


Validation of the Dutch Quality of Life in Hand Eczema Questionnaire (QOLHEQ)*

J.A.F. Oosterhaven ¹, R.F. Ofenloch ² and M.L.A. Schuttelaar¹

¹Department of Dermatology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

²Occupational Dermatology, Department of Dermatology, University Hospital Heidelberg, Heidelberg, Germany

Linked Comment: Shen and Chen. *Br J Dermatol* 2020; **183**:8–9.

Summary

Correspondence

M.L.A. Schuttelaar.

E-mail: m.l.a.schuttelaar@umcg.nl

Accepted for publication

16 September 2019

Funding sources

None.

Conflicts of interest

None declared.

*Plain language summary available online

DOI 10.1111/bjd.18558

Background Measurement instruments should be validated for use in the population for which they are intended. The Quality of Life in Hand Eczema Questionnaire (QOLHEQ) has been developed to measure impairment of health-related quality of life in patients with hand eczema.

Objectives To assess validity, reproducibility, responsiveness and interpretability of the Dutch version of the QOLHEQ.

Methods This was a prospective validation study in adult patients with hand eczema. At three time points (T₀, baseline; T₁, after 1–3 days; T₂, after 4–12 weeks), data from the QOLHEQ and multiple reference instruments were collected. Scale structure was assessed using item response theory analysis and structural equation modelling (SEM). Single-score validity and responsiveness were tested with hypotheses on correlations with reference instruments. Concerning reproducibility, intraclass correlation coefficients (ICC_{agreement}) and standard error of agreement (SEM_{agreement}) were checked. Regarding interpretability, bands for severity of quality-of-life impairment were proposed. Also, smallest detectable change (SDC) and minimally important change (MIC) were determined.

Results At T₀, 300 individuals participated in the study (54% were male, mean age 45 years). Rescoring of the scale structure fitted the Rasch model and the SEM. The ICC_{agreement} was 0.91 (95% confidence interval 0.85–0.94) and the SEM_{agreement} was 5.2 points. Of the a priori formulated hypotheses, 80% (single-score validity) and 64% (change scores for responsiveness) were confirmed. The SDC was 14.4 points and the MIC was 11.5 points.

Conclusions The Dutch version of the QOLHEQ has a good structural validity and reproducibility and has a high single-score validity and moderate responsiveness. An improvement of ≥ 15 points should be regarded as a real, important change within the Dutch population.

What's already known about this topic?

- The Quality of Life in Hand Eczema Questionnaire (QOLHEQ) measures impairment of health-related quality of life (HRQoL) in patients with hand eczema.
- The QOLHEQ was validated in Germany and Japan, but the validity and interpretability of the Dutch version are unknown.

What does this study add?

- This study shows that the Dutch QOLHEQ is a valid instrument to measure HRQoL impairment in Dutch patients with hand eczema, demonstrating good reliability and moderate responsiveness.
- Methods of item response theory are applied to assess and refine the scoring structure.

- Severity gradings to interpret single and change scores, specifically in Dutch patients, are proposed.

What are the clinical implications of this work?

- The Dutch QOLHEQ can now be used to measure HRQoL impairment in Dutch patients with hand eczema.

Hand eczema is a disease that is associated with an impaired quality of life.^{1,2} Until recently, this was measured using generic (nonspecific) health measurement instruments [such as the EuroQol (EQ)-5D questionnaire]³ or skin-specific instruments [such as the Dermatology Life Quality Index (DLQI)].⁴ Although the use of these instruments might provide some insight into global quality-of-life impairment in patients with hand eczema, one might wonder whether these instruments indicate the true extent of the impairment.^{5,6} In order to assess this issue properly, the disease-specific Quality of Life in Hand Eczema Questionnaire (QOLHEQ) was designed by an international group. In 2014, the German version of the QOLHEQ was validated in a sample of patients with hand eczema. It was found to be valid, reliable and reproducible in a German population.⁷ Translations into several languages were made and a cross-cultural international validation study was performed to make international comparison possible.⁸ However, when translating a measurement instrument and applying it to a new population, it is necessary for such an instrument to be validated for use in that new population. In this study we will report on the scale structure, single-score validity, reproducibility, change-score validity (responsiveness) and interpretability of the Dutch QOLHEQ.

Patients and methods

This study was performed according to a previously published guideline, which adheres to the guidelines developed by the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) group.⁹ We briefly describe our methods below. The QOLHEQ is a 30-item questionnaire with five response categories (never, rarely, sometimes, often, all the time) assessing impairment of health-related quality of life (HRQoL) overall and concerning four subscales, i.e. Symptoms, Emotions, Functioning, Treatment and Prevention. It was translated into Dutch using a six-step method, including forward and backward translations and pilot testing for content validity.¹⁰ The final Dutch version is provided in File S1 (see Supporting Information). A longitudinal design was used to assess the studied measurement properties. Patients were asked to complete the QOLHEQ and reference instruments at three time points, while their hand eczema was also clinically evaluated (Fig. 1).

Study population

Patients were included if they were ≥ 18 years of age and had hand eczema for a duration of at least 1 week, which had

been diagnosed by a dermatologist. Patients with concomitant skin disease on other parts of the body were also eligible for inclusion. Patients with other dermatological hand disease and/or who were unable to complete questionnaires by themselves were excluded. Recruitment was performed between March 2017 and December 2018, and took place at the dermatology department of the University Medical Center Groningen (UMCG). The Medical Ethical Review Board of the UMCG confirmed that this study did not fall under the scope of the Medical Research Involving Human Subjects Act (reference METc 2014/391).

