Received: 2 July 2019 | Revised: 5 August 2019 | Accepted: 22 August 2019

DOI: 10.1111/pai.13120

ORIGINAL ARTICLE

WILEY

Measurement properties of quality-of-life outcome measures for children and adults with eczema: An updated systematic review

Michaela Gabes[1] | Christina Tischer[1] | Christian Apfelbacher[1,2] | for the quality of life working group of the Harmonising Outcome Measures for Eczema (HOME) initiative

[1]Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany

[2]Institute of Social Medicine and Health Economics, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany

Correspondence
Christian Apfelbacher, Medical Sociology, Department of Epidemiology and Preventive Medicine, University of Regensburg, Regensburg, Germany.
Email: christian.apfelbacher@med.ovgu.de

Edited by: Motohiro Ebisawa

Abstract

Objective: The aim of this updated systematic review was to systematically assess the measurement properties of previously discussed and new quality-of-life patient-reported outcome measures (PROMs) in children and adults with eczema using the new COSMIN guideline.

Methods: A systematic literature search was conducted in PubMed and EMBASE. Eligible studies reported on measurement properties of quality-of-life PROMs for children and adults with eczema. The methodological quality of selected already known PROMs and new evidence identified through the literature search was assessed with the COSMIN Risk of Bias checklist. The adequacy of included PROMs was judged with updated quality criteria, and the quality of evidence of the summarized results was graded. Finally, PROMs were placed in a recommendation category (A-C).

Results: In total, 133 measurement properties of nine different PROMs were assessed. No PROM could be placed in category A due to a lack of validation studies. Only the DLQI fulfilled the criteria for category C and therefore should not be recommended for use. All other PROMs were placed in category B, that is, they still have the opportunity to be recommended, but need further validation.

Conclusions: Currently, no PROM for quality of life can be recommended for use in children and adults with eczema. Further validation is needed. The DLQI cannot be recommended for future use.

KEYWORDS

atopic dermatitis, eczema, measurement properties, patient-reported outcome measures, quality of life

Abbreviations: ABS-A, Atopic Dermatitis Burden Scale for Adults; CADIS, Childhood Atopic Dermatitis Impact Scale; CDLQI, Children's Dermatology Life Quality Index; CIAD, Childhood Impact of Atopic Dermatitis; DIELH, *Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen* [German instrument for the assessment of quality of life in skin diseases]; DLQI, Dermatology Life Quality Index; FLQA-c, Freiburg Life Quality Assessment core module; FLQA-d, Freiburg Life Quality Assessment for Dermatoses; IDQoL, Infants' Dermatitis Quality of Life Index; InToDermQoL, Infants and Toddlers Dermatology Quality of Life; ISDL, Impact of Chronic Skin Disease on Daily Life; QoLIAD, Quality of Life Index for Atopic Dermatitis.

1 | BACKGROUND

Eczema (also called atopic dermatitis or atopic eczema) is a chronic inflammatory skin condition that affects up to 25% of children and 2%-3% of adults. Being one of the most common symptoms, pruritus is responsible for much of the skin disease burden.[1] Eczema has profound impacts on the quality of life (QoL) of both affected children, their families, and adults.[2] Quality of life is one of the four core outcome domains defined by the Harmonising Outcome Measures for Eczema (HOME) initiative (www.homeforeczema.org) and should be measured and reported in every clinical trial.[3] Quality of life is measured by self- or proxy-reported questionnaires, referred to patient-reported outcome measures (PROMs). PROMs are used in clinical trials to reflect the patient's perspective.[4] In 2016, Heinl and colleagues published two systematic reviews, one for adults[5] and one for children,[6] in which measurement properties of quality-of-life instruments for eczema were assessed. To this end, the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist[7-9] was used, as well as content comparison, adequacy criteria, best evidence synthesis, and four degrees of recommendation (A-D) based on the three criteria of the OMERACT (Outcome Measures in Rheumatology) filter (ie, truth, discrimination, and feasibility).[10] Since those systematic reviews of Heinl et al,[5,6] the methodology of the assessment of measurement properties of patient-reported outcome measures has been further developed. The new COSMIN Risk of Bias checklist was published in 2017 exclusively for use in systematic reviews. A new ordering of the measurement properties and new labels for the rating system were established, and some standards on missing data, on the sample size, as well as the translation process, were removed.[11] Criteria for good measurement properties were updated, a grading of the quality of evidence was established, and degrees of recommendation was formulated. Altogether, a whole guideline for systematic reviews of patient-reported outcome measures was developed.[12] For this reason, we looked at those previously discussed PROMs again using this new approach including the recently developed COSMIN Risk of Bias checklist, updated criteria for good measurement instruments, the GRADE approach, and a new categorization of the recommendation categories (A-C) proposed by the COSMIN group.[12] Furthermore, the initial systematic reviews were updated by rerunning the systematic literature search and applying the new approach to the methodological assessment of the new evidence.

Key Message

This updated systematic review gives an overview of all currently existing quality-of-life patient-reported outcome measures for children and adults with atopic dermatitis. An assessment of the methodological quality of those instruments is used to formulate recommendations for future use.

2 | MATERIALS AND METHODS

2.1 | Literature search

A systematic, librarian-assisted literature search in PubMed and EMBASE was performed on January 22, 2019. The initial search strategies (for children and adults) of Heinl and Apfelbacher[13,14] were merged, and the search string for PubMed is shown in detail in Appendix 1.

