both outpatients and hospitalized patients and between DLQI scores and the time willing to be spent on an effective treatment for outpatients. Comparisons among diagnoses, controlling for multiple comparisons, showed that DLQI scores for atopic dermatitis patients were significantly higher ($p < 0.01$) than scores for patients with acne, viral warts, collagenosis, erythroderma, urticaria and other eczemas. Psoriasis patients had significantly higher DLQI scores ($p < 0.05$) than patients with collagenosis and the group of patients with other skin diseases. DLQI predicted time willing to be spent on treatment (R~0.13; $p < 0.01$)	Indeterminate	High
Measurement invariance			
Responsiveness			
DLQI-I			
Content validity			
Structural validity	EFA: 4 factors (CFI=0.998, TLI=0.992) CFA: unidimensional scale (CFI=0.994, TLI=0.986)	Sufficient	High
Internal consistency	Cronbach's alpha=0.83	Sufficient	High
Reliability			
Measurement error			
Construct validity	The correlations between the DLQI and the PDI, Skindex-29 functioning, emotions and symptoms scales were 0.81 (+), 0.72 (+), 0.64 (+) and 0.56 (+), respectively.	Sufficient	High
Measurement invariance			
Responsiveness			
DQLI-N			
Content validity			
Structural validity	Unidimensional scale	Indeterminate	Moderate (due to risk of bias)
Internal consistency	Cronbach's alpha=0.9	Sufficient	Moderate(duetoindeterminate structural validity)
Reliability			
Measurement error			
Construct validity	Higher levels of disease severity (PASI) are significantly related to poor QoL (total and individual DLQI item score)	Sufficient	Moderate (due to risk of bias)
Measurement invariance			
Responsiveness			

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
DLQI-S			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha: symptoms & feelings=0.561; daily activities=0.741; leisure=0.687; personal relationships=0.442	Insufficient	High
Reliability	Kappa=0.83	Sufficient	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	Results 83% consistent with a priori hypotheses	Sufficient	Low (due to risk of bias)
Measurement invariance			
Responsiveness			
DLQI-T			
Content validity			
Structural validity	2 factor model	Indeterminate	Low (due to risk of bias and imprecision)
Internal consistency	Cronbach's alpha=0.9 (2 factor model, but only one total score)	Sufficient	Very low (due to risk of bias and imprecision)
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
DLQI-U			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha=0.81–0.86	Sufficient	Very low (due to risk of bias and imprecision)
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
DQOLs			
Content validity	Relevance (?), comprehensiveness (?), comprehensibility (?)	Indeterminate	Very low (due to risk of bias)
Structural validity	4 factors	Indeterminate	Moderate (due to risk of bias)
Internal consistency	Cronbach's alpha: psychosocial items=0.92; activity items=0.83	Sufficient	Very low (due to risk of bias)
Reliability	ICC=0.84 for psychosocial and activity scores	Sufficient	Low (due to risk of bias)
Measurement error			
Construct validity	Results 50% consistent with a priori hypotheses. Total mean psychosocial scores were relatively high for eczema, acne and disorders of keratinization (mainly ichthyosis or palmar-plantar keratoderma) and lowest for benign and malignant tumours, though the differences were not statistically significant (–). Total activities score independent of age (age-adjusted means calculated using regression analysis) was significantly higher for subjects with disorders of keratinization and eczema (i.e. more restrictions on activities due to skin problems) than subjects with benign and malignant tumours ($p=0.006$ and 0.002, respectively) and pigmentary disorders ($p=0.01$ and 0.002, respectively) (+)	Insufficient	Low (due to risk of bias)
Measurement invariance			
Responsiveness			
DSQL			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	5 factors	Indeterminate	High
Internal consistency	Cronbach's alpha ≥0.8	Sufficient	High
Reliability	ICC ≥0.82 for all subscales	Sufficient	Low (due to risk of bias)
Measurement error			