Reference instruments

The following reference instruments were used. The questions for the hand-eczema-specific assessment and the assessment of change were pilot tested for content validity prior to this study.^{9,10}

Hand-eczema-specific assessment (in Dutch, here freely translated), each with the response categories 'not at all', 'slightly', 'moderately', 'strongly' and 'very strongly':

- Global anchor question: How did your hand eczema bother you in your overall health state in the past 7 days?
- Symptoms subscale anchor: How did the symptoms of your hand eczema (such as pain, itch, fissuring, redness) bother you in the past 7 days?
- Emotions subscale anchor: How strongly did your hand eczema affect your emotional well-being (e.g. making you angry, frustrated or anxious about the future) in the past 7 days?
- Functioning subscale anchor: How strongly did your hand eczema affect your functioning (e.g. performing your homework/work or doing hobbies) in the past 7 days?
- Treatment and Prevention subscale anchor: How did treatment and prevention of your hand eczema bother you in the past 7 days?

Skin-specific HRQoL instruments:

- DLQI: comprising 10 items scored on a 4-point scale, with six dimensions (symptoms and feelings, daily activities, leisure, work and school, personal relationships, treatment).⁴
- Skindex-29: comprising 29 items (or technically 30 items) scored on a 5-point scale, with three dimensions (symptoms, emotions and functioning).¹¹

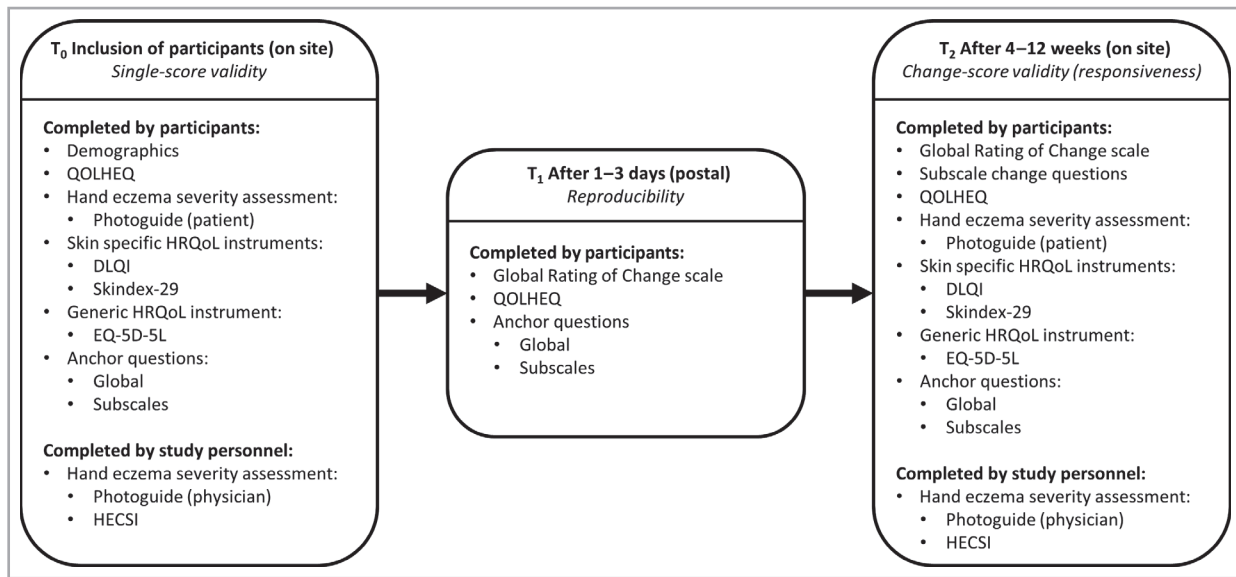


Fig 1. Overview of the longitudinal study design. DLQI, Dermatology Life Quality Index; EQ-5D, quality-of-life questionnaire of the EuroQol Group; HECSI, Hand Eczema Severity Index; HRQoL, Health-Related Quality of Life; QOLHEQ, Quality of Life in Hand Eczema Questionnaire.

Generic:

- EQ-5D-5L: comprising five items scored on a 5-point scale, and a visual analogue scale ranging from 0 to 100.¹²

Severity (morphological signs and extent):

- The photographic guide for severity of hand eczema ('Photoguide'): an instrument to measure the clinical severity of hand eczema (by study personnel or patient) on a 5-point scale (clear, almost clear, moderate, severe, very severe).¹³
- The Hand Eczema Severity Index: a continuous scale ranging from 0 to 360, assessed by study personnel.¹⁴

Assessment of change:

- A Global Rating of Change (GRC) scale was used to assess which patients were unchanged compared with baseline at T₁ and which patients had changed (worsened or improved) at T₂. Patients responded to the question 'Overall, has there been any change in how your hand eczema bothers you since the last time you completed the QOLHEQ?' using the following seven categories: much improvement, moderate improvement, minor improvement, no change, minor deterioration, moderate deterioration, much deterioration.
- Subscale change questions: similar questions were asked to assess changes in the subscales at T₂, but phrased as 'Has there been any change in how (insert subscale) bothers you since the last time you completed the QOLHEQ?' (much improvement, moderate improvement, minor improvement, no change, minor deterioration, moderate deterioration, much deterioration).

