‘A box the shape of me’:

the challenge of developing and evaluating patient-centred outcomes for use in eczema clinical trials

Laura Howells

BSc (Hons), MSc, MBPsS

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

June 2019

Abstract

Background:

Eczema is a chronic, itchy skin condition with onset typically in early childhood. Clinical trials compare health-related outcomes between two or more groups receiving different interventions. The Harmonising Outcome Measures in Eczema (HOME) initiative is an international collaboration working together to develop a core outcome set of outcome measures to be included in all eczema clinical trials to improve the comparability of outcome measures across trials.

Aims & objectives:

This thesis aims to inform two of the outcome domains recommended for measurement in eczema clinical trials by the HOME initiative group: long-term control of eczema and patient-reported symptoms. For the long-term control of eczema domain, the key objectives were to understand the concept of long-term control of eczema from key stakeholders' perspectives and to develop an instrument to measure eczema control. For the patient-reported symptoms domain the key objectives were to explore the measurement model of the Patient-Oriented Eczema Measure (POEM) and inform interpretability of the scores of the POEM.

Methods:

To understand the concept of long-term control of eczema from key stakeholders' perspectives, qualitative methods using online focus groups with people with eczema and caregivers of children with eczema across six countries, and an online survey of the HOME membership was conducted. To develop an instrument to measure eczema control, development of a conceptual framework, focus groups, expert panel design and feedback, cognitive interviews, and piloting of items was conducted. The measurement model of the POEM was explored using thought experiments and theoretical discourse. Anchor-based and distribution-based

statistical approaches to examining the minimally important change score and the smallest detectable change were used to assess the interpretability of the POEM scores.

Results:

With regards to progressing the long-term control of eczema domain of the HOME initiative; 'eczema control' is suggested to be a multi-faceted experience. A patient-reported outcome measure of eczema control called Recap of atopic eczema (RECAP) has been developed and initial testing shows that it is appropriate for use in eczema clinical trials and routine care.

With regards to progressing the patient-reported symptoms domain of the HOME initiative; mental experimentation suggests POEM was developed using approaches consistent with a formative measurement model. Interpretation of POEM scores has been improved by understanding that a change in score that is two points or under is consistent with potential measurement error, and that changes in scores should be three points or more before the change is deemed to be clinically important. From the data available, there was no evidence that POEM scores require different interpretation according to disease severity, age, sex and ethnicity.

Conclusions:

RECAP and POEM are both patient-reported outcome measures that are fit for purpose for measuring long-term control of eczema and patient-reported symptoms (respectively) in eczema clinical trials. This body of work has contributed towards informing evidence-based consensus decisions by the HOME initiative.

Publications, presentations and prizes arising from this work

A. Published journal articles

HOWELLS, L. M., CHALMERS, J. R., COWDELL, F., RATIB, S., SANTER, M. & THOMAS, K. S. 2017. 'When it goes back to my normal I suppose': a qualitative study using online focus groups to explore perceptions of 'control' among people with eczema and parents of children with eczema in the UK. *BMJ Open,* 7**,** e017731.

HOWELLS, L., RATIB, S., CHALMERS, J. R., BRADSHAW, L., THOMAS, K. S. 2018. How should minimally important change scores for the Patient-Oriented Eczema Measure be interpreted? A validation using varied methods. *British Journal of Dermatology,* 178**,** 1135-1142.

CHALMERS, J. R., THOMAS, K. S., APFELBACHER, C., WILLIAMS, H. C., PRINSEN, C. A., SPULS, P. I., SIMPSON, E., GERBENS, L. A. A., BOERS, M., BARBAROT, S., STALDER, J. F., ABUABARA, K., AOKI, V., ARDELEANU, M., ARMSTRONG, J., BANG, B., BERENTS, T. L., BURTON, T., BUTLER, L., CHUBACHI, T., CRESSWELL-MELVILLE, A., DELOZIER, A., ECKERT, L., EICHENFIELD, L., FLOHR, C., FUTAMURA, M., GADKARI, A., GJERDE, E. S., VAN HALEWIJN, K. F., HAWKES, C., **HOWELLS, L.,** HOWIE, L., HUMPHREYS, R., ISHII, H. A., KATAOKA, Y., KATAYAMA, I., KOUWENHOVEN, W., LANGAN, S. M., LESHEM, Y. A., MERHAND, S., MINA-OSORIO, P., MUROTA, H., NAKAHARA, T., NUNES, F. P., NYGAARD, U., NYGARDAS, M., OHYA, Y., ONO, E., REHBINDER, E., ROGERS, N. K., ROMEIJN, G. L. E., SCHUTTELAAR, M. L. A., SEARS, A. V., SIMPSON, M. A., SINGH, J. A., SROUR, J., STUART, B., SVENSSON, A., TALMO, G., TALMO, H., TEIXEIRA, H. D., THYSSEN, J. P., TODD, G., TORCHET, F., VOLKE, A., VON KOBYLETZKI, L., WEISSHAAR, E., WOLLENBERG, A. & ZANIBONI, M. 2018. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *British Journal of Dermatology.*

HOWELLS, L., THOMAS, K., SEARS, A. V., NASR, I., WOLLENBERG, A., SCHUTTELAAR, M. A., ROMEIJN, G. L. E., PALLER, A. S., MUELLER, K., DOYTCHEVA, K., KATAOKA, Y., DAGUZE, J., BARBAROT, S., KOBYLETZKI, L. V., BECKMAN, L., RATIB, S., COWDELL, F., SANTER, M. & CHALMERS, J. 2018. Defining and measuring “eczema control”: An international qualitative study to explore the views of those living with and treating atopic eczema. *Journal of the European Academy of Dermatology and Venereology*.