2.2 | Eligible studies

Eligible studies reported on dermatology- or disease-specific QoL instruments for children or adults with eczema. We only included full-text articles about the development and/or validation of measurement properties. In case of mixed patient samples, at least 50% of the patients had to be eczema patients or subgroup analyses for eczema patients were available.

2.3 | Study selection

Two independent reviewers judged titles and abstracts found in the literature search. Furthermore, the same reviewers applied the eligibility criteria to the relevant abstracts. In case of disagreement, consensus was reached by discussion within the research team.

2.4 | Data extraction

Data of previously discussed PROMs and data of the new evidence since 2015 were extracted by two independent reviewers. The selection of previously discussed instruments was based on decisions of the HOME initiative and aspects on feasibility (eg, PROMs which were not free of charge were excluded). Regarding the first systematic review by Heinl et al[5] on adult QoL PROMs, only articles reporting on the DLQI (Dermatology Life Quality Index) and Skindex were used for data extraction. PROMs, such as DIELH (*Deutsches Instrument zur Erfassung der Lebensqualität bei Hauterkrankungen* [German instrument for the assessment of quality of life in skin diseases]),[15,16] FLQA-c (Freiburg Life Quality Assessment core module),[17] FLQA-d (Freiburg Life Quality Assessment for Dermatoses),[18] and ISDL (Impact of Chronic Skin Disease on Daily Life),[19] were not further considered for a potential recommendation since they were formerly placed in category C or D,[5] and they were not considered as candidate instruments by the HOME initiative.[20] Articles reporting on the QoLIAD (Quality of Life Index for Atopic Dermatitis)[21] were not further assessed since this PROM is not available free of charge by the authors and, therefore, it cannot be recommended for use due to reasons of feasibility. For the same reason, the CIAD (Childhood Impact of Atopic Dermatitis)[22] for children was not further assessed. Thus, regarding the second systematic review by Heinl et al (2016)[13] on children with eczema, articles reporting

TABLE 1 Boxes of the COSMIN Risk of Bias checklist

Box 1	PROM development	
Box 2	Content validity	Content validity
Box 3	Structural validity	Internal structure
Box 4	Internal consistency	
Box 5	Cross-cultural validity\ Measurement invariance	
Box 6	Reliability	Remaining measurement properties
Box 7	Measurement error	
Box 8	Criterion validity	
Box 9	Hypotheses testing for construct validity	
Box 10	Responsiveness	

on the IDQoL (Infants' Dermatitis Quality of Life Index), CDLQI (Children's Dermatology Life Quality Index), CADIS (Childhood Atopic Dermatitis Impact Scale), and DISABKIDS were used for data extraction.

In a first step, the recently developed COSMIN Risk of Bias checklist was used to assess the methodological quality of the measurement properties of the single studies.[11] The 10 boxes of the COSMIN Risk of Bias checklist are presented in Table 1. Content validity is considered to be the most important measurement property since it should be clear that all items of a PROM are relevant, comprehensive, and comprehensible regarding the construct of interest and the target population.[23]

In a second step, the quality of the measurement properties was assessed using updated criteria for good measurement properties (based on Terwee et al[24]) to see whether the respective measurement property of the PROM is sufficient (+), insufficient (–), or indeterminate (?).[25]

In a third step, the quality of evidence was graded using the GRADE approach. If there were concerns about the trustworthiness of a result, the quality of evidence of the summarized results was downgraded per measurement property per PROM. Downgrading was possible for risk of bias, inconsistency, imprecision, and/or indirectness. The quality of evidence was judged as either high, moderate, low, or very low. If an overall rating was indeterminate or inconsistent, no grading of the quality of evidence was given.[25]

Finally, each PROM was placed in a recommendation category according to its adequacy and quality of evidence. The COSMIN group proposed three categories of recommendation.[12] A PROM is placed in category A if there is sufficient content validity (any level) and at least low-quality evidence for sufficient internal consistency. A PROM is placed in category C if there is high-quality evidence for an insufficient measurement property. PROMs that can be neither categorized in A nor in C are placed in category B. PROMs in category A can be recommended for use, and results obtained with these PROMs can be seen as trustworthy. PROMs in category B need further validation; however, they still have the opportunity to be recommended for use. PROMs in category C should not be recommended for use. If only PROMs of category B are found, the PROM with the best evidence for content validity can be preliminarily recommended for use, until further evidence is given.[25]

3 | RESULTS

3.1 | Literature search

The systematic literature search identified 1944 records in PubMed and 1035 records in EMBASE. After deduplication, 2503 titles and abstracts were screened. Of 41 screened full texts, 10 papers fulfilled the eligibility criteria and were considered for data extraction (see Figure 1). Three papers reported on new PROMs (after 2015), and seven papers reported on previously discussed PROMs (before 2015).

3.2 | Data extraction

In total, 34 papers (24 papers before 2015 and 10 papers after 2015) reporting on nine PROMs were analyzed. The methodological quality of 133 single studies was assessed (see Figure 2). Almost two thirds of all studies (60.9%) reported on the family of instruments DLQI, CDLQI, and IDQoL. Fifty (37.6%) measurement properties were rated as having very good methodological quality, twelve (9.0%) as having adequate, 48 (36.1%) as having doubtful, and 23 (17.3%) as having inadequate methodological quality.

Five of the nine PROMs were developed for children and four PROMs for adults. Relevant characteristics of the nine different PROMs, such as number of items, recall period, response categories, and target population, are presented in Table 2.

