Measurement properties of eczema-specific measures of health-related quality of life: systematic review

HOME III meeting, San Diego, 06-07/04/2013

Dr. sc. hum. Christian Apfelbacher MSc PhD
Institute of Epidemiology and Preventive Medicine
University Hospital Regensburg

Universität Regensburg

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Background

HOME II, Amsterdam 2011: HrQoL added to core set of outcome domains (Schmitt 2012 Allergy)

=> Need to identify appropriate instrument for the measurement of HrQoL

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Quality of life in eczema: domains

Mood, sleep

Clothing

Activities

Relationships

Family life

Education

Employment

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Objectives

i. To *systematically assess* measurement properties of eczema-specific measures of HrQoL
ii. To *identify* outcomes measures for eczema-specific HrQoL
    i. that meet predefined criteria (OMERACT filter)
    ii. have potential to be recommended in the future (depending on further validation studies)
    iii. that do NOT meet predefined criteria and therefore should NOT be used any more
iii. *Provide evidence-base* for further consensus-based standardization and prioritization

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Methods: literature search

- systematic literature search, using Terwees highly sensitive PubMed search filter (Terwee 2009)
- search terms: health-related quality of life/quality of life/health status/bother
- Eligibility criteria:
  - measures specifically designed to measure HrQoL in eczema (children/adults/parents?)
  - papers describe the development and/or validation of an eczema-specific measure of HrQoL
- not eligible: measures that measure HrQoL in skin disease in general (such as DLQI or Skindex)

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Methods: Evidence tables

i. *Characteristics of studies*: reference, geographical location, setting, study type, key characteristics of study subjects, name of measure, domains measured and number of items

ii. Assessment of *methodological quality of studies*: COSMIN checklist, applying worst score method for each of the measurement properties (yielding a quality rating per measurement property per study: “poor”, “fair”, “good”, “excellent”)

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Methods: Quality assessment of measures

- Predefined criteria for rating the quality of measures as recommended by the COSMIN group will be used (Terwee 2007)
- Criteria in accordance with OMERACT filter adopted by HOME (Schmitt 2012) and with previous SR on outcomes measures for AE (Schmitt 2007)

Quality of each measurement property rated as
+: positive rating
?: indeterminate rating
-: negative rating
0: no information available

Table 1
Quality criteria for measurement properties of health status questionnaires

Property	Definition	Quality criteria[a,b]
1. Content validity	The extent to which the domain of interest is comprehensively sampled by the items in the questionnaire	+A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection; ?A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method; −No target population involvement; 0No information found on target population involvement.
2. Internal consistency	The extent to which items in a (sub)scale are intercorrelated, thus measuring the same construct	+Factor analyses performed on adequate sample size (7 * # items and ≥100) AND Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95; ?No factor analysis OR doubtful design or method; −Cronbach's alpha(s) <0.70 or >0.95, despite adequate design and method; 0No information found on internal consistency.
3. Criterion validity	The extent to which scores on a particular questionnaire relate to a gold standard	+Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥0.70; ?No convincing arguments that gold standard is "gold" OR doubtful design or method; −Correlation with gold standard <0.70, despite adequate design and method; 0No information found on criterion validity.
4. Construct validity	The extent to which scores on a particular questionnaire relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured	+Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses; ?Doubtful design or method (e.g., no hypotheses); −Less than 75% of hypotheses were confirmed, despite adequate design and methods; 0No information found on construct validity.
5. Reproducibility		
5.1. Agreement	The extent to which the scores on repeated measures are close to each other (absolute measurement error)	+MIC < SDC OR MIC outside the LOA OR convincing arguments that agreement is acceptable; ?Doubtful design or method OR (MIC not defined AND no convincing arguments that agreement is acceptable); −MIC ≥ SDC OR MIC equals or inside LOA, despite adequate design and method; 0No information found on agreement.
5.2. Reliability	The extent to which patients can be distinguished from each other, despite measurement errors (relative measurement error)	+ICC or weighted Kappa ≥ 0.70; ?Doubtful design or method (e.g., time interval not mentioned); −ICC or weighted Kappa < 0.70, despite adequate design and method; 0No information found on reliability.
6. Responsiveness	The ability of a questionnaire to detect clinically important changes over time	+SDC or SDC < MIC OR MIC outside the LOA OR RR > 1.96 OR AUC ≥ 0.70; ?Doubtful design or method; −SDC or SDC ≥ MIC OR MIC equals or inside LOA OR RR ≤ 1.96 OR AUC < 0.70, despite adequate design and methods; 0No information found on responsiveness.
7. Floor and ceiling effects	The number of respondents who achieved the lowest or highest possible score	+≤15% of the respondents achieved the highest or lowest possible scores; ?Doubtful design or method; −>15% of the respondents achieved the highest or lowest possible scores, despite adequate design and methods; 0No information found on interpretation.
8. Interpretatability	The degree to which one can assign qualitative meaning to quantitative scores	+Mean and SD scores presented of at least four relevant subgroups of patients and MIC defined; ?Doubtful design or method OR less than four subgroups OR no MIC defined; 0No information found on interpretation.