Construct validity	100% consistent with a priori hypotheses (DSQL significantly correlated with patient-rated severity (–0.59). The means of the 5 DSQL scale score values, and the 2 SF-36 scale values were significantly larger ($p<0.05$) among patients classified with severe CD and with acne scarring than those in the less severe disease group (+).	Sufficient	High
Measurement invariance			
Responsiveness	*t*-test analysis (not shown) of the within-group mean DSQL change scores which correspond to the effect size statistics was conducted testing the hypothesis that the population mean DSQL change score is 0, vs the 2-sided alternative that the mean is different from 0. For each DSQL dimension score and the total score, the *p*-values were <0.01 for each treatment group, with the exception of the change in work/school limitation among placebo, which was $p=0.628$ (+).	Sufficient	High
FLQA-d			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity			
Internal consistency	Cronbach's alpha: physical complaints=0.8 (+); everyday life=0.89 (+); social life=0.86 (+); emotions=0.86 (+); treatment=0.69 (–); satisfaction=0.83 (+)	Sufficient	High
Reliability	'Test retest': physical complains=0.89; everyday life=0.91; social life=0.9; emotions=0.87; treatment=0.88; satisfaction=0.86	Indeterminate	Low (due to risk of bias)

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
Measurement error			
Construct validity	Discriminant validity of the FLQA-d can be assumed since there were significant differences between patients with psoriasis and atopic dermatitis in 5 of 6 scales. A lack of discrimination in the scale "therapy" is plausible since the treatment in the groups was almost identical. Convergent validity: In comparable scales, elevated correlations between the FLQA and the QCSD, ALLTAG and DLQI were found. For example, the FLQA scale "emotional status" showed elevated correlations with the QCSD scale "anxious-depressive mood" (r=0.73), with the DLQI sum score (r=0.68) and with the ALLTAG scale "psyche" (r=0.80). In all scales, patients with atopic dermatitis, psoriasis and urticaria showed significantly reduced QoL compared with healthy people.	Sufficient	High
Measurement invariance			
Responsiveness	Reactivity to change: There was good sensitivity to change in all scales ($p < 0.001$)	Sufficient	High
PBI			
Content validity	Relevance (±), comprehensiveness (+), comprehensibility (±)	Inconsistent	Low (due to risk of bias)
Structural validity			
Internal consistency	Cross sectional study=Cronbach's alpha<0.91; longitudinal study=Cronbach's alpha=0.97	Indeterminate (as there is no structural validity study)	Low (due to risk of bias)
Reliability	Test retest reliability (r=0.68); cross sectional study=item-total correlations ranged from 0.32 to 0.68); longitudinal study=item-total correlations between 0.50 and 0.86	Indeterminate	Low (due to risk of bias)
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness	Convergent validity (r>0.5)	Sufficient	Moderate (due to risk of bias)
PeDeSI			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	Unidimensional scale	Indeterminate	Moderate (due to risk of bias)
Internal consistency	Cronbach's alpha=0.9	Sufficient	Moderate (due to risk of bias)
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
PRISM			
Content validity			
Structural validity			
Internal consistency			
Reliability			
Measurement error			
Construct validity	PRISM was significantly correlated with Skindex-29 and DLQI	Sufficient	High
Measurement invariance			
Responsiveness	PRISM showed a reduction in patients' burden of illness during hospitalization ($p < 0.001$). Cohen effect size (d=0.67)	Sufficient	Low (due to imprecision)
SF-QES			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	EFA: 5 factors (RMSEA ≥0.06; CFI<0.95)	Insufficient	High
Internal consistency	Cronbach's alpha: self-esteem - 0.82; retreat=0.76; experienced refusal=0.84; concealment=0.76; composure=0.69	Sufficient	Moderate (due to risk of bias)
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
Skindex			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	7 factors	Indeterminate	High
Internal consistency	Internal consistency was tested on subscales hypothesised by authors rather than on those identified by the factor analysis	Indeterminate	Very low (due to risk of bias)
