University of Groningen

Interpretability of the Quality Of Life in Hand Eczema Questionnaire

Oosterhaven, Jart A F; Ofenloch, Robert F; Schuttelaar, Marie L A

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Interpretability of the Quality Of Life in Hand Eczema Questionnaire

Jart A.F. Oosterhaven[1], Robert F. Ofenloch[2] and Marie L.A. Schuttelaar[1]

The Quality Of Life in Hand Eczema Questionnaire (QOLHEQ) is used to measure impairment of health-related quality of life in hand eczema. Here, we prospectively studied the interpretability of international QOLHEQ scores at three time points: baseline, after 1–3 days (T$_1$), and after 4–12 weeks (T$_2$). Adult patients with hand eczema completed the QOLHEQ and anchor questions for overall assessment of health-related quality of life impairment. Interpretability of single scores was assessed at baseline by defining severity bands based on agreement with the anchor questions. Smallest detectable change was calculated at T$_1$. Minimally important change of improvement was calculated at T$_2$ using three methods: mean cut-off, receiver operating characteristic, and 95% limit. A total of 294 adult patients were included (160 males, mean age 44.9 years). The final proposed severity band of overall QOLHEQ single scores (κ-coefficient of agreement, 0.431) was not at all, 0–10; slightly, 11–39; moderately, 40–61; strongly, 62–86; and very strongly, ≥87. Separate overall severity bands were proposed for males and females and the four subscales of the QOLHEQ. The smallest detectable change in 166 unchanged patients was 18.6 points. The preferred minimally important change, obtained with the receiver operating characteristic method, was 21.5 points. An overall QOLHEQ score of ≥22 is recommended as cut-off for a minimally important, real change.

INTRODUCTION

The Quality Of Life in Hand Eczema Questionnaire (QOLHEQ) is a measurement instrument that was developed with a group of international experts in cooperation with patients (Ofenloch et al., 2014). It is used to measure impairment of health-related quality of life (HRQoL) in patients with hand eczema. The QOLHEQ was validated for national use in Germany and Japan (Minamoto et al., 2018; Ofenloch et al., 2014). Furthermore, a cross-cultural validation was performed involving six countries (Ofenloch et al., 2017), classifying the QOLHEQ as a third generation measurement instrument for assessing HRQoL (Nijsten, 2012). Now that the measurement properties of the QOLHEQ are extensively studied, it is useful to obtain an understanding of what QOLHEQ scores actually mean, especially regarding international scores from the cross-cultural validation study. Also, it is useful to investigate what changes on the QOLHEQ and its subscales constitute an important improvement to aid clinical decision making and sample size calculation of clinical trials. Interpretability is defined as "the degree to which one can assign qualitative meaning (i.e., clinical or commonly understood connotations) to an instrument's quantitative scores or change in scores" (Mokkink et al., 2010). In this study, we aimed to assess the interpretability of international QOLHEQ scores.

RESULTS

Interpretability of single scores

At baseline, 294 patients were included in this study; see Table 1 for basic characteristics. Clinical severity of hand eczema was determined at baseline according to the photographic guide for severity (Coenraads et al., 2005): almost clear ($n = 77$, 26.2%), moderate ($n = 114$, 38.8%), severe ($n = 81$, 27.6%), and very severe ($n = 22$, 7.5%). The distribution of QOLHEQ overall scores stratified by Global anchor categories for HRQoL impairment is shown in Figure 1. All anchor questions correlated > 0.50 with the QOLHEQ overall and subscale scores. In total, 18 bands for HRQoL impairment were tested for the overall QOLHEQ; 8 bands were tested both for the Symptoms and Emotions subscales, 4 bands for the Functioning subscale, and 18 bands for the Treatment and Prevention subscale. See Table 2 for the final chosen bands and Supplementary Tables S2–S15 and Supplementary Figures S2–S5 for details on the testing of single scores.

