Allergy EUROPEAN JOURNAL OF ALLERGY AND CLINICAL IMMUNOLOGY EAACI

Measurement properties of adult quality-of-life measurement instruments for eczema: a systematic review

Journal:	*Allergy*
Manuscript ID	ALL-2015-00658
Wiley - Manuscript type:	Original Article: Epidemiology and Genetics
Date Submitted by the Author:	01-Oct-2015
Complete List of Authors:	Heinl, Daniel; Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg Prinsen, Cecilia; Department of Epidemiology and Biostatistics, EMGO+ Institute for Health and Care Research, VU University Medical Center Deckert, Stefanie; Centre for Evidence-based Healthcare, Medizinische Fakultät Carl Gustav Carus, TU Dresden Chalmers, Joanne; Centre of Evidence Based Dermatology, University of Nottingham Drucker, Aaron; Division of Dermatology, University Health Network Ofenloch, Robert; Department of Clinical Social Medicine, University Hospital Heidelberg Humphreys, Rosemary; Budleigh Salterton, Devon Sach, Tracey; Norwich Medical School, University of East Anglia Chamlin, Sarah; Ann and Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine Apfelbacher, Christian; Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg; Division of Public Health and Primary Care, Brighton and Sussex Medical School
Keywords :	quality-of-life, HOME initiative, measurement properties, core outcome set, atopic dermatitis

Measurement properties of adult quality-of-life measurement instruments for eczema: a systematic review

Daniel Heinl[1], Cecilia A.C. Prinsen[2] PhD, Stefanie Deckert[3] MPH, Joanne R. Chalmers[4] PhD, Aaron M. Drucker[5] MD, Robert Ofenloch[6] PhD, Rosemary Humphreys[7] MA, Tracey Sach[8] PhD, Sarah L. Chamlin[9] MD, Jochen Schmitt[3] MD, Christian Apfelbacher[1,10§] PhD

1 Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg, Germany
2 VU University Medical Center, Department of Epidemiology and Biostatistics, EMGO+ Institute for Health and Care Research, Amsterdam, The Netherlands
3 Centre for Evidence-based Healthcare, Medizinische Fakultät Carl Gustav Carus, TU Dresden, Germany
4 Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, United Kingdom
5 Division of Dermatology, University Health Network, Toronto, Canada
6 Department of Clinical Social Medicine, University Hospital Heidelberg, Heidelberg, Germany
7 Budleigh Salterton, Devon, United Kingdom
8 Norwich Medical School, University of East Anglia, Norwich, United Kingdom
9 Ann and Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, United States of America
10 Division of Public Health and Primary Care, Brighton and Sussex Medical School, Falmer, United Kingdom

Short title: Eczema quality of life measures: systematic review

Corresponding author:
Christian Apfelbacher
Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg, Germany
Dr.-Gessler-Str. 17
93051 Regensburg, Germany
Phone: +49 941 944 5231
Fax:
e-mail: christian.apfelbacher@klinik.uni-regensburg.de

Abstract

Background

The Harmonising Outcomes Measures for Eczema (HOME) initiative has identified QoL as a core outcome domain to be evaluated in every eczema trial. It is unclear which of the existing QoL instruments is most appropriate for this domain. Thus, the aim of this review was to systematically assess the measurement properties of existing measurement instruments developed and/or validated for the measurement of QoL in adult eczema.

Methods

We conducted a systematic literature search in PubMed and Embase identifying studies on measurement properties of adult eczema QoL instruments. For all eligible studies, we assessed the adequacy of the measurement properties and the methodological quality with the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist. A best evidence synthesis summarizing findings from different studies was the basis to assign four degrees of recommendation (A-D).

Results

15 articles reporting on 17 instruments were included. No instrument fulfilled the criteria for category A. Six instruments were placed in category B, meaning that they have the potential to be recommended depending on the results of further validation studies. Three instruments had poor adequacy in at least one required adequacy criterion and were therefore put in category C. The remaining eight instruments were minimally validated and were thus placed in category D.

Conclusions

Currently, no QoL instrument can be recommended for use in adult eczema. The Quality of Life Index for Atopic Dermatitis (QoLIAD) and the Dermatology Life Quality Index (DLQI) are recommended for further validation research.

Key words

Core outcome set; eczema; HOME initiative; measurement properties; quality of life

Abbreviations

COS: Core outcome set; COSMIN: COnsensus-based Standards for the selection of health Measurement Instruments; DIF: Differential item functioning; HOME: Harmonising Outcome Measures in Eczema; HrQoL: health-related quality of life; ICC: Intraclass Correlation Coefficient; IRT: Item response theory; MIC: Minimal important change; MID: Minimal important difference; OMERACT: Outcome Measures in Rheumatology; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; QoL: quality of life

Word count body text: 3,653
Manuscript table count: 7
Manuscript figure count: 1

Introduction

Eczema (synonymous with atopic eczema, atopic dermatitis) is a common, chronic, relapsing skin disease that affects both children and adults. Recent studies suggest that eczema prevalence rates in adults are in excess of 10% (1, 2). There are numerous treatments for eczema, many of which have been studied in randomized controlled trials. However, the lack of standardization of eczema outcome measurement instruments in clinical trials currently limits the possibility to compare and synthesize results in order to determine the best treatments, hampering evidence-based decision making and rendering the generation of treatment recommendations difficult.

Therefore, the Harmonising Outcome Measures in Eczema (HOME) initiative (www.homeforeczema.org) set out to define a core outcome set (COS) to be applied in all future eczema trials. A COS is an agreed minimum set of outcomes that should be measured and reported in all clinical trials of a specific disease or trial population (3). Clinical signs, symptoms, long-term control of flares and quality of life have been identified as the core outcome domains by the HOME initiative (4-6).

In accordance with the HOME roadmap (7), we set out to perform a systematic review of the measurement properties of all instruments that were developed and validated to measure QoL in eczema patients.

Methods

Protocol and registration

This systematic review was developed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (8). A study protocol was published beforehand (9) and has also been registered in the International Prospective Register of Systematic Reviews (PROSPERO): CRD42015017138.

Literature search

On 9 January 2015, we performed a systematic literature search in PubMed and EMBASE, as described in detail in the previously published protocol (9).

It was supplemented by hand searching of reference lists of included studies and key articles on this topic. We also searched the PROQOLID database (http://www.proqolid.org).

Eligible studies

The eligibility criteria laid out in the protocol were applied (9). In accordance with a consensus-based decision of the HOME initiative (10), only disease- or dermatology-specific, and not generic QoL measurement instruments, were eligible.

Assessment of the methodological quality of included studies

The COnsensus-based standards for the selection of health Measurement Instruments (COSMIN) checklist was used to evaluate the methodological quality of the included studies (11-14).

Assessment of measurement properties and further characteristics of QoL instruments

We assessed all measurement properties from the COSMIN checklist in this review, with the exception of criterion validity since no gold standard exists for QoL. Interpretability and feasibility data were collected where available. With the exception of content comparison and instrument characteristics, we regarded different language versions of the same questionnaire separately because we consider these to be distinct instruments. Our main reason for this approach was the fact that it cannot be assumed that different language versions of measurement instruments show the same measurement properties. Strictly speaking, it is the measurements that are valid, reliable and responsive and not the instruments *per se.*

Content comparison

We compared the content of each instrument at content domain level. In QoL questionnaires, subsets of items belonging together based on their content are often referred to as content domains. The original development paper for each instrument was consulted to obtain this information. We largely adopted the domains mentioned therein.

Adequacy of the measurement properties

The predefined criteria for rating the adequacy of measurement properties recommended by the COSMIN group were used in a slightly modified version (15) (Table 1). Hypothesis testing was split into the aspects convergent/divergent (defined as the correlation between instruments measuring similar/different constructs (16)) and discriminative validity (defined as the ability of a measurement instrument to distinguish between different subgroups of patients (16)) throughout the review. Findings from both aspects were integrated into an overall rating in the end (see also 'Differences between protocol and review'). Where studies applied item response theory (IRT) methods in the evaluation of measurement properties, rather than in the development of measurement instruments, we were able to evaluate the adequacy and methodological quality of internal consistency, construct validity, structural validity, and cross-cultural validity.

Table 1: Adequacy criteria for measurement properties adapted from (15) and (17)

Property	Rating	Adequacy criteria
Reliability		
Internal consistency (CTT methods applied)	+ ? -	Cronbach's alpha(s) ≥ 0.70 Cronbach's alpha not determined Cronbach's alpha(s) < 0.70
Internal consistency (IRT methods applied)	+ ? -	Person Separation Index ≥ 0.70 Person Separation Index not determined Person Separation Index < 0.70
Measurement error	+ ? -	MIC > SDC OR MIC outside the LoA MIC not defined MIC <= SDC OR MIC equals or inside LoA
Reliability	+ ? -	ICC/weighted Kappa >=0.70, OR Pearson's r >= 0.80 Neither ICC/weighted Kappa, nor Pearson's r determined ICC/weighted Kappa < 0.70 OR Pearson's r < 0.80
Validity		
Content validity	+ ? -	All items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement AND the questionnaire is considered to be comprehensive Not enough information available Not all items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Construct validity		
Structural validity (CTT methods applied)	+ ? -	Factors should explain at least 50% of the variance Explained variance not mentioned Factors explain < 50% of the variance
Structural validity (IRT methods applied)	+ ? -	Residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37, item scalability >0.30, IRT model fit: G2 >0.01, no DIF for important subject characteristics (such as age, gender, education): McFadden's R2 < 0.02, OR no non-uniform DIF Important statistics not reported Residual correlations among the items after controlling for the dominant factor ≥ 0.20 OR Q3's ≥ 0.37, item scalability ≤0.30, IRT model fit: G2 ≤0.01, important DIF for important subject characteristics (such as age, gender, education): McFadden's R2 ≥0.02, OR non-uniform DIF
Hypothesis testing (convergent/divergent validity)	+ ? -	Correlations with instruments measuring the same construct >=0.50 OR at least 75% of the results are in accordance with the hypotheses AND correlation with related constructs is higher than with unrelated constructs Solely correlations determined with unrelated constructs Correlations with instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance

		with the hypotheses OR correlation with related constructs is lower than with unrelated constructs
Hypothesis testing (discriminative validity)	+ ? -	Differences in scores on the measurement instrument for all evaluated patient subgroups are statistically significant OR ≥75% of results in accordance with hypotheses Some differences statistically significant, others not Differences in scores on the measurement instrument for all evaluated patient subgroups are not statistically significant OR <75% of results in accordance with hypotheses
Cross-cultural validity	+ ? -	No differences in factor structure OR no important DIF between language versions Multiple group factor analysis not applied AND DIF not assessed Differences in factor structure OR important DIF between language versions
Responsiveness		
Responsiveness	+ ? -	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70 AND correlations with changes in related constructs are higher than with unrelated constructs Solely correlations determined with unrelated constructs Correlations with changes on instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlations with changes in related constructs are lower than with unrelated constructs

Abbreviations: *DIF* = Differential item functioning; *ICC* = Intraclass correlation coefficient; *IRT* = Item response theory; LoA = Limits of agreement; *MIC* = Minimal important change; *SDC* = Smallest detectable change.