Statistical analysis

Sample size

We used an item/participant ratio of 1 : 10. The QOLHEQ has 30 items, which resulted in a sample size of 300 participants.⁹

Scale structure

We used techniques of modern test theory to check the scale structure (structural validity) of the Dutch QOLHEQ. An item response theory (IRT) analysis was performed in order to test whether the subscales of the Dutch QOLHEQ fit the assumed unidimensional Rasch model, using RUMM2030 (RummLab Pty Ltd., Duncraig, Australia). As we obtained a significant likelihood ratio ($P < 0.001$) in all four subscales of the QOLHEQ, we applied a model with an unrestricted parameterization where the thresholds can differ across items, i.e. the partial credit model (a two-parameter logistic model for polytomous response categories). Fit to the Rasch model was determined using the χ^2 -statistic over the item-trait interaction for each item and subscale. Also, means and SDs of fit residuals for the item-person interaction were checked. Individual item fit was also tested using a χ^2 -test. To check for differential item functioning (DIF) an analysis of variance (ANOVA) was performed according to sex and age group (median split of the study population). DIF was assumed to be clinically relevant if a mean difference of 0.5 logits was found for an item.

Furthermore, we tested whether the QOLHEQ fitted a predefined structural equation model (SEM) with confirmatory factor analysis (CFA) using Amos Version 23.0 (IBM, Armonk, NY, U.S.A.). This predefined model was built to assess a second-order construct, HRQoL, measured using four latent factors (subdomains), i.e. Symptoms, Emotions, Functioning, Treatment and Prevention.⁷ These subdomains were measured using the 30 items of the QOLHEQ. Owing to a multivariate kurtosis of the data (Mardia's coefficient = 172.8), various fit indices were calculated using the unweighted least squares method, which is robust against violations of the assumptions of a multivariate normal distribution.¹⁵

Measures of internal consistency of each subscale were reported using Cronbach's α and the Person Separation Index (PSI), calculated using RUMM2030. For both of these calculations, values between 0.70 and 0.95 were considered as evidence for good internal consistency.

Single-score validity and responsiveness (change-score validity)

Tests on the correlation between the Dutch QOLHEQ and the reference instruments were performed on single scores (at T_0) and change scores (at T_2) using Pearson's correlation coefficient (r). Strong correlation (+++) was defined as $r > 0.7$; moderate correlation (++) as $0.7 > r > 0.4$; and weak correlation (+) as $0.4 > r > 0.2$. For the change scores, correlation differences of a minimum of 0.10 were seen as relevant. Furthermore, as recommended by COSMIN, it was tested whether correlations of changes in QOLHEQ score with changes in instruments measuring similar constructs were ≥ 0.50 , and additionally whether correlations of changes in QOLHEQ score with changes in instruments measuring related but dissimilar constructs were lower, i.e. 0.30–0.50.¹⁶ Validity was considered to be high if $< 25\%$ of hypotheses were rejected, moderate if 25–50% were rejected, and poor if $> 50\%$ were rejected.

Reproducibility

Measurement error was reported with the standard error of measurement ($SEM_{\text{agreement}}$) between participants at T_0 and unchanged participants at T_1 . Reliability (test–retest) was reported in the same patients with the intraclass correlation coefficient ($ICC_{\text{agreement}}$), using a two-way mixed effects model for absolute agreement.¹⁷ An $ICC_{\text{agreement}}$ value of > 0.70 was considered acceptable.¹⁸

Interpretability

For single scores, cut-off values for bands indicating how hand eczema affects HRQoL were calculated using the weighted kappa (κ) coefficient of agreement between QOLHEQ scores and the global anchor and subscale anchors. In order not to underestimate the burden for patients when using the banding, we investigated the bands within 0.01 of the highest κ -values. The final band chosen was the band for which the amount of patients reporting a higher impairment according to the anchor question, compared with the band, was lowest. For change scores, the smallest detectable change (SDC) was calculated using the formula $SCD = 1.96 \times \sqrt{2} \times SEM_{\text{agreement}}$.¹⁷ The minimally important change (MIC) for improvement was determined using three different anchor-based methods (File S2 provides details for change-score interpretability; see Supporting Information). For deterioration, no MIC was determined because too few patients deteriorated to allow sound conclusions to be drawn.

Missing values

In eight cases, the QOLHEQ was missing one item. For these cases, the value 0 was imputed.⁷ At T_1 , four cases had skipped a whole page, containing 10 QOLHEQ items. These four cases were excluded from the analyses for reproducibility. One case was missing one DLQI item; here, the value 0 was imputed.⁴ In four cases, the Skindex was missing one item and in one case it was missing two items. For the calculation of the total score, these cases were divided by 28 and 27, respectively. Analyses were performed using SPSS Statistics for Windows, version 23.0 (IBM).

Results

Overall, 300 patients were included in the study at baseline (T_0). A study flowchart is provided in Figure 2. Of the 294 patients included in the T_0 analyses, 54.4% were male and the mean age was 44.9 years. While the rating of hand eczema severity did not differ between sexes, female patients indicated significantly more impairment in HRQoL than male patients on the total QOLHEQ, on all subscales and on the DLQI. Detailed characteristics of the study population and mean T_0 values of the reference instruments are reported in Table 1.

Scale structure

When running the Rasch analysis we found disordered thresholds for 10 items across all subscales. These items were mostly affected by the categories 'rarely' and 'sometimes', indicating that the Dutch population may have problems differentiating between these categories in general. Therefore, we combined these categories for all items, which resulted in a scoring structure of 0-1-1-2-3 for the whole QOLHEQ. This structure fitted the Rasch model for all subscales. However, we still found relevant disordered thresholds for item 26 (Costs). To fix this, we rescored this item to 0-1-1-1-2 (see Table 2 for detailed item characteristics). Rasch analysis of the subscales then revealed the following:

- Symptoms: overall $\chi^2 = 28.7$, degrees of freedom (d.f.) = 28; $P > 0.43$. A PSI of 0.85 and a Cronbach's α of 0.86 indicated a good internal consistency.
- Emotions: overall $\chi^2 = 42.7$, d.f. = 32; $P > 0.09$. A PSI of 0.86 and a Cronbach's α of 0.89 indicated a good internal consistency.
- Functioning: overall $\chi^2 = 39.8$, d.f. = 32; $P > 0.16$. A PSI of 0.86 and a Cronbach's α of 0.89 indicated a good internal consistency.
- Treatment and Prevention: overall $\chi^2 = 33.8$, d.f. = 28; $P > 0.20$. A PSI of 0.78 and a Cronbach's α of 0.78 indicated a good internal consistency.