Reliability	Pearson's test-retest conducted on subscales hypothesised by authors rather than on those identified by the factor analysis	Indeterminate	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	The mean scores of patients with inflammatory dermatoses such as eczema, psoriasis, or acne were significantly higher than mean scores of those with isolated skin lesions such as moles or skin cancers ($p = 0.05$) for all scales except physical limitations ($p = 0.09$). In patients with isolated benign noninflammatory skin lesions, inflammatory skin diseases, and in the subgroup with eczematous dermatitis, there was no correlation between physicians' judgments of the clinical severity of skin disease and Skindex scale scores (p~0.3). In the 21 patients with acne vulgaris, however, judgements of clinical severity correlated with 2 of the 8 scales: the physical limitations scale and the embarrassment scale (values of r=0.51 and 0.47, respectively, $p < 0.04$). Only 66% of a priori hypotheses met.	Insufficient	Low (due to risk of bias)

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
Measurement invariance			
Responsiveness	34 patients who reported that overall their skin condition was better than 6 months previously, the mean scores of 7 Skindex scales decreased, and 3 of them decreased significantly ($p=0.05$). In 27 patients who reported that overall their skin condition had remained the same in the previous 6 months, only 1 scale score (Fear) changed significantly ($p>0.05$). In 6 patients who responded that overall their skin condition was worse, all 8 scale scores increased, and 4 increased significantly ($p=0.05$). **Only some scales were significantly responsive**. Less than 75% of a priori hypotheses met.	Insufficient	Very low (due to risk of bias)
Skindex-29			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	3 factors: emotions, functioning, symptoms	Indeterminate	High
Internal consistency	Cronbach's alpha: symptoms=0.87 (+); emotions=0.94(+); functioning=0.96 (+)	Sufficient	High
Reliability	Pearson's correlation coefficient: symptoms=0.91; emotions=0.88; functioning=0.92 (?)	Indeterminate	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	The categorization of both the emotions and functioning subscales of the Skindex-29 correlated very well with the psychosocial component of the Skindex-17 (r=0.89 and 0.93, respectively) and moderately with the symptom scale of the Skindex-17 (r=0.56 and 0.52, respectively). The symptom subscales of both instruments correlated extremely well (r=0.93). Participants with inflammatory dermatoses were significantly more likely to have higher levels of impact on the scales of the Skindex-29, except that patients with seborrhoeic dermatitis reported significantly less psychosocial impairment. Overall, patients with psoriasis reported the highest levels of impact, especially on the symptom-related items. As expected, alopecia and vitiligo had a substantial effect on the psychosocial, but significantly less on the symptom domain. Although about 10% of patients with naevi reported an effect on their lives, it was significantly lower compared with the other diseases for all the Skindex-29 scales. For all subscales, mean scores of patients with psoriasis or eczema were significantly higher than mean scores of those with benign skin lesions or non-melanoma skin cancer ($p<0.001$). Skindex-29 significantly, but modestly correlated with physician rated clinical severity ($p\leq0.001$).	Sufficient	High
Measurement invariance			
Responsiveness	t-tests found mean scores were significantly improved for patients who had improved, significantly worse for patients who had worsened and stable for stable patients	Sufficient	Very low (due to risk of bias)
S29-C			
Content validity	Relevance (–), comprehensiveness (?), comprehensibility (–)	Insufficient	Very low (due to risk of bias)
Structural validity	3 factors: SRMS=0.06, RMSEA=0.11; CFI=0.96	Insufficient	Moderate
Internal consistency	Cronbach's alpha=0.97, but result should be ignored because there should be a test for each of the 3 scales	Indeterminate	Very low (due to risk of bias)
Reliability			
Measurement error			