+ positive rating, ? indeterminate rating, - negative rating

Best evidence synthesis

Where an instrument was evaluated in multiple studies, the findings were synthesized provided the characteristics of the included studies were sufficiently similar and the methodological quality of the included studies was sufficient (18). The criteria for best evidence synthesis are outlined in Table 2.

Table 2: Levels of evidence for the overall adequacy of a measurement property, adapted from (19)

Level	Rating	Criteria
Strong	+++, ? (strong) or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++, ? (moderate) or --	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+, ? (limited) or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
Unknown	Weak	Only studies of poor methodological quality

+ positive rating, ? indeterminate rating, - negative rating

Generating recommendations for the use of QoL measurement instruments for eczema

For each reviewed instrument, a standardized recommendation for usage or required future validation work was made depending on the adequacy of the instrument and the methodological quality of the included studies.

Four categories of recommendation were made (9):

A. QoL measurement instrument meets all requirements and is recommended for use.
B. QoL measure meets two or more adequacy items, but performance in all other required adequacy items is unclear, so that the outcome measure has the potential to be recommended in the future depending on the results of further validation studies.
C. QoL measure has low adequacy in at least one required adequacy criterion (≥1 rating of 'minus') and therefore is not recommended to be used any more.
D. QoL measure has (almost) not been validated. Its performance in all or most relevant adequacy items is unclear so that it is not recommended to be used until further validation studies clarify its adequacy.

Finally, we aimed to identify one most appropriate (currently available) instrument to assess QoL in adults with eczema.

Differences between protocol and review

In this manuscript, we specified that generic instruments are not eligible for our review. Unlike what was planned in the original protocol (9), we did not perform a content comparison at item level because the resulting comparison table would have been too large and thus not informative. Instead, we compared the content of the different QoL instruments at content domain level.

For reasons of clarity, we decided to use the term "adequacy of the measurement properties" instead of "quality of the measurement properties". For studies applying IRT methods, only internal consistency, construct validity, structural, and cross-cultural validity were evaluated, if applicable. In addition, as the review was conducted it was clear that some minor alterations were required to the adequacy criteria presented in table 3 of the protocol and table 1 of this review, respectively:

- For internal consistency, the indeterminate rating ('?') was changed from "Dimensionality not known OR Cronbach's alpha not determined" to "Cronbach's alpha not determined" in order to avoid redundancy between the adequacy criteria and the COSMIN criteria for methodological quality. Adequacy criteria for IRT methods were added.
- Although the adequacy criteria for content validity refer to a questionnaire's target population (which may be other than eczema), we applied the same inclusion criteria for content validity studies like for the other measurement properties, i.e. at least 50% eczema patients in the sample or subgroup analysis for eczema patients presented, because we were interested in the instruments' content validity in eczema patients.
- The IRT criteria for structural validity were amended with information on differential item functioning (DIF) (20). A positive rating can now also be obtained if a study shows that there is no non-uniform DIF. Occurrence of non-uniform DIF results in a negative rating according to the new criteria.
- The criteria suggested by Terwee et al. for hypothesis testing were only applied to convergent and divergent validity. Self-developed criteria for discriminative validity, which is another aspect of hypothesis testing, were added. The adequacy criteria for interpretability were omitted since interpretability is not considered to be a formal measurement property by the COSMIN initiative (12).

The best evidence synthesis ratings were complemented by an indeterminate rating for strong, moderate and limited levels of evidence each. This was done for scenarios where a QoL instrument would obtain an indeterminate rating for a certain measurement property. An indeterminate rating was assigned where no clear evidence was available for either a positive or negative rating.

In order to obtain an overall rating for hypothesis testing, findings from best evidence synthesis for convergent/divergent and discriminative validity were synthesized according to the following criteria: in case of conflicting ratings, the worse rating determined the overall rating for hypothesis testing; if either convergent/divergent or discriminative validity obtained an indeterminate rating, the rating for the other aspect of hypothesis testing determined the overall rating for hypothesis testing.

Results

In total, we found 16 eligible articles (21-36) (Figure 1). Of these, we were able to obtain 15 full text papers. One manuscript pertaining to the Ukrainian versions of the Dermatology Life Quality Index (DLQI) and the Skindex-16 could not be procured and was thus excluded (25).

Most of the included studies reported on the DLQI (n=6) (23, 24, 28-30, 35) and the Quality of Life Index for Atopic Dermatitis (QoLIAD, n=3) (31, 34, 36). Two studies presented information on the Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen (DIELH) (32, 33). Skindex-29 was evaluated in two studies (22, 26). One study each was available for the Freiburg Life Quality Assessment core module (FLQA-c) (21), the Freiburg Life Quality Assessment for Dermatoses (FLQA-d) (23) and the Impact of Chronic Skin Disease on Daily Life (ISDL) (27). An overview of the content of these different instruments is shown in table 3. Symptoms and emotions are captured by six out of seven questionnaires whereas all other content domains are included in a lower number of instruments. Four instruments (DIELH, DLQI, FLQA-c and FLQA-d) share the most content domains whereas the QoLIAD does not have any content domains in common with the other QoL instruments. Other characteristics of the included instruments are shown in table 4. The number of items ranges from 10 to 54. Almost all instruments use a 4- or 5-point Likert scale. Only the ISDL applies a visual analogue scale (VAS) in addition, whereas the QoLIAD has a dichotomous response format.

Table 3: Comparison of the content of the different QoL instruments on domain level.

Domain	DIELH	DLQI	FLQA-c	FLQA-d	ISDL*	QoLIAD	Skindex-29†
Symptoms	X	X	X	X	X		X
Emotions	X	X	X	X	X		X
Activities of daily living	X	X	X	X	X		
Leisure	X	X					
Work/study	X	X					
Social life	X	X	X	X	X		
Treatment	X	X	X	X			
Functioning							X
Satisfaction			X	X			
Stigmatization					X		
Illness cognitions					X		
Need for mental and emotional stimulation						X	
Need for physical and emotional stability						X	
Need for security						X	
Need to share and belong						X	
Esteem needs						X	
Need for personal development and fulfillment						X	

*The ISDL distinguishes several higher level domains that contain a number of subordinate domains each. The subordinate domains were used for this content comparison. The exact domains are (subordinate domains in brackets): physical functioning (skin status; physical symptoms of itch, pain and fatigue; scratch response), psychological functioning (anxiety; negative mood; positive mood), stressors (disease impact on daily life; stigmatization), illness cognitions (helplessness; acceptance; perceived benefits), social support (perceived support; social network).

†Content comparison of Skindex-29 is based on dimensions empirically derived from factor analysis and not on content-related domains.

Abbreviations: *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen; *DLQI* = Dermatology Life Quality Index; *FLQA-c* = Freiburg Life Quality Assessment core module; *FLQA-d* = Freiburg Life Quality Assessment for Dermatoses; *ISDL* = Impact of Chronic Skin Disease on Daily Life; *QoLIAD* = Quality of Life Index for Atopic Dermatitis

Table 4: Characteristics of the different instruments.

Characteristic	DIELH	DLQI	FLQA-c	FLQA-d	ISDL	QoLIAD	Skindex-29
Target population	Patients with any dermatological condition	Patients with skin disease	Patients with skin disease	Patients with chronic inflammatory skin disease	Patients with chronic skin disease	Eczema patients	Patients with skin disease
Number of items	36	10	28	54	32	25	29
Number of subscales	7	6	6	6	5	None	3
Number/type of response categories	5-point Likert scale (and 'not applicable')	4-point Likert scale	5-point Likert scale	ND	10-cm-VAS for physical symptoms; 5-point Likert scale for positive and negative mood; 4-point Likert scale for all other scales	Dichotomous (true/not true)	5-point Likert scale
Scoring algorithm	Calculation of a sum score, range 0-180	Calculation of a sum score, range 0-30	Calculation of a scale score by averaging the answers within a scale, range 1-5; no total score	ND	Calculation of subscale scores by summing up the subscales' items scores	Calculation of a sum score, range 0-25	Calculation of a scale score by averaging responses to items in a given scale
Recall period in the items	ND	1 week	1 week	ND	ND	ND	ND
Administration costs	ND	No charge for unfunded studies; $9.50 per patient for pharmaceutical companies (37)	ND	ND	ND	No charge for non-commercial studies; Administration fee of £100 for commercial studies (38)	ND
Available translations	German	More than 90 (37)	German	German	Dutch	Dutch, English (UK), English (US), French, German, Italian, Japanese, Spanish (39)	16 language versions (40)

Abbreviations: *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen; *DLQI* = Dermatology Life Quality Index; *FLQA-c* = Freiburg Life Quality Assessment core module; *FLQA-d* = Freiburg Life Quality Assessment for Dermatoses; *ISDL* = Impact of Chronic Skin Disease on Daily Life; *ND* = Not described; *QoLIAD* = Quality of Life Index for Atopic Dermatitis; *UK* = United Kingdom; *US* = United States (of America); *VAS* = visual analogue scale

Characteristics of the included studies

Table 5 contains information on the settings and the study populations in the included studies. All included studies were conducted in Europe with the exception of the validation studies of the US versions of the QoLIAD and the Skindex-29. Most studies recruited their participants in a secondary care setting while primary care patients were included in only two studies. Additionally, there was significant variation with respect to sample size, with 15 patients being the smallest and 286 patients the largest sample size of a single study.

Table 5: Important characteristics of the included development and validation studies.