- Three further manuscripts are currently in prepartion.

B. Published conference abstracts:

HOWELLS, L., THOMAS, K. S., RATIB, S. & CHALMERS, J. 2016. Is the Patient Oriented Eczema Measure (POEM) fit for purpose as the core instruemnt for patient reported eczema symptoms? A series of linked validation studies to inform the Harmonising Outcome Measures for Eczema (HOME) initiative. *British Journal of Dermatology,* 175**,** 48.

HOWELLS, L., RATIB, S., CHALMERS, J., BRADSHAW, L. E. & THOMAS, K. S. 2017. The minimum clinically important difference (MCID) of the patient oriented eczema measure (POEM): do different methods of calculation give different MCIDs? *Trials,* 18.

C. Presentations and meetings attended

Date	Event attended	Location	Work presented
5-7 July 2016	British Association of Dermatologists (BAD) Annual Meeting 2016	Birmingham, UK	POSTER: What Predicts Distress in Psoriatic Arthritis? (presented research conducted prior to PhD but conference contained content relevant to the PhD)
13 Oct 2016	Skin Deep – 20 Years of Research – British Skin Foundation	London, UK	E-POSTER: Is the Patient Oriented Eczema Measure (POEM) fit for purpose as the core instrument for patient reported eczema symptoms? A series of linked validation studies to inform the Harmonising Outcome Measures for Eczema (HOME) initiative
10-11 Nov 2016	Core Outcome Measures in Effectiveness Trials (COMET) VI Meeting	Amsterdam, The Netherlands	POSTER: Reflections on the value of on-line discussion groups to gather patients' views during the development of a core outcome set for eczema
9-10 Jan 2017	Cochrane Skin CS-Cousin Annual Meeting, 2017	Berlin, Germany	n/a
25 Jan 2017	Research Education Advice and Communication in Health (REACH) Meeting	Southampton, UK	ORAL: A qualitative study exploring patient and carer experiences of long-term control of eczema
12-14 June 2017	Harmonising Outcome Measures in Eczema (HOME) V consensus meeting	Nantes, France	ORAL: International online focus groups to understand "long-term control of eczema" from patient and parent perspectives
17 May 2017	Pint of Science 2017 Life on the Border: Our Bodies' Barriers	Nottingham, UK	ORAL: "Eczema can feel like a rollercoaster" (public engagement talk)
7-10 May 2017	4th International Clinical Trials Methodology Conference and Society for Clinical Trials 38th Annual Meeting	Liverpool, UK	POSTER: How much change on the Patient-Oriented Eczema Measure (POEM) is important?

17th May 2017	Eczema and contact dermatitis: An Evidence Based Update	Nottingham, UK	n/a
28 June 2017	The Medical and Health Sciences Faculty Postgraduate Research Forum	Nottingham, UK	ORAL AND POSTER: Could improvements in how clinical trials measure eczema symptoms lead to better treatments for eczema?
11 Jan 2018	Core Outcome Measures in Effectiveness Trials (COMET) UK meeting for ongoing COS developers	London, UK	n/a
15-16 Jan 2018	Cochrane Skin CS-Cousin Annual Meeting 2018	Amsterdam, The Netherlands	n/a
11 April 2018	HOME VI Meeting	Utrecht, The Netherlands	n/a
11-13 April 2018	10th Georg Rajka International Symposium on Atopic Dermatitis (ISAD)	Utrecht, The Netherlands	ORAL: Interpreting change in Patient-Oriented Eczema Measure scores: Calculating the smallest detectable change and the minimally important change
17 Oct 2018	Sue Watson Postgraduate Presentation, School of Medicine, University of Nottingham	Nottingham, UK	ORAL: Measuring eczema control in clinical trials: a potential solution
12-13 Dec 2018	UK Society of Behavioural Medicine (UKSBM) Annual Scientific Meeting	Birmingham, UK	ORAL: What is 'eczema control'? An international qualitative approach to exploring views of people living with or treating eczema
7-8 Jan 2019	Cochrane Skin CS-Cousin Annual Meeting 2019	Paris, France	n/a
4 March 2019	Academic Orthopaedics, Trauma and Sports Medicine Lunchtime Seminar Programme, University of Nottingham	Nottingham, UK	ORAL: Deciding how to measure outcomes in clinical trials, using atopic eczema as an example"
8-10 April 2019	HOME VII Meeting	Tokyo, Japan	ORAL: Development and validation of Recap of atopic eczema (RECAP)

D. Awards

- School of Medicine Doctoral Programmes Committee Funding Award of £600 to attend the HOME VII meeting in Tokyo, Japan (April 2019)
- University of Nottingham Graduate School Travel Prize of £600 to attend the HOME VII meeting in Tokyo, Japan (April 2019)

Table of Contents

List of Abbreviations

BATHE	Bath additives for the Treatment of cHildhood Eczema (trial)
CI	Confidence Interval
CLOTHES	Clothes for the relief of Eczema trial
COMET	Core Outcome Measures in Effectiveness Trials initiative
COMET (trial)	Choice of Moisturiser for Eczema Treatment (trial)
COSMIN	COnsensus-based Standards for the selection of health Measurement Instrument (initiative)
COSMIN/COMET	COnsensus-based Standards for the selection of health Measurement Instrument/Core Outcome Measures in Effectiveness Trials (initiative)
CREAM	ChildRen with Eczema Antibiotic Management (trial)
CTT	Classical Test Theory
EASI	Eczema Area and Severity Index
ICC	Intraclass Correlation Coefficient
IGA	Investigator Global Assessment
IgE	Immunoglobin E
ISAAC	International Study of Asthma and Allergies in Childhood
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
ISRCTN	International Standard Randomised Controlled Trials Number
MAcAD	methotrexate versus azathioprine for severe atopic dermatitis (trial)
MCID	Minimal clinically important change
MIC	Minimally important change
MID	Minimally important difference
NESS	Nottingham Eczema Severity Score
NHS	National Health Service
NICE	National Institute for Health and Care Excellence