3.3 | Children—Summary of findings of previously discussed PROMs

3.3.1 | IDQoL—Infants Dermatitis Quality of Life[27-35]

The IDQoL showed high-quality evidence for sufficient reliability and construct validity and very-low-quality evidence for sufficient

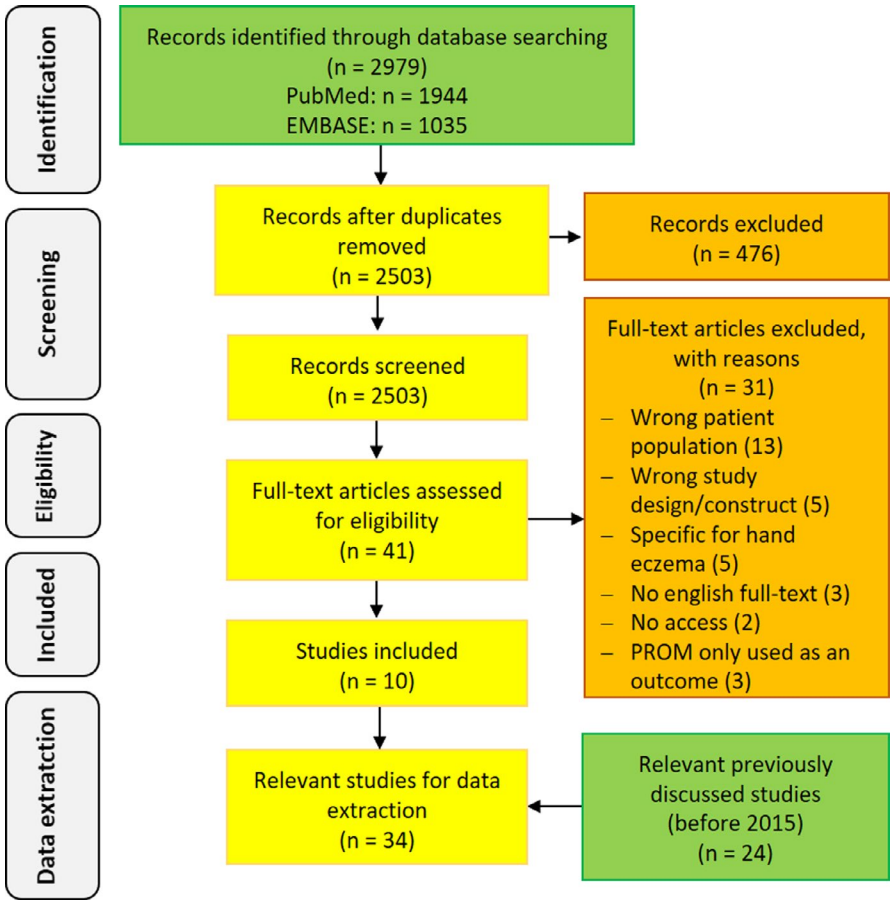

FIGURE 1 Adapted Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Protocols 2009 flow diagram.[26] For more information, visit www.prisma-statement.org

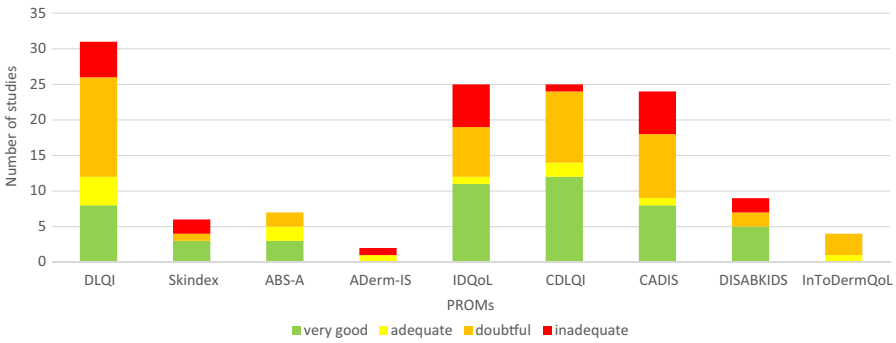

FIGURE 2 Distribution of the quality of the single studies per PROM

content validity and responsiveness. The overall rating for internal consistency was indeterminate since no study on structural validity was available, and therefore, the criteria for "at least low evidence for sufficient structural validity" were not met (see Table 3).

3.3.2 | CDLQI—Children's Dermatology Life Quality Index[31,32,34,36-40]

The quality of evidence for sufficient content validity and construct validity of the CDLQI was moderate. Internal consistency was indeterminate for the same reason as for the IDQoL. No study reported intraclass correlation coefficients (ICCs) or a weighted kappa for reliability; thus, the overall rating was indeterminate (see Table 3).

3.3.3 | CADIS—Childhood Atopic Dermatitis Impact Scale[41-45]

We found low to moderate quality of evidence for sufficient content validity, reliability, and responsiveness of the CADIS. Studies on structural validity of the CADIS were available, but not all information for a sufficient rating was reported. Thus, the overall rating for structural validity was indeterminate and, therefore, the overall rating for internal consistency as well. Regarding hypotheses testing for construct validity, the results did not show a clear pattern. Less than one third of our hypotheses could be confirmed. Since inconsistency could not be resolved, the overall rating was chosen to be inconsistent (see Table 3).