Construct validity	Patients with inflammatory dermatosis had significantly higher scale and total scores of Skindex-29 and Skindex-16 compared with patients with isolated lesions (Table SIV) (+). It was demonstrated that moderate to good correlations between Skindex-29 and DLQI, Skindex-16 and DLQI (Skindex-29, r=0.43–0.84; Skindex-16, r=0.39–0.83)	Sufficient	High
Measurement invariance			
Responsiveness			
S29-G			
Content validity			
Structural validity	3 factors	Indeterminate	High
Internal consistency	Cronbach's alpha: functioning=0.93–0.94; emotions=0.92; symptoms=0.85–0.87	Sufficient	High
Reliability			
Measurement error			
Construct validity	Correlations between the 3 Skindex-29 scales and corresponding scales of the other skin disease-specific questionnaires were high and significant	Sufficient	High
Measurement invariance			
Responsiveness	At the time of discharge, patients reported lower HRQOL on the Skindex-29 compared with the time of their admission to the hospital. The differences were significant on all 3 scales in both samples	Sufficient	Very low (due to risk of bias)
S29-S			
Content validity	Relevance (–), comprehensiveness (?), comprehensibility (?)	Indeterminate	Very low (due to risk of bias)
Structural validity			
Internal consistency	Cronbach's alpha: emotions=0.885–0.898; symptoms=0.752–0.922; functioning=0.901–0.922	Sufficient	High
Reliability	ICC: emotions=0.468; symptoms=0.348; functioning=0.498	Insufficient	Low (due to risk of bias)
Measurement error			
Construct validity	The scores on all 3 domains of Skindex-29 from both ratings correlated significantly and positively with the DLQI scores in both ratings. The patients with more severe skin diseases, such as acne vulgaris and psoriasis, scored significantly higher in the emotional and functioning domains. However, this was not the case in the symptoms domain. Less than 75% of *a priori* hypotheses met.	Insufficient	High
Measurement invariance			
Responsiveness			

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
S29-Sp			
Content validity	Relevance (−), comprehensiveness (−), comprehensibility (−)	Insufficient	High
Structural validity			
Internal consistency	Cronbach's alpha: functioning=0.78–0.89; emotional=0.87–0.91; symptoms=0.7–0.84	Sufficient	High
Reliability	ICC: global=0.94; functioning =0.93; emotions=0.92; symptoms=0.85	Sufficient	Very low (due to risk of bias and precision)
Measurement error			
Construct validity	Patients with skin problems (groups A + I) had significant scores higher than healthy people (group S) in the 3 scales of the Spanish version of Skindex-29. Similarly, the patients with inflammatory dermatoses (group I) had global scores and by significantly higher scales that patients with skin lesions isolated (group A) ($p<0.005$). Patients with skin problems had significantly higher scores than healthy people for all 3 scales of the Spanish version of Skindex-29 ($p<0.01$) (+). Similarly, patients with inflammatory dermatoses had significantly higher scale scores than patients with isolated lesions ($p<0.01$) (+)	Sufficient	High
Measurement invariance			
Responsiveness	There were differences significant between the scores of the 2 tests of the 40 patients who reported having improved their disease and the size of the overall effect was 0.76. There were not any in the case of the 10 patients that were clinically the same in the second visit and the size of the overall effect in this group was 0.11.	Sufficient	Very low (due to risk of bias and imprecision)
Skindex-16			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	3 factors	Sufficient	Moderate (due to risk of bias)
Internal consistency	Cronbach's alpha: symptoms=0.86, emotions=0.93, functioning=0.88	Sufficient	Moderate (due to risk of bias)
Reliability			
Measurement error			
Construct validity	For all 3 scales, mean scores of patients with inflammatory dermatoses were significantly higher than mean scores of those with isolated skin lesions ($p<0.0001$)	Sufficient	Low (due to risk of bias)
Measurement invariance			