		Study characteristics				
					Study population	
QoL instrument	**Number of studies**	**Geographic location(s)**	**Language(s)**	**Setting(s)**	**Number of participants per study**	**Age range (years)**
DIELH	2 (32, 33)	Germany	German	Secondary care	85 (32)	ND
					ND (33)	ND (oldest: 88)
Danish DLQI	1 (30)	Denmark	Danish	Secondary care	66	ND
English DLQI (UK)	3 (28, 29, 35)	United Kingdom	English (UK)	Secondary care (28)	13 (28)	ND (28)
				Primary care (29)	56 (29)	16-53 (29)
				Community (35)	146 (35)	20-82 (35)
German DLQI	1 (23)	Germany	German	Tertiary care	80	ND
Spanish DLQI	1 (24)	Spain	Spanish	Secondary care	114	ND
FLQA-c	1 (21)	Germany	German	Tertiary care	253	17-75
FLQA-d	1 (23)	Germany	German	Tertiary care	80	ND
ISDL	1 (27)	Netherlands	Dutch	Secondary care	128	16-77
Dutch QoLIAD	1 (36)	Netherlands	Dutch	Secondary care	15 (item generation)	ND
					20 (field testing)	ND
					46 (validation)	16-67
English QoLIAD (UK)	2 (34, 36)	United Kingdom	English (UK)	Community (34)	146 (34)	20-82 (34)
				Community and secondary care (36)	36 (item generation) (36)	ND (36)
					21 (field testing) (36)	ND (36)
				Community (36)	286 (validation) (36)	16-86 (36)
English QoLIAD (US)	1 (36)	United States of America	English (US)	ND	ND (item generation)	ND
				Secondary care	20 (field testing)	ND
					178 (validation)	16-78
French QoLIAD	1 (36)	France	French	ND	ND (item generation)	ND
				Secondary care	ND (field testing)	ND
				Community	213 (validation)	18-86
German QoLIAD	1 (36)	Germany	German	ND	ND (item generation)	ND
				Secondary care	17 (field testing)	ND
				Community and secondary care	187 (validation)	17-77
Italian QoLIAD	1 (36)	Italy	Italian	Secondary care	14 (item generation)	ND
					15 (field testing)	

		Study characteristics				
					Study population	
QoL instrument	**Number of studies**	**Geographic location(s)**	**Language(s)**	**Setting(s)**	**Number of participants per study**	**Age range (years)**
Spanish QoLIAD	1 (31, 36)*	Spain	Spanish	ND (36)	ND (36)	ND (36)
				Community and secondary care (36)	20 (field testing) (36)	ND (36)
				Secondary care (31, 36)	83 (validation) (31, 36)	16-81 (31, 36)
English Skindex-29 (US)	1 (26)	United States of America	English (US)	Primary and secondary care	103	ND
German Skindex-29	1 (22)	Germany	German	Tertiary care	76	ND

Abbreviations: *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen; *DLQI* = Dermatology Life Quality Index; *FLQA-c* = Freiburg Life Quality Assessment core module; *FLQA-d* = Freiburg Life Quality Assessment for Dermatoses; *ISDL* = Impact of Chronic Skin Disease on Daily Life; *ND* = Not described; *QoLIAD* = Quality of Life Index for Atopic Dermatitis; *UK* = United Kingdom; *US* = United States (of America)

*Two articles on the Spanish QoLIAD were included but regarded as one study due to duplicate publication. From de Lucas 2003, only validation data not presented in Whalley 2004 was taken into account.

Validity of the instruments and recommendations

The number of studies assessing the different measurement properties of each QoL instrument identified is given in table 6. From the 15 included studies, we were able to rate the methodological quality of 67 measurement properties. One measurement property (1%) was rated as having excellent, 18 (27%) as having good, 31 (46%) as having fair and 17 (25%) as having poor methodological quality according to the COSMIN checklist. Our synthesis of the results and level of evidence for the properties of each instrument is presented in table 7. There was no instrument for which all measurement properties of interest have been examined. As a result, none of the instruments complied with all of our pre-specified requirements of truth, discrimination and feasibility. Detailed results for every single instrument and study included are available as an online appendix to this publication (tables E1-E55).

Table 6: Number of studies assessing the measurement properties of QoL instruments for adults with eczema

Measurement property	DIELH	Danish DLQI	English DLQI (UK)	German DLQI	Spanish DLQI	FLQA-c	FLQA-d	ISDL	Dutch QoLIAD	English QoLIAD (UK)	English QoLIAD (US)	French QoLIAD	German QoLIAD	Italian QoLIAD	Spanish QoLIAD	English Skindex-29 (US)	German Skindex-29
Internal consistency	/	/	1 (35)	/	/	/	/	1 (27)	1 (36)	2 (34, 36)	1 (36)	1 (36)	1 (36)	/	1 (36)	/	/
Measurement error	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Reliability	/	/	/	/	1 (24)	/	/	/	1 (36)	1 (36)	1 (36)	1 (36)	1 (36)	/	1 (36)	/	/
Content validity	/	/	1 (29)	/	/	/	/	1 (27)	1 (36)	1 (36)	1 (36)	/	1 (36)	1 (36)	1 (36)	/	/
Structural validity	1 (33)	/	1 (35)	/	/	/	/	/	/	1 (34)	/	/	/	/	/	/	/
Hypothesis testing	2 (32, 33)	1 (30)	2 (28, 29)	1 (23)	/	1 (21)	1 (23)	1 (27)	1 (36)	1 (36)	1 (36)	1 (36)	1 (36)	/	1 (31, 36)	1 (26)	1 (22)
Cross-cultural validity	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Responsiveness	/	/	/	/	1 (24)	1 (21)	/	1 (27)	/	/	/	/	/	/	/	/	/

Abbreviations: *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen; *DLQI* = Dermatology Life Quality Index; *FLQA-c* = Freiburg Life Quality Assessment core module; *FLQA-d* = Freiburg Life Quality Assessment for Dermatoses; *ISDL* = Impact of Chronic Skin Disease on Daily Life; *QoLIAD* = Quality of Life Index for Atopic Dermatitis; *UK* = United Kingdom; *US* = United States (of America)

Table 7: Summary of measurement properties of QoL instruments for adults with eczema

Measurement property	DIELH	Danish DLQI	English DLQI (UK)	German DLQI	Spanish DLQI	FLQA-c	FLQA-d	ISDL	Dutch QoLIAD	English QoLIAD (UK)	English QoLIAD (US)	French QoLIAD	German QoLIAD	Italian QoLIAD	Spanish QoLIAD	English Skindex-29 (US)	German Skindex-29
Internal consistency	/	/	--	/	/	/	/	Weak	Weak	++	+	+	+	/	Weak	/	/
Measurement error	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Reliability	/	/	/	/	+	/	/	/	Weak	? (limited)	? (limited)	? (limited)	? (limited)	/	? (limited)	/	/
Content validity	/	/	-	/	/	/	/	Weak	++	++	++	/	++	+++	++	/	/
Structural validity	? (limited)	/	--	/	/	/	/	/	/	? (moderate)	/	/	/	/	/	/	/
Hypothesis testing	+	Weak	+	Weak	/	? (limited)	Weak	-	-	+	+	+	+	/	+	++	? (limited)
Cross-cultural validity	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
Responsiveness	/	/	/	/	+	Weak	/	Weak	/	/	/	/	/	/	/	/	/
Recommen-dation	**D**	**D**	**C**	**D**	**B**	**D**	**D**	**C**	**C**	**B**	**B**	**B**	**B**	**D**	**B**	**D**	**D**

Abbreviations: *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen; *DLQI* = Dermatology Life Quality Index; *FLQA-c* = Freiburg Life Quality Assessment core module; *FLQA-d* = Freiburg Life Quality Assessment for Dermatoses; *ISDL* = Impact of Chronic Skin Disease on Daily Life; *QoLIAD* = Quality of Life Index for Atopic Dermatitis; *UK* = United Kingdom; *US* = United States (of America)

Recommendations are defined as follows: *A,* QoL measurement instrument meets all requirements and is recommended for use; *B,* QoL measure meets two or more quality items, but performance in all other required quality items is unclear, so that the outcome measure has the potential to be recommended in the future depending on the results of further validation studies; *C,* QoL measure has low quality in at least one required quality criterion (≥1 rating of 'minus') and therefore is not recommended to be used any more; *D,* QoL measure has (almost) not been validated. Its performance in all or most relevant quality items is unclear so that it is not recommended to be used until further validation studies clarify its quality.

+++, ++, +, positive rating indicating adequate measurement property; *? (moderate), ? (limited),* intermediate rating indicating intermediate measurement property; *-, --,* negative rating indicating inadequate measurement property (please refer to table 2 for further details); Weak = measurement property was assessed only in studies of poor methodological quality; / = not assessed

Internal consistency was good for most language versions of the QoLIAD, with Cronbach's α ranging from 0.88 to 0.94 (36). In a population of 146 eczema patients, the Person Separation Index of the DLQI amounted to 0.63, resulting in a negative rating for internal consistency (35). For all other instruments, there was either no evidence on internal consistency or only evidence from methodologically poor studies. Measurement error was not assessed for any of the included instruments. An indeterminate rating was found for most language versions of the QoLIAD in terms of reliability. Of the other instruments, reliability information was available for the Spanish DLQI only; with an Intraclass Correlation Coefficient (ICC) of 0.77 between the two administrations, this instrument showed good reliability (24).

There was moderate evidence of good content validity for most QoLIAD versions. There was strong evidence that the Italian QoLIAD has good content validity. Content validity was found to be limited for the UK version of the DLQI in a population of 56 eczema patients; these patients considered the DLQI not comprehensive and found some items irrelevant (for instance, items 1 and 9 were not considered relevant by any patient in that study) (29). Likewise, structural validity of the UK version of the DLQI was found to be poor due to non-uniform differential item functioning (DIF) of items 6 and 7 with respect to gender and age, respectively. Moreover, 2/10 items showed uniform DIF with respect to gender, 3/10 items exhibited uniform DIF with respect to age, and there was disease-specific DIF for 5/10 items when patients with eczema and psoriasis were compared. Item residual statistics were indicative of a misfit to the Rasch model, although item-trait interaction suggested that the DLQI fits a Rasch model for eczema patients (35). Structural validity of the UK version of the QoLIAD as well as of DIELH is unclear. With data available for 15/17 QoL instruments, hypothesis testing (i.e. construct validity) was the measurement property most frequently assessed. Good construct validity was found for the DIELH and most QoLIAD versions. Correlations between QoLIAD (except Dutch and Italian) and DLQI were moderate to high (r=0.58-0.77) with most values being above 0.70. Similar but lower correlations were found between QoLIAD (except Dutch and Italian) and the Psychological General Well-Being Schedule (PGWB) (r=0.55-0.79) (36). Good convergent validity was also demonstrated for the UK version of the DLQI (29). With the exception of the Dutch and the Spanish QoLIAD versions, patients could be clearly discriminated according to perceived severity, current flares of symptoms and general health using the QoLIAD (36). The ISDL and the Dutch QoLIAD got negative ratings for hypothesis testing. While convergent validity of the Dutch QoLIAD was adequate, its discriminative validity was poor and resulted in a negative rating (36). The English Skindex-29 (US version) had good discriminative validity (26). For the remaining questionnaires, hypothesis testing assessments either led to an indeterminate rating or were conducted methodologically poorly.

