SYSTEMATIC REVIEW

Measurement properties of patient-reported outcome measures for eczema control: a systematic review

B.L. Stuart,¹,* L. Howells,² R.L. Pattinson,³ J.R. Chalmers,² D. Grindlay,² N.K. Rogers,² E. Grinich,⁴ T. Pawlitschek,⁴ E.L. Simpson,⁴ K.S. Thomas²

¹Aldermoor Health Centre, University of Southampton, Southampton, UK
²Centre of Evidence Based Dermatology, School of Medicine, University of Nottingham, Nottingham, UK
³College of Biomedical and Life Sciences, School of Healthcare Sciences, Cardiff University, Cardiff, UK
⁴Oregon Health and Science University, Portland, OR, USA
*Correspondence: Beth Stuart. E-mail: bls1@soton.ac.uk

Abstract
Atopic eczema (herein referred to as ‘eczema’) is a skin disease characterized by remitting and relapsing symptoms. The Harmonising Outcome Measures for Eczema (HOME) initiative was developed to establish a core outcome set (COS) for eczema to be measured for all future eczema trials. The core outcome set for atopic eczema clinical trials includes the domain for patient-reported eczema control, but a review of the validation of available eczema control instruments was lacking. We aimed to review the literature and systematically assess the measurement properties of validated patient-reported outcome instruments that capture eczema control. PubMed and Ovid EMBASE were searched up to 24 January 2020 for any study that reported on PROM instrument development or validation. The COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) criteria were used to assess the quality of eligible studies. We screened 12 036 titles and abstracts and 58 full texts. A total of 12 papers were included, reporting on seven PROMS. These were assessed with respect to development, reliability, construct validity and responsiveness. Two instruments, Recap of Atopic Eczema (RECAP) and the Atopic Dermatitis Control Tool (ADCT), have been developed and validated to a sufficient standard to support their recommendation as patient-reported outcome instruments for measuring control of atopic eczema as part of the HOME Core Outcome Set.

Introduction
Rationale
Atopic eczema (herein referred to as ‘eczema’) is a common chronic, inflammatory skin disease characterized by itching and dry skin. It typically develops in children aged two years or under, but adult onset may also occur.

Many people experience relapsing and remitting symptoms, with periods of ‘flare’, during which their eczema worsens.¹

Given this episodic nature, it is important to capture whether patients are able to get and maintain control of their disease.

The Harmonising Outcome Measures in Eczema (HOME) initiative aims to create a Core Outcome Set to be used in all eczema clinical trials. During the HOME II consensus meeting in Amsterdam in 2011, long-term control of eczema was included through consensus vote as one of the four domains deemed important to measure in all trials of eczema. Additional domains included clinician-reported signs, patient-reported...
symptoms and quality of life.² Multiple HOME meetings subse-
sequently took place, but challenges arose regarding how eczema
control should be defined.

An international qualitative study including patients, parents
and clinicians suggests eczema control is a multifaceted con-
struct involving changes in disease activity, the treatment and
management of the condition and psychological, social and
physical functioning.³ It was agreed by consensus at HOME V
that the long-term control domain represents something in addition
to repeated measures of signs, symptoms, quality of life and recom-
manded addition of a patient-reported global instrument that cap-
tures eczema control.⁴ This review sought to identify suitable
instruments for capturing 'eczema control'. The HOME domain of
long-term control is conceptualized as repeated measurement of
eczema control over time, in addition to the other core outcome set
domains.

A previous systematic review was conducted to explore which
strategies were available for measuring the HOME domain of
long-term control.⁶ However, that review was not specific to
eczema control and did not assess the methodological qualities
of the included measurement tools. The current review was
designed to identify instruments suitable for capturing the construct
of 'global eczema control' as recommended at the HOME V meet-
ing and evaluated the measurement properties of the identified
instruments.