PEER	Paediatric Eczema Elective Registry
PGA	Patient/Parent Global Assessment
POEM	Patient-Oriented Eczema Measure
PRO	Patient Reported Outcome
HOME	Harmonizing Outcome Measures in Eczema (initiative)
SD	Standard deviation
S.e	Standard error
SEM	Standard error of measurement
SPSS	Statistical Package for the Social Sciences
SWET	Softened Water Eczema Trial
FDA	Food and Drug Administration

List of Tables

List of Figures

Acknowledgements

I would like to start by saying thank you to the British Skin Foundation for funding this PhD.

Thank you to my supervisors, Professor Kim Thomas, Dr Jo Chalmers and Dr Sonia Gran for their support and mentorship. I have learnt so much from you all, including the importance of patient involvement in research and the spirit of collaborative working, and I couldn't think of better people to have worked alongside. You have all been so kind. Kim, Jo, Sonia, thanks for everything.

The work in this thesis has been the product of collaboration with numerous individuals who have been a pleasure to work with. Thank you to Lucy Bradshaw, Beth Stuart, Miriam Santer, Fiona Cowdell, Matt Ridd, Daisy Gaunt, Nick Francis, Chao Huang, Phyllis Spuls, Louise Gerbens, Christian Apfelbacher, Laura Beate von Kobyletski, Linda Beckman, Amy Paller, Kristina Doytcheva, Kelley Mueller, Alison Sears, Sandra Lawton, Tim Burton, Amina Ahmed, Lynita Howie, Ibrahim Nasr, Andreas Wollenberg, Marie-Louise Schuttlaar, Margreet Romeijn, Justine Daguzé, Sebastien Barbarot, and Natasha Rogers (and thank you for all the design input Natasha). I would also like to thank the following individuals who all took the time to share their expertise and advice with me and for that I remain truly thankful; Anne Copay, Berend Terluin, Jessica de Boer, Louise Sabir, Hywel Williams, Carolyn Charman, Paul Leighton and my supportive mentor who inspires me endlessly, Christine Bundy.

A huge thank you to everyone who took part in the studies within this thesis. This work would not have been possible without you.

A final thank you to my friends, family members, and colleagues, with a special mention to my partner Chris and sister Bethan, for all the sunshine they bring to my life.

Dedication

I would like to dedicate this thesis to my parents, Geraint and Elizabeth Howells, whose love is ever-present. Mum, Dad, you are truly the best.

Statement of Contribution

This thesis was written in entirety by the PhD candidate, Laura M Howells (LMH) with comment and revisions from PhD supervisors Kim S Thomas (KST), Joanne R Chalmers (JRC) and Sonia Gran nee Ratib (SG). However, other individuals collaborated on projects within the thesis and in some cases jointly-authored publications based on studies that are presented within this thesis. Contributions for each chapter have been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 3

Candidate's contributions: Contributed to the conception of the study, took a lead role in designing the study protocol, responsible for recruitment, took a lead role in facilitating online focus groups, took a lead role in the data analysis and interpretation of findings, drafted and finalised the chapter (and associated journal article).

Collaborators' contributions: KST, JRC and SG contributed to the conception of the study. KST, JRC, SG, Fiona Cowdell (FC) and Miriam Santer (MS) contributed to the design of the work. KST, JRC, SG and FC were involved in data collection. FC and MS were key advisors and secondary coders in the analysis process, while KST, JRC and SG were also involved in the interpretation of data. KST, JRC, SG, FC and MS revised the manuscript critically for important intellectual content. All authors gave approval for the final version to be published. KST, JRC and SG reviewed this thesis chapter.

Chapter 4

Candidate's contributions: Contributed to the conception of the study, took a lead role in designing the study protocol, contributions to data collection and analysis of UK online focus groups outlined above, took a lead role in co-ordinating the team

and advising data collection methods for team members conducting online focus groups in other countries, took a lead role in the data analysis and interpretation, drafted and finalised the chapter (and associated journal article).

Collaborator's contributions: KST, JRC and SG contributed to the conception of the study. KST, JRC, SG, FC and MS contributed to the design of the online focus group part of the study. KST, JRC, SG, Alison Sears (AS), Ibrahim Nasr (IN) and Andreas Wollenberg (AW) contributed to the design of the online survey part of the study. KST, JRC, SG, FC and MS contributions to data collection and analysis of UK online focus groups outlined above. Satellite teams for the online focus group part of the study were responsible for adapting the UK online focus group design, data collection, and analysing the data collected against the UK developed framework analysis [USA: Amy Paller (AP), Kelley Mueller (KM), Kristina Doytcheva (KT), the Netherlands: Marie-Louise Schuttelaar (MLS), Margreet Romeijn (MR), Japan: Yoko Kataoka (YK), France: Justine Daguzé (JD), Sebastien Bararot (SB), Sweden: Linda Beckman (LB), Laura von Kokyletzki (LVK)]. All authors revised the manuscript critically for important intellectual content. All authors gave approval for the final version to be published. KST, JRC and SG reviewed this thesis chapter.