Overview of QOLHEQ overall scores falling outside the proposed banding

Of the study group, 29 patients (9.9%) had a Global anchor score >1 point outside of that predicted by the final overall QOLHEQ band. There were 81 (27.6%) patients with an actual Global anchor score 1 point lower than the final QOLHEQ band predicted. In this group, there were significantly more females ($P < 0.05$). There were 51 patients (17.3%) with an actual Global anchor score 1 point higher than the final overall QOLHEQ band predicted. In this group,

[1]Department of Dermatology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; and [2]Occupational Dermatology, Department of Dermatology, University Hospital Heidelberg, Heidelberg, Germany

Correspondence: Marie L. A. Schuttelaar, Department of Dermatology, University Medical Center Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands. E-mail: m.l.a.schuttelaar@umcg.nl

there were significantly more males ($P < 0.01$). Because of this, we decided to propose separate bandings for males and females for the overall QOLHEQ (Table 2). For subscales, there were no sex and age distribution differences between patients falling within the proposed banding and those falling outside it. Only for the Treatment and Prevention scale were there significantly more females in the group with an anchor score 1 point lower than the final band predicted and significantly more males in the group with an anchor score 1 point higher. For the sake of clarity and simplicity, we decided not to propose separate bandings for males and females for only this subscale.

Interpretability of change scores

Smallest detectable change (SDC). The SDC of the QOL-HEQ was calculated based on the analysis of 166 unchanged patients on the Global anchor question of change between T_0 and T_1. This gave the following results:

- QOLHEQ overall: standard error of measurement ($SEM_{agreement}$) = 6.7 points; SDC = 18.6 points;
- Symptoms subscale: $SEM_{agreement}$ = 2.0 points; SDC = 5.5 points;
- Emotions subscale: $SEM_{agreement}$ = 2.3 points; SDC = 6.4 points;
- Functioning subscale: $SEM_{agreement}$ = 2.6 points; SDC = 7.3 points;
- Treatment and Prevention subscale: $SEM_{agreement}$ = 1.8 points; SDC = 4.9 points.

Minimally important change (MIC). The correlation between the change in QOLHEQ score and the change in the Global anchor question for change in HRQoL impairment between T_0 and T_2 was $r = 0.51$. The correlations for the subscales were as follows:

- Symptoms: $r = 0.55$;
- Emotions: $r = 0.44$;
- Functioning: $r = 0.52$;
- Treatment and Prevention: $r = 0.28$.

Therefore, the anchor questions were considered to be acceptable anchors for determination of the MIC, except for the anchor question for the Treatment and Prevention subscale (Cella et al., 2002; Revicki et al., 2008). We did determine the MIC for this scale, but the lack of good correlation between the change anchor and the change score must be seriously considered when using the values for the Treatment and Prevention subscale. The correlations were not influenced by sex or age, except for the correlation between sex and the Emotions subscale where there was a significant difference between males and females ($P < 0.01$). For the sake of clarity and simplicity, we decided not to calculate separate MICs for males and females for only this subscale. Calculation of the MIC of improvement according to our three used methods resulted in the values shown in Table 3. The distribution of the overall QOLHEQ change scores was visualized as anchor-based distribution for improved and unchanged patients, along with the three MIC values. For this, along with details on calculation of the MIC for the