Responsiveness in eczema patients was investigated for only 3 questionnaires but only the Spanish DLQI was proven responsive (24).

Values for the minimal important change (MIC) or minimal important difference (MID) were not available for any of the included questionnaires. Data on floor and ceiling effects (i.e. ≥15% of patients having the lowest/the highest score) were available from one study for the QoLIAD. Only the US version of the QoLIAD showed some floor effects with 17.1% and 18.5% of respondents having the lowest score for visits 1 and 2, respectively. No floor or ceiling effects were observed for the other QoLIAD versions (36). In a sample of 56 eczema patients, the English DLQI (UK) exhibited no ceiling effects (29). Likewise, there were no floor or ceiling effects in the 13 eczema patients taking

part in the development study of the English DLQI (UK) (28). Completion time for the Spanish QoLIAD was found to be 5 minutes or less (36).

Discussion

In this systematic review, the measurement properties of seven different adult eczema QoL instruments were evaluated. None of these instruments fulfilled all predefined filter criteria for truth, discrimination and feasibility, indicating the need for further validation work.

Currently, no QoL instrument can be highly recommended. In general, more validation research on all QoL questionnaires included in this review would be desirable. The QoLIAD (36) in several language versions was placed in category B, meaning that it has the potential to be recommended in the future depending on the results of further validation studies. The same is true for the Spanish language version of the DLQI (24), although less information is available for this instrument compared to the QoLIAD. For the majority of the questionnaires, i.e. DIELH (33), Danish DLQI (30), German DLQI (23), FLQA-c (21), FLQA-d (23), Italian QoLIAD (36), English Skindex-29 (US) (26) and German Skindex-29 (22), further usage cannot be recommended until more validation data is available since the performance of these instruments is largely unclear. Three instruments, the English DLQI (UK version) (28), ISDL (27) and Dutch QoLIAD (36), were found to have low adequacy in at least one required adequacy criterion and therefore are considered problematic for further use in eczema patients.

The QoLIAD, in several language versions, is a valid and internally consistent QoL instrument applying a needs-based model. According to this model, QoL is determined by an individual's ability and capacity to satisfy their needs, with high QoL when most needs and lowest QoL when few or none of the needs are met. Consequently, instruments based on this model assess the overall impact of a disease and its treatment. This is also reflected by the fact that the QoLIAD is the only instrument that does not have any content domains in common with the other instruments. As a result, the QoLIAD may not cover some of the aspects that clinicians might consider important in clinical practice. Floor or ceiling effects of the 25-item questionnaire were almost not observed and it was quickly completed. Although good construct validity was shown for most language versions of the QoLIAD, the negative rating for hypothesis testing for the Dutch QoLIAD indicates that the QoLIAD's construct validity should be further examined. Reliability, structural validity and cross-cultural validity of the QoLIAD are unclear and should also be further investigated. Measurement error and responsiveness of the QoLIAD have not yet been investigated. Moreover, interpretability data (i.e., definition and ranges of the QoLIAD that represent mild, moderate and severe QoL impairments in eczema) are not available.

The Spanish DLQI is a 10-item QoL instrument that was shown reliable and responsive in eczema patients. The validity of this DLQI version has not yet been tested. Even though plenty of information concerning floor and ceiling effects as well as other interpretability data is available for other language versions of the DLQI in populations other than eczema, respective data of the Spanish DLQI obtained in eczema patients are not available.

We found the English (UK) version of the DLQI to have poor internal consistency, content and structural validity in eczema patients. Thus, the English DLQI (UK version) is not suggested to assess QoL in eczema patients. Likewise, the ISDL and the Dutch version of the QoLIAD are not suggested for use either because of a lack of construct validity.

As we included a number of instruments that are dermatology-specific and thus were not specifically developed for patients with eczema, content validity of those instruments in eczema patients is of

great importance. Dermatology-specific instruments are more likely to miss issues that eczema patients consider important simply because they were developed for patients with skin disease in general. Whereas good content validity was shown for the QoLIAD, an eczema-specific instrument, content validity of the included dermatology-specific instruments in eczema patients was almost not investigated. One study found limited content validity of the English DLQI (UK) in eczema patients. This finding challenges the applicability of the DLQI to eczema patients and raises the question whether other language versions of this instrument may have better content validity. Particularly for the Spanish DLQI, shown to be adequately reliable and responsive, a thorough examination of its content validity in eczema patients is needed.

As most data on interpretability were not gathered in eczema samples, only little information on interpretability was available for the included instruments. For instance, a MIC of 4 points has been proposed for the DLQI, but the corresponding studies did not meet our eligibility criteria (41, 42). Banding systems to assign clinical meaning to the scores have been suggested both for the DLQI (43) and the Skindex-29 (44-46), but none of these studies was found eligible. Thus, future validation studies should also look at interpretability in eczema patients.

Strengths and limitations of this review

We registered and published a protocol prior to our systematic review and highlighted differences between the protocol and final review. A validated, precise search filter was used to identify all possibly eligible articles of any language indexed in PubMed, EMBASE or both (47). Aiming to find the best evidence for eczema patients, we used predefined and strict eligibility criteria. We applied the COSMIN checklist to rate the study quality and gather information on interpretability and feasibility (11-14). At least two reviewers were involved in every step of the review process assure quality. Frequent discussions took place within the research team in order to resolve discrepancies.

A potential limitation of our systematic review is that we only searched PubMed and EMBASE, thus possibly missing articles listed elsewhere. However, we were not able to find any further eligible articles through a thorough hand search. We were not able to retrieve one eligible article providing information on the measurement properties of the Ukrainian versions of the Skindex-16 and the DLQI (25).

Recommendations to researchers, clinicians and decision makers

This review suggests that currently only the QoLIAD and the DLQI have the potential to be recommended for use depending on the results of further validation studies. These validation studies should investigate several language versions of the QoLIAD and the DLQI, also including the versions that were found inadequate for use in eczema patients in order to possibly confirm the findings of previous studies, thus strengthening the evidence base for the recommendations presented in this systematic review. The Dutch QoLIAD, the ISDL and the UK version of the DLQI are not suggested for use in eczema trials unless future validation studies show – in contrast to the existing evidence – adequate measurement properties for these instruments.

Clinicians and researchers should include a QoL measurement instrument in every future eczema trial because QoL is one of the core outcome domains of the proposed COS. As no instrument for measuring adult QoL in eczema trials can be highly recommended at the moment, the HOME initiative suggests using any QoL instrument that is at least valid, reliable and feasible in eczema patients (48). Unfortunately, we found in our review that currently no such instrument is available.

An ideal solution to this quandary does not exist. Clinicians and researchers need to balance validity, reliability and feasibility. We suggest that researchers should include one of the two instruments from category B, i.e. the QoLIAD or the DLQI, in their trials.

Figure legends

Figure 1

Figure 1: Diagram of article flow during literature search and article screening according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards.

Acknowledgments

The authors are very grateful for the help of Valeria Aoki (Sao Paulo, Brazil), Simon Bein (Regensburg, Germany) and Viktoria Eleftheriadou (Coventry, United Kingdom) with foreign language articles.

Authors' contributions

DH conceptualized the research plan for the systematic review, screened search results, extracted data, assessed measurement properties, coordinated the work of the other reviewers, wrote the manuscript and reviewed it for important intellectual content. CP screened search results, extracted data, assessed measurement properties, wrote the manuscript and reviewed it for important intellectual content. SD performed the literature search, extracted data, assessed measurement properties, helped with the methodology section and reviewed the manuscript for important intellectual content. JC extracted data, assessed measurement properties, and reviewed the manuscript for important intellectual content. AD screened search results, extracted data, assessed measurement properties, wrote the manuscript and reviewed it for important intellectual content. RO screened search results, extracted data, assessed measurement properties and reviewed the manuscript for important intellectual content. RH contributed to the study design and reviewed the manuscript for important intellectual content. TS screened search results, extracted data, assessed measurement properties and reviewed the manuscript for important intellectual content. SC screened search results, extracted data, assessed measurement properties and reviewed the manuscript for important intellectual content. JS conceptualized the research plan for the systematic review, extracted data, assessed measurement properties and reviewed the manuscript for important intellectual content. CA conceptualized the research plan for the systematic review, extracted data, assessed measurement properties, wrote the manuscript and reviewed it for important intellectual content.

Conflicts of Interest Statement

Funding sources

Tracey Sach is funded by a Career Development Fellowship from the National Institute for Health Research. Rosemary Humphreys receives funding from the Harmonising Outcome Measures in Eczema (HOME) initiative for travel and accommodation costs for meetings. The funding sources had no involvement in study design, collection, analysis and interpretation of data, writing of the manuscript, or the decision to submit the manuscript for publication.

Conflicts of interest

Dr. Drucker is involved with the development of a novel quality of life assessment instrument for atopic dermatitis that is as yet unpublished. Jochen Schmitt and Christian Apfelbacher are members of the executive committee of the HOME initiative. All authors are ordinary members of the HOME initiative. The authors declare that there are no other conflicts of interests.