Chapter 5

Candidate's contributions: Contributed to the conception of the study, took a lead role in designing the study protocol, responsible for recruitment, co-facilitated the focus group, took a lead role in co-ordination of the expert panel, conducted the interviews and designed the online survey, took a lead role in the data analysis and interpretation of findings, drafted and finalised the chapter (and associated journal article).

Collaborators contributions: KST, JRC and SG contributed to the conception of the study. The expert panel included KST, JRC, SG, AVS, LVK, Matt Ridd (MR), Phyllis Spuls (PS), Sandra Lawton (SL), Natasha Rogers (NR), Tim Burton (TB), Lynita Howie

(LH), Amina Ahmed (AA) and Christian Apfelbacher (CA). All team members contributed to the design of the study. Paul Leighton (PL) was involved as advisor for qualitative data collection and as a secondary coder for data analysis of the focus group. JRC and AVS were involved in secondary coding of the cognitive interview data. All team members contributed to the design of the questionnaire. KST, JRC and SG were involved in the data analysis process of the interviews. SG was involved in the analysis process of the online survey. All team members were involved in the interpretation of the data. All authors revised the manuscript critically for important intellectual content. All authors gave approval for the final version to be published. KST, JRC and SG reviewed this thesis chapter.

Chapter 6

Candidate's contributions: Contributed to the conceptualisation of the study, conducted literature searches, drafted and finalised the chapter.

Collaborators' contributions: KST, JRC and SG were involved in the conceptualisation of this chapter and revised it. Hywel Williams (HW) and Carolyn Charman (CC) provided important insights and comments as the original developers of the POEM.

Chapter 7 – study 1

Candidate's contributions: Contributed to the conceptualisation and design of the study. Conducted the analysis and took a lead role in the interpretation of the results. Drafted a manuscript for publication. Wrote the thesis chapter.

Collaborators' contributions: SG, JRC, KST, MR, PS, Lucy Bradshaw (LB), Beth Stuart (BS), Miriam Santer (MS), Daisy Gaunt (DG), Louise Gerbens (LG), Chao Huang (CH) and Nick Francis (NF) were involved in the designed the study. All authors were involved in the interpretation of the results. All authors revised the manuscript critically for important intellectual content. This thesis chapter was revised by KST, JRC and SG.

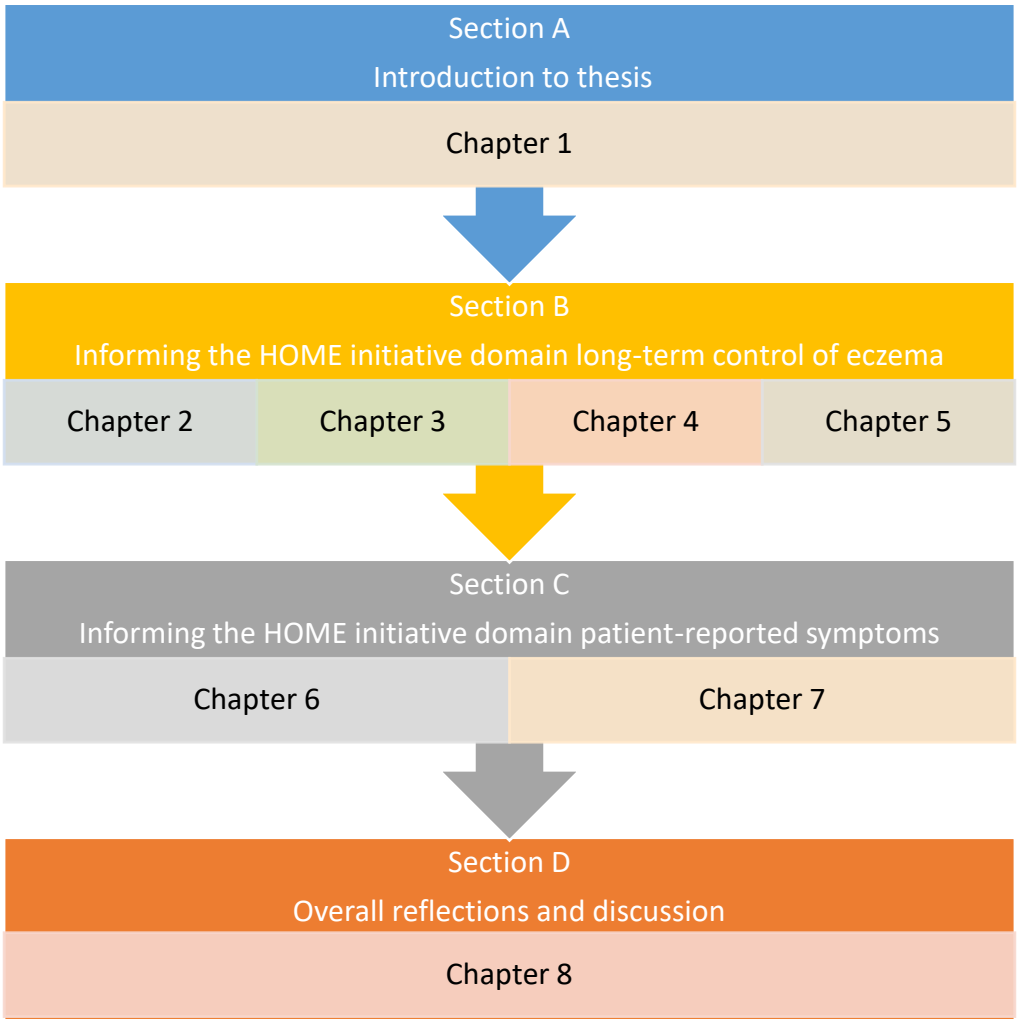
Chapter 7 – study 2

Candidate's contributions: Contributed to the conceptualisation and design of the study. Conducted the analysis and took a lead role in the interpretation of the results. Drafted a manuscript for publication. Wrote the thesis chapter.