Responsiveness	The mean scores of all 3 scales of the brief version of Skindex remained stable or changed appropriately in patients who responded that their skin was the same or had improved compared with the first time they responded to the instrument (a year previously). The number of patients who responded that the condition of their skin had worsened in the last year was too few to demonstrate statistical significance.	Indeterminate	Very low (due to risk of bias and imprecision)
S16-A			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha: whole=0.93, emotions=0.92, functioning=0.87, symptoms=0.81	Sufficient	High
Reliability			
Measurement error			
Construct validity	Patients with skin problems had significantly higher scores than healthy people for the global score as well as each of the subscales of the Arabic version of Skindex-16 ($p<0.02$). Similarly, patients with inflammatory dermatoses had significantly higher global, emotional and functional scale scores than patients with isolated lesions ($p<0.001$); however, for the symptoms subscale there was no significant difference.	Sufficient	Moderate (due to risk of bias)
Measurement invariance			
Responsiveness			
S16-BP			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha: symptoms=0.87; emotions=0.93; functioning=0.89	Sufficient	High
Reliability	ICC: symptoms=0.95; emotions=0.86; functioning=0.85	Sufficient	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	H1: All 3 Skindex-16 scales exhibited strong correlation with DLQI scores (rho=0.664, 0.766 and 0.712 for the domains symptoms, emotions and functioning, respectively). H2: Relative to the HADS-A and the HADS-D, the Skindex-16 domain with the highest correlation coefficient was emotions, with moderate correlation (0.4±0.6); H3: values for the domains symptoms and functioning were low (close to 0.4). DV: H1: The Kruskal-Wallis test detected differences among the 3 groups (emotions: $p<0.001$; symptoms: $p=0.002$; functioning: $p<0.001$). Analysis using the Mann-Whitney test with Bonferroni correction showed differences in all 3 Skindex-16 domains between the mild and moderate skin disease groups (emotions: $p<0.001$; symptoms: $p=0.049$; functioning: $p<0.001$) and between the mild and severe skin disease groups (emotions: $p=0.002$; symptoms: $p=0.001$; functioning: $p=0.002$). H2: Patients with inflammatory dermatosis presented higher scores on emotions ($p=0.016$) and functioning ($p=0.056$), but not in symptoms ($p=0.298$), when compared with patients with localised lesions.	Sufficient	High

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
Measurement invariance			
Responsiveness			
S16-C			
Content validity	Relevance (-), comprehensiveness (?), comprehensibility (-)	Insufficient	Very low (due to risk of bias)
Structural validity	3 factors: RMSEA=15, CFI=0.95, SRMR=0.07	Insufficient	High
Internal consistency	Cronbach's alpha=0.96 but should be calculated for each subscale rather than the whole scale	Indeterminate	Very low
Reliability			
Measurement error			
Construct validity	Patients with inflammatory dermatosis had significantly higher scale and total scores of Skindex-29 and Skindex-16 compared with patients with isolated lesions. It was demonstrated that moderate to good correlations between Skindex-29 and DLQI, Skindex-16 and DLQI (Skindex-29, r=0.43–0.84; Skindex-16, r=0.39–0.83).	Sufficient	High
Measurement invariance			
Responsiveness			
S16-J			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha=0.92; symptoms=0.87; emotions=0.89; functioning=0.83	Sufficient	High
Reliability			
Measurement error			
Construct validity	Patients with skin diseases showed significantly higher scores than those of healthy people for all 3 scales ($p<0.001$). Similarly, patients with inflammatory dermatoses showed significantly higher scores than patients with isolated lesions ($p<0.001$).	Sufficient	High
Measurement invariance			
Responsiveness			
S16-M			
Content validity	Relevance (–), comprehensiveness (?), comprehensibility (?)	Indeterminate	Very low (due to risk of bias)
Structural validity			
Internal consistency	Cronbach's α values for the entire instrument and for the individual scales were high (a=0.83, 0.81, 0.71 and 0.82 for the entire instrument and the 7-item emotional, the 5-item functional and the 4-item symptom scales, respectively).	Sufficient	High
Reliability	ICC interobserver ranged from 0.80 to 0.85; test-retest ranged from 0.93 to 0.96	Sufficient	Low (due to risk of bias and imprecision)
Measurement error			
Construct validity	The EQ-5D correlated negatively with all scales	Sufficient	High
Measurement invariance			
Responsiveness			
S16-U			
Content validity			
Structural validity			
Internal consistency	Cronbach's alpha: emotions=0.87–0.89; symptoms=0.83–0.84; functioning=0.89	Sufficient	Moderate (due to imprecision)
Reliability			
Measurement error			
Construct validity			
Measurement invariance			
Responsiveness			
TQL			
Content validity	Relevance (±), comprehensiveness (?), comprehensibility (–)	Inconsistent	Low (due to risk of bias)
Structural validity	6 factors	Indeterminate	Moderate (due to risk of bias)
Internal consistency	Cronbach's alpha: total 0.82; social life 0.77; emotional 0.78; daily activities 0.78; cognitive 0.79; symptom 0.80; sexual life 0.84	Sufficient	Moderate (due to risk of bias)
Reliability	ICC: total 0.88; social life 0.88; emotional 0.84; daily activities 0.86; cognitive 0.83; symptom 0.63; sexual life 0.71	Insufficient	Very low (due to risk of bias and imprecision)
Measurement error			
Construct validity	There was a significant correlation between the physician's judgments of the clinical severity of skin disease and the six scale scores of the QOL instrument (r=0.25–0.38; $p=0.001$). All the mean scores of the severe group were significantly higher than the mean scores of the not-severe group. The mean scores of patients were significantly higher than those of the healthy controls (Table II). The overall mean TQL score for the patients was 14.69 (±10.28) and for the controls was 4.1 (±7.16). The mean scores of the patients with vitiligo, acne vulgaris, urticaria, hirsutismus and psoriasis were higher ($p<0.001$) than the mean scores of those with isolated skin lesions such as benign skin tumours and verruca.	Sufficient	Moderate (due to risk of bias)
Measurement invariance			
Responsiveness			
VQ-Dermato			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Very low (due to risk of bias)
Structural validity	7 factors	Indeterminate	Moderate (due to risk of bias)

Table SV. (contd.)

	Summary or pooled results	Overall rating	Quality of evidence
Internal consistency	Cronbach's alpha: DI (self-perception)=0.88, D2 (daily living activity)=0.86, D3 (mood state)=0.88, D4 (social functioning)=0.84, D5 (leisure activity)=0.85, D6 (treatment-induced restriction)=0.77, D7 (physical discomfort)=0.63	Insufficient	Moderate (due to risk of bias)
Reliability	Pearson's r>0.8, $p<0.001$	Indeterminate	Low (due to risk of bias)
Measurement error			
Construct validity	Four of the 7 dimensions of the VQ-Dermato were strongly correlated with the values from the visual analogue scale for self-assessment of QoL. For D1, D2, D3, D4, D5, D6 and D7, r values were 0.55, 0.49, 0.53, 0.48, 0.41, 0.38 and 0.26, respectively. A strong correlation was found between some dimensions of the VQ-Dermato and some dimensions of the SF36: the physical symptoms dimension from the VQ-Dermato ("discomfort") with the equivalent dimension from the SF36 ("bodily pain"); the physical capacity dimension of the VQ-Dermato ("daily living activity") with the equivalent dimensions of the SF36 ("physical functioning", "role-physical" and "vitality") but also with the physical symptoms dimension ("bodily pain") of the SF36; the psychological dimension from the VQ-Dermato ("mood state") with the equivalent dimension of the SF36 ("mental health") and to a lower degree with "vitality" and "social functioning"; the social dimension from the VQ-Dermato ("social functioning") with the equivalent dimension ("social functioning") from the SF36; "self-perception" from the VQ-Dermato with "social functioning" and "mental health" from the SF36. This strong correlation between the SF36 and VQ-Dermato assessments of closely related aspects of health confirms that the proposed labels of the VQ-Dermato dimensions were in agreement with what they actually measured. The links were inconsistent and weak between VQ-Dermato dimensions and sociodemographic criteria (age, sex), extension of lesions (percent of skin surface involved) and physician's judgement about the severity of disease. Scores were significantly lower in females than in males for "selfperception", "mood state" and "leisure activity" ($p<0.001$), but there was no significant difference between sexes, for the other 5 dimensions. Age was found to be weakly correlated only with "daily living activity" (r=0.30, $p<0.001$) and "treatment-induced restriction" (r=0.24, $p<0.01$). The impact on "daily living activity" and on "treatment-induced restrictions" was significantly lower in people under 65 years than in people older than 65 years. The percentage of skin surface involved correlated weakly only with "self-perception" (r=0.14, $p<0.05$), "leisure activity" (r=0.22, $p<0.05$) and "discomfort" (r=0.19, $p<0.05$). All the dimensions but "mood state" and "discomfort" were related to the physician's judgement about the severity of disease using both the "same disease scale" (r ranges between 0.17 and 0.28) and "all skin diseases scale" (r ranges between 0.26 and 0.31). The impact on the 8 VQ-Dermato dimensions was significantly higher ($p<0.05$) when there was a function restriction. The only significantly different VQ-Dermato dimension between visible (face and hands) and invisible diseases (no involvement of face and hands) was "self-perception" ($p<0.05$). When a sex-dependent definition of visible diseases (face and hands in men, face, hands and legs in women) was chosen, scores were higher in visible diseases for 3 dimensions: "self-perception", "daily living activity" and "social functioning". No significant link was found between duration of the disease and any of the VQ-Dermato dimensions. The scores were significantly different ($p<0.01$) in each category of chronic skin disease for all VQ-Dermato dimensions except "social functioning". Most of our a priori hypotheses about the different categories of disease (see methods) were confirmed: (1) higher "self-perception" score in dermatitis of the face, psoriasis, chronic eczema and chronic urticaria, dyschromic disorders and alopecia; (2) high "daily living activity" scores in severe trophic disorders of the legs and hand and foot dermatitis and low scores in alopecia, dermatitis of the face and multiple carcinomas; (3) high "physical discomfort" scores in chronic eczema and chronic urticaria, hand and foot dermatitis and severe trophic disorders of the legs and low scores in alopecia, dermatitis of the face and multiple carcinomas.	Sufficient	High
Measurement invariance			
Responsiveness	At the final visit, the patients considered as clinically improved had significantly lower scores (better QoL) than the patients unchanged for 7 of the 8 VQ-Dermato dimensions.	Sufficient	Moderate (due to risk of bias)

ACSD: Adjustment to Chronic Skin Diseases Questionnaire; CDLQI: Children's Dermatology Life Quality Index; CDLQI-C: Children's Dermatology Life Quality Index Cantonese; CQI-CSD: Consumer Quality Index Chronic Skin Disease; DIS: Dermatology Intimacy Scale; DLQI: Dermatology Life Quality Index; DLQI-B: Dermatology Life Quality Index Brazilian-Portuguese; DLQI-C: Dermatology Life Quality Index Chinese; DLQI-D: Dermatology Life Quality Index Danish; DLQI-I: Dermatology Life Quality Index Italian; DLQI-N: Dermatology Life Quality Index Norwegian; DLQI-S: Dermatology Life Quality Index Sinhala; DLQI-T: Dermatology Life Quality Index Turkish; DLQI-U: Dermatology Life Quality Index Ukrainian; DQOLs: Dermatology Quality of Life scales; DSQL: Dermatology-specific Quality of Life; FLQA-d: Freiburg Life Quality Assessment; PBI: Patient Benefit Index; PeDeSI: Person-Centred Dermatology Self-Care Index; PRISM: Pictorial Representation of Illness & Self-Measure; SF-QES: Short-form of the Questionnaire on Experience with Skin Complaints; S29-C: Skindex-29-Chinese; S29-G: Skindex-29-German; S29-S: Skindex-29-Serbian; S29-Sp: Skindex-29-Spanish; S16-A: Skindex-16-Arabic; S16-BP: Skindex-16-Brazilian-Portuguese; S16-C: Skindex-16-Chinese; S16-J: Skindex-16-Japanese ; S16-M: Skindex-16-Moroccan-Arabic; S16-U: Skindex-16-Ukrainian; TQL: Turkish Quality of Life Instrument for skin disease; SF36: Short Form-36; QoL: quality of life.