TABLE 2 Characteristics of the included PROMs

	PROMs for children					PROMs for adults			
	IDQoL	**CADIS**	**CDLQI**	**DISABKIDS-ADM**[a]	**InToDermQoL**	**DLQI**	**Skindex (16/29)**	**ABS-A**	**ADerm-IS**
Target population	Children with eczema (0-4 y)	Children with eczema (0-6 y)	Children with skin diseases (4-16 y)	Children/adolescents with eczema (8-16 y)	Children with skin diseases (0-4 y)	Adults with skin diseases	Adults with skin diseases	Adults with eczema	Adults with moderate-to-severe eczema
Recall period	1 wk	4 wk	1 wk	4 wk	1 wk	1 wk	1 wk/4 wk	1 wk	24 h and 1 wk
Completion time[b]	2-3 min	6 min	1-2 min	10-20 min	?	2-8 min	5 min	?	?
Number of items	10	45/43/41/33[c]	10	12	10/12/15	10	16/29	18	3 (daily), 7 (weekly)
Type of response categories	4-point Likert scale	5-point Likert scale	4-point Likert scale	5-point Likert scale	4-point Likert scale	4-point Likert scale	6/7-point Likert scale	6-point Likert scale	11-point Likert scale
Validated language versions	Arabic, Dutch, English, Italian, Swedish, Serbian	English, Italian, Japanese	Danish, English, Malay, Serbian, Spanish, Swedish	At least two European Languages[d], Portuguese	Croatian, German, Greek, Maltese, Polish, Romanian, Ukrainian, French, Danish, Spanish	English, German, Spanish, Danish, Chinese, Ukrainian	English, German, Ukrainian	French	English
Other available language versions	Over 20 languages	Spanish	Over 50 languages	Danish, Dutch, English, French, German, Greek, Norwegian, Swedish	-	Over 90 languages	Over 30 languages	English, Italian, Spanish	-

[a]Proxy-report and self-report version have the same structure.
[b]Information from various sources, times may not be entirely reliable.
[c]Number of items depends on the language version (US-English version: 45 items, Japanese version: 43 items, Italian long-version: 41 items, Italian short-version: 33 items).
[d]Only self-report version. Languages not specified in publication, no information upon request.
? = no information available.

TABLE 3 Summary of findings' table of the IDQoL, CDLQI, CADIS, and DISABKIDS

	Summary or pooled result	Overall rating	Quality of evidence
IDQoL			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (±)	Inconsistent → results based on the majority of results: sufficient	Very low (due to risk of bias and inconsistency)
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Reliability	0.89-0.95	Sufficient	High
Hypotheses testing for construct validity	18 out of 30 hypotheses confirmed	Sufficient	High (since most of the unconfirmed hypotheses came from inadequate studies)
Responsiveness	1 out of 1 hypothesis confirmed, n = 25	Sufficient	Very low (due to risk of bias and imprecision)
CDLQI			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Moderate (due to risk of bias)
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Reliability	ICC or weighted kappa not reported	Indeterminate	–
Hypotheses testing for construct validity	11 out of 21 hypotheses confirmed	Inconsistent → inconsistency resolved (based on the results of very good and adequate studies): sufficient	Moderate (due to inconsistency)
CADIS			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Low (due to risk of bias)
Structural validity	Not all information for a sufficient rating reported	Indeterminate	–
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Reliability	0.89-0.96	Sufficient	Moderate (due to risk of bias)
Hypotheses testing for construct validity	11 out of 38 hypotheses confirmed, 7 out of 37 hypotheses partly confirmed	Inconsistent → inconsistency could not be resolved	–
Responsiveness	1 out of 1 hypothesis confirmed, n = 228	Sufficient	Low (due to risk of bias)
DISABKIDS			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Moderate (due to risk of bias)
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Hypotheses testing for construct validity	1 out of 4 hypotheses confirmed, 1 out of 4 hypotheses partly confirmed	Sufficient	High (unconfirmed hypotheses came from inadequate studies)
InToDermQoL			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Moderate (due to risk of bias)
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Hypotheses testing for construct validity	1 out of 1 hypothesis confirmed	Sufficient	Moderate (due to risk of bias)

3.3.4 | DISABKIDS—Atopic Dermatitis Module[46,47]

There was moderate quality of evidence for sufficient content validity and high quality of evidence for sufficient construct validity of the DISABKIDS. Internal consistency was indeterminate due to a lack of studies on structural validity (see Table 3).

3.4 | Children—Summary of findings of new PROMs since 2015

3.4.1 | InToDermQoL—Infants and Toddlers Dermatology Quality of Life[48]

The InToDermQoL showed moderate quality of evidence for sufficient content validity and construct validity. Internal consistency was rated as indeterminate since no study on structural validity was performed (see Table 3).

3.5 | Adults—Summary of findings of previously discussed PROMs

3.5.1 | DLQI—Dermatology Life Quality Index[18,37,49-55]

The DLQI was the only PROM with high-quality evidence for an insufficient measurement property, namely structural validity. This means there was at least one study of very good quality with a sample size larger than 100 patients available which reported this insufficient measurement property. Thus, internal consistency could not be determined. Even for cross-cultural validity/measurement error, there was moderate quality of evidence for insufficiency, since differential item functioning was found. Content validity, reliability, and responsiveness had a sufficient overall rating (with moderate to high quality of evidence). The results for hypotheses testing showed inconsistent results, and inconsistency could not be resolved (see Table 4).