Collaborators' contributions: SG, JRC, KST, and LB were involved in the conceptualisation and design of the study. All authors were involved in the interpretation of the results. All authors revised the manuscript critically for important intellectual content. This thesis chapter was revised by KST, JRC and SG.

Overview of thesis

Outcomes of clinical trials are the pinnacle of evidence-based medicine, yet historically outcomes have received scant attention in eczema research. This thesis assembles a series of studies that work towards a common goal of improving how patient-centred outcomes are measured in eczema clinical trials. The thesis is comprised of four sections. A schematic outline of how the sections fit together is illustrated below:



Section A is a narrative literature review designed to provide a rationale for the focus on improving patient-centred outcome measures for eczema clinical trials and serve as a comprehensive background to the field to equip the reader with an

understanding of the research landscape that the studies within this thesis are situated.

Section B comprises three studies that build from defining eczema control to the development of a patient-reported outcome measure of eczema control. Section C comprises three studies that work towards ensuring that the Patient-Oriented Eczema Measure (POEM) is fit for purpose for use in eczema clinical trials. Together, Sections B and C have informed the Harmonising Outcome Measures in Eczema (HOME) initiative's mission to establish a consensus-derived minimum set of outcomes, known as a 'core outcome set', for use in all eczema clinical trials worldwide.

Section D provides reflective comment on the thesis as a whole and the contributions made to the field before looking beyond the thesis towards future directions.

Section A

Introduction to thesis

This section is a narrative literature review. Firstly, the chapter will explain why eczema is a long-term condition that should be a global research priority. Secondly, it will review theory, methodology and practice relating to the development and evaluation of measurement instruments. Thirdly, it will assess the current use of outcome measures in eczema trials and comment on some of the shortcomings that led to the development of the Harmonising Outcome Measures in Eczema (HOME) initiative.

Chapter 1

The need for consistent, valid and patient-relevant outcomes in eczema clinical trials: a literature review

1.1 Eczema

1.1.1 Clinical presentation and diagnostic criteria

Eczema is a chronic, inflammatory skin disease that is characterised by itching and dry skin. It is a common condition that typically develops in children aged 2 or under (Herd et al., 1996b). Eczema lesions can appear red or dark (depending on skin colour), and can result in symptoms such as weeping, scaling, crusting and blistering (Figure 1-1).

Many people experience relapsing and remitting of symptoms, with periods of worsened eczema often described as 'flares' (Schofield et al., 2009). However, others will experience a continuous disease course over a long period of time (Weidinger and Novak, 2016). Eczema can occur all over the body, but typically the form and distribution of lesions follows age-related patterns (Weidinger and Novak, 2016). Staphylococcus aureus, bacteria that can be found on skin, tends to colonise on the skin of people with eczema and can get into the skin when eczematous lesions are present, leaving them prone to developing infections (Weidinger and Novak, 2016).

