Eczema Trials: Quality of Life Instruments Used and Their Relation to Patient-Reported Outcomes. A Systematic Review

Daniel HEINL1, Joanne CHALMERS2, Helen NANKERVIS3 and Christian J. APFELBACHER1
1Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany, and 2Centre for Evidence Based Dermatology, University of Nottingham, Nottingham, United Kingdom

It is unclear which quality of life instruments have thus far been used in eczema trials. Therefore, we aimed to identify these instruments. We searched the Global Resource of Eczema Trials (GREAT) database for reports of randomized controlled trials. Information on patient-reported outcomes, particularly quality of life, was extracted from eligible studies. Two-hundred and eighty-seven full texts reporting on 303 trials and 72 abstracts were included. Of the 303 studies, 63 (20.8%) assessed quality of life and used 18 named and 4 unnamed instruments. The Dermatology Life Quality Index (DLQI), the Children’s Dermatology Life Quality Index (CDLQI), the Infants’ Dermatitis Quality of Life Index (IDQOL), and the Dermatitis Family Impact (DFI) were the most common measures in adults, children, infants, and caregivers, respectively. In conclusion, only about one fifth of eczema trials include a quality of life measure as outcome. Many different instruments are used, limiting the possibilities of comparing and synthesising individual trials’ findings. Key words: eczema; atopic dermatitis; quality of life; patient-reported outcomes; HOME initiative.

Accepted Dec 15, 2015; Epub ahead of print Dec 17, 2015

Christian Apfelbacher MSc, PhD, Medical Sociology, Institute of Epidemiology and Preventive Medicine, University of Regensburg, Dr.-Gessler-Str. 17, DE-93051 Regensburg, Germany. E-mail: Christian.apfelbacher@klinik.uni-regensburg.de

Eczema (synonym: atopic dermatitis (AD)) is a common skin disease that affects both children and adults. It exerts a negative impact on the quality of life (QoL) of the patients and their families and places a considerable financial burden on patients and society (1, 2). The disease is characterized by a chronic or chronically relapsing course, with pruritus being the main symptom (3). The prevalence of eczema has increased over recent years (4). Despite a multitude of available treatment options, important uncertainties remain in the treatment of eczema requiring the conduct of high quality randomized controlled trials (RCTs) (5, 6). The use of non-standar-

dized and inadequately validated outcome measurement instruments (OMIs) in eczema trials hampers evidence-based decision making because treatment effects may be over- or underestimated. Furthermore, comparison and evidence synthesis is rendered difficult when outcome measurement is not standardized.

Therefore, the Harmonising Outcome Measures for Eczema (HOME) initiative set out to define a core outcome set (COS) that should be assessed in all eczema trials in the future (7). A COS is a consensus-derived minimum set of outcomes to be assessed in a specific situation (8). HOME agreed to consider clinical signs, symptoms, long term control and QoL as core outcome domains (9). For each of these domains an adequate OMI needs to be identified. For the signs domain, this process has been completed and the Eczema Area and Severity Index (EASI) has been identified as the currently most adequate measurement instrument to assess clinical signs in eczema (10). To standardize processes and to provide a standard for COS development in dermatology, the HOME initiative has published a roadmap (11). According to this roadmap, the first step for each core outcome domain is a comprehensive review of what OMIs have actually been used.

QoL, as one of these core outcome domains, is usually classified as a patient-reported outcome (PRO). A PRO is defined as any report coming directly from patients, without interpretation by physicians or others, about how they function or feel in relation to a health condition and its therapy (12). Where clinical trials use a PRO as primary endpoint, this is a reflection of the importance that study authors place on PROs. Thus, with QoL being a PRO, we hypothesized that authors who used PROs as primary endpoints in eczema trials would also be more likely to apply QoL questionnaires.