Table 1. Basic Characteristics of Study Population

Characteristic	Males (n = 160)	Females (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
QOLHEQ overall			
Mean (SD)	38.8 (21.1)[1]	48.8 (22.3)[1]	43.3 (22.2)
Range	2−96	0−103	0−103
Symptoms subscale			
Mean (SD)	10.9 (5.3)[1]	13.1 (5.4)[1]	11.9 (5.5)
Range	0−23	0−25	0−25
Emotions subscale			
Mean (SD)	9.6 (6.6)[1]	12.2 (7.1)[1]	10.8 (7.0)
Range	0−27	0−30	0−30
Functioning subscale			
Mean (SD)	9.9 (7.0)[1]	13.2 (7.7)[1]	11.4 (7.5)
Range	0−28	0−30	0−30
Treatment and Prevention subscale			
Mean (SD)	8.4 (4.7)[1]	10.4 (4.8)[1]	9.3 (4.9)
Range	0−20	0−21	0−21
HRQoL impaired − Global			
Not at all, n (%)	32 (20.0)	30 (22.4)	62 (21.1)
Slightly, n (%)	56 (35.0)	31 (23.1)	87 (29.6)
Moderately, n (%)	42 (26.3)	42 (31.3)	84 (28.6)
Strongly, n (%)	24 (15.0)	25 (18.7)	49 (16.7)
Very strongly, n (%)	6 (3.8)	6 (4.5)	12 (4.1)
HRQoL impaired − Symptoms[1]			
Not at all, n (%)	3 (1.9)	6 (4.5)	9 (3.1)
Slightly, n (%)	55 (34.4)	22 (16.4)	77 (26.2)
Moderately, n (%)	40 (25.0)	39 (29.1)	79 (26.9)
Strongly, n (%)	43 (26.9)	53 (39.6)	96 (32.7)
Very strongly, n (%)	19 (11.9)	14 (10.4)	33 (11.2)
HRQoL impaired − Emotions			
Not at all, n (%)	68 (42.5)	45 (33.6)	113 (38.4)
Slightly, n (%)	39 (24.4)	33 (24.6)	72 (24.5)
Moderately, n (%)	29 (18.1)	28 (20.9)	57 (19.4)
Strongly, n (%)	18 (11.3)	23 (17.2)	41 (13.9)
Very strongly, n (%)	6 (3.8)	5 (3.7)	11 (3.7)
HRQoL impaired − Functioning			
Not at all, n (%)	38 (23.8)	24 (17.9)	62 (21.1)
Slightly, n (%)	44 (27.5)	31 (23.1)	75 (25.5)
Moderately, n (%)	41 (25.6)	29 (21.6)	70 (23.8)
Strongly, n (%)	24 (15.0)	34 (25.4)	58 (19.7)
Very strongly, n (%)	13 (8.1)	16 (11.9)	29 (9.9)
HRQoL impaired − Treatment and Prevention			
Not at all, n (%)	56 (35.0)	42 (31.3)	98 (33.3)
Slightly, n (%)	52 (32.5)	37 (27.6)	89 (30.3)
Moderately, n (%)	33 (20.6)	35 (26.1)	68 (23.1)
Strongly, n (%)	13 (8.1)	17 (12.7)	30 (10.2)
Very strongly, n (%)	6 (3.8)	3 (2.2)	9 (3.1)

Abbreviations: HRQoL, health-related quality of life; QOLHEQ, Quality Of Life in Hand Eczema Questionnaire.
[1]Significant difference between males and females ($P < 0.01$).

QOLHEQ and subscales, see Supplementary Tables S16−S20 and Supplementary Figures S6−S10.

Floor and ceiling effects. For the overall QOLHEQ, the Symptoms subscale, and the Treatment and Prevention subscale, <9% of patients scored the highest score or the lowest
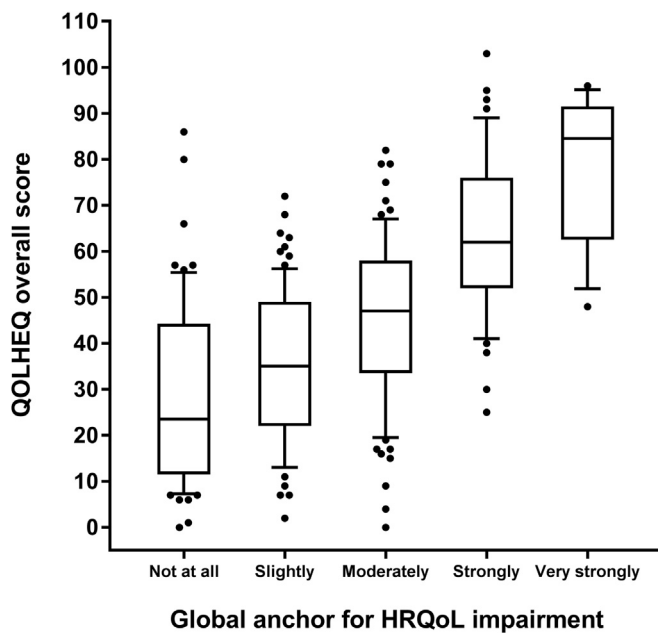
Figure 1. Distribution of the QOLHEQ overall score by Global anchor for HRQoL impairment. Boxes represent 25th–75th percentile with the middle line representing the median. The ends of the whiskers represent 10th–90th percentile. Outliers are plotted as circles beyond the whiskers. HRQoL, health-related quality of life; QOLHEQ, Quality Of Life in Hand Eczema Questionnaire.