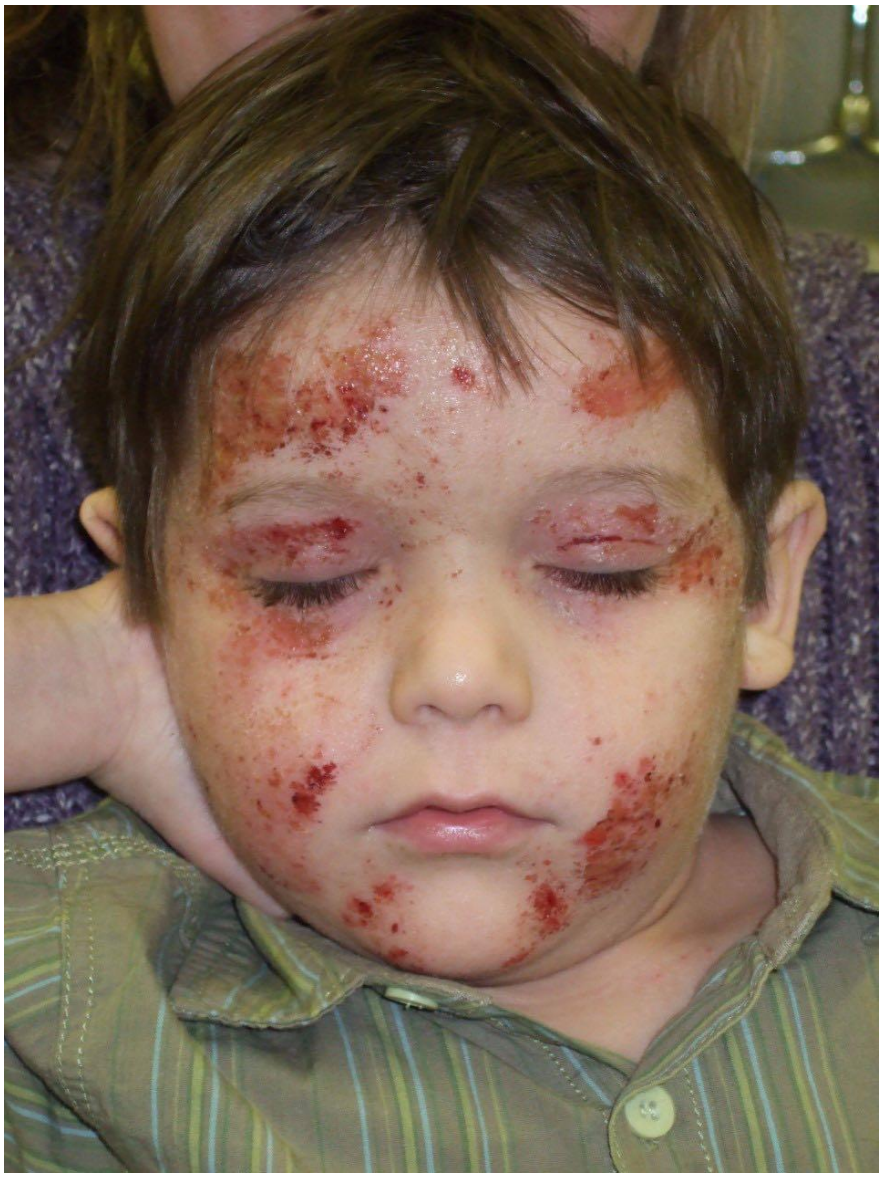


Figure 1-1 Typical appearance of severe eczema

Note. Used with consent and permission.

Hanifin and Rajka provided the first list of clinical symptoms of eczema for use as a comprehensive diagnostic criterion (Hanifin and Rajka, 1980). The UK working group simplified this list to create a minimum list of reliable discriminators that could be more easily used in epidemiological and clinical studies. The UK working group's diagnostic criteria has higher sensitivity and specificity than the original list (Williams et al., 1994a). A systematic review of diagnostic criteria recommended intervention

studies use the UK working group diagnostic criteria as they were the most extensively validated (Brenninkmeijer et al., 2008). The diagnostic criteria are itchy skin (or parental reporting of rubbing or scratching) plus three of the following characteristics: history of involvement of skin creases, history of asthma or hay fever (or history of eczema in first degree relative if > 4 years old), generally dry skin in the last year, visible flexural eczema (or visible dermatitis of the cheeks and extensor surfaces if under 18 months) and onset under aged 2 (only applies if ≥ 4 years old) (Williams et al., 1994a).

1.1.2 Terminology: What is in a name?

The terms 'eczema', 'atopic eczema' and 'atopic dermatitis' are often used interchangeably. The World Allergy Organisation proposed the use of the term eczema when referring to skin diseases with the clinical characteristics that involve a genetically determined skin barrier defect (Johansson et al., 2004). They also recommended the term 'atopic eczema' should only be used if the immunological mechanism has been assessed and it is confirmed that the underlying inflammation is dominated by an Immunoglobin E (IgE) antibody (Johansson et al., 2004). Current evidence suggests eczema is associated with elevated blood serum IgE levels in two-thirds of individuals with clinically diagnosed eczema (Flohr et al., 2004), however the proportion is lower in population-based samples (Abuabara et al., 2019).

A recent meta-analysis looked at the frequency of the use of terms and suggested due to 'atopic dermatitis' being used most frequently this should be the recommended term (Kantor et al., 2016). However, the review also highlighted differential term use across languages and locations and 46.9% used the term eczema. The International Eczema Council (IEC) felt that eczema was too imprecise and achieved a consensus among 77 of their international researchers and clinicians to use the prefix 'atopic' followed by either 'dermatitis' or 'eczema' in all publications, presentations and discussions (Silverberg et al., 2017). However, the

authors recognised the challenge this poses for discussing the condition with patients who are used to using the term ‘eczema’.

For purpose of this thesis the terms eczema, atopic eczema and atopic dermatitis will be viewed as synonymous and ‘eczema’ will be used, as this term is understood by patients and commonly used in the UK.

1.1.3 Epidemiology of eczema

Overall, eczema is estimated to affect up to 20% of children and 3% of adults worldwide (Nutten, 2015). However, prevalence estimates for eczema are variable, and a systematic review found that the annual prevalence of eczema in general practice populations ranged from 1.8% - 9.5%, whereas the annual prevalence of eczema in the open population surveys by the International Study of Asthma and Allergies in Childhood (ISAAC) ranged from 11.4% - 24.2% (Pols et al., 2016). Open population surveys with self-report from patients may overestimate disease, whilst general practice diagnosis may underestimate cases due to a threshold where patients will choose to visit their healthcare services as well as incomplete coding of eczema on the healthcare system (Pols et al., 2016).

The evidence for clearance of eczema with age has been mixed. It has typically been reported that eczema has a 50-70% clearance rate by 12 years (Ballardini et al., 2012, Ellis et al., 2012, Wuthrich, 1999). These studies may overestimate clearance with age as eczema is a relapsing condition, therefore participants in remission at the time of measurement in studies may subsequently relapse, and recent cohort studies suggest that eczema may be a lifelong condition for many. A recent prospective study using the US based Paediatric Eczema Elective Registry (PEER) found at every age between 2 and 26 years more than 80% of participants had symptoms of eczema and/or were using medication to treat their eczema (Margolis et al., 2014). A systematic review including seven longitudinal birth cohorts found the prevalence of eczema was similar in childhood and adolescence/early adulthood

(Abuabara et al., 2017). Regarding severity, the review reported that some people's eczema got better, others got worse and some remained at a similar level of disease (Abuabara et al., 2017). However, the lack of data meant that more firm conclusions about changes in severity over time could not be drawn.