DIF analysis showed significant uniform DIF for only one item in the Functioning subscale. Item 3 (Home duties) showed that women have a slightly higher chance (+0.6

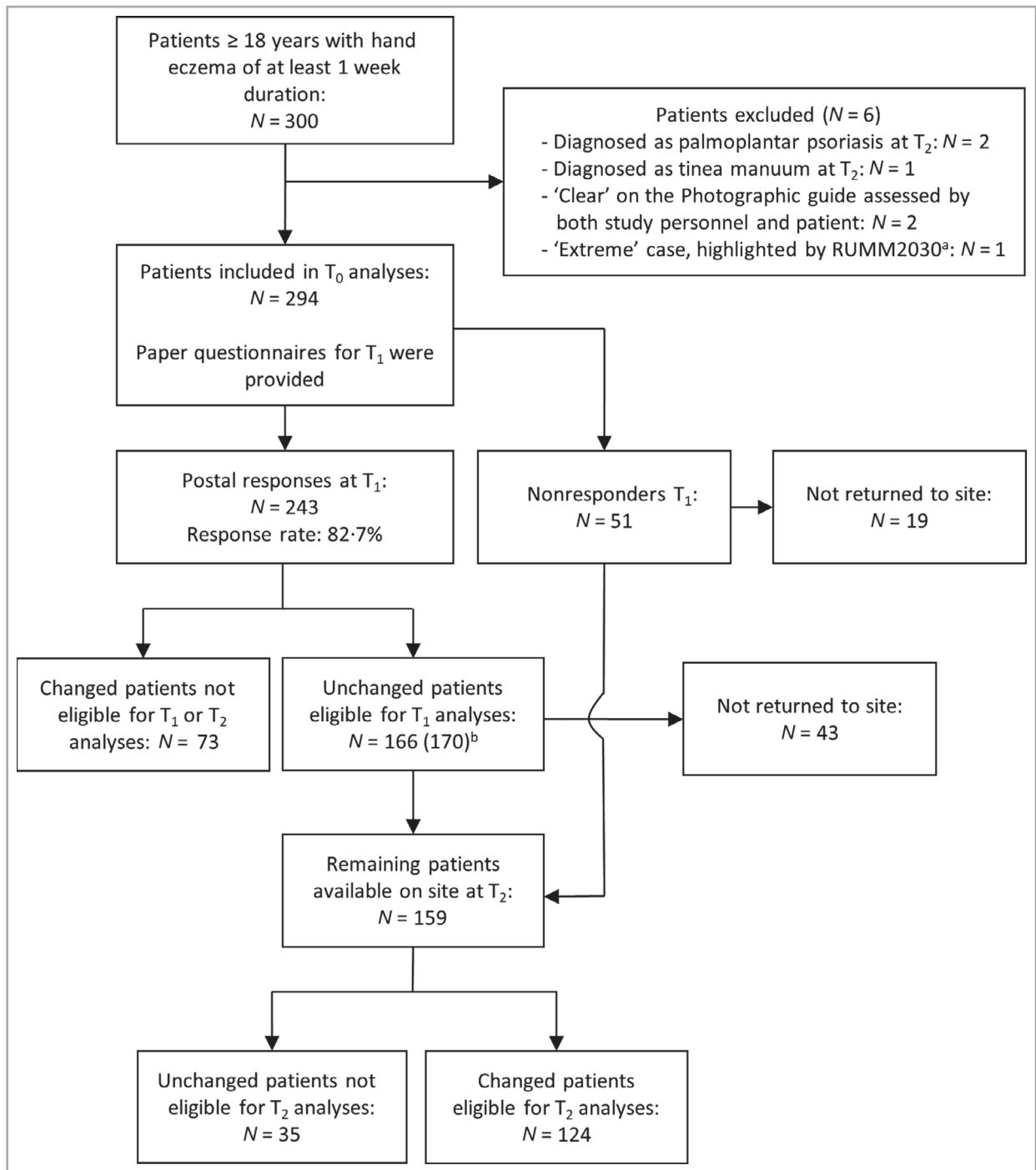


Fig 2. Study flow diagram. QOLHEQ, Quality of Life in Hand Eczema Questionnaire. ^aThis case had almost clear hand eczema (assessed by both study personnel and patient) but had answered all items of the QOLHEQ with ‘always’. ^bFour patients were excluded for T₁ analyses, because they had skipped a whole page of the QOLHEQ, thus bringing the total to 166.

logits) of being impaired for this item. This seems plausible, as women are still more often involved in performing home duties than men, and corresponds to what was found in the German validation study of the QOLHEQ.⁷

The CFA showed that the Dutch QOLHEQ was a good fit for the proposed SEM (Table 3). The total maximum scores

that can be obtained with the Dutch QOLHEQ within the Dutch population are now as follows: total score = 89; Symptoms = 21; Emotions = 24; Functioning = 24; Treatment and Prevention = 20. An SPSS syntax to recode the QOLHEQ to Dutch scores can be found in File S3 (see Supporting Information).

