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 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 remitting and relapsing 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. It was agreed by consensus at the/Home VII meeting in Japan that eczema control should not be measured by flares or well-controlled weeks. 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) assessment 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. 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 undertook a survey of the membership of the HOME membership and contacted experts in the field to identify any relevant instruments that were in development.

**Data collection and data items**

The review followed the process set out in the COSMIN guidance for systematic reviews of PROMs. 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 adjudication by others in the study team, who had not reviewed the instrument, in case of disagreement (JC, KT and ES). Measurement properties were determined as sufficient, insufficient, incomplete or not applicable.
indeterminate, not assessed or not applicable as outlined in step two of the COSMIN assessment.8 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.10 Internal consistency and structural validity are not appropriate assessments for multi-item scales developed using a formative model.8,11 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.9,12

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.8

**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 are8 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.3,13–23 (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 (AESC) 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 the study.

Table 1 Characteristics of included studies

| Author/Year          | Instrument                                      | Country/language                  | Recall period | Target population   | Number of items | Sample size |
|----------------------|--------------------------------------------------|-----------------------------------|---------------|---------------------|-----------------|-------------|
| Howells et al. 2019  | Recap of Atopic Eczema (RECAP)                   | UK/English                        | 1 week        | Adults and children | 7               | 97          |
| Howells et al. 2020  |                                                  |                                   |               |                     |                 |             |
| Bhanot et al. 2020   |                                                  |                                   |               |                     |                 |             |
| Simpson et al. 2019  | Atopic Dermatitis Control Tool (ADCT)           | United States/ English            | 1 week        | Adults              | 6               | 1010        |
| Pariser et al. 2020  |                                                  |                                   |               |                     |                 |             |
| Blome et al. 2016    | Patient Benefit Index (PBI) 2.0                  | English, German                   | N/A           | Adults              | 24              | 16          |
| Topp et al. 2019     |                                                  |                                   |               |                     |                 | 64          |
| Vakharia et al. 2018 | Single-item                                      | United States/ English            | N/A           | Adults              | 1               | 265         |
| Silverberg et al. 2018 |                                              |                                   |               |                     |                 | 602         |
| Arents et al. 2019   | Atopic Eczema Score of Emotional Consequences (AESEC) | 9 European countries (Czech Republic, Denmark, France, Germany, Italy, Netherlands, Spain, Sweden, UK)/English | N/A           | Adults              | 28              | 1189        |
| Evers et al. 2008    | Impact of Chronic Skin Disease on Daily Life (ISDL) | Netherlands/English             | N/A           | Adults              | 16              | 128         |
| Darrigade et al. 2019 | Atopic Dermatitis Score 7 (ADS7)               | Belgium/English                  | 1 day         | Adults and children | 2               | 81          |

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 the study.
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?’ and was similarly downgraded for content validity to inadequate. This instrument may have appropriate content validity for patient-reported disease severity.

The study assessing responsiveness for PBI 2.0 reported that not all hypotheses for testing responsiveness were met; therefore, responsiveness is scored ‘inconsistent’.

Where they were reported, reliability, responsiveness and hypothesis testing tended to be sufficient for all instruments, with moderate to high-quality evidence.

**Recommendations**

Based on the risk of bias, the overall rating and the quality of the evidence, the RECAP and ADCT scored ‘A’, suggesting that they could be considered for inclusion in the core outcome set. The two instruments were developed independently but are similar; both multi-items scales with a recall period of one week, validated in English only. The questionnaires cover similar domains with similar response values scored 0-4, though RECAP has an additional question separating itch from intense itch. Whilst RECAP is validated for use in adults and children, ADCT is currently validated only for use in adults. Recommended cut-offs for defining eczema control are available for ADCT but not for RECAP.

In April 2019, a preliminary version of these results, based on scoping searches, was presented at HOME VII and used to inform initial decisions about a recommended PROM for the long-term control domain. Some results were amended following further independent COSMIN assessments and changes to reporting of some studies following peer review (See Appendix S1 for summary of changes). HOME provisionally included RECAP and ADCT in their core outcome set, subject to further research and assessment. This review supports that recommendation. Whilst the single-item patient global assessment also was considered, HOME decided that the response options did not adequately capture the concept of eczema control and so this could not be recommended as a global measure of eczema control.

The AESEC and ADS7 did not have sufficient content validity for further assessment. Based on the currently available published literature, they are not suitable measures for eczema control, though they may capture other important aspects of the patient’s experience of eczema. Similarly, whilst the ISDL had some aspects that were sufficient, overall evidence was low for content validity, which made it unlikely to be suitable as a core measure of eczema control.

**Strengths and limitations**

This was a formal systematic review which followed the COSMIN methodology. This is a robust process that aims to provide the best evidence for decisions about the validation of measurement instruments.