ISAAC studies have conducted international epidemiological studies that have suggest eczema is a global problem (Williams et al., 1999). A systematic review of epidemiological studies between 1990 and 2010 found the prevalence of eczema was increasing in Africa, eastern Asia, western Europe and parts of Northern Europe (including the UK) (Deckers et al., 2012).

Generally, the severity distribution of eczema is predominantly mild. One large cross-sectional survey of 1760 children aged 1-5 years found the 1 year period prevalence of eczema for patients registered with a GP in Nottingham was 84% mild, 14% moderate and 2% severe eczema, as classified by a dermatologist (Emerson et al., 1998). The majority of eczema patients were seen in primary care with only 6% of patients referred to secondary care (Emerson et al., 1998). Relatively little research has been done to assess severity distribution in eczema populations, but this information may be useful for resource allocation decisions (Nankervis et al., 2016).

1.1.4 Pathogenesis of eczema

The two hallmarks of eczema are a defective epidermis (skin barrier) and type 2 immunity (Eyerich et al., 2015). The epidermis does not contain enough lipids (fats), which allows moisture to leave the skin, as well as allowing allergens to enter through the skin barrier. Mutations in the gene encoding for filaggrin can contribute to epidermis impairment (Weidinger et al., 2006, Smith et al., 2006). A meta-analysis of 24 studies involving patients recruited in secondary care found that having a mutation in the gene encoding for filaggrin was associated with more severe eczema compared to those without this mutation (Rodriguez et al., 2009). People with eczema may also have a T helper 2 cell dominant immune response, which makes

them more sensitive to the allergens entering than those without eczema. In a lot of cases, this increased sensitivity leads to an increased production of the antibody IgE, which leads to the allergic symptoms.

1.1.5 Causes of eczema

There is a strong body of evidence suggesting eczema is caused by a complex interplay of genetic and environmental factors. A recent systematic review of population-based twin studies found that monozygotic (identical) twins had three times higher concordant rates of eczema than dizygotic (fraternal) twins (Elmose and Thomsen, 2015). The overall heritability of eczema estimated from the included studies was around 75% (Elmose and Thomsen, 2015).

Epidemiological studies provide evidence that there are environmental and lifestyle factors associated with an increased risk of developing eczema. Eczema has been observed to be more frequent in wealthier families (Williams et al., 1994b). Living in an urban area has also been associated with a higher risk of eczema (Schram et al., 2010). Studies have shown that migrants tend to have an eczema prevalence similar to the population risk in the community they have moved to, with a higher risk associated with western industrialised countries (Williams and Burrell-Morrus, 2000). Findings that children in larger families were less likely to develop eczema lead to the "hygiene hypothesis" which postulates that a lack of exposure to allergens in early childhood increases the likelihood of developing eczema (Strachan, 1989). Subsequent studies have suggested that it may not be as straightforward as the hygiene hypothesis suggests (Flohr et al., 2005). Whilst it is clear that environment plays a role in the development of eczema, the role of specific environmental factors in the development of eczema is not yet fully understood (Thomas et al., 2014).

Many factors are anecdotally recalled and believed to worsen their eczema by people living with the condition, but there is limited evidence on what causes

worsening of eczema. A systematic review found some evidence that certain foods, house dust mite, stress and seasonal factors may cause worsening of eczema in some sub-groups of patients (Langan and Williams, 2006). One study followed up children aged 0-15 years daily for 9 months to assess the association between environmental "triggers" and disease flares reported by the participants (Langan et al., 2006a). Stress, damp and heat were all associated with flaring of eczema. However, this study was an exploratory study that may not have been large enough (N = 25) or long enough in duration (28 days) to capture all relevant associations (Langan et al., 2006a). A larger study (N=60) suggested that nylon clothing, dust, unfamiliar pets, sweating and shampoos play a direct role in the worsening of childhood eczema (Langan et al., 2009). These studies prospectively collected both eczema flares and environmental exposures, which is important when relying on patient self-report as retrospective studies are susceptible to recall bias.

1.1.6 Treatment and management of eczema

There is currently no known cure for eczema, therefore research on eczema treatments tends to be focused on improving management of the symptoms and on the secondary prevention of exacerbation in eczema to try and reduce the burden of the condition (Thomas et al., 2015).

The current National Institute for Health and Care Excellence (NICE) guideline for treatment of eczema in children under 12 in the UK recommends a holistic approach to assessment that takes skin severity along with quality of life and psychosocial wellbeing into account (NICE, 2007). It also proposes a stepped approach to treatment, which means treatment is tailored to the severity of the eczema (NICE, 2007). For mild eczema, emollients and mild potency topical corticosteroids are recommended; for moderate eczema, emollients, moderate potency topical corticosteroids, topical calcineurin inhibitors and bandages are recommended; for severe eczema, emollients, potent topical corticosteroids, topical calcineurin inhibitors, bandages, phototherapy and systemic therapy are recommended (NICE,

2007). There are currently no NICE guidelines for treatment of eczema in patients over 12 years old.

The burden of using treatments for people with eczema and their caregivers can be significant. Non-adherence to treatment regimens continues to be a prevalent problem in eczema (Santer et al., 2013). Barriers to treatment adherence in childhood eczema are caregivers' beliefs about the treatment, treatments being perceived as too time-consuming and the child resisting the treatment (Santer et al., 2013). Caregivers report dissatisfaction with the 'trial and error' approach to treatment often experienced in primary care and perceive it as dismissive (Santer et al., 2012).