MIC = minimal important change; SDC = smallest detectable change; LOA = limits of agreement; ICC = Intraclass correlation; SD, standard deviation.
[a] + = positive rating; ? = indeterminate rating; − = negative rating; 0 = no information available.
[b] Doubtful design or method = lacking of a clear description of the design or methods of the study, sample size smaller than 50 subjects (should be at least 50 in every (subgroup) analysis), or any important methodological weakness in the design or execution of the study.

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Content validity

+	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection
?	A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method
-	No target population involvement
0	No information found on target population

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Doubtful design or method

"Lack of a clear description of design or methods of study, sample size < 50 (should be >= 50 for every (subgroup) analysis), or any important methodological weakness in design or execution of study" (Terwee 2007)

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Internal consistency

+	Factor analyses performed on adequate sample size (7 * # items and >100) AND Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95
?	No factor analysis OR doubtful design or method
-	Cronbach's alpha(s) < 0.70 or > 0.95, despite adequate design and method
0	No information found on internal consistency

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Criterion validity

+	Convincing arguments that gold standard is ‘‘gold’’ AND correlation with gold standard >0.70
?	No convincing arguments that gold standard is ‘‘gold’’ OR doubtful design or method
-	Correlation with gold standard < 0.70, despite adequate design and method
0	No information found on criterion validity

! There is no gold standard for the measurement of HrQoL !

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Construct validity

+	Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses
?	Doubtful design or method (e.g., no hypotheses);
-	Less than 75% of hypotheses were confirmed, despite adequate design and methods
0	No information found on construct validity.

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Reproducibility / agreement

+	MIC < SDC OR MIC outside the LOA OR convincing arguments that agreement is acceptable
?	Doubtful design or method OR (MIC not defined AND no convincing arguments that agreement is acceptable)
-	MIC >= SDC OR MIC equals or inside LOA, despite adequate design and method
0	No information found on agreement

MIC: minimal important change
SDC: smallest detectable change
LOA: limits of agreement

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Reproducibility / reliability

+	ICC or weighted Kappa >= 0.70
?	Doubtful design or method (e.g., time interval not mentioned)
-	ICC or weighted Kappa < 0.70, despite adequate design and method
0	No information found on reliability

ICC: Intraclass correlation coefficient

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Responsiveness

+	SDC or SDC < MIC OR MIC outside the LOA OR RR > 1.96 OR AUC >= 0.70
?	Doubtful design or method
-	SDC or SDC >= MIC OR MIC equals or inside LOA OR RR <= 1.96 OR AUC < 0.70, despite adequate design and methods;
0	No information found on responsiveness

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Floor and ceiling effects

+	<15% of the respondents achieved the highest or lowest possible scores
?	Doubtful design or method
-	>15% of the respondents achieved the highest or lowest possible scores, despite adequate design and methods
0	No information found on interpretation

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Interpretability

+	Mean and SD scores presented of at least four relevant subgroups of patients and MIC defined
?	Doubtful design or method OR less than four subgroups OR no MIC defined
-	
0	No information found on interpretation

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Best evidence synthesis

Level	Rating	Criteria
Strong	+++ or ---	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or --	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
Unknown	?	Only studies of poor methological quality

Furlan 2009 Spine, Schellingerhout 2012 Quality of Life Research

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Recommendation: Categories

A) Outcome measure meets all requirements to be recommended for use

B) Outcome 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) Outcome measure has low quality in at least one required quality criteria (≥1 rating of "minus") and therefore is not recommended to be used any more.

D) Outcome 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

Universität Regensburg

Dr. sc. hum. Christian Apfelbacher PhD
Medizinische Soziologie
Fakultät Medizin

Required ratings for recommendation

Quality item (name)	Inclusion in OMERACT filter	Required rating for recommendation
Content validity	Truth	+
Construct validity	Truth	+
Internal consistency	Discrimination	+
Test-retest reliability	Discrimination	+
Sensitivity to change / Responsiveness	Discrimination	+
Floor or ceiling effects	Not required by OMERACT filter	Not required for recommendation
Interpretability	Feasibility	+
Acceptability / Ease of use	Feasibility	+