Table 1 Baseline (T_0) characteristics

	Male patients (n = 160)	Female patients (n = 134)	Total (n = 294)
Age (years)			
Mean (SD)	45.0 (14.5)	44.8 (17.5)	44.9 (15.9)
Range	18–74	18–83	18–83
Photoguide severity (patient)			
Mean (SD)	2.8 (0.9)	2.7 (0.8)	2.7 (0.8)
Range	1–5	1–5	1–5
Photoguide severity (physician)			
Mean (SD)	3.2 (0.9)	3.1 (0.9)	3.2 (0.9)
Range	2–5	2–5	2–5
DLQI			
Mean (SD)	7.5 ^a (6.1)	9.3 ^a (6.7)	8.4 (6.4)
Range	0–27	0–26	0–27
Skindex-29			
Mean (SD)	33.8 (20.7)	38.0 (22.0)	35.7 (21.4)
Range	0–91	0–96	0–96
EQ-5D-5L			
Mean value score (SD)	0.77 (0.21)	0.74 (0.25)	0.76 (0.23)
Range	–0.16–1.00	–0.24–1.00	–0.24–1.00
Mean VAS score (SD)	72.9 (17.1)	73.2 (19.1)	73.1 (18.1)
Range	10–100	9–100	9–100
HECSI			
Mean (SD)	48.6 (41.1)	40.9 (35.5)	45.1 (38.8)
Range	3–192	2–144	2–192
QOLHEQ			
Total, mean (SD)	29.0 (15.2) ^a	36.3 (16.8) ^a	32.3 (16.3)
Range	3–75	0–82	0–82
Symptoms, mean (SD)	8.8 (4.1) ^a	10.5 (4.4) ^a	9.6 (4.3)
Range	0–19	0–21	0–21
Emotions, mean (SD)	6.8 (4.6) ^a	8.5 (5.2) ^a	7.6 (5.0)
Range	0–21	0–24	0–24
Functioning, mean (SD)	6.8 (4.6) ^a	8.9 (5.4) ^a	7.7 (5.1)
Range	0–20	0–22	0–22
Treatment and Prevention, mean (SD)	6.7 (3.7) ^a	8.3 (3.8) ^a	7.4 (3.9)
Range	0–17	0–17	0–17

DLQI, Dermatology Life Quality Index; EQ-5D, quality-of-life questionnaire of the EuroQol Group; HECSI, Hand Eczema Severity Index; QOLHEQ, Quality of Life in Hand Eczema Questionnaire. ^aDifferences between male patients and female patients are significant ($P < 0.05$) according to Student's *t*-test.

Single-score validity and responsiveness (change-score validity)

Of the a priori formulated hypotheses for single-score validity, 80% were confirmed, indicating high validity of the Dutch QOLHEQ (Table 4). In the analysis of responsiveness, 124 cases were included because these participants indicated that they had changed at T_2 according to the GRC scale, while being unchanged at T_1 or when compared with baseline if they were nonresponders at T_1 . Therefore these cases represented patients who had 'really changed'. In these patients, 64% of the a priori formulated hypotheses for change scores were confirmed, indicating a moderate responsiveness of the Dutch QOLHEQ (Table 5).

Reproducibility

There were 166 cases included in the analysis for reproducibility. This concerns the unchanged patients at T_1

according to the GRC scale. The $SEM_{\text{agreement}}$ of the complete QOLHEQ was 5.2 points. The $ICC_{\text{agreement}}$ was 0.91 [95% confidence interval (CI) 0.85–0.94], indicating good reproducibility. For the subscales, we found the following values, which indicated good reproducibility for all four subscales:

- Symptoms: $SEM_{\text{agreement}} = 1.6$ points; $ICC_{\text{agreement}} = 0.88$ (95% CI 0.84–0.91).
- Emotions: $SEM_{\text{agreement}} = 1.8$ points; $ICC_{\text{agreement}} = 0.88$ (95% CI 0.82–0.92).
- Functioning: $SEM_{\text{agreement}} = 1.9$ points; $ICC_{\text{agreement}} = 0.88$ (95% CI 0.80–0.92).
- Treatment and Prevention: $SEM_{\text{agreement}} = 1.5$ points; $ICC_{\text{agreement}} = 0.86$ (95% CI 0.80–0.89).

Interpretability

For single scores, several bands for severity of HRQoL impairment were tested for the overall QOLHEQ score and subscales.