In accordance with the HOME roadmap (11), the primary aim of this systematic review was to identify the QoL measurement instruments used in eczema trials from the year 2000 onwards. We were also interested in whether there were any time trends in their usage. A secondary aim was to find out whether the consideration of a PRO as a primary endpoint in eczema trials was related to the inclusion of a QoL instrument as an outcome measure.
METHODS

Sample article selection
To obtain a comprehensive selection of eczema trials, we searched the Global Resource of Eczema Trials (GREAT) Database (13), which includes records of all RCTs of eczema treatments (14). An article was considered eligible if it was an eczema treatment trial published since 2000, was indexed in the GREAT database by 31 May 2014 and if a full text or an abstract was available in either English or German language. We did not consider any other sources of eczema trials besides the GREAT database.

Outcomes
The outcomes of interest were: i) the proportion of articles that assessed a QoL outcome, ii) the proportion of articles that assessed a PRO, iii) whether the inclusion of a QoL measure was related to whether the primary endpoint was a PRO, iv) what QoL instruments were used, v) the number of QoL instruments per study and vi) the number of studies published and the proportion including QoL instruments over time.

A patient-reported outcome (PRO) was defined according to Patrick et al. (12) as any report coming directly from patients, without interpretation by physicians or others, about how they function or feel in relation to a health condition and its therapy. Any outcome used in the included studies that fulfilled this definition was recorded as a PRO. We did not categorize PROs according to content or type. The term ‘composite index’ was used to describe any score or index that is composed of both a PRO and a non-PRO part.

Data extraction
Data extraction was carried out by D.H. To ensure consistency in the data extraction, guidelines on what information should be gathered and how this information should be evaluated was agreed on beforehand by D.H. and C.A. Where these guidelines were not applicable to certain studies, the whole team decided about how to evaluate the information from those studies. A second data extraction was performed for a random sample of 10% of the papers by J.C. as a measure of quality assurance. For this sample, results were compared between the first and the second data extraction and discrepancies were resolved within the whole team. Where a resolution of a discrepancy within this random sample meant that changes were necessary to the initially extracted information by D.H., these changes were also made in the data extractions of the rest of the studies where applicable. Where the study was reported only in an abstract, only data on QoL was extracted.

Data analysis
Statistical data analysis was split into a descriptive and an analytical part.

Descriptive analysis
We determined the absolute number and the percentage of articles assessing QoL. To get an overview of the most common QoL measures, we recorded which questionnaires were used by how many studies. Moreover, we determined the number of QoL instruments used per article (only regarding articles which assessed QoL) and explored changes in the usage of QoL measures over time, which we visualized in diagrams created with Microsoft Excel.

We also calculated the proportion of PROs (with/without composite indices) in relation to all outcomes, the absolute number and percentage of articles assessing PROs (with/without composite indices) in relation to all articles, the median number (and interquartile range (IQR)) of outcomes per article, the median number (and IQR) of PROs per article (with/without composite indices; only regarding articles which assess PROs) and the number of composite indices per article (only regarding articles which assess composite indices). Furthermore, we analysed changes in the total number of outcomes over time and depicted our findings in a diagram.

Analytical analysis
We hypothesized that authors who are generally in sympathy with the integration of PROs in eczema trials would also be more likely to apply QoL questionnaires. Therefore, we computed the absolute and relative frequency of articles assessing QoL in articles with PROs as primary endpoint, in articles with composite indices as primary endpoint, in articles which did not specify their primary endpoint and in articles with non-PROs as primary endpoint. A chi-square test was conducted to test our hypothesis about the connection between a study’s primary endpoint and QoL assessment. The results of this chi-square test were presented in a contingency table. Level of significance was set at 5%.

For all analyses, IBM SPSS 22.0 was used. Data was extracted and figures were designed with Microsoft Excel 2013. We used EndNote X6 to manage references.

RESULTS
Our search yielded 378 papers that were published since 2000 and indexed in the GREAT database by no later than 31 May 2014. References to these papers can be found in Appendix S1. Nineteen articles were not eligible. The reasons for exclusion were: paper was protocol only (n = 9), no English or German abstract or full text was available (n = 5), paper reported on a study already included (n = 4), paper was conference publication and not available as abstract or full text (n = 1). Of the 359 eligible articles, we were able to obtain the full text for 287 papers, and an abstract only for the remaining 72 articles. The 287 full text papers reported on 303 studies. The distribution of publications over time is shown in Fig. 1. As can be seen from Fig. 1, 2011 saw the highest number of trials. Despite some minor differences, similar numbers of studies were found eligible for every publication year.