values obtained in different countries when these values are (re)scored using the scoring structure from the cross-cultural validation study (Ofenloch et al., 2017). The QOLHEQ can also be used to assess HRQoL impairments in four specific domains. For these domains (subscales), interpretability values were reported separately.

Differences were found between the three values obtained for the MIC in this study. The best method to define the MIC has not yet been determined. However, of the anchor-based methods, the receiver operating characteristic method was suggested as a good choice because it is aimed at minimizing misclassification of patients who are importantly improved and patients who are unchanged (Terwee et al., 2010).

The receiver operating characteristic method gave a MIC value of 21.5 points for the overall QOLHEQ, whereas the SDC was 18.6 points. This means that a change in QOLHEQ score of 22 or higher can be considered an important change for patients, as well as a real change beyond measurement error. We therefore recommend using a score of ≥ 22 as cutoff for a minimally important, real change. For subscales, the cut-offs are based on the SDC, because the established MIC was smaller than the SDC. Therefore, the MIC is meaningless and the SDC is the smallest value beyond measurement error closest to it. Thus, the cut-offs are the following: Symptoms, 6 points; Emotions, 7 points; Functioning, 8 points; and Treatment and Prevention, 5 points.

Although international values can now be compared, still only in three countries (Germany, Japan, and the Netherlands) have validation studies been performed to assess measurement properties and precision of QOLHEQ measurements specifically for that country (Minamoto et al., 2018; Ofenloch et al., 2014; Oosterhaven et al., 2019). To use the QOLHEQ in a new language version, a validation study should still be performed (Oosterhaven et al., 2017). Ultimately, results from national studies should report both national values and international values.

It is important to realize that the values presented in this study are particularly valid for Germany, Japan, and the Netherlands, because for Sweden, Finland, and Turkey several subscales were adjusted for differential item functioning (Ofenloch et al., 2017). This means that the interpretability values presented here might be slightly deviating from true values for these countries. However, until interpretability studies are performed with patients specifically

score at all three time points. For the Emotions subscale and the Functioning subscale, 17.6% and 19.5% of patients, respectively, reached the lowest score at T_2. On the anchor question for Emotions at T_2, all these patients indicated that they were not at all impaired in their HRQoL. For the Functioning anchor at T_2, only one patient indicated a slight impairment in HRQoL; all other patients reported they were not at all impaired. Therefore, we consider these scores as truly low and do not consider these subscales to have a floor effect. There were no ceiling effects for the Emotions and Functioning subscales. For details, see Supplementary Table S21.