3.5.2 | Skindex[50,56,57]

The Skindex was the only PROM with sufficient internal consistency since there is one paper available which performed a confirmatory factor analysis on the Skindex and confirmed its structural validity.[58] The quality of evidence was downgraded for one level (from high to moderate) due to imprecision because the total sample size was ≤100. For construct validity, inconsistent results could not be resolved (see Table 4).

3.6 | Adults—Summary of findings of new PROMs since 2015

3.6.1 | ABS-A—Atopic Dermatitis Burden Scale for Adults[59]

The quality of evidence for sufficient content validity and reliability of the ABS-A was low since there was only one study of doubtful quality available. There was high-quality evidence for sufficient construct validity because most of the confirmed hypotheses came from studies of very good quality. The study on structural validity reported not all information for a sufficient rating. Thus, the overall rating of structural validity as well as the overall rating for internal consistency was indeterminate (see Table 4).

3.6.2 | ADerm-IS—Atopic Dermatitis Impact Scale[60]

The development study of the ADerm-IS was published in 2018, and until now, content validity is the only measurement property which has been assessed. This content validity study was the only content validity study with high quality of evidence for sufficient content validity since the study was of adequate quality (see Table 4).

3.7 | Degrees of recommendation

For all included PROMs, there was evidence for sufficient content validity in patients with eczema except for the Skindex. So far, no content validity study had been conducted in our target population. However, Skindex was the only PROM with at least low-quality evidence for sufficient internal consistency because its structural validity had been confirmed via confirmatory factor analysis.[58] Since no PROM fulfilled the two criteria, evidence for sufficient content validity and at least low-quality evidence for sufficient internal consistency, no candidate instrument could be placed in category A. The DLQI was the only instrument with high-quality evidence for an insufficient measurement property and fulfilled therefore the criterion for category C (see Table 5).

All PROMs expect for the DLQI need further validation; however, they still have the opportunity to be placed in category A and therefore recommended for use. The DLQI did not comply with the COSMIN guideline and cannot be recommended for future use due to insufficient measurement properties.

4 | DISCUSSION

This update of two systematic reviews assessed the measurement properties of nine different quality-of-life patient-reported outcome measures for children and adults with eczema. None of the included PROMs could be placed in category A and therefore recommended for use due to a lack of validation studies of good quality. At least one study of very good quality reported insufficient structural validity of the DLQI according to the COSMIN guidelines.[25] As a result, the DLQI was the only PROM for adults which cannot be recommended for future use.

In comparison with the former approach, the new COSMIN guideline for systematic reviews of patient-reported outcome measures[12] seems to be less strict and slightly more sympathetic toward candidate PROMs. In the initial systematic reviews, four degrees of recommendation (A-D) were used in accordance with

TABLE 4 Summary of findings' table of the DLQI and Skindex (the red row indicates the criterion for a placement in category C and therefore no recommendation for use)

	Summary or pooled result	Overall rating	Quality of evidence
DLQI			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Moderate (due to risk of bias)
Structural validity	Misfit/poor fit to the Rasch model, violation of unidimensionality (–)	Insufficient	High
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Cross-cultural validity/measurement error	DIF was found, n = 927	Insufficient	Moderate (due to risk of bias)
Reliability	0.77	Sufficient	Moderate (due to risk of bias)
Hypotheses testing for construct validity	10 out of 17 hypotheses confirmed	Inconsistent → inconsistency could not be resolved	–
Responsiveness	3 out of 3 hypotheses confirmed, n = 611	Sufficient	High
Skindex			
Internal consistency	0.84-0.89; n = 63; at least low evidence for sufficient structural validity[a]	Sufficient	Moderate (due to imprecision)
Hypotheses testing for construct validity	6 out of 11 hypotheses confirmed	Inconsistent → inconsistency could not be resolved	–
ABS-A			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	Low (due to risk of bias)
Structural validity	Not all information for a sufficient rating reported	Indeterminate	–
Internal consistency	Criteria for "at least low evidence for sufficient structural validity" not met	Indeterminate	–
Reliability	0.89	Sufficient	Low (due to risk of bias)
Hypotheses testing for construct validity	3 out of 4 hypotheses confirmed	Sufficient	High (most of the confirmed hypotheses came from very good studies)
ADerm-IS			
Content validity	Relevance (+), comprehensiveness (+), comprehensibility (+)	Sufficient	High

[a]See He et al., 2014, table 9.[58]

the OMERACT filter since there was international consensus that the OMERACT quality criteria need to be met for eczema instruments as well.[14] Almost all PROMs were categorized in category C or D except for the CADIS and the Spanish DLQI which reached category B. To be formerly placed in category A, a PROM had to meet all requirements of the OMERACT filter, such as truth (content validity and construct validity), discrimination (reliability, internal consistency, and sensitivity to change), and feasibility (interpretability and ease of use). This meant a sufficient rating (+) for each measurement property which is almost impossible.[14] With the new COSMIN degrees of recommendation, only sufficient content validity and at least low-quality evidence for sufficient internal consistency are required to be recommended as the most suitable PROM for the construct and population of interest. A PROM is now solely placed in category C if there is high-quality evidence for an insufficient measurement property. Thus, more candidate PROMs have the potential to be recommended if further validation is given. With the new methodological approach, most of the studies (37.6%) were rated as "very good," the best possible rating. In the previous systematic reviews, no measurement property of PROMs for children and only 1% of the measurement properties of PROMs for adults obtained the best possible COSMIN rating. Not only the degrees of recommendations are less strict, but also the COSMIN Risk of Bias checklist is more benevolent toward the single studies.