Table 2 Results of Rasch analysis

Item number in questionnaire	Description	Location ^a	Range ^b	Fit residual	χ^2 -test	P-values ^c	Scoring structure
Symptoms							
23	Bleeding	1.40	-1.51-3.58	-0.69	2.76	0.60	0-1-1-2-3
9	Causing loss of sleep	1.15	-0.55-2.35	1.07	4.79	0.31	0-1-1-2-3
11	Fissuring	-0.09	-2.25-1.77	0.25	2.98	0.56	0-1-1-2-3
1	Painful	-0.30	-3.01-1.75	-0.69	7.05	0.13	0-1-1-2-3
20	Redness	-0.40	-2.98-1.39	1.41	1.46	0.83	0-1-1-2-3
6	Itching	-0.58	-3.27-1.90	0.73	2.30	0.68	0-1-1-2-3
28	Dryness	-1.19	-3.71-1.16	-0.32	7.31	0.12	0-1-1-2-3
Emotions							
30	Nervous	1.01	-1.07-2.47	-2.45	9.25	0.06	0-1-1-2-3
19	Sad/depressed	0.84	-1.51-2.53	-2.37	9.34	0.05	0-1-1-2-3
27	Embarrassed	0.81	-1.09-2.25	-0.39	1.48	0.83	0-1-1-2-3
10	Anxious about future	0.36	-1.23-1.28	1.55	4.25	0.37	0-1-1-2-3
21	Irritated	0.13	-2.66-2.18	0.25	1.78	0.78	0-1-1-2-3
16	Hide my hands	-0.04	-1.62-0.86	3.02	8.51	0.07	0-1-1-2-3
5	Frustrated	-1.28	-4.22-0.89	-0.68	1.18	0.88	0-1-1-2-3
8	Annoyed	-1.84	-4.90-0.15	1.30	6.99	0.14	0-1-1-2-3
Functioning							
17	Avoiding contact with people	1.33	-0.29-2.53	1.06	6.00	0.20	0-1-1-2-3
25	Affecting friendships	1.06	-0.64-2.09	0.90	0.48	0.98	0-1-1-2-3
29	Touching partner	0.81	-1.15-2.62	0.91	0.73	0.95	0-1-1-2-3
15	Dressing myself	0.48	-1.79-2.05	-2.27	11.12	0.03	0-1-1-2-3
12	Leisure time/hobbies	-0.51	-3.12-1.27	-0.98	9.70	0.05	0-1-1-2-3
14	Washing myself	-0.83	-2.82-0.70	-0.84	2.24	0.69	0-1-1-2-3
2	Restricting job	-0.94	-3.57-1.02	-0.26	6.54	0.16	0-1-1-2-3
3	Home duties	-1.40	-4.19-0.48	-0.41	2.95	0.57	0-1-1-2-3
Treatment and Prevention							
26	Costs	0.97	-0.20-2.13	1.82	10.48	0.03	0-1-1-1-2 ^d
24	Side-effects	0.60	-0.89-1.35	-0.60	7.45	0.11	0-1-1-2-3
18	Visiting physician	0.57	-1.25-1.52	-0.28	2.44	0.66	0-1-1-2-3
7	Time-consuming	0.05	-2.17-1.50	-0.63	4.14	0.39	0-1-1-2-3
4	Wearing gloves	-0.33	-1.52-0.57	0.33	3.56	0.47	0-1-1-2-3
22	Avoiding certain things	-0.67	-2.62-0.64	-0.07	1.71	0.79	0-1-1-2-3
13	Using creams	-1.19	-3.33-0.28	0.35	2.87	0.58	0-1-1-2-3

^aLocation sorted by severity; items assessing most severe impairment are on top of each domain/subscale. ^bRange of thresholds of each item. ^cAccording to a χ^2 -test; misfit was considered significant if $P < 0.007$ or $P < 0.008$ (dependent on number of items in the subscale). ^dScoring structure adjusted additionally.

Table 3 Fit indices for the structural equation model of the Quality of Life in Hand Eczema Questionnaire, consisting of four subscales (factors) loading on a higher order factor measuring health-related quality of life

Fit index	Complete second-order model	Model fit	Recommendation for good fit ^a	Recommendation for acceptable fit ^a
SRMR	0.067	Acceptable	< 0.05	$0.05 < \text{SRMR} \leq 0.10$
GFI	0.980	Good	> 0.95	$0.90 \leq \text{GFI} < 0.95$
AGFI	0.976	Good	> 0.90	$0.85 \leq \text{AGFI} < 0.90$
NFI	0.976	Good	> 0.95	$0.90 \leq \text{NFI} < 0.95$
RFI	0.974	Good	> 0.95	$0.90 \leq \text{RFI} < 0.95$

^aAccording to guidelines by Schermelleh-Engel *et al.*¹⁵ SRMR, Standardized Root Mean Residual; GFI, Goodness of Fit Index; AGFI, Adjusted GFI; NFI, Normed Fit Index; RFI, Relative Fit Index.

For the overall QOLHEQ, we propose separate bands for male patients and female patients. The final band chosen for the overall QOLHEQ had a κ -value of 0.430 (not at all, 0–13; slightly, 14–28; moderately, 29–44; strongly, 45–64; very strongly, ≥ 65). All proposed bands and details on the

calculation of single-score interpretability are provided in File S4 (see Supporting Information).

The SCD in 166 unchanged patients at T₁ was 14.4 points for the overall QOLHEQ. The preferred MIC, obtained using the receiver operating characteristic method, was 11.5. The

Table 4 Single-score validity (at T₀) correlations between the Quality of Life in Hand Eczema Questionnaire (QOLHEQ) and reference instruments

	Correlation hypothesized ^a	Correlation found	R ²	Hypotheses confirmed?
Reference measure				
DLQI	+++	0.77	0.59	Yes
Skindex-29	+++	0.80	0.64	Yes
Global anchor	++	0.59	0.35	Yes
EQ-5D-5L (VAS) ^b	++	-0.33	0.11	No
EQ-5D-5L (Value) ^b	++	-0.57	0.32	Yes
Photoguide (patient)	++	0.47	0.22	Yes
Photoguide (physician)	+	0.43	0.18	No
HECSI	+	0.37	0.14	Yes
Subscales^c				
Symptoms anchor	+++	0.70	0.49	Yes
Emotions anchor	+++	0.71	0.50	Yes
Functioning anchor	+++	0.72	0.52	Yes
Treatment and Prevention anchor	+++	0.58	0.34	No
Skindex-29 Symptoms subscale	+++	0.77	0.59	Yes
Skindex-29 Emotions subscale	+++	0.85	0.73	Yes
Skindex-29 Functioning subscale	+++	0.70	0.49	Yes
Specific comparisons				
QOLHEQ Symptoms subscale – Photoguide (physician)	++	0.52	0.27	Yes
QOLHEQ Symptoms subscale - HECSI	++	0.46	0.21	Yes

DLQI, Dermatology Life Quality Index; EQ-5D, quality-of-life questionnaire of the EuroQol Group; HECSI, Hand Eczema Severity Index; VAS, visual analogue scale. ^aA priori defined: strong correlation (+++) $r > 0.7$; moderate correlation (++) $0.7 > r > 0.4$; weak correlation (+) $0.4 > r > 0.2$. ^bNegative value, because the EQ-5D-5L is scored inversely to the QOLHEQ. ^cCorrelation between QOLHEQ subscale score and reference instrument.