DISCUSSION
In this study, we presented values to aid interpretability of single scores and change scores of international QOLHEQ values. This will contribute to the comparison of QOLHEQ

Table 2. Final Bands for HRQoL Impairment Expressed as Single Scores on the QOLHEQ

	HRQoL impaired					Coefficient of agreement (κ) with anchor question	Correlation (*r*) with anchor question
	Not at all	Slightly	Moderately	Strongly	Very strongly		
QOLHEQ overall	0–10	11–39	40–61	62–86	≥87	0.431	0.58
QOLHEQ overall (M)	0–10	11–35	36–53	54–83	≥84	0.461	0.61
QOLHEQ overall (F)	0–13	14–40	41–61	62–86	≥87	0.407	0.54
Symptoms subscale	0	1–8	9–13	14–19	≥20	0.529	0.69
Emotions subscale	0–6	7–12	13–20	21–26	≥27	0.531	0.71
Functioning subscale	0–2	3–9	10–15	16–24	≥25	0.564	0.73
Treatment and Prevention subscale	0–6	7–9	10–14	15–18	≥19	0.440	0.58

Abbreviations: F, female; HRQoL, health-related quality of life; M, male; QOLHEQ, Quality Of Life in Hand Eczema Questionnaire.
See Supplementary Tables S2–S15 for all tested bands.

Table 3. MIC Indicated as Points Improvement on the QOLHEQ

MIC method	QOLHEQ overall	Symptoms	Emotions	Functioning	Treatment and Prevention[1]
Mean cut-off	12.6	3.0	2.9	4.5	0.5
ROC curve	21.5	3.5	3.5	5.5	2.5
95% limit	22.5	7.6	9.6	9.3	6.5

Abbreviations: MIC, minimally important change; ROC, receiver operating characteristic; QOLHEQ, Quality Of Life in Hand Eczema Questionnaire.
MIC values are presented for the overall QOLHEQ and its subscales.
[1]Anchor question of change correlated <0.30 with the Treatment and Prevention subscale change score.

from these countries, we advise the use of the values presented in this study for international comparison.

A strength of this study is that the determination of important in the MIC was given solely by patients and not by physicians. By indicating whether patients had changed and subsequently indicating whether this change was important to them, the MIC is considered from the perspective of the patient. For patient-reported outcomes, this is highly recommended (de Vet et al., 2011).

A limitation of this study is that only one anchor question was used for each domain, whereas it is now recommended to use multiple anchors (Prinsen et al., 2018). This interpretability study was part of a larger validation study. The use of only one anchor was chosen to minimize the burden on patients who had to complete many questionnaires. Another limitation is that some correlations between the anchor questions and scores were only modest, similar to what was found for other patient-reported measures in dermatology (Charman et al., 2013; Hongbo et al., 2005; Vakharia et al., 2018). This most likely means that there are aspects to the anchor questions that are not fully covered by the items in the QOLHEQ. A different choice of anchor or a different wording of the anchor question might yield different results. However, for all but one subscale (the Treatment and Prevention change score), the correlations were within recommended values to perform a proper interpretability study. Therefore, clinicians can be fairly confident that the bands accurately reflect the degree of impact experienced by the patient. A third limitation is that patients with very severe hand eczema are underrepresented in this study (7.5%). However, this adequately reflects our clinical experience, as we encounter very severe hand eczema less often than milder severities. A final limitation is that answers to the anchor questions were only provided by Dutch patients. In the cross-cultural validation study, it was already assessed whether patients with the same level of HRQoL impairment respond similarly to the items from the QOLHEQ. However, the same might not necessarily be true for the anchor questions. This needs to be taken into consideration when using the values presented in this study, as it may influence generalizability.

In conclusion, we suggest using the QOLHEQ as measurement instrument for patient-reported outcomes in studies involving patients with hand eczema, as it is extensively validated and now tested for interpretability. In this regard, the QOLHEQ might also be considered for incorporation into a core outcome set for hand eczema, which is currently under development (Rönsch et al., 2019).