Objectives

1 To identify all validated patient-reported outcome instru-
   ments that capture control of eczema.
2 To systematically assess the measurement properties of those
   instruments with respect to validity, consistency, reliability,
   responsiveness and measurement error, guided by the COS-
   MIN guidelines.
3 To provide an evidence base for future recommendations by
   HOME for instruments measuring long-term control of
   eczema in clinical trials to be included in a Core Outcome
   Set.

Methods

Protocol and registration
The study was both conducted and reported in line with the
PRISMA guidelines and was preregistered on 26 May 2020 on
PROSPERO (CRD42020162312).

Eligibility criteria
We included any study that reported on patient-reported out-
come measurement (PROM) instrument development or valida-
tion. The instruments had to be designed to capture patient-
reported control of eczema of any severity in either adults or
children. The papers had to either explicitly state that the instru-
ment measured eczema control or both reviewers had to agree that

the instrument met the following definition of eczema control 'a
multifaceted construct involving changes in the signs and symp-
toms of eczema, psychological, social and physical functioning,
and the treatment and management of the condition'.³ Only
instruments pertaining to atopic eczema were included.

This review was designed to inform HOME consensus decisions
at the HOME VII meeting in Japan.⁶ It was agreed at the HOME V
meeting in France that eczema control should not be measured by
flares or well-controlled weeks⁴ Therefore, papers that sought to
validate these measures were excluded. Quality of Life is a separate
domain for HOME purposes and therefore papers which define
control in terms of quality of life were excluded. Similarly, 'itch' is
a subdomain of the symptoms domain, and so papers developing
or validating measures of itch were not included. Finally, it is not
possible to undertake The CONsensus-based Standards for the
selection of health Measurement Instruments (COSMIN) assess-
ment process (see below) where only an abstract is available.
Therefore, conference abstracts were excluded.

Information sources and searches
Searches were run in Ovid Embase and PubMed on 24 January
2020 as these are the two databases for which the COSMIN filter
is available for validation studies.⁷ The search strategies (see
Appendix S1) combine alternative free text terms and subject
headings for three search concepts, to be combined using the
Boolean operator AND: (1) atopic dermatitis, (2) disease control
and (3) the relevant COSMIN filter for the database concerned.
There were no language restrictions.

In addition to the search of the databases above, we under-
took a survey of the membership of the HOME membership and
contacted experts in the field to identify any relevant instru-
ments that were in development.

Data collection and data items
The review followed the process set out in the COSMIN guid-
ance for systematic reviews of PROMs.⁸

As per the COSMIN guidelines, we assessed the following
measurement properties:

- Content validity
- Internal consistency
- Structural validity
- Hypotheses testing (construct validity)
- Cross-cultural validity
- Reliability
- Measurement error
- Responsiveness

Data were extracted in an Excel spreadsheet developed for the
review. Data extraction was performed independently by two
paired reviewers (BS, LH, RP, JC, EG, TP and ES) with adjudica-
tion by others in the study team, who had not reviewed the
instrument, in case of disagreement (JC, KT and ES). Measure-
ment properties were determined as sufficient, insufficient,
indeterminate, not assessed or not applicable as outlined in step two of the COSMIN assessment. Measurements that had insufficient content validity were deemed not to capture the construct of interest and were therefore not assessed further.

PROMs can be developed based on a reflective or a formative model. Internal consistency and structural validity are not appropriate assessments for multi-item scales developed using a formative model. Therefore, these have not been assessed for PROMs developed using this approach.

Extracted information for each paper included:

- Study characteristics including author, year, country of origin and study design.
- Characteristics of the PROM including construct being measured, the target population, the number of items and response categories.
- Measurement properties of the instruments including content validity, ease and usefulness of interpretation and item fit statistics.

Risk of bias

The COSMIN checklist was used to evaluate the methodological quality of included studies. This is undertaken first for each study individually using the COSMIN Risk of Bias checklist.

For each of the measurement properties, the COSMIN checklist consists of 5–18 items covering methodological standards (organized in nine boxes for the nine measurement properties). In addition, each item can be scored on a four-point scale (i.e. ‘inadequate’, ‘doubtful’, ‘adequate’ and ‘very good’). Taking the lowest rating for each item in one box, an overall score is obtained for each measurement property separately.