This study has several strengths and limitations: A strength of this updated systematic review is the fact that the initial search string was used to rerun the literature search. The two initial search strings for adults and children respectively were combined by our

TABLE 5 Degrees of recommendation

PROMs	Category A		Category C	Recommendation
	Evidence for sufficient content validity (any level)	At least low-quality evidence for sufficient internal consistency	High-quality evidence for an insufficient measurement property	
IDQoL	✓	✗	✗	B
CDLQI	✓	✗	✗	B
CADIS	✓	✗	✗	B
DISABKIDS	✓	✗	✗	B
InToDermQoL	✓	✗	✗	B
DLQI	✓	✗	✓	C
Skindex	✗	✓	✗	B
ABS-A	✓	✗	✗	B
ADerm-IS	✓	✗	✗	B

academic librarian. Furthermore, at least two independent reviewers were involved in every single step of the assessment. One reviewer (MG) completed every step of the review process to ensure consistency. Discrepancies were discussed frequently and resolved within the research team. As in the initial systematic reviews, only two databases, PubMed and EMBASE, were searched which can be listed here again as a limitation. Furthermore, reference lists of included studies were not hand-searched and no other small databases were searched for relevant studies.

Despite the loosened requirements for a recommendation, no PROM could be placed in category A. Future validation research should focus on the PROMs of category B and fill in the validation gaps.

ACKNOWLEDGMENTS

We are very grateful to Helge Knüttel, the academic librarian of the University Hospital Regensburg, who revised and reran the initial literature search. We are indebted to Peter Werkmann for the screening of all titles, abstracts, and full texts as a second reviewer.

CONFLICT OF INTEREST

Christian Apfelbacher has received institutional funding from Dr Wolff GmbH and consultancy fees from Dr Wolff GmbH and Sanofi Genzyme. He is a member of the executive committee of the Harmonising Outcome Measures for Eczema (HOME) initiative.

ORCID

Michaela Gabes https://orcid.org/0000-0003-1672-7549
Christian Apfelbacher https://orcid.org/0000-0003-3805-8219

REFERENCES

1. Eichenfield LF, Tom WL, Chamlin SL, et al. Guidelines of care for the management of atopic dermatitis: section 1. Diagnosis and assessment of atopic dermatitis. *J Am Acad Dermatol*. 2014;70(2):338-351.
2. Drucker AM, Wang AR, Li WQ, Sevetson E, Block JK, Qureshi AA. The burden of atopic dermatitis: summary of a report for the National Eczema Association. *J Invest Dermatol*. 2017;137(1):26-30.
3. Schmitt J, Spuls P, Boers M, et al. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. *Allergy*. 2012;67(9):1111-1117.
4. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract*. 2006;12(5):559-568.
5. Heinl D, Prinsen CA, Deckert S, et al. Measurement properties of adult quality-of-life measurement instruments for eczema: a systematic review. *Allergy*. 2016;71(3):358-370.
6. Heinl D, Prinsen C, Sach T, et al. Measurement properties of quality-of-life measurement instruments for infants, children and adolescents with eczema: a systematic review. *Br J Dermatol*. 2017;176(4):878-889.
7. Mokkink LB, Terwee CB, Patrick 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. *Quality Life Res: Int J Quality Life Aspects Treatment Care Rehab*. 2010;19(4):539-549.
8. Mokkink LB, Terwee CB, Patrick 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.
9. 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. *Quality Life Res: Int J Quality Life Aspects Treatment Care Rehab*. 2012;21(4):651-657.
10. Schmitt J, Langan S, Deckert S, et al. Assessment of clinical signs of atopic dermatitis: a systematic review and recommendation. *J Allergy Clin Immunol*. 2013;132(6):1337-1347.
11. Mokkink LB, de Vet H, Prinsen C, COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality Life Res: Int J Quality Life Aspects Treatment Care Rehab*. 2018;27(5):1171-1179.
12. Prinsen C, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality Life Res*. 2018;27(5):1147-1157.
13. Heinl D, Prinsen CA, Drucker AM, et al. Measurement properties of quality of life measurement instruments for infants, children and