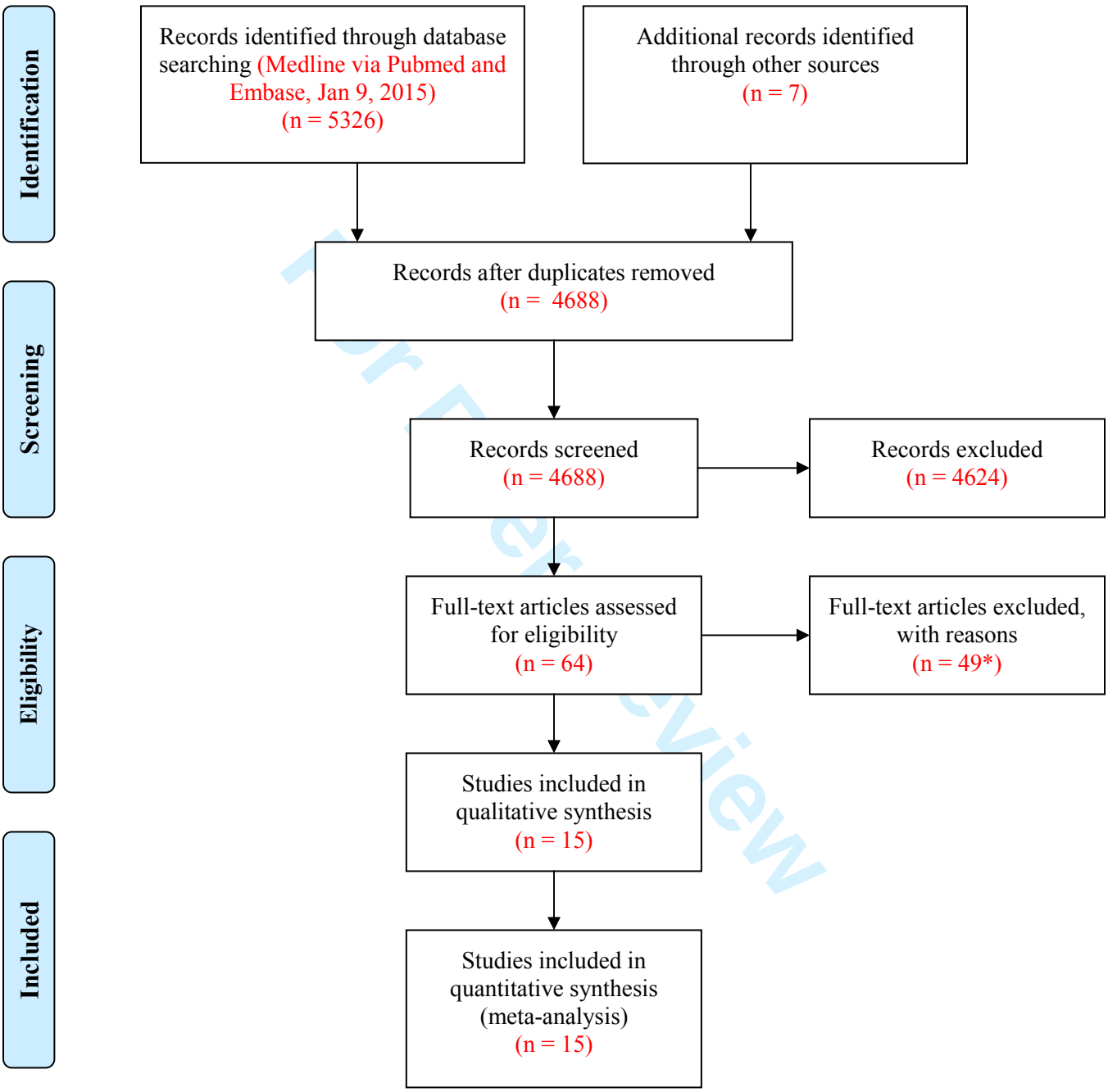
References

1. Silverberg JI, Hanifin JM. Adult eczema prevalence and associations with asthma and other health and demographic factors: a US population-based study. *J Allergy Clin Immunol* 2013;**132**(5):1132-1138.
2. Vinding GR, Zarchi K, Ibler KS, Miller IM, Ellervik C, Jemec GB. Is adult atopic eczema more common than we think? - A population-based study in Danish adults. *Acta Derm Venereol* 2014;**94**(4):480-482.
3. Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;**8**:39.
4. Schmitt J, Langan S, Stamm T, Williams HC, Harmonizing Outcome Measurements in Eczema Delphi p. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. *J Invest Dermatol* 2011;**131**(3):623-630.
5. Schmitt J, Spuls P, Boers M, Thomas K, Chalmers J, Roekevisch E, et al. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. *Allergy* 2012;**67**(9):1111-1117.
6. Schmitt J, Williams H, Group HD. Harmonising Outcome Measures for Eczema (HOME). Report from the First International Consensus Meeting (HOME 1), 24 July 2010, Munich, Germany. *Br J Dermatol* 2010;**163**(6):1166-1168.
7. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol* 2015;**135**(1):24-30.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;**6**(7):e1000097.
9. Apfelbacher CJ, Heinl D, Prinsen CA, Deckert S, Chalmers J, Ofenloch R, et al. Measurement properties of adult quality-of-life measurement instruments for eczema: protocol for a systematic review. *Syst Rev* 2015;**4**(1):48.
10. Minutes of the HOME III Meeting. Harmonising Outcome Measures in Eczema (HOME) initiative; 2013. Available from: http://www.nottingham.ac.uk/homeforeczema/documents/minutes-home-iii-final.pdf (accessed at 14/09/2015)
11. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010;**10**:22.
12. 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**(7):737-745.
13. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;**19**(4):539-549.
14. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res* 2012;**21**(4):651-657.
15. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;**60**(1):34-42.
16. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: a practical guide: Cambridge University Press; 2011.
17. PROMIS® Instrument Development and Validation Scientific Standards Version 2.0. 2013 [cited; Available from: http://www.nihpromis.org/%28X%281%29S%28l4ziabyqjykgamhigno4zi5p%29%29/Documents/PROMISStandards_Vers2.0_Final.pdf

18. Furlan AD, Pennick V, Bombardier C, van Tulder M, Editorial Board CBRG. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine (Phila Pa 1976)* 2009;**34**(18):1929-1941.
19. Schellingerhout JM, Verhagen AP, Heymans MW, Koes BW, de Vet HC, Terwee CB. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. *Qual Life Res* 2012;**21**(4):659-670.
20. PROMIS Methodology. Patient Reported Outcomes Measurement Information System (PROMIS) Standards 2015. [cited 14/07/2015]; Available from: http://www.nihpromis.org/science/methodology
21. Augustin M, Lange S, Wenninger K, Seidenglanz K, Amon U, Zschocke I. Validation of a comprehensive Freiburg Life Quality Assessment (FLQA) core questionnaire and development of a threshold system. *Eur J Dermatol* 2004;**14**(2):107-113.
22. Augustin M, Wenninger K, Amon U, Schroth MJ, Kuster W, Chren M, et al. German adaptation of the Skindex-29 questionnaire on quality of life in dermatology: validation and clinical results. *Dermatology* 2004;**209**(1):14-20.
23. Augustin M, Zschocke I, Lange S, Seidenglanz K, Amon U. [Quality of life in skin diseases: methodological and practical comparison of different quality of life questionnaires in psoriasis and atopic dermatitis]. *Hautarzt* 1999;**50**(10):715-722.
24. Badia X, Mascaro JM, Lozano R. Measuring health-related quality of life in patients with mild to moderate eczema and psoriasis: clinical validity, reliability and sensitivity to change of the DLQI. The Cavide Research Group. *Br J Dermatol* 1999;**141**(4):698-702.
25. Chernyshov PV. Health related quality of life in adult atopic dermatitis and psoriatic patients matched by disease severity. *G Ital Dermatol Venereol* 2014.
26. 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**(11):1433-1440.
27. Evers AW, Duller P, van de Kerkhof PC, van der Valk PG, de Jong EM, Gerritsen MJ, et al. The Impact of Chronic Skin Disease on Daily Life (ISDL): a generic and dermatology-specific health instrument. *Br J Dermatol* 2008;**158**(1):101-108.
28. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994;**19**(3):210-216.
29. Herd RM, Tidman MJ, Ruta DA, Hunter JA. Measurement of quality of life in atopic dermatitis: correlation and validation of two different methods. *Br J Dermatol* 1997;**136**(4):502-507.
30. Holm EA, Wulf HC, Stegmann H, Jemec GB. Life quality assessment among patients with atopic eczema. *Br J Dermatol* 2006;**154**(4):719-725.
31. Lucas Rd, González-Enseñat MA, Giménez Anna, Puig L, Vicente A, Díaz JM, et al. Validación de las versiones españolas de los cuestionarios Parents' Index of Quality of Life in Atopic Dermatitis (PIQoL-AD) y Quality of Life Index for Atopic Dermatitis (QoLIAD): un estudio internacional. *Actas Dermo-Sifiliográficas* 2003;**94**(8):539-548.
32. Schafer T, Staudt A, Ring J. [German instrument for the assessment of quality of life in skin diseases (DIELH). Internal consistency, reliability, convergent and discriminant validity and responsiveness]. *Hautarzt* 2001;**52**(7):624-628.
33. Schafer T, Staudt A, Ring J. [Development of the German Scale for Assessing Quality of Life in Skin Diseases]. *Hautarzt* 2001;**52**(6):492-498.
34. Twiss J, McKenna SP. Comparing the impact of psoriasis and atopic dermatitis on quality of life: co-calibration of the PSORIQoL and QoLIAD. *Qual Life Res* 2015;**24**(1):105-113.
35. Twiss J, Meads DM, Preston EP, Crawford SR, McKenna SP. Can we rely on the Dermatology Life Quality Index as a measure of the impact of psoriasis or atopic dermatitis? *J Invest Dermatol* 2012;**132**(1):76-84.
36. Whalley D, McKenna SP, Dewar AL, Erdman RA, Kohlmann T, Niero M, et al. A new instrument for assessing quality of life in atopic dermatitis: international development of the Quality of Life Index for Atopic Dermatitis (QoLIAD). *Br J Dermatol* 2004;**150**(2):274-283.

37. Dermatology Quality of Life Index (DLQI). [cited 15/06/2015]; Available from: http://www.cardiff.ac.uk/dermatology/quality-of-life/dermatology-quality-of-life-index-dlqi/
38. Administration costs of QoLIAD. In. Inquiry about administration costs of QoLIAD by e-mail. ed.
39. Measures database, Galen Research. [cited 15/06/2015]; Available from: http://www.galen-research.com/measures-database/
40. Skindex (Skindex / Skindex-29 / Skindex-16). [cited 15/06/2015]; Available from: http://www.proqolid.org/instruments/skindex_skindex_skindex_29_skindex_16
41. Basra MK, Salek MS, Camilleri L, Sturkey R, Finlay AY. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology* 2015;**230**(1):27-33.
42. Khilji FA, Gonzalez M, Finlay AY. Clinical meaning of change in Dermatology Life Quality Index scores. *Br J Dermatol* 2002;**147**(Suppl 62):50.
43. 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**(4):659-664.
44. Dunbar S, McCombs K, DeLong L, Alexander M. Banding and the clinical interpretation of Skindex-29 scores using an anchor based technique. *J Am Acad Dermatol* 2010;**62**(3):AB68.
45. Nijsten T, Sampogna F, Abeni D. Categorization of Skindex-29 scores using mixture analysis. *Dermatology* 2009;**218**(2):151-154.
46. Prinsen CA, Lindeboom R, Sprangers MA, Legierse CM, de Korte J. Health-related quality of life assessment in dermatology: interpretation of Skindex-29 scores using patient-based anchors. *J Invest Dermatol* 2010;**130**(5):1318-1322.
47. Terwee CB, Jansma EP, Riphagen, II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009;**18**(8):1115-1123.
48. Chalmers J, Deckert S, Schmitt J, Harmonising Outcome Measures for Atopic Dermatitis Initiative Executive B. Reaching clinically relevant outcome measures for new pharmacotherapy and immunotherapy of atopic eczema. *Curr Opin Allergy Clin Immunol* 2015;**15**(3):227-233.

PRISMA 2009 Flow Diagram



*One article was found eligible, but could not be procured and was thus excluded.