Descriptive analysis
Overall, only 63 (20.8%) studies assessed QoL. The QoL instruments that were applied in these studies are listed in Table I for adults and children and families, respectively. In adults, the Dermatology Life Quality Index (DLQI) (15) was the most frequently used self-reported QoL measure; in children, the Children’s Dermatology Life Quality Index (CDLQI) (16) was the most popular self-reported questionnaire and the

1http://www.medicaljournals.se/acta/content/?doi=10.2340/00015555-2322

Acta Derm Venereol 96
Infants’ Dermatitis Quality of Life Index (IDQOL) (17) was the proxy-reported instrument most often used. QoL of carers of children was predominantly assessed with the Dermatitis Family Impact (DFI) questionnaire (18). Altogether, 18 named and 4 unnamed QoL questionnaires were used; of these, 4 were infant- or children-specific measures, 4 assessed the QoL of carers and 16 instruments were applicable to adult patients with eczema. Six instruments, 4 of which are validated, were eczema-specific.

Of the 63 trials that assessed QoL, we found that the majority of studies (n = 41, 65.1%) used only one QoL measurement instrument. Two QoL instruments were applied in 16 studies (25.4%) and the remaining 6 studies (9.5%) included 3 QoL measurement instruments. Analysis over time showed that although there were fluctuations from year to year, the proportion of trials that include QoL measures has remained largely static since 2000 (Fig. 2). For instance, none of the 22 studies that were published in 2005 included a QoL instrument whereas studies from 2006 with an inclusion rate of QoL measurement instruments of 31% are even above average. The highest percentage of studies assessing QoL (50%) was observed in 2014; however, this finding needs to be put into context as only 4 studies from 2014 were included in total.

Similarly, we could not observe any clear trends towards increased or reduced usage of the most frequently applied specific QoL instruments (Fig. S11).

In most years, less than 10% of the included full texts

![Fig. 1. Distribution of included full text studies over time.](image)

**Table I. Quality of life instruments used in adults and children and families/carers**

<table>
<thead>
<tr>
<th>Instrument, Ref.</th>
<th>Studies n (%), Ref.</th>
<th>Type</th>
<th>Full name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adults (n = 36)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLQI (15)</td>
<td>20 (56) (19–38)</td>
<td>Dermatology-specific</td>
<td>Dermatology Life Quality Index</td>
</tr>
<tr>
<td>Skindex-29 (39)</td>
<td>2 (6) (40, 41)</td>
<td>Dermatology-specific</td>
<td></td>
</tr>
<tr>
<td>EDLQ (42)</td>
<td>1 (3) (43)</td>
<td>Generic</td>
<td>Everyday Life Questionnaire/Alltagsleben</td>
</tr>
<tr>
<td>EQ-5D (44)</td>
<td>1 (3) (45)</td>
<td>Generic</td>
<td>EuroQol-5D</td>
</tr>
<tr>
<td>SF-36 (46)</td>
<td>1 (3) (47)</td>
<td>Generic</td>
<td>Short form 36</td>
</tr>
<tr>
<td>SIP (48)</td>
<td>1 (3) (49)</td>
<td>Generic</td>
<td>Sickness Impact Profile</td>
</tr>
<tr>
<td>WTP (50)</td>
<td>1 (3) (49)</td>
<td>Generic</td>
<td>Willingness To Pay</td>
</tr>
<tr>
<td>DIEHL (51)</td>
<td>1 (3) (52)</td>
<td>Dermatology-specific</td>
<td>Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen</td>
</tr>
<tr>
<td>DLQI (modified)* (15)</td>
<td>1 (3) (53)</td>
<td>Dermatology-specific</td>
<td>Dermatology Life Quality Index</td>
</tr>
<tr>
<td>FLQA-d (54)</td>
<td>1 (3) (49)</td>
<td>Dermatology-specific</td>
<td>Freiburg Life Quality Assessment for Dermatoses</td>
</tr>
<tr>
<td>ISDL (modified)* (55)</td>
<td>1 (3) (36)</td>
<td>Dermatology-specific</td>
<td>Impact of Chronic Skin Disease on Daily Life</td>
</tr>
<tr>
<td>Self-provided*</td>
<td>1 (3) (56)</td>
<td>Dermatology-specific</td>
<td>Unnamed</td>
</tr>
<tr>
<td>Skindex-16 (57)</td>
<td>1 (3) (58)</td>
<td>Dermatology-specific</td>
<td></td>
</tr>
<tr>
<td>Skindex-17 (59)</td>
<td>1 (3) (60)</td>
<td>Dermatology-specific</td>
<td></td>
</tr>
<tr>
<td>EDI* (61)</td>
<td>1 (3) (62)</td>
<td>Eczema-specific</td>
<td>Eczema Disability Index</td>
</tr>
<tr>
<td>QoLIAD (63)</td>
<td>1 (3) (26)</td>
<td>Eczema-specific</td>
<td>Quality of Life Index for Atopic Dermatitis</td>
</tr>
<tr>
<td><strong>Children and families/carers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Proxy-reported instruments</strong> (n = 15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDQOL (17)</td>
<td>14 (93) (36, 64–76)</td>
<td>Eczema-specific</td>
<td>Infants’ Dermatitis Quality of Life Index</td>
</tr>
<tr>
<td>Self-provided*</td>
<td>1 (7) (77)</td>
<td>Eczema-specific</td>
<td>Unnamed</td>
</tr>
<tr>
<td><strong>Self-reported instruments</strong> (n = 20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDLQI (16)</td>
<td>18 (90) (28, 65, 69, 72–75, 78–88)</td>
<td>Dermatology-specific</td>
<td>Children’s Dermatology Life Quality Index</td>
</tr>
<tr>
<td>CDLQI (modified)* (16)</td>
<td>2 (10) (89, 90)</td>
<td>Dermatology-specific</td>
<td>Children’s Dermatology Life Quality Index</td>
</tr>
<tr>
<td><strong>Instruments assessing the quality of life of adult carers of children with eczema</strong> (n = 20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFI (18)</td>
<td>14 (70) (45, 64–67, 69–73, 81, 86, 87, 91)</td>
<td>Dermatology-specific</td>
<td>Dermatitis Family Impact</td>
</tr>
<tr>
<td>Questionnaire by Rüden et al. (92)</td>
<td>3 (15) (43, 93, 94)</td>
<td>Eczema-specific</td>
<td>Unnamed</td>
</tr>
<tr>
<td>PIFIqoL-AD (95)</td>
<td>2 (10) (88, 96)</td>
<td>Eczema-specific</td>
<td>Parents’ Index of Quality of Life in Atopic Dermatitis</td>
</tr>
<tr>
<td>Self-provided*</td>
<td>1 (5) (83)</td>
<td>Unknown</td>
<td>Unnamed</td>
</tr>
</tbody>
</table>

*Instruments marked with an asterisk have not been validated at all.

Proxy-reported means that the (primary) caregiver of an infant fills in a questionnaire that assesses the quality of life of the infant. Proxy-reported instruments are often used in infants and younger children because they cannot report on their quality of life themselves due to their inability to read and a lack of understanding. *Self-reported instruments are used in older children. These questionnaires are filled in by the children themselves, not by their caregiver.
applied the DLQI (15), the IDQOL (17), the CDLQI (16) or the DFI (18). In 2014, 25% of the included full texts used the 4 instruments depicted in Fig. S1; however, this result should not be interpreted as a recent rise in usage of these measures since only 4 full texts from 2014 were included in our systematic review.

For the studies which were reported in abstract form only, only 4 out of 72 (6%) assessed QoL. The CDLQI (16) was used in 2 abstracts and the DFI (18) in 1 abstract. Three further QoL instruments were reported, but were not named.

The assessment of the full text articles revealed that a total of 2,633 outcomes were assessed of which 809 (30.7%) were PROs or composite indices (i.e. a scale that is composed of both a PRO and a non-PRO part). Of these, 633 (24.0%) were PROs and 176 (6.7%) were composite indices. The majority of studies (281, 92.7%) included at least one PRO and/or a composite index as any endpoint (primary, secondary, other). A total of 230 (85.9%) studies included at least one PRO, whereas composite indices were assessed in just over half of studies (164, 54.1%).

The median number of outcomes per study was 7 (IQR: 5–11). In studies that assessed PROs and/or composite indices, the median number of PROs was 2 (IQR 1–4). The same values were found when looking at PROs without composite indices. For studies using composite indices, the vast majority of 153 (93.3%) studies included only one composite index, 10 studies (6.1%) two composite indices and only a single study applied 3 composite indices. Analysis over time showed that the median number of outcomes per study has plateaued since 2000. The highest median number of outcomes per study was 9.5 in 2006 whereas the lowest number was 5 in 2013. A median of 6 outcomes was found for 4 years; the same is true for a median of 7 outcomes. In 3 years, the median number of outcomes per study amounted to 8.

### Analytical analysis

We were able to categorise the endpoints for 302 studies: 32 studies (10.6%) chose a PRO as primary endpoint, 58 trials (19.2%) a composite index, 81 studies (26.8%) had a non-PRO as primary endpoint and 131 studies (43.4%) did not specify their primary endpoint. The endpoint for one study could not be categorised.

There was a statistically significant association between the type of primary endpoint (PRO, composite index, non-PRO, not specified) and the assessment of QoL (yes/no) \( (p = 0.002, \text{Table II}) \). Studies with a PRO as primary endpoint were more likely to measure QoL, followed by studies with a composite index as primary endpoint. Likewise, study authors that used a distinct non-PRO as primary endpoint were more likely to include QoL measurement than study authors that did not specify their primary endpoint in more detail.

### DISCUSSION

QoL is considered particularly relevant for chronic skin diseases such as eczema. Inclusion of QoL instruments in RCTs is of great importance, given the fact that the patient’s perspective on the efficacy of a certain treatment often deviates from clinicians’ assessments (97). Our study demonstrated that the majority of studies (approximately 90%) include at least one PRO. This is in contrast to previous findings that only about 25% of dermatology trials included a participant efficacy outcome (98). However, we did not apply any limitations concerning the type of PRO which may explain these differences. Despite the fact that most studies included a PRO, it was clear from this review that the majority of outcomes (approximately three-quarters) reported are non-PROs.

Even though the QoL of patients and their family is greatly impacted by eczema, respective outcome measures often seem to only play a minor part in eczema trials (99). One study, however, noted a substantial increase in the usage of QoL instruments in eczema trials from 1985 to 2010 (100). We were not able to verify this trend since our findings suggest that the inclusion of QoL measurement instruments has changed very little over time. A reason for this result may be that we looked at the relative frequency of studies assessing QoL instead of absolute numbers, taking into account...
the larger quantity of trials in our observation period, compared to the observation period investigated by Rehal & Armstrong (100). Different time intervals in that study and our review may present a further explanation of this discrepancy since the idea of QoL in dermatology emerged in the early 1990s and the development of the first QoL instruments for dermatological conditions falls also in this time period. Consequently, a broader inclusion of QoL measurement instruments in trials did not start until the late 1990s. Rehal & Armstrong (100) reported that 14 different QoL instruments were used in eczema trials from 1985 to 2010. In contrast, we found that from 2000 to 2014, study authors applied 22 different instruments, suggesting a growing number of existing QoL OMIs. Nonetheless, findings on the most frequently applied QoL instruments were similar in both reviews.

The QoL instruments that were mostly used in clinical trials, i.e. the DLQI (15), the IDQOL (17), the CDLQI (16) and the DFI (18), all have been developed at one academic medical centre (see http://www.cardiff.ac.uk/dermatology/quality-of-life/). Reasons for the widespread use of these instruments may be that they are available in many language versions and that they are easy to use. All 4 instruments fit on one A4 page whereas other questionnaires are often longer. A critical review recommends the Skindex-29 (39) rather than the DLQI as dermatology-specific QoL measure (101), but we identified only 2 trials in which it was actually used.

With respect to the lack of “hard” outcomes such as mortality in eczema, QoL measures could fill this gap and provide the necessary evidence to judge the effectiveness and appropriateness of interventions from the patients’ perspective. Against this backdrop, it is surprising that only 1 out of 5 eczema trials include QoL instruments, particularly since similar figures are obtained for different diseases in other fields of medicine. For example, one study found that 16% of drug clinical trials published in 2005 in 5 high quality journals included QoL measures (102). However, most of these studies reported on heart disease, cancer or other serious illnesses where “hard” outcomes are available. Authors of future eczema trials should therefore consider the inclusion of a QoL measurement in their trials.

We could show that authors who chose PROs as primary endpoints in their trials were also more likely to include QoL measures than researchers that decided to use any other endpoints. Surprisingly, this observation holds also for composite indices: A significantly higher proportion of studies with a composite index as primary endpoint assessed QoL than did studies with a non-PRO or a not specified endpoint. This implies that the measurement of QoL is not so much dependent on individual characteristics of a trial when opting for or against the inclusion of QoL instruments; instead, the general attitude of study authors towards PROs appears to determine whether or not QoL is measured as well. In addition, this finding may also explain why there was no increase in QoL measurement over time despite ongoing efforts to promote the use of QoL instruments. Moreover, there seems to be an association between how well researchers report on their study results and the measurement of QoL since authors that did not specify their primary endpoint were least likely to use QoL instruments. However, further research is warranted to find out why QoL measurement instruments are not included in more trials.

Some further attention must also be drawn to the high proportion of studies that did not specify their primary endpoint (43.4%). In contrast, Nassar et al. (103) found that only 20% of the RCTs on non-neoplastic skin diseases that were published in 2009 did not state their primary outcome. However, they restricted their search to journals with an impact factor of at least 2. As they also showed that a clear definition of the primary endpoint was significantly associated with a higher journal impact factor, this result may present an explanation for our findings because we did not narrow down our eligibility criteria to high quality journals.

In conclusion, we could demonstrate that a high proportion of trials include some sort of PROs but that QoL was only assessed in about one fifth of all trials. Even though a range of QoL measurement instruments have been used in RCTs of eczema, most studies applied the DLQI (15) for adults, the CDLQI (16) for older children, the IDQOL (17) for infants or the DFI (18) for adult carers of children with eczema.

We provide an up-to-date review on QoL OMIs used for eczema. A strength of this study was the use of the GREAT database, which searches 6 databases, including 3 specialist databases. A recent study showed a high sensitivity of the GREAT database, with 94% of trials cited in systematic reviews on eczema treatments listed in the GREAT database (104). The GREAT database therefore is considered a primary and comprehensive source to identify eczema RCTs. We did not consider any other study designs for inclusion in this systematic review.

Limitations of our study were the language restriction to English and German and our focus on the time interval from 2000 to 2014. In this way, QoL questionnaires in other languages may have been missed or underestimated and older QoL instruments may be underrepresented in our review. Also, we did not consider ongoing trials for this review. As a result, we cannot rule out the possibility that different findings would be obtained when regarding studies that are currently under way.

Results on the number of reported outcomes, the number of reported PROs, the number of reported composite indices, the proportion of validated indices used, the assessment of adverse events and additional safety assessments will be reported elsewhere.
The aim of this systematic review was not to critically appraise the measurement properties of the available QoL scales for eczema patients. Instead, this systematic review is intended to form the basis for further research on the appropriateness of the mentioned QoL instruments for eczema patients. As the use of so many different QoL instruments in eczema trials limits the possibility to synthesize their findings in meta-analyses and systematic reviews, the HOME initiative aims to define a COS including one distinct QoL instrument. A critical appraisal of the measurement properties of existing QoL instruments is the prerequisite for doing so and will be subject to a further systematic review. Our review is the first step to reach the goal of including a QoL instrument in the COS.

Conflict of interest: CA is a member of the HOME executive committee. DH and JC are members of the HOME initiative. The authors declare that they have no further conflicts of interest. The authors did not receive any financial funding to conduct this study.

REFERENCES

Immunol 2012; 130: 925–931 e924.
What quality of life instruments have been used in eczema trials?

603