Summary measures and strategy for data synthesis

Due to the measurement properties being assessed in this review, data have been synthesized qualitatively in accordance with the methodology outlined in the COSMIN guidelines.

For each included PROM, the data were combined into an overall score of sufficient, insufficient, indeterminate, not assessed or not applicable for the measurement property.

Risk of bias across studies

An overall GRADE of quality for that instrument has been produced (High, Moderate, Low or Very Low).

Taking into account this rating, alongside issues of interpretability and feasibility, if relevant, a recommendation was then made on the PROM or PROMs with the best validity for use in measuring control in eczema patients. The criteria set out by COSMIN are as follows:

- A – Evidence for sufficient content validity and at least low-quality evidence of sufficient internal validity if applicable (PROM can be recommended).
- B – PROMs that may have the potential to be recommended, but further validation studies are needed.
- C – High-quality evidence for an insufficient measurement property (PROM should not be recommended).

Results

Study selection

We identified 14 272 papers and after removing duplicates 12 036 were screened by two reviewers (BS and LH) for eligibility. One additional instrument was identified from the 106 responses received from the survey of HOME members. We obtained 58 full texts and identified 12 eligible papers reporting on seven instruments (Fig. 1).

Table 1 sets out the key characteristics of all included scales. All scales were developed and validated in English, with the exception of the Patient Benefit Index (PBI) 2.0 which also had a German version. The English version of the PBI 2.0 was the one assessed in this review. There was one single-item patient global severity measure and five multi-item instruments. Only RECAP, ADCT and Atopic Dermatitis Score 7 (ADS7) specified a recall period. RECAP and ADCT were validated over a 1-week period. ADS7 has two questions which are completed daily for 7 days and then a total score calculated for the 7-day period based on the daily recall. RECAP and ADS7 were validated for use in both adults and children whilst the others were for use in adults only. Review authors involved in the development/validation of an eligible control scale were not permitted to assess their own instruments.

Risk of bias

We judged all PROMs to have been developed using a formative model, and therefore, structural validity and internal consistency were not assessed. None of the studies assessed measurement error or cross-cultural validity. Table 2 sets out the ratings.

The quality of the validation studies of the Atopic Eczema Score of Emotional Consequences (A ESEC) scored as doubtful and ADS7 and the Impact of Chronic Skin Disease on Daily Life (ISDL) as inadequate, on the PROM development. These were downgraded due to the concept elicitation aspect. The methods used to explore the relevance and the comprehensiveness of the included questions were not conducted in a manner consistent with the COSMIN recommendations on qualitative interviewing.

The quality of the RECAP and ADCT validation studies were rated as ‘very good’ across all assessed domains.

Methodological quality and quality of evidence across studies

The methodological quality, rated as sufficient, insufficient or inconsistent and the GRADE rating (high/moderate/low/very low) for each instrument is set out in Table 3.
The AESEC was assessed as insufficient with respect to content validity. AESEC was in part developed using free-text responses to a question via a social media platform. It is unclear whether this approach to PROM development would elicit all the key aspects of eczema control, nor that the final questionnaire items had been tested with the population of interest for...
comprehensibility. As such, it was not further assessed. While lacking content validity for eczema control, it may have adequate content for emotional consequence of the disease – its intended purpose.

Similarly, the ADS7 was assessed as insufficient with respect to content validity, as it was not clear how the PROM content had been developed or validated. As such, it was also not further assessed.

The ISDL was inconsistent with respect to content validity. The relevance and the comprehensiveness were very good, but it was unclear whether the response options matched the question or were clearly understood by the intended population. The overall PROM development was rated as inadequate, and therefore, the GRADE rating of the evidence was low.