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Detailed results: Rating of measurement properties of outcomes instruments of quality of life of adult eczema patients and assessment of the methodological quality of the included studies

Table of contents

Abbreviations and symbols used

+ positive rating

? indeterminate rating

- negative rating

AD = atopic dermatitis; *AE* = atopic eczema; *ANOVA* = analysis of variance; *COSMIN* = COnsensus-based Standards for the selection of health status Measurement INstruments; *DIELH* = Deutsches Instrument zur Erfassung der Lebensqualität; *DIF* = Differential item functioning; *DLQI* = Dermatology Life Quality Index; *EASI* = Eczema Area and Severity Index; *GWBI* = General Well-Being Index; *INVAS* = Investigator overall assessment of disease severity; *QoL* = quality of life; *MCS* = Mental component score ; *NL* = Netherlands; *PCS* = Physical component score; *PGI* = Patient-Generated Index; *PGWB* = Psychological General Well-Being Index; *PRUVAS* = subjective measure of pruritus severity; *PTVAS* = subjective measure of eczema severity; *SCORAD* = SCORing Atopic Dermatitis; *SF-36* = Short form 36; *TCS* = topical corticosteroids; *UK* = United Kingdom; *US* = United States of America

1. Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen (DIELH)

Table E1: Structural validity of the DIELH

Author	Structural validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E1)	Principal components analysis within the single diagnostic groups (including AE) performed; questions were included if they did not load >0.7 on more than one factor	Not given	?	Number of AE patients unknown	fair
Conclusion: One study assessed structural validity of the DIELH and indicated unclear structural validity as a QoL instrument for eczema ➔ Structural validity of the DIELH as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E2: Discriminative validity of the DIELH

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E2)	Comparison of the sum scores of different diagnostic groups (Kruskal-Wallis test); hypothesis: Patients with chronic inflammatory dermatoses (like AE) have an higher impact on QoL	Median total score for AE 75.5 (highest value of all diagnostic groups); statistically significant ($p<0.0001$)	+	85 AE patients	fair
Conclusion: One study assessed discriminative validity of the DIELH and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the DIELH as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

2. Dermatology Life Quality Index (DLQI) – Danish version

Table E3: Convergent/divergent validity of the Danish DLQI

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E3)	Spearman correlation coefficients between DLQI and 8 dimensions/PCS/MCS of the SF-36; Spearman correlation coefficients between DLQI, PRUVAS, PTVAS and INVAS; Wilcoxon rank scores between DLQI and SCORAD	The spearman correlation coefficients between DLQI and 8 dimensions/PCS/MCS of the SF-36 range between -0.54 (General health) and -0.11 (Bodily pain); most correlations <0.5 Spearman correlation coefficients for DLQI were 0.62 with PRUVAS, 0.81 with PTVAS and 0.82 with INVAS. DLQI was significantly (P < 0.0001) associated with objective SCORAD.	-	66 patients with eczema	poor
Conclusion: One study assessed convergent/divergent validity of the Danish DLQI, but due to poor methodological study quality no conclusion can be drawn ➔ Convergent/divergent validity of the Danish DLQI as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E4: Discriminative validity of the Danish DLQI

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E3)	Discriminative was assessed (using Wilcoxon rank scores) by seeing how well the QOL measures could discriminate between groups of participants according to clinical assessed SCORAD	Differences in DLQI scores between patients with mild and moderate AD (according to objective SCORAD) were statistically significant (P<0.0001).	+	66 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the Danish DLQI, but due to poor methodological study quality no conclusion can be drawn ➔ Discriminative validity of the Danish DLQI as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

3. Dermatology Life Quality Index (DLQI) – English version (UK)

Table E5: Internal consistency of the English DLQI (UK)

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E4)	Person Separation Index (PSI)	0.63 for eczema patients (considered low by the author)	-	146 patients with eczema	good
Conclusion: One study assessed internal consistency of the UK version of the DLQI and indicated inadequate internal consistency as a QoL instrument for eczema ➔ Internal consistency of the English DLQI (UK) as a measurement of QoL: **inadequate** ➔ Quality of evidence: **good**					

Table E6: Content validity of the English DLQI (UK)

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E5)	Comparison of the areas/activities in the DLQI and those that were mentioned by the patients in the PGI; hypothesis: patients would include a broader range of affected areas in their responses to the PGI than those included in the DLQI	36 patients (64%) mentioned areas or activities not part of the DLQI, 20 patients identified only areas included in the DLQI; DLQI item 1 not mentioned by any patient	-	56 patients with eczema	fair
Conclusion: One study assessed content validity of the UK version of the DLQI and indicated inadequate content validity as a QoL instrument for eczema ➔ Content validity of the English DLQI (UK) as a measurement of QoL: **inadequate** ➔ Quality of evidence: **fair**					

Table E7: Structural validity of the English DLQI (UK)

Author	Structural validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E4)	Examination of overall fit to the Rasch model by reference to the overall item-trait interaction χ^2-fit value and via Item and Person interaction statistics; examination of DIF by ANOVA of standardized residuals.	DLQI does not fit a Rasch model for the overall sample, but fits a Rasch model for AD patients (item-trait interaction = 0.460); item residual statistics indicative of model misfit for the AD patients; 5/10 items showed DIF for different parameters (age and/or gender) in the AD sample. 5/10 items showed disease-specific DIF in the overall sample. A single item (item 4, P=0.048) showed misfit to the model. Items 4 (P=0.010) and 7 (P=0.043) showed uniform DIF by gender, and item 6 (P=0.012) exhibited nonuniform DIF by gender. Items 2 (P=0.010), 4 (P=0.020), 7 (P<0.001), and 10 (P=0.028) showed uniform DIF by age, and item 7 (p<0.001) showed nonuniform DIF by age.	-	292 patients (overall sample, 146 psoriasis, 146 eczema) 146 patients (eczema sample)	good
Conclusion: One study assessed structural validity of the UK version of the DLQI and indicated inadequate structural validity as a QoL instrument for eczema ➔ Structural validity of the English DLQI (UK) as a measurement of QoL: **inadequate** ➔ Quality of evidence: **good**					

Table E8: Convergent/divergent validity of the English DLQI (UK)

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E5)	Correlation between DLQI and PGI and individual DLQI questions were calculated. The mean PGI scores of those who scored 0 on items of the DLQI were compared, using a t-test, with the scores of those who scored 1-3 in each item of the DLQI. Calculation of correlations between the DLQI and the costs of eczema; hypothesis: patients with poor QoL incur high total costs, health service costs and personal costs	Total correlation between DLQI and PGI -0.52 ($p<0.001$) --> positive rating; Questions 1-5 of the DLQI were correlated with the PGI but only question 2 had a correlation of >0.5. Questions 6-10 were not statistically significant correlated. Correlations between DLQI and total costs -0.34 ($p<0.01$); correlation between DLQI and health service costs -0.47 ($p<0.001$); no correlation with personal costs Positive rating because correlation with a QoL measure (PGI) is higher than these correlations	+	56 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the UK version of the DLQI and indicated adequate convergent validity as QoL instrument for eczema ➔ Convergent/divergent validity of the English DLQI (UK) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E9: Discriminative validity of the English DLQI (UK)

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E6)	Comparison of DLQI scores between patients with atopic eczema, pruritus and psoriasis with patients with acne, basal cell carcinoma and viral warts	Scores for patients with atopic eczema, generalized pruritus and psoriasis were higher than for patients with acne, basal cell carcinoma and viral warts (P<0.001)	+	13 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the UK version of the DLQI, but due to poor methodological study quality no conclusion can be drawn ➔ Discriminative validity of the English DLQI (UK) as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

4. Dermatology Life Quality Index (DLQI) – German version

Table E10: Discriminative validity of the German DLQI

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E7)	Discriminative validity: Comparison of mean and subscale scores between patients with psoriasis and AD; t-test to determine statistical significance	Differences in mean score statistically significant (p<0.01); Differences in all subscale scores statistically significant except for leisure/sport and relationships	?	80 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the German DLQI, but due to poor methodological study quality no conclusion can be drawn ➔ Discriminative validity of the German DLQI as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

5. Dermatology Life Quality Index (DLQI) – Spanish version

Table E11: Reliability of the Spanish DLQI

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E8)	Test retest using Intraclass Correlation Coefficient (ICC) between two administrations	ICC between the two administrations was 0.77 (95% CI) for eczema patients	+	45 patients with eczema	fair
Conclusion: One study assessed reliability of the Spanish DLQI and indicated adequate reliability as a QoL instrument for eczema ➔ Reliability of the Spanish DLQI as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E12: Responsiveness of the Spanish DLQI

Author	Responsiveness				
	Method	Result	Interpretation	Study base	COSMIN score
(E8)	Change in scores over three visits after starting TCS	V1 = 4.53, V2 = 2.80, V3 = 1.64. Change between V1 and V3 was statistically significant (p=<0.001); change between V1 and V2 not statistically significant	?	69 patients with eczema	fair
(E8)	Sensitivity to change - effect size (ES) statistic	ES for change in overall DLQI score between visits 1 and 3 was 0.82.	+	69 patients with eczema	fair
Conclusion: One study assessed responsiveness of the Spanish DLQI and indicated adequate responsiveness as a QoL instrument for eczema ➔ Responsiveness of the Spanish DLQI as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

6. Freiburg Life Quality Assessment – core module (FLQA-c)

Table E13: Convergent/divergent validity of the FLQA-c

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E9)	FLQA scores compared to SCORAD severity scores using Pearson correlation coefficient	Low and moderate correlations between severity score and FLQA scales; between r = 14 and r = 34 in atopic dermatitis patients (p108, 2nd column)	?	253 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the FLQA-c and indicated unclear convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the FLQA-c as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E14: Discriminative validity of the FLQA-c

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E9)	Comparison of scores between AD and psoriasis patients (ANOVA for independent samples)	Differences between AD and psoriasis patients statistically significant (p<0.001) for 5/6 subscales	?	253 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the FLQA-c and indicated unclear discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the FLQA-c as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E15: Responsiveness of the FLQA-c

Author	Responsiveness				
	Method	Result	Interpretation	Study base	COSMIN score
(E9)	Comparison of patient scores after 4 weeks of treatment (paired t-test)	Changes in scores on all subscales statistically significant (p<0.001) for AD patients	+	Number of AD patients unknown	poor
Conclusion: One study assessed responsiveness of the FLQA-c, but due to poor methodological study quality no conclusion can be drawn ➔ Responsiveness of the FLQA-c as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