MATERIALS AND METHODS

Study population and design

This prospective study was performed at the department of Dermatology in the University Medical Center Groningen, a tertiary referral center for hand eczema. The study was part of a longitudinal validation study of the Dutch version of the QOLHEQ, for which we previously published a guideline (Oosterhaven et al., 2017). The design is in accordance with guidelines by the consensus-based standards for the selection of health measurement instruments (COSMIN) group (Mokkink et al., 2010; de Vet et al., 2011). In short, we included adult patients ($\geq$18 years) with hand eczema of all clinical severities, lasting at least one week, as diagnosed by a dermatologist according to current guidelines (Diepgen et al., 2015; Menné et al., 2011). Patients completed the QOLHEQ and multiple anchor questions at three time points: at baseline (T_0); after 1—3 days (T_1) to identify as many unchanged patients as possible; and after 4—12 weeks (T_2) to obtain a sample of deteriorated, unchanged, and improved patients. In the time between T_0 and T_2, patients were allowed to use any form of treatment for their hand eczema. See Supplementary Figure S1 for a study flow chart. Recruitment was done between March 2017 and December 2018. The Medical Ethical Review Board of the University Medical Center Groningen confirmed that this study did not fall under the scope of the Medical Research Involving Human Subjects Act (reference: METc 2014/391).

Studied instrument

The QOLHEQ is an instrument with 30 questions covering four subscales (Ofenloch et al., 2014). It measures overall HRQoL impairment and for each domain (subscale) separately: Symptoms (7 questions); Emotions (8 questions); Functioning (8 questions); and Treatment and Prevention (7 questions). Response categories are never, rarely, sometimes, often, and always. These are initially scored as 0, 1, 2, 3, and 4, respectively. In the cross-cultural validation study, a rescoring of certain items was defined using methods within the framework of item response theory. Also, question number 18 from the Treatment and Prevention subscale, concerning visiting physicians, was removed from the scoring (Ofenloch et al., 2017). See Supplementary Table S1 for the rescoring, which was also used in this study.

Anchors

The following single-score anchor questions were used at T_0 to assess the degree of HRQoL impairment perceived by patients (Oosterhaven et al., 2017). In a cognitive interview pilot, these anchor questions were tested for content validity in patients with hand eczema with varying disease severity. Answer categories were equal for all questions: not at all, slightly, moderately, strongly, and very strongly.

- Global (overall): How did your hand eczema bother you in your overall health state in the past seven days?

- Symptoms subscale: How did the symptoms of your hand eczema (like pain, itch, fissuring, redness) bother you in the past seven days?
- Emotions subscale: How strong did your hand eczema affect your emotional well-being (e.g. making you angry, frustrated, or anxious about the future) in the past seven days?
- Functioning subscale: How strong did your hand eczema affect your functioning (e.g. performing your (home)work or doing hobbies) in the past seven days?
- Treatment and Prevention subscale: How did treatment and pre-vention of your hand eczema bother you in the past seven days?

At T_1 and T_2, anchor questions for change in impairment were asked. Answer categories were much improvement, moderate improvement, minor improvement, no change, minor deterioration, moderate deterioration, and much deterioration.

- Global (overall): Overall, has there been any change in how your hand eczema bothers you since the last time you completed the QOLHEQ?
- Symptoms subscale: Has there been any change in how the symptoms of your hand eczema (like pain, itch, fissuring, redness) bother you since the last time you completed the QOLHEQ?
- Emotions subscale: Has there been any change in how strongly your hand eczema affects your emotional well-being (e.g., making you angry, frustrated, or anxious about the future) since the last time you completed the QOLHEQ?
- Functioning subscale: Has there been any change in how strong your hand eczema affects your functioning (e.g., performing your (home)work or doing hobbies) since the last time you completed the QOLHEQ?
- Treatment and Prevention subscale: Has there been any change in how the treatment and prevention of your hand eczema bother you since the last time you completed the QOLHEQ?

For each of the change anchors, a follow-up question was asked to determine the importance of a change:

- If you indicated a change (improvement or deterioration), was this change *important* for you? (Answer categories: no, yes).