The single-item patient global severity measure asks 'Would you describe your AD or eczema as mild, moderate or severe?'. It was similarly downgraded for content validity to inadequate. While it was not always clear whether a formative or reflective approach to development was used. We have only been able to assess those aspects of validation that have been reported in the published papers. It was not always reported whether a formative or reflective approach to development was used. We have had to use our

<table>
<thead>
<tr>
<th>Instrument</th>
<th>PROM development</th>
<th>Reliability</th>
<th>Hypothesis testing/</th>
<th>Responsiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECAP</td>
<td>Very good</td>
<td>Very good</td>
<td>Very good</td>
<td>Very good</td>
</tr>
<tr>
<td>ADCT</td>
<td>Very good</td>
<td>Very good</td>
<td>Very good</td>
<td>Very good</td>
</tr>
<tr>
<td>PBI 2.0</td>
<td>Adequate</td>
<td>Not assessed</td>
<td>Very good</td>
<td>Very good</td>
</tr>
<tr>
<td>Single-item</td>
<td>Adequate</td>
<td>Not assessed</td>
<td>Very good</td>
<td>Very good</td>
</tr>
<tr>
<td>AESEC</td>
<td>Doubtful</td>
<td>Not assessed</td>
<td>Very good</td>
<td>Not assessed</td>
</tr>
<tr>
<td>ISDL</td>
<td>Inadequate</td>
<td>Not assessed</td>
<td>Very good</td>
<td>Very good</td>
</tr>
<tr>
<td>ADS7</td>
<td>Inadequate</td>
<td>Not assessed</td>
<td>Very good</td>
<td>Not assessed</td>
</tr>
</tbody>
</table>

Table 2 COSMIN Risk of bias checklist (Very good, adequate, doubtful and inadequate)
judgement to determine this. It is possible that PROMs we felt had taken a formative approach in fact used a reflective approach. Moreover, no papers assessed measurement error or cross-cultural validity, which are aspects that are likely to be important for PROMs that are included in a core outcome set. All PROMs would benefit from further validation work.

Clinical and research implications

Whilst further work is required to assess aspects of validity for the recommended instruments, HOME has now recommended a full Core Outcome Set, including two of the instruments assessed in this review. The aim is for these to now be adopted in all clinical trials in eczema which measure eczema control.

This review suggests that the recommended instruments have been developed and validated in a robust process. Whilst a preliminary version of this review was presented to help the attendees at the HOME VII meeting reach a decision, this full review supports the decision made during the meeting. However, it also highlights a number of areas for each instrument that have not been fully validated. As these instruments become more widely used in clinical trials, researchers may wish to build in further validation work, for example on measurement error, cross-cultural validity and interpretability.

There is also further work to be done to understand how these PROMS relate to one another – we do not yet know whether a single-item patient global measure would perform as well as the multi-item scales or whether these multi-item scales ultimately are very similar and just one could be recommended or whether both tools need to remain in the Core Outcome Set to fully capture this domain. Similarly, we do not know to what extent these multi-item scales capture a construct that is truly distinct to repeated measures of signs/symptoms/quality of life. As studies adopt the Core Outcome Set, it may be useful to undertake secondary analyses of the collected data to explore the extent to which these domains are truly distinct underlying constructs. This may have the potential to reduce the Core Outcome Set instruments and therefore reduce participant burden in future trials.

Conclusion

RECAP and ADCT have been developed and validated to a sufficient standard to support their recommendation as PROMS for measuring control of atopic eczema.

Ethical approval

Not required.

Authorship statement

BS, LH, JC, DG, NKR and KT drafted the protocol and PROSPERO registration. DG developed the search terms with input from all authors and BS and LH acquired the data. BS, LH, RP, JC, EG, TP and ES extracted and evaluated the data. BS and LH
drafted the manuscript and all authors contributed to and approved the final version.

**Data sharing**
No additional data are available.

**References**

1. Schofield JK, Grindlay DJC, Williams HC. Skin conditions in the UK: a health care needs assessment. Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK, 2009.


**Supporting information**

Additional Supporting Information may be found in the online version of this article:

**Appendix S1.** PROMs for eczema control - validation studies review.