7. Freiburg Life Quality Assessment for Dermatoses (FLQA-d)

Table E16: Discriminative validity of the FLQA-d

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E7)	Comparison of subscale scores between patients with psoriasis and AD; t-test to determine statistical significance	Differences in all subscale scores statistically significant ($p<0.01$) except for social life and treatment --> 4/6 statistically significant different --> indeterminate rating (in contrast to DLQI no data on mean scores)	?	80 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the FLQA-d, but due to poor methodological study quality no conclusion can be drawn ➔ Discriminative validity of the FLQA-d as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

8. Impact of Chronic Skin Disease on Daily Life (ISDL)

Table E17: Internal consistency of the ISDL

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E10)	Cronbach's alpha but poorly described	Ranged from 0.64 - 0.93	+	128 patients with eczema	poor
Conclusion: One study assessed internal consistency of the ISDL, but due to poor methodological study quality no conclusion can be drawn ➔ Internal consistency of the ISDL as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E18: Content validity of the ISDL

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E10)	Items based on literature had to be relevant for the construct to be measured; health professionals and patients with chronic skin diseases evaluated the initial item pool, resulting in 30 eligible items	Normal distributions of all items in pilot study	?	Item generation: unknown Pilot study: 65 psoriasis and 77 AD patients	poor
Conclusion: One study assessed content validity of the ISDL, but due to poor methodological study quality no conclusion can be drawn ➔ Content validity of the ISDL as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E19: Convergent/divergent validity of the ISDL

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E10)	Convergent validity of ISDL assessed with patients rating of disease activity on a 4-point Likert scale (extent and severity of skin involvement of main disease characteristics for each body area), DLQI, anxiety scale (SCL), depression scale (SCL) and neuroticism scale (EPQ). Calculated Pearson's correlation coefficient.	Too many individual results to list. Moderate (0.30-0.50) to relatively high (>0.50) correlations in expected directions. More correlations <0.5 than above 0.5, see table 3 in paper.	-	128 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the ISDL and indicated inadequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the ISDL as a measurement of QoL: **inadequate** ➔ Quality of evidence: **fair**					

Table E20: Discriminative validity of the ISDL

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E10)	Comparison of scores of AD and psoriasis patients	AD patients had significantly higher scores for itch (t=3.27, p<0.001), scratch response (conscious t=4.95, p<0.001; automatic t=6.40, p<0.001) and daily-life impact (t=4.14, p<0.001); differences in scores on all other subscales not statistically significant	?	128 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the ISDL, but due to poor methodological study quality no conclusion can be drawn ➔ Discriminative validity of the ISDL as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E21: Responsiveness of the ISDL

Author	Responsiveness				
	Method	Result	Interpretation	Study base	COSMIN score
(E10)	Effect study of 5-session of cognitive behavioural group therapy programme where patients learn to cope with itch and reduce scratching to assess sensitivity to change.	Physical functioning: skin status (t=3.85), itch (t=5.07), conscious scratching (t=5.47), automatic scratching (t=4.80) - all p<0.001, pain (t=3.62, p<0.01), fatigue (t=1.89, p<0.07). Daily life impact: t=4.31, p<0.001, helplessness (t=2.70, p<0.01), acceptance (t= -3.52, p<0.01), perceived benefits (t= -3.59, p<0.01), anxiety (t=2.43, p=0.02). No significant changes for negative and positive mood, stigmatization and social support. So 11/16 showed some correlation.	?	49 patients with eczema	poor
Conclusion: One study assessed responsiveness of the ISDL, but due to poor methodological study quality no conclusion can be drawn ➔ Responsiveness of the ISDL as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

9. Quality of Life Index for Atopic Dermatitis (QoLIAD) – Dutch version

Table E22: Internal consistency of the Dutch QoLIAD

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.88 (time 1) and 0.89 (time 2)	+	39 patients with eczema	poor
Conclusion: One study assessed internal consistency of the Dutch QoLIAD, but due to poor methodological study quality no conclusion can be drawn ➔ Internal consistency of the Dutch QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E23: Reliability of the Dutch QoLIAD

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.80	?	17 patients with eczema	poor
Conclusion: One study assessed reliability of the Dutch QoLIAD, but due to poor methodological study quality no conclusion can be drawn ➔ Reliability of the Dutch QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E24: Content validity of the Dutch QoLIAD

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Interviews (15 NL, 65 total) to explore the effect AD has on the patient to generate wording for items. Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42. Final version had 25 items - fit to Rasch model. Local dependency between items was minimal - minimal item redundancy	+	Item generation and selection: 15 patients with eczema Field testing: 20 patients with eczema	good
Conclusion: One study assessed content validity of the Dutch QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the Dutch QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

Table E25: Convergent/divergent validity of the Dutch QoLIAD

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.79 (time 1) and 0.58 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.63 (time 1) and 0.47 (time 2).	+	39 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the Dutch QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the Dutch QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E26: Discriminative validity of the Dutch QoLIAD

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Dutch measure was not statistically significant for all 3 assessment groups. May be due to small sample size in Netherlands.	-	39 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the Dutch QoLIAD and indicated inadequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the Dutch QoLIAD as a measurement of QoL: **inadequate** ➔ Quality of evidence: **fair**					

10. Quality of Life Index for Atopic Dermatitis (QoLIAD) – English version (UK)

Table E27: Internal consistency of the English QoLIAD (UK)

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.91 (time 1) and 0.94 (time 2)	+	269 patients with eczema	fair
(E12)	Internal consistency was assessed using Person Separation Index (PSI)	The PSI given in table 2 indicate there is a good level of internal reliability as they were greater than 0.7 (0.91 for initial fit of QoLIAD and 0.82 when 2 items removed).	+	146 patients with eczema	good
Conclusion: Two studies assessed internal consistency of the UK version of the QoLIAD and indicated adequate internal consistency as a QoL instrument for eczema ➔ Internal consistency of the English QoLIAD (UK) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair** to **good**					

Table E28: Reliability of the English QoLIAD (UK)

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.86	?	269 patients with eczema	fair
Conclusion: One study assessed reliability of the UK version of the QoLIAD and indicated unclear reliability as a QoL instrument for eczema ➔ Reliability of the English QoLIAD (UK) as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E29: Content validity of the English QoLIAD (UK)

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Interviews (36 UK, 65 total) to explore the effect AD has on the patient to generate wording for items. Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42 items. Final version had 25 items - fit to Rasch model. Local dependency between items was minimal - minimal item redundancy	+	Item generation and selection: 36 patients with eczema Field testing: 21 patients with eczema	good
Conclusion: One study assessed content validity of the UK version of the QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the English QoLIAD (UK) as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

Table E30: Structural validity of the English QoLIAD (UK)

Author	Structural validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E12)	Examination of overall fit to the Rasch model by reference to the overall item-trait interaction χ^2-fit value and via Item and Person interaction statistics; examination of DIF by ANOVA of standardized residuals.	QoLIAD fits the Rasch model (item-trait interaction = 0.28), although there is evidence for marginal multidimensionality. No clear item misfit found. Authors do not refer to DIF in the results section (except for disease, but not statement whether DIF was uniform or non-uniform)	?	146 patients with eczema	good
Conclusion: One study assessed structural validity of the UK version of the QoLIAD and indicated unclear structural validity as a QoL instrument for eczema ➔ Structural validity of the English QoLIAD (UK) as a measurement of QoL: **unclear** ➔ Quality of evidence: **good**					

Table E31: Convergent/divergent validity of the English QoLIAD (UK)

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.69 (time 1) and 0.77 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.55 (time 1) and 0.55 (time 2).	+	269 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the UK version of the QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the English QoLIAD (UK) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E32: Discriminative validity of the English QoLIAD (UK)

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Differences in scores for all 3 assessment groups statistically significant ($p<0.001$ for severity and general health, $p<0.01$ for flares)	+	269 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the UK version of the QoLIAD and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the English QoLIAD (UK) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

11. Quality of Life Index for Atopic Dermatitis (QoLIAD) – English version (US)

Table E33: Internal consistency of the English QoLIAD (US)

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.93 (time 1) and 0.92 (time 2)	+	170 patients with eczema	fair
Conclusion: One study assessed internal consistency of the US version of the QoLIAD and indicated adequate internal consistency as a QoL instrument for eczema ➔ Internal consistency of the English QoLIAD (US) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E34: Reliability of the English QoLIAD (US)

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.90	?	170 patients with eczema	fair
Conclusion: One study assessed reliability of the US version of the QoLIAD and indicated unclear reliability as a QoL instrument for eczema ➔ Reliability of the English QoLIAD (US) as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E35: Content validity of the English QoLIAD (US)

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42 items Final version had 25 items - fit to Rasch model. Local dependency between items was minimal - minimal item redundancy	+	Item generation and selection: not described Field testing: 20 patients with eczema	good
Conclusion: One study assessed content validity of the US version of the QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the English QoLIAD (US) as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

Table E36: Convergent/divergent validity of the English QoLIAD (US)

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.74 (time 1) and 0.75 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.55 (time 1) and 0.67 (time 2).	+	170 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the US version of the QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the English QoLIAD (US) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E37: Discriminative validity of the English QoLIAD (US)

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Differences in scores for all 3 assessment groups statistically significant (p<0.001 for severity and general health, p<0.01 for flares)	+	170 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the US version of the QoLIAD and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Convergent validity of the English QoLIAD (US) as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

12. Quality of Life Index for Atopic Dermatitis (QoLIAD) – French version

Table E38: Internal consistency of the French QoLIAD

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.90 (time 1) and 0.92 (time 2)	+	200 patients with eczema	fair
Conclusion: One study assessed internal consistency of the French QoLIAD and indicated adequate internal consistency as a QoL instrument for eczema ➔ Internal consistency of the French QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E39: Reliability of the French QoLIAD

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.89	?	200 patients with eczema	fair
Conclusion: One study assessed reliability of the French QoLIAD and indicated unclear reliability as a QoL instrument for eczema ➔ Reliability of the French QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E40: Convergent/divergent validity of the French QoLIAD

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.65 (time 1) and 0.71 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.63 (time 1) and 0.66 (time 2).	+	200 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the French QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the French QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E41: Discriminative validity of the French QoLIAD

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Differences in scores for all 3 assessment groups statistically significant (p<0.001)	+	200 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the French QoLIAD and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the French QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

13. Quality of Life Index for Atopic Dermatitis (QoLIAD) – German version

Table E42: Internal consistency of the German QoLIAD

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.91 (time 1) and 0.92 (time 2)	+	178 patients with eczema	fair
Conclusion: One study assessed internal consistency of the German QoLIAD and indicated adequate internal consistency as a QoL instrument for eczema ➔ Internal consistency of the German QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E43: Reliability of the German QoLIAD