SDC and MIC of the subscales and further details on calculations are provided in File S2 (see Supporting Information).

Discussion

In this study, we tested various measurement properties of the Dutch QOLHEQ. We proposed a scoring structure that fitted a Rasch model, and demonstrated good validity and reproducibility, and moderate responsiveness. An improvement of ≥ 15 points within the Dutch population should be regarded as a real and important improvement.

Compared with the German version, the Dutch QOLHEQ had to be substantially rescored. A possible explanation for this is that the Dutch translation for the item 'sometimes' ('nu en dan') was not optimal; possibly, in Dutch, it was too similar to the category 'rarely' ('zelden'). A future validation study could assess whether another translation (e.g. changing 'nu en dan' to 'soms') may yield a better discrimination on the lower end of the Dutch QOLHEQ scale. Still, the Dutch translation of the QOLHEQ as presented here fulfils the rigorous requirements of modern test theory including IRT and SEM. Therefore, it is ready to be used in any study assessing HRQoL impairments in a Dutch population of patients with hand eczema. However, when reporting QOLHEQ results for Dutch patients in future studies, both the national and international values, which were obtained for six languages in a cross-cultural validation study,⁸ should be reported for the sake of international comparison.

The most problematic issues in the analysis pertained to the item 'Costs'. Participants could not be distinguished based on

this item. The health insurance companies in the Netherlands reimburse the treatment of hand eczema, including several emollients and protective gloves. As a result, the out-of-pocket costs for hand eczema are often low. This may offer a good explanation as to why a large group of patients ($n = 161$) chose 'never' for this item. We chose to keep this item in the instrument as it may still be important for a small subgroup of patients. However, if efforts were to be made in the future to reduce the amount of items in the QOLHEQ, for example to increase its ease of use, this item should be the first to be considered for removal.

Most of our a priori stated hypotheses were confirmed in the analyses for single-score validity. For the single scores, the Photoguide, as scored by a physician, correlated moderately with the QOLHEQ (0.43), where we had expected it to be only weakly correlated (< 0.4). However, the Photoguide, as scored by the patient, still correlated more strongly with the QOLHEQ than the physician score, which was as we had hypothesized. Therefore, we do not consider this to be an issue. For the change scores, the QOLHEQ showed a higher or comparable responsiveness when compared with the reference instruments, indicating that the QOLHEQ was sensitive to detect change in HRQoL in patients with hand eczema.

This validation study was performed using a paper version of the QOLHEQ. In recent decades, digital questionnaires have been increasingly used for capturing patient-reported outcomes, mainly because they provide direct integration into medical health records and research databases. If a paper questionnaire is adapted to an electronic version, this may alter the measurement properties of the questionnaire.¹⁹ However, this is not always

Table 5 Responsiveness (change-score validity) in changed patients between T₀ and T₂

	Correlations found	Hypotheses confirmed?
Hypothesis on correlations		
Change QOLHEQ – GRC > Change DLQI – GRC	0.46 vs. 0.38	No
Change QOLHEQ – GRC > Change Skindex-29 – GRC	0.46 vs. 0.33	Yes
Change QOLHEQ – GRC > Change EQ-5D value – GRC	0.46 vs. –0.27 ^a	Yes
Change QOLHEQ – GRC > Change EQ-5D VAS – GRC	0.46 vs. –0.25 ^a	Yes
Change QOLHEQ – change Photoguide (physician) > Change DLQI – change Photoguide (physician)	0.46 vs. 0.45	No
Change QOLHEQ – change Photoguide (physician) > Change Skindex-29 – change Photoguide (physician)	0.46 vs. 0.46	No
Change QOLHEQ – change Photoguide (physician) > Change EQ-5D Value – change Photoguide (physician)	0.46 vs. –0.36 ^a	Yes
Change QOLHEQ – change Photoguide (physician) > Change EQ-5D VAS – change Photoguide (physician)	0.46 vs. –0.23 ^a	Yes
Change QOLHEQ – change HECSI > Change DLQI – change HECSI	0.40 vs. 0.35	No
Change QOLHEQ – change HECSI > Change Skindex-29 – change HECSI	0.40 vs. 0.33	No
Change QOLHEQ – change HECSI > Change EQ-5D Value – change HECSI	0.40 vs. –0.27 ^a	Yes
Change QOLHEQ – change HECSI > Change EQ-5D VAS – change HECSI	0.40 vs. –0.15 ^a	Yes
Hypothesis on subscale correlations		
Change QOLHEQ Symptoms – Symptoms change question > Change Skindex-29 Symptoms – Symptoms change question	0.49 vs. 0.44	No
Change QOLHEQ Emotions – Emotions change question > Change Skindex-29 Emotions – Emotions change question	0.47 vs. 0.30	Yes
Change QOLHEQ Functioning – Functioning change question > Change Skindex-29 Functioning – Functioning change question	0.54 vs. 0.37	Yes
Hypothesis according to COSMIN		
Instruments measuring similar constructs		
Change QOLHEQ – GRC	0.46	No
Change QOLHEQ – Change DLQI	0.56	Yes
Change QOLHEQ – Change Skindex-29	0.63	Yes
Instruments measuring related, but dissimilar constructs		
Change QOLHEQ – Change Photoguide (physician)	0.46	Yes
Change QOLHEQ – Change HECSI	0.40	Yes
Change QOLHEQ – Change EQ-5D value	–0.48 ^a	Yes
Change QOLHEQ – Change EQ-5D VAS	–0.26 ^a	No
DLQI, Dermatology Life Quality Index; EQ-5D, quality-of-life questionnaire of the EuroQol Group; GRC, Global Rating of Change scale; HECSI, Hand Eczema Severity Index; QOLHEQ, Quality of Life in Hand Eczema Questionnaire; VAS, visual analogue scale. ^a Negative value, because the EQ-5D-5L is scored inversely to the QOLHEQ.		

the case,^{20,21} but it is something that should be taken into consideration in future studies involving the QOLHEQ.