Interpretability assessments

Single scores. We used an anchor-based method to define severity bands (stratified scores indicating thresholds of severity cat-egories) for the QOLHEQ with scores obtained at T_0, using the single-score anchor questions. Several numerical cut-offs of the QOLHEQ score and subscale scores were tested against the impairment indi-cated on the anchor questions using a weighted kappa (κ) coefficient of agreement to determine level of agreement between these. Nu-merical cut-off points were considered based on QOLHEQ scores that corresponded to a one-step increase in mean, median, and/or mode on the anchor (see Supplementary Materials and Methods for addi-tional information). Sensitivity analysis consisted of tests for differ-ences in sex and age distribution between patients whose severity could be predicted based on the final chosen band and those whose scores disagreed with the predicted severity according to that band (Charman et al., 2013; Chopra et al., 2017; Hongbo et al., 2005). In order not to underestimate the burden for patients when using the banding, we investigated the bands with the highest κ-values and those within a distance of 0.01. The final band chosen was that for which the number of patients reporting a higher impairment accord-ing to the anchor question compared with the band was lowest.

Change scores. SDC is defined as "the smallest change in score that can be detected by the instrument, beyond measurement error" (de Vet et al., 2011). The SDC for the QOLHEQ and subscales was determined in unchanged patients at T_1, as identified using the Global change an-chor. For this, $SEM_{agreement}$ was obtained using the square root of the within-subject total variance of an analysis of variance. The SDC was then calculated by using the following formula (de Vet et al., 2006):

$$SDC = 1.96 \times \sqrt{2} \times SEM_{agreement}.$$

MIC is defined as "the smallest change in the construct to be measured which patients perceive as important" (de Vet et al., 2011). The MIC for the QOLHEQ score and subscales was determined in changed patients at T_2, as identified using the anchor questions for change. The anchor questions were deemed appropriate to use as anchor for determination of the MIC if their correlation with the QOLHEQ score was at least >0.30, but preferably >0.50 (Cella et al., 2002; Revicki et al., 2008). Change scores were calculated for the QOLHEQ and anchor questions by sub-tracting the score at T_0 from the score at T_2. Thus, negative scores correspond to deterioration and positive scores to improvement in HRQoL. Patients were stratified according to their degree of change, taking into account the indication of their change as important/not important, as mentioned previously. Three MIC values were determined for the QOLHEQ overall and subscale scores:

- The MIC based on the mean change in QOLHEQ value that cor-responds with a one-step important change on the anchor ques-tions for change;
- The MIC of the receiver operating characteristic cut-off point, indicating the point closest to the upper left corner, where the sum of percentages of correctly classified patients is highest; and
- The MIC based on the 95% upper limit cut-off point of the un-changed (or not importantly changed) patients, which corresponds to $mean_{change} + 1.645 \times$ standard $deviation_{change}$ of this group (or strictly to the $mean_{difference}$ and standard $deviation_{difference}$ because this concerns theoretically unchanged patients).

These values for the MIC of the QOLHEQ change score were graphically presented using the anchor-based MIC distribution method (de Vet et al., 2007). We only determined the MIC for improved patients. Too few patients ($n < 50$) deteriorated compared with baseline to draw sound conclusions about the MIC for deterioration.

Floor and ceiling effects. It was determined what proportion of patients achieved the highest and lowest possible scores on the QOLHEQ. Floor and ceiling effects were considered to be present if the lowest or highest QOLHEQ score was achieved by >15% of patients (McHorney and Tarlov, 1995).

Data analysis

No sample size calculation was performed. A general recommendation for interpretability studies is to use a minimum of 50, but preferably at least 100, patients with a minimum of 50 patients in the smallest sub-group to calculate the MIC using the receiver operating characteristic method (de Vet et al., 2011). This study meets these recommendations. Spearman's rho (r) was used to calculate correlation. Chi-squared test and Student t-test were used to calculate differences between groups. In eight cases, the QOLHEQ was missing one item. For these, the value 0 was imputed (Ofenloch et al., 2014). At T_1, four cases had skipped a whole page, containing 10 QOLHEQ items. These four cases were excluded

from analysis. Analyses were performed with IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp. Armonk, NY), and GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla, CA, www.graphpad.com).