adolescents with eczema: protocol for a systematic review. *Syst Rev*. 2016;5:25.
14. Apfelbacher CJ, Heinl D, Prinsen CA, et al. Measurement properties of adult quality-of-life measurement instruments for eczema: protocol for a systematic review. *Syst Rev*. 2015;4:48.
15. 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]. *Der Hautarzt; Zeitschrift fur Dermatologie, Venerologie, und verwandte Gebiete*. 2001;52(7):624-628.
16. Schafer T, Staudt A, Ring J. [Development of the German scale for assessing quality of life in skin diseases]. *Der Hautarzt; Zeitschrift fur Dermatologie, Venerologie, und verwandte Gebiete*. 2001;52(6):492-498.
17. 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: EJD*. 2004;14(2):107-113.
18. 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]. *Der Hautarzt; Zeitschrift fur Dermatologie, Venerologie, und verwandte Gebiete*. 1999;50(10):715-722.
19. Evers AW, Duller P, van de Kerkhof PC, 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.
20. Chalmers JR, Simpson E, Apfelbacher CJ, et al. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol*. 2016;175(1):69-79.
21. Whalley D, McKenna SP, Dewar AL, 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.
22. McKenna SP, Doward LC, Meads DM, Tennant A, Lawton G, Grueger J. Quality of life in infants and children with atopic dermatitis: addressing issues of differential item functioning across countries in multinational clinical trials. *Health Qual Life Outcomes*. 2007;5:45.
23. Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for assessing the content validity of Patient-Reported Outcome Measures (PROMs). User manual. 2017.
24. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
25. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome. Measures (PROMs) - user manual. 2018.
26. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Open Med: Peer-Reviewed Independent Open-Access J*. 2009;3(3):e123-e130.
27. Alzolibani AA. Cultural adaptation of the Arabic version of the Infants' Dermatitis Quality of Life Index. *Saudi Med J*. 2013;34(5):518-524.
28. Alzolibani AA. Impact of atopic dermatitis on the quality of life of Saudi children. *Saudi Med J*. 2014;35(4):391-396.
29. Baranzoni NS, Mantovani LG, De Portu S, Monzini MS, Giannetti A. Validation of the Italian version of the Infants Dermatitis Quality of Life and Family Dermatitis Indexes. *Giornale Italiano di Dermatologia e vererologia*. 2007;142:423-432.
30. Beattie PE, Lewis-Jones MS. An audit of the impact of a consultation with a paediatric dermatology team on quality of life in infants with atopic eczema and their families: further validation of the Infants' Dermatitis Quality of Life Index and Dermatitis Family Impact score. *Br J Dermatol*. 2006;155(6):1249-1255.
31. Boccardi D, D'Auria E, Turati F, et al. Disease severity and quality of life in children with atopic dermatitis: PO-SCORAD in clinical practice. *Minerva Pediatr*. 2017;69(5):373-380.
32. Ganemo A, Svensson A, Lindberg M, Wahlgren CF. Quality of life in Swedish children with eczema. *Acta dermato-venereologica*. 2007;87(4):345-349.
33. Lewis-Jones MS, Finlay AY, Dykes PJ. The Infants' Dermatitis Quality of Life Index. *Br J Dermatol*. 2001;144(1):104-110.
34. Raznatovic Djurovic M, Jankovic J, Tomic Spiric V, Jankovic S. Health-related Quality of Life in Children with Moderate to Severe Atopic Dermatitis. *Acta dermatovenerologica Croatica: ADC*. 2015;23(3):178-184.
35. van Valburg RW, Willemsen MG, Dirven-Meijer PC, Oranje AP, van der Wouden JC, Moed H. Quality of life measurement and its relationship to disease severity in children with atopic dermatitis in general practice. *Acta dermato-venereologica*. 2011;91(2):147-151.
36. Aziah MS, Rosnah T, Mardziah A, Norzila MZ. Childhood atopic dermatitis: a measurement of quality of life and family impact. *Med J Malaysia*. 2002;57(3):329-339.
37. Holm EA, Wulf HC, Stegmann H, Jemec GB. Life quality assessment among patients with atopic eczema. *Br J Dermatol*. 2006;154(4):719-725.
38. Lewis-Jones MS, Finlay AY. The Children's Dermatology Life Quality Index (CDLQI): initial validation and practical use. *Br J Dermatol*. 1995;132(6):942-949.
39. Maksimovic N, Jankovic S, Marinkovic J, Sekulovic LK, Zivkovic Z, Spiric VT. Health-related quality of life in patients with atopic dermatitis. *J Dermatol*. 2012;39(1):42-47.
40. Ramirez-Anaya M, Macias ME, Velazquez-Gonzalez E. Validation of a Mexican Spanish version of the Children's Dermatology Life Quality Index. *Pediatric Dermatol*. 2010;27(2):143-147.
41. Chamlin SL, Cella D, Frieden IJ, et al. Development of the Childhood Atopic Dermatitis Impact Scale: initial validation of a quality-of-life measure for young children with atopic dermatitis and their families. *J Invest Dermatol*. 2005;125(6):1106-1111.
42. Chamlin SL, Frieden IJ, Williams ML, Chren M-M. Effects of atopic dermatitis on young American children and their families. *Pediatrics*. 2004;114(3):607-611.
43. Chamlin SL, Lai JS, Cella D, et al. Childhood Atopic Dermatitis Impact Scale: reliability, discriminative and concurrent validity, and responsiveness. *Arch Dermatol*. 2007;143(6):768-772.
44. Neri E, Agostini F, Gremigni P, et al. Italian validation of the Childhood Atopic Dermatitis Impact Scale: a contribution to its clinical application. *J Invest Dermatol*. 2012;132(11):2534-2543.
45. Yamaguchi C, Futamura M, Chamlin SL, Ohya Y, Asano M. Development of a Japanese Culturally Modified Version of the Childhood Atopic Dermatitis Impact Scale (JCMV-CADIS). *Allergology Int*. 2016;65(3):312-319.
46. Baars RM, Atherton CI, Koopman HM, Bullinger M, Power M. The European DISABKIDS project: development of seven condition-specific modules to measure health related quality of life in children and adolescents. *Health Qual Life Outcomes*. 2005;3(1):70.
47. Deon KC, Santos DM, Bullinger M, Santos CB. Preliminary psycometric assessment of the Brazilian version of the DISABKIDS Atopic Dermatitis Module. *Rev Saude Publica*. 2011;45(6):1072-1078.
48. Chernyshov PV, Boffa MJ, Corso R, et al. Creation and pilot test results of the dermatology-specific proxy instrument: the Infants and Toddlers Dermatology Quality of Life. *J Eur Acad Dermatol Venereol*. 2018;32(12):2288-2294.
49. 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.
50. Chernyshov PV. Health-related quality of life in adult atopic dermatitis and psoriatic patients matched by disease severity. *Giornale*