Many people with eczema and caregivers of children with eczema are interested in, and often prefer, non-pharmacological treatments (Thomas et al., 2011, Santer et al., 2012). One reason for this interest is likely due to the perceived side effects of long-term use of topical corticosteroids such as skin thinning (Thomas et al., 2011). However, the only non-pharmacological treatment that currently has a reasonable evidence base is the use of educational interventions, although the exact content and structure of such interventions is still unclear and likely to be country-specific (Nankervis et al., 2016). Furthermore, many educational interventions likely in part focus on education relating to treatment use. Whilst the current management of eczema is focused on secondary prevention of escalation in disease, there is also a growing body of studies investigating primary prevention of eczema (Abrahamsson et al., 2007, Simpson et al., 2014).

1.1.7 Economic cost of eczema

Substantial costs for society and healthcare services are incurred due to eczema being such a common condition (Sach et al., 2016). There are also a number of financial costs for people with eczema and their families such as a loss of financial earnings, non-propriety treatments, special clothes and bedding, extra house

cleaning, diets and travelling to access healthcare services (Lewis-Jones, 2006). These costs have been shown to directly relate to the severity of the disease and to be higher than other chronic childhood illnesses such as asthma and diabetes (Lewis-Jones, 2006). The financial burden is particularly problematic for low income families (Lewis-Jones, 2006).

Based on a sample of patients in a semi-rural area of Scotland, it was estimated that in the financial year of 1995-1996 that the total annual cost of eczema in the UK was £465 million (Herd et al., 1996a). It was estimated that health service cost was £125 million, the cost to patients was £297 million and the cost to society due to loss of working days was £43 million (Herd et al., 1996a). During the same year, an evaluation using population census data and the results of a cross-sectional study in the Nottingham area estimated that the annual UK cost of eczema in children 1-5 years was £47 million (Emerson et al., 2001). Costs to the state were estimated as £30 million for the National Health Service (NHS), which was mainly due to consultation and prescription costs (Emerson et al., 2001). The majority of consultations were in primary care and the majority of prescription costs were emollients and emollient bath oils (Emerson et al., 2001). The remaining estimated £17 million costs were family spending for the child's eczema (Emerson et al., 2001). The higher cost estimates in the study by Herd et al. (1996a) is likely due to the inclusion of both adults and children, and adults tend to have more chronic disease and a larger surface area to treat (Emerson et al., 2001). These estimates were based on relatively small samples to estimate the UK population and are now somewhat outdated with changes to costs of services and new treatments available such as biologics, therefore there is a need for updated cost estimates (Sach et al., 2016).

Understanding the economic cost of eczema on a global scale presents further challenges. A comparison of costs across countries can be difficult due to different types of costs encountered (Schuttelaar et al., 2011). For example, in the Netherlands health insurance pays for prescriptions of emollients and protective dressings, but these costs are incurred by the family in Italy (Schuttelaar et al., 2011).

Given that eczema is a global problem, this is an important area to enable global cost-effectiveness studies.

1.1.8 Impact on people living with eczema

Eczema can have profound effects on both children and adults living with the condition. For children with eczema, it has been suggested that the impairment to quality of life is greater than other common childhood diseases such as asthma, diabetes, enuresis and cystic fibrosis (Lewis-Jones, 2006). The widespread effects on families with a child with eczema including sleep loss, psychological pressures on the parents, impacting daily activities such as housework, diet, social life, holidays, school and financial difficulties has been documented (Lawson et al., 1998).

Children with eczema have reported more sleep problems due to their skin than psoriasis or vitiligo patients (Manzoni et al., 2012). Other studies have found that lower sleep efficiency, more frequent night-time awakenings, more difficulty getting to sleep and greater difficulty awakening for school are observed in eczema patients compared to controls when using objective measures of sleep (Fishbein et al., 2015). Sleep disturbance of caregivers was also found to be greater in parents of children with eczema than children with asthma and was associated with greater maternal anxiety and depression and greater paternal anxiety (Moore et al., 2006).

Eczema also poses a significant burden for those who continue to live with the condition in adulthood and can have a significant emotional impact (Drucker et al., 2017). A population-based study of adults in the US found if eczema patients had a sleep problem (fatigue, sleepiness and insomnia) there was higher odds of poor health status, number of sick days and doctor visits than having eczema without sleeping problems or sleeping problems without eczema (Silverberg et al., 2015).

A retrospective course of life questionnaire found that in young adults aged 18-30 years, patients with severe eczema were not achieving the same social milestones as peers with moderate eczema or peers without eczema (Brenninkmeijer et al., 2009).

These milestones have been found to be important for adjustment to adult life. This highlights the potentially profound effects for those with severe eczema (Brenninkmeijer et al., 2009). The continued burden that eczema has on individuals and society provides a clear rationale for the need to improve understanding of which interventions are the safest and most efficacious in this population.

1.2 Outcome measures in clinical trials

The WHO (World Health Organisation, accessed 1/3/2019) defines clinical trials as "any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes.". Outcome measures (syn. endpoints) are the measures chosen to assess the impact of the intervention.

Traditionally, clinical trials have focused on biomedically defined outcome measures (Feinstein, 1987). This has been due to a predominant view that such data is hard, scientific data, however Feinstein argues that clinical examination is required to provide a humanistic approach to the data collected "as the improvements shown with clinimetric indexes are what most people seek in clinical care" (Feinstein, 1987).

This argument has since been extended beyond clinician-reported outcomes to patient-reported outcomes, due to recognition that clinician-reported outcomes can still miss important information about outcomes that are important to patients (Fitzpatrick et al., 1998). An illustrative example is that many eczema trials have used clinician-reported itch, and it is easy to see why