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

ORCIDs
Jart A. F. Oosterhaven: https://orcid.org/0000-0001-7214-7528
Robert F. Ofenloch: https://orcid.org/0000-0002-3532-6110
Marie L. A. Schuttelaar: https://orcid.org/0000-0002-0766-4382

REFERENCES

Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. Qual Life Res 2002;11:207—21.

Charman CR, Venn AJ, Ravenscroft JC, Williams HC. Translating Patient-Oriented Eczema Measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. Br J Dermatol 2013;169:1326—32.

Chopra R, Vakharia PP, Sacotte R, Patel N, Immaneni S, White T, et al. Severity strata for Eczema Area and Severity Index (EASI), modified EASI, Scoring Atopic Dermatitis (SCORAD), objective SCORAD, Atopic Dermatitis Severity Index and body surface area in adolescents and adults with atopic dermatitis. Br J Dermatol 2017;177:1316—21.

Coenraads PJ, Van Der Walle H, Thestrup-Pedersen K, Ruzicka T, Dreno B, De La Loge C, et al. Construction and validation of a photographic guide for assessing severity of chronic hand dermatitis. Br J Dermatol 2005;152: 296—301.

de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. Qual Life Res 2007;16:131—42.

de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59:1033—9.

de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. 1st ed. New York: Cambridge University Press; 2011.

Diepgen TL, Andersen KE, Chosidow O, Coenraads PJ, Elsner P, English J, et al. Guidelines for diagnosis, prevention and treatment of hand eczema. J Dtsch Dermatol Ges 2015;13:e1—22.

Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: what do dermatology life quality index scores mean? J Invest Dermatol 2005;125:659—64.

McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4: 293—307.

Menné T, Johansen JD, Sommerlund M, Veien NK. Danish Contact Dermatitis Group. Hand eczema guidelines based on the Danish guidelines for the diagnosis and treatment of hand eczema. Contact Dermatitis 2011;65:3—12.

Minamoto K, Diepgen TL, Sato K, Noguchi H, Yamashita N, Yoshimura K, et al. Quality Of Life in Hand Eczema Questionnaire: validation of the Japanese version of a disease-specific measure of quality of life for hand eczema patients. J Dermatol 2018;45:1301—5.

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737—45.

Nijsten T. Dermatology Life Quality Index: time to move forward. J Invest Dermatol 2012;132:11—3.

Ofenloch RF, Weisshaar E, Dumke AK, Molin S, Diepgen TL, Apfelbacher C. 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.

Ofenloch RF, Oosterhaven JAF, Susitaival P, Svensson Å, Weisshaar E, Minamoto K, et al. Cross-cultural validation of the Quality Of Life in Hand Eczema Questionnaire (QOLHEQ). J Invest Dermatol 2017;137:1454—60.

Oosterhaven JAF, Ofenloch RF, Schuttelaar MLA. Validation of the Dutch Quality Of Life in Hand Eczema Questionnaire (QOLHEQ) [Epub ahead of print]. Br J Dermatol 2019. https://dx.doi.org/10.1111/bjd.18558. Accessed 30 October 2019.

Oosterhaven JAF, Schuttelaar MLA, Apfelbacher C, Diepgen TL, Ofenloch RF. Guideline for translation and national validation of the Quality Of Life in Hand Eczema Questionnaire (QOLHEQ). Contact Dermatitis 2017;77: 106—15.

Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147—57.

Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102—9.

Rönsch H, Apfelbacher C, Brans R, Matterne U, Molin S, Ofenloch R, et al. Which outcomes have been measured in hand eczema trials? A systematic review. Contact Dermatitis 2019;80:201—7.

Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol 2010;63:524—34.

Vakharia PP, Chopra R, Sacotte R, Patel N, Immaneni S, White T, et al. Severity strata for five patient-reported outcomes in adults with atopic dermatitis. Br J Dermatol 2018;178:925—30.