However, 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 2** COSMIN Risk of bias checklist (Very good, adequate, doubtful and inadequate)

| Instrument | PROM development | Reliability | Hypothesis testing/construct validity | Responsiveness |
|------------|------------------|-------------|--------------------------------------|----------------|
| RECAP      | Very good        | Very good   | Very good                            | Very good      |
| ADCT       | Very good        | Very good   | Very good                            | Very good      |
| PBI 2.0    | Adequate         | Not assessed | Very good                            | Very good      |
| Single-item| Adequate         | Not assessed | Very good                            | Very good      |
| AESEC      | Doubtful         | Not assessed | Very good                            | Not assessed   |
| ISDL       | Inadequate       | Not assessed | Very good                            | Very good      |
| ADS7       | Inadequate       | Not assessed | Very good                            | Not assessed   |

---

JEADV 2021 © 2021 The Authors. Journal of the European Academy of Dermatology and Venereology published by John Wiley & Sons Ltd on behalf of European Academy of Dermatology and Venereology
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 IK, Grundlay DJC, Williams HC. Skin conditions in the UK: a health care needs assessment. Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK, 2009.

2. 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: Eur J Allergy Clin Immunol* 2012; 67: 1111–1117.

3. Howells L, Thomas KS, Sears AV et al. Defining and measuring 'eczema control': an international qualitative study to explore the views of those living with and treating atopic eczema. *J Eur Acad Dermatol Venereol* 2019; 33: 1124–1132.

4. Chalmers JR, Thomas KS, Apfelbacher C et al. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol* Blackwell Publishing Ltd 2018; 178: e332–e341.

5. Barbarot S, Rogers NK, Abuabara K et al. Strategies used for measuring long-term control in atopic dermatitis trials: A systematic review. *J Am Acad Dermatol* 2016; 75: 1038–1044.

6. Thomas KS, Apfelbacher CA, Chalmers JR et al. Recommended core outcome instruments for health-related quality of life, long-term control and itch intensity in atopic eczema trials: results of the HOME VII consensus meeting. *Br J Dermatol* 2021; https://doi.org/10.1111/bjd.19751.

7. Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res* 2009; 18: 1115–1123.

8. Prinsen CAC, Mokkink LB, Bouter LM et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018; 27: 1147–1157.

9. Terwee CB, Prinsen CAC, Chiarotto A et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018; 27: 1159–1170.

10. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: A practical guide. Cambridge University Press, Cambridge, 2011. https://doi.org/10.1017/CBO9780511996214.

11. Streiner DL. Being inconsistent about consistency: When coefficient alpha does and doesn’t matter. *J Pers Assess* 2003; 80: 217–222.

12. Mokkink LB, de Vet HCW, Prinsen CAC et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018; 27: 1171–1179.

13. Arens BWM, Mensing U, Seitz IA et al. Atopic eczema score of emotional consequences—a questionnaire to assess emotional consequences of atopic eczema. *Allergy J Int* 2019; 28: 277–288.

14. Blome C, von Usslar K, Augustin M. Feasibility of using qualitative interviews to explore patients’ treatment goals: experience from dermatology. *Patient* 2016; 9: 261–269.

15. Topp J, Augustin M, von Usslar K et al. Measuring patient needs and benefits in dermatology using the patient benefit index 2.0: A validation study. *Acta Dermatol Venereol* 2019; 99: 211–217.

16. Howells LM, Chalmers JR, Gran S et al. Development and initial testing of a new instrument to measure the experience of eczema control in adults and children: Recap of atopic eczema (RECAP). *Br J Dermatol* 2020; 183: 524–536.

17. Pariser DM, Simpson EL, Gadkari A et al. Evaluating patient-perceived control of atopic dermatitis: design, validation, and scoring of the Atopic Dermatitis Control Tool (ADCT). *Curr Med Res Opin* 2020; 36: 367–376.

18. Simpson E, Eckert L, Gadkari A et al. Validation of the Atopic Dermatitis Control Tool (ADCT©) using a longitudinal survey of biologic-treated patients with atopic dermatitis. *BMC Dermatol* 2019; 19: 15.

19. Vakharia PP, Chopra R, Sacotte R et al. Validation of patient-reported global severity of atopic dermatitis in adults. *Allergy Eur J Allergy Clin Immunol* 2018; 73: 451–458.

20. Silverberg JI, Chiesa Fuxench ZC, Gelfand JM et al. Content and construct validity, predictors, and distribution of self-reported atopic dermatitis severity in US adults. *Ann Allergy, Asthma Immunol* 2018; 121: 729–734.e4.

21. Evers AWM, Duller P, Van De Kerkhof PCM et al. The Impact of Chronic Skin Disease on Daily Life (ISDL): A generic and dermatology-specific health instrument. *Br J Dermatol* 2008; 158: 101–108.

22. Darrigade A, Colmant C, Montjoye L et al. Atopic Dermatitis Score 7 (ADS7?): A promising tool for daily clinical assessment of atopic dermatitis. *Allergy* 2020; 75: 1264–1266.

23. Bhanot A, Peters T, Ridd M. Assessing the validity, responsiveness and reliability of the RECAP measure of eczema long-term control. *Br J Dermatol* Submitted 2020; 184: 955–957 Ref BJID-2019-2243.R3

**Supporting information**

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

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