italiano di dermatologia e venereologia: organo ufficiale, Societa italiana di dermatologia e sifilografia. 2016;151(1):37-43.
51. 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.
52. 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.
53. Liu Y, Li T, An J, Zeng W, Xiao S. Rasch analysis holds no brief for the use of the Dermatology Life Quality Index (DLQI) in Chinese neurodermatitis patients. *Health Qual Life Outcomes*. 2016;14:17.
54. Patel KR, Singam V, Vakharia PP, et al. Measurement properties of three assessments of burden used in atopic dermatitis in adults. *Br J Dermatol*. 2019;180(5):1083-1089.
55. 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.
56. Augustin M, Wenninger K, Amon U, et al. German adaptation of the Skindex-29 questionnaire on quality of life in dermatology: validation and clinical results. *Dermatology (Basel, Switzerland)*. 2004;209(1):14-20.
57. 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.
58. He Z, Lu C, Chren MM, et al. Development and psychometric validation of the Chinese version of Skindex-29 and Skindex-16. *Health Qual Life Outcomes*. 2014;12:190.
59. Taieb A, Boralevi F, Seneschal J, et al. Atopic dermatitis burden scale for adults: development and validation of a new assessment tool. *Acta dermato-venereologica*. 2015;95(6):700-705.
60. Foley C, Tundia N, Simpson E, Teixeira HD, Litcher-Kelly L, Bodhani A. Development and content validity of new patient-reported outcome questionnaires to assess the signs and symptoms and impact of atopic dermatitis: The Atopic Dermatitis Symptom Scale (ADerm-SS) and the Atopic Dermatitis Impact Scale (ADerm-IS). *Curr Med Res Opin*. 2019;35(7):1139-1148.

How to cite this article: Gabes M, Tischer C, Apfelbacher C; for the quality of life working group of the Harmonising Outcome Measures for Eczema (HOME) initiative. Measurement properties of quality-of-life outcome measures for children and adults with eczema: An updated systematic review. *Pediatr Allergy Immunol*. 2020;31:66–77. https://doi.org/10.1111/pai.13120

APPENDIX 1

Search string Medline (via PubMed)

#1: (modified precision search terms by Terwee et al 2009) (instrumentation[sh] OR Validation Studies[pt] OR "reproducibility of results"[MeSH Terms] OR reproducib*[tiab] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tiab] OR clinometr*[tiab] OR "observer variation"[MeSH] OR observer variation[tiab] OR "discriminant analysis" [MeSH] OR reliab*[tiab] OR valid*[tiab] OR coefficient[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR "item correlation"[tiab] OR "item correlations"[tiab] OR "item selection"[tiab] OR "item selections"[tiab] OR "item reduction"[tiab] OR "item reductions"[tiab] OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR intertechnician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab]) OR (intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR "coefficient of variation"[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group" [tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimensionality[tiab] OR subscale*[tiab] OR "multitrait scaling analysis"[tiab] OR "multitrait scaling analyses"[tiab] OR "item discriminant"[tiab] OR "interscale correlation"[tiab] OR "interscale correlations"[tiab]) OR ((error[tiab] OR errors[tiab]) AND (measure*[tiab] OR correlat*[tiab] OR evaluat*[tiab] OR accuracy[tiab] OR accurate[tiab] OR precision[tiab] OR mean[tiab])) OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR "variability analysis"[tiab] OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR (small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "minimal important change"[tiab] OR "minimal important difference"[tiab]) OR ("minimally important change"[tiab] OR "minimally important difference"[tiab] OR "minimal detectable change"[tiab] OR "minimal detectable difference"[tiab] OR "minimally detectable change"[tiab] OR "minimally detectable difference"[tiab] OR "minimal real change"[tiab] OR "minimal real difference"[tiab] OR "minimally real change"[tiab] OR "minimally real difference"[tiab] OR "ceiling effect" [tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item

bank"[tiab] OR "cross-cultural equivalence"[tiab] OR accepta*[tiab] OR "ease of use"[tiab] OR practica*[tiab] OR feasib*[tiab]) #2: Document types to be excluded and animal-only studies; erroneous NOT changed to OR ("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) OR ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])#3:#1 NOT #2#4:(quality of life[MH] OR quality of life[TW] OR health status[MH] OR health status[TW] OR "activities of daily living"[MH] OR activities of daily living[TW] OR life quality* OR daily life[TW]OR health level[TW] OR level of health[TW] OR patient reported outcome[TW] OR CDLQI[TW] OR IDQOL[TW] OR Skindex[TW] OR Eczema Disability Index[TW])#5("dermatitis, atopic"[MeSH] OR atopic dermatitis[tiab] OR atopic eczema[tiab] OR eczema[MeSH] OR eczema[tiab] OR "neurodermatitis"[MeSH] OR Neurodermatitis[tiab] OR skin diseases[MH] OR skin disease*[tiab] OR dermatology[tiab])#6:#3 AND #4 AND #5#7: Only records from 2015 onwards#6 AND (2015:2010[mhda] OR 2015:2020[edat] OR 2015:2020[crdt] OR 2015:2020[dp] OR 2015:2020[epdat] OR 2015:2020[ppdat])