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.86	?	178 patients with eczema	fair
Conclusion: One study assessed reliability of the German QoLIAD and indicated unclear reliability as a QoL instrument for eczema ➔ Reliability of the German QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E44: Content validity of the German QoLIAD

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42 items. Final version had 25 items - fit to Rasch model. Local dependency between items was minimal - minimal item redundancy	+	Item generation and selection: not described Field testing: 17 patients with eczema	good
Conclusion: One study assessed content validity of the German QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the German QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

Table E45: Convergent/divergent validity of the German QoLIAD

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.70 (time 1) and 0.73 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.64 (time 1) and 0.68 (time 2).	+	178 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the German QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the German QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E46: Discriminative validity of the German QoLIAD

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Differences in scores for all 3 assessment groups statistically significant (p<0.001)	+	178 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the German QoLIAD and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the German QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

14. Quality of Life Index for Atopic Dermatitis (QoLIAD) – Italian version

Table E47: Content validity of the Italian QoLIAD

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Interviews (14 Italy, 65 total) to explore the effect AD has on the patient to generate wording for items. Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42 items.	+	Item generation and selection: 14 patients with eczema Field testing: 15 patients with eczema	excellent
Conclusion: One study assessed content validity of the Italian QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the Italian QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **excellent**					

15. Quality of Life Index for Atopic Dermatitis (QoLIAD) – Spanish version

Table E48: Internal consistency of the Spanish QoLIAD

Author	Internal consistency				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Cronbach's coefficient	0.88 (time 1) and 0.90 (time 2)	+	80 patients with eczema	poor
Conclusion: One study assessed internal consistency of the Spanish QoLIAD, but due to poor methodological study quality no conclusion can be drawn ➔ Internal consistency of the Spanish QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E49: Reliability of the Spanish QoLIAD

Author	Reliability				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Test-retest. Patients completed the QoLIAD twice, 2 weeks apart.	Spearman's correlation coefficient = 0.88	?	80 patients with eczema	fair
Conclusion: One study assessed reliability of the Spanish QoLIAD and indicated unclear reliability as a QoL instrument for eczema ➔ Reliability of the Spanish QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

Table E50: Content validity of the Spanish QoLIAD

Author	Content validity				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Tested for cultural applicability across countries. Patients completed questionnaire and interviewed to identify and remove problematic items. Field testing to further reduce items.	All needs affected by AD identified (not listed here) resulting in 76 item scale. 20 removed and 11 modified after cross cultural validation to yield a 56 item version for field testing. At the field testing stage 14 items were removed and two modified leaving 42. Final version had 25 items - fit to Rasch model. Local dependency between items was minimal - minimal item redundancy	+	Item generation and selection: not described Field testing: 20 patients with eczema	good
Conclusion: One study assessed content validity of the Spanish QoLIAD and indicated adequate content validity as a QoL instrument for eczema ➔ Content validity of the Spanish QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

Table E51: Convergent/divergent validity of the Spanish QoLIAD

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Patients completed the QoLIAD, DLQI and PGWB/GWBI (general) twice, 2 weeks apart. Assessed correlation between scales. Ranges predicted 0.6-0.8 for DLQI and 0.5-0.7 for PGWB/GWBI using Spearman's rank correlation coefficients	Correlations between QoLIAD and DLQI 0.76 (time 1) and 0.75 (time 2). Correlations between QoLIAD and PGWB/GWBI 0.79 (time 1) and 0.76 (time 2).	+	80 patients with eczema	fair
Conclusion: One study assessed convergent/divergent validity of the Spanish QoLIAD and indicated adequate convergent validity as a QoL instrument for eczema ➔ Convergent/divergent validity of the Spanish QoLIAD as a measurement of QoL: **adequate** ➔ Quality of evidence: **fair**					

Table E52: Discriminative validity of the Spanish QoLIAD

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E11)	Assessed ability of scale to discriminate between i) patient perceived severity (mild / moderate and quite/very severe AD, ii) flare or no flare and iii) patient perceived general health (excellent, good, fair or poor)	Spanish measure was not statistically significant for flare. Differences on scores between the two other assessment groups were statistically significant (p<0.001)	?	80 patients with eczema	fair
(E13)	Calculated differences in scores, compared QoLIAD and body parts affected (face/hands, face, hands), QoLIAD and treatment because of the symptoms; tested for statistical significance using Mann-Whitney-U and Kruskal-Wallis test; no hypotheses	QoLIAD and body parts: p=0.004 for face affected, p=0.114 for face/hands, p=0.052 for hands --> QoLIAD could distinguish patients whose face was affected QoLIAD and treatment: p=0.392 1/4 statistically significant	?	79 patients with eczema	poor
Conclusion: One study assessed discriminative validity of the Spanish QoLIAD and indicated unclear discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the Spanish QoLIAD as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

16. Skindex-29 – English version (US)

Table E53: Convergent/divergent validity of the English Skindex-29 (US)

Author	Hypothesis testing (convergent/divergent validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E14)	Determination of correlations between scores on the instrument and physician's judgment of severity of the skin disease using Pearson's correlation coefficients	Significant correlation with the emotion scale (r=0.29, P<0.01); correlations for the two other scales not statistically significant	?	102 patients with eczema	poor
Conclusion: One study assessed convergent/divergent validity of the English Skindex-29 (US), but due to poor methodological study quality no conclusion can be drawn ➔ Convergent/divergent validity of the English Skindex-29 (US) as a measurement of QoL: **unclear** ➔ Quality of evidence: **poor**					

Table E54: Discriminative validity of the English Skindex-29 (US)

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E14)	Comparison of scales scores of eczema patients with patients with isolated lesions (benign growths, nonmelanoma skin cancer) using the Wilcoxon rank-sum test Hypothesis: Patients with inflammatory dermatoses would have higher scale scores than patients with isolated lesions	Mean scores of patients with eczema were significantly higher than those with benign skin lesions or nonmelanoma skin cancer (P<0.001) for all 3 subscales	+	102 patients with eczema	good
Conclusion: One study assessed discriminative validity of the English Skindex-29 (US) and indicated adequate discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the English Skindex-29 (US) as a measurement of QoL: **adequate** ➔ Quality of evidence: **good**					

17. Skindex-29 – German version

Table E55: Discriminative validity of the German Skindex-29

Author	Hypothesis testing (discriminative validity)				
	Method	Result	Interpretation	Study base	COSMIN score
(E15)	Pearson correlation coefficients for each Skindex-29 subscale with EASI scores	Correlation between EASI/Skindex-29: functioning 0.73, emotion 0.61, symptoms 0.72 (all statistically significant)	?	13 patients with eczema	poor
(E15)	Pearson correlation coefficients for each Skindex-29 subscale with self-ratings of skin symptoms, itch and sleep disturbance	Correlation between patient ratings of severity and Skindex-29: functioning 0.54-0.59, emotion 0.35-0.40, symptoms 0.62-0.71 (all statistically significant)	?	63 patients with eczema	fair
Conclusion: One study assessed discriminative validity of the German Skindex-29 and indicated unclear discriminative validity as a QoL instrument for eczema ➔ Discriminative validity of the German Skindex-29 as a measurement of QoL: **unclear** ➔ Quality of evidence: **fair**					

References

E1. Schafer T, Staudt A, Ring J. [Development of the German Scale for Assessing Quality of Life in Skin Diseases]. *Hautarzt* 2001;**52**(6):492-498.

E2. Schafer T, Staudt A, Ring J. [German instrument for the assessment of quality of life in skin diseases (DIELH). Internal consistency, reliability, convergent and discriminant validity and responsiveness]. *Hautarzt* 2001;**52**(7):624-628.

E3. Holm EA, Wulf HC, Stegmann H, Jemec GB. Life quality assessment among patients with atopic eczema. *Br J Dermatol* 2006;**154**(4):719-725.

E4. Twiss J, Meads DM, Preston EP, Crawford SR, McKenna SP. Can we rely on the Dermatology Life Quality Index as a measure of the impact of psoriasis or atopic dermatitis? *J Invest Dermatol* 2012;**132**(1):76-84.

E5. Herd RM, Tidman MJ, Ruta DA, Hunter JA. Measurement of quality of life in atopic dermatitis: correlation and validation of two different methods. *Br J Dermatol* 1997;**136**(4):502-507.

E6. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994;**19**(3):210-216.

E7. Augustin M, Zschocke I, Lange S, Seidenglanz K, Amon U. [Quality of life in skin diseases: methodological and practical comparison of different quality of life questionnaires in psoriasis and atopic dermatitis]. *Hautarzt* 1999;**50**(10):715-722.

E8. Badia X, Mascaro JM, Lozano R. Measuring health-related quality of life in patients with mild to moderate eczema and psoriasis: clinical validity, reliability and sensitivity to change of the DLQI. The Cavide Research Group. *Br J Dermatol* 1999;**141**(4):698-702.

E9. Augustin M, Lange S, Wenninger K, Seidenglanz K, Amon U, Zschocke I. Validation of a comprehensive Freiburg Life Quality Assessment (FLQA) core questionnaire and development of a threshold system. *Eur J Dermatol* 2004;**14**(2):107-113.

E10. Evers AW, Duller P, van de Kerkhof PC, van der Valk PG, de Jong EM, Gerritsen MJ, et al. The Impact of Chronic Skin Disease on Daily Life (ISDL): a generic and dermatology-specific health instrument. *Br J Dermatol* 2008;**158**(1):101-108.

E11. Whalley D, McKenna SP, Dewar AL, Erdman RA, Kohlmann T, Niero M, et al. A new instrument for assessing quality of life in atopic dermatitis: international development of the Quality of Life Index for Atopic Dermatitis (QoLIAD). *Br J Dermatol* 2004;**150**(2):274-283.

E12. Twiss J, McKenna SP. Comparing the impact of psoriasis and atopic dermatitis on quality of life: co-calibration of the PSORIQoL and QoLIAD. *Qual Life Res* 2015;**24**(1):105-113.

E13. Lucas Rd, González-Enseñat MA, Giménez Anna, Puig L, Vicente A, Díaz JM, et al. Validación de las versiones españolas de los cuestionarios Parents' Index of Quality of Life in Atopic Dermatitis (PIQoL-AD) y Quality of Life Index for Atopic Dermatitis (QoLIAD): un estudio internacional. *Actas Dermo-Sifiliográficas* 2003;**94**(8):539-548.

E14. 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**(11):1433-1440.

E15. Augustin M, Wenninger K, Amon U, Schroth MJ, Kuster W, Chren M, et al. German adaptation of the Skindex-29 questionnaire on quality of life in dermatology: validation and clinical results. *Dermatology* 2004;**209**(1):14-20.