A limitation of this study was that the identification of unchanged patients at T₁ and changed patients at T₂ was based on patient memory. A certain amount of recall bias cannot be ruled out, especially for T₂, which was assessed 4–12 weeks following T₀. Another limitation was that between T₀ and T₁ many patients (n = 73) had already indicated a change in the impairment that they perceived to be a result of their hand eczema, limiting the sample size for reproducibility and responsiveness, although the numbers are still acceptable.²² This clearly reflects the variable course that is often associated with the disease. A final limitation could have been the short time between T₀ and T₁, in which participants may have been able to recall their answers from T₀. However, as was apparent from the number of patients who had quickly changed at T₁, this short period was needed to ensure a sufficient number of

eligible participants at T₁. Also, we believe that answers given on the 30-item long QOLHEQ would have been hard to recall, even after 1–3 days.

In conclusion, the Dutch version of the QOLHEQ has been shown to be a valid, reproducible and responsive instrument in the Dutch hand eczema population. We recommend its use to measure impairment of HRQoL in Dutch patients with hand eczema.

References

- 1 Moberg C, Alderling M, Meding B. Hand eczema and quality of life: a population-based study. *Br J Dermatol* 2009; **161**:397–403.
- 2 Agner T, Andersen KE, Brandao FM *et al.* Hand eczema severity and quality of life: a cross-sectional, multicentre study of hand eczema patients. *Contact Dermatitis* 2008; **59**:43–7.
- 3 Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37**:53–72.

- 4 Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)—a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994; **19**:210–16.
- 5 Coenraads PJ, Bouma J, Diepgen TL. [Quality of life of patients with occupationally-induced hand eczema]. *Hautarzt* 2004; **55**:28–30 (in German).
- 6 Ahmed A, Shah R, Papadopoulos L, Bewley A. An ethnographic study into the psychological impact and adaptive mechanisms of living with hand eczema. *Clin Exp Dermatol* 2015; **40**:495–501.
- 7 Ofenloch RF, Weisshaar E, Dumke AK *et al.* The Quality of Life in Hand Eczema Questionnaire (QOLHEQ): validation of the German version of a new disease-specific measure of quality of life for patients with hand eczema. *Br J Dermatol* 2014; **171**:304–12.
- 8 Ofenloch RF, Oosterhaven JAF, Susitaival P *et al.* Cross-cultural validation of the Quality of Life in Hand Eczema Questionnaire (QOLHEQ). *J Invest Dermatol* 2017; **137**:1454–60.
- 9 Oosterhaven JAF, Schuttelaar MLA, Apfelbacher C *et al.* Guideline for translation and national validation of the Quality of Life in Hand Eczema Questionnaire (QOLHEQ). *Contact Dermatitis* 2017; **77**:106–15.
- 10 Wild D, Grove A, Martin M *et al.* Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005; **8**:94–104.
- 11 Chren MM, Lasek RJ, Flocke SA, Zyzanski SJ. Improved discriminative and evaluative capability of a refined version of Skindex, a quality-of-life instrument for patients with skin diseases. *Arch Dermatol* 1997; **133**:1433–40.
- 12 Herdman M, Gudex C, Lloyd A *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; **20**:1727–36.
- 13 Coenraads PJ, Van Der Walle H, Thestrup-Pedersen K *et al.* Construction and validation of a photographic guide for assessing severity of chronic hand dermatitis. *Br J Dermatol* 2005; **152**:296–301.
- 14 Held E, Skoet R, Johansen JD, Agner T. The hand eczema severity index (HECSI): a scoring system for clinical assessment of hand eczema. A study of inter- and intraobserver reliability. *Br J Dermatol* 2005; **152**:302–7.
- 15 Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods Psychol Res Online* 2003; **8**:23–74.
- 16 Prinsen CAC, Mokkink LB, Bouter LM *et al.* COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018; **27**:1147–57.
- 17 de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006; **59**:1033–9.
- 18 Terwee CB, Bot SD, de Boer MR *et al.* Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**:34–42.
- 19 Juniper EF, Langlands JM, Juniper BA. Patients may respond differently to paper and electronic versions of the same questionnaires. *Respir Med* 2009; **103**:932–4.
- 20 Bjorner JB, Rose M, Gandek B *et al.* Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. *J Clin Epidemiol* 2014; **67**:108–13.
- 21 Robles N, Rajmil L, Rodriguez-Arjona D *et al.* Development of the web-based Spanish and Catalan versions of the Euroqol 5D-Y (EQ-5D-Y) and comparison of results with the paper version. *Health Qual Life Outcomes* 2015; **13**:72.
- 22 de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. New York, NY: Cambridge University Press, 2011.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

File S1. Dutch Quality of Life in Hand Eczema Questionnaire.

File S2. Interpretability of Dutch change scores.

File S3. SPSS syntax to recode the Dutch Quality of Life in Hand Eczema Questionnaire.

File S4. Interpretability of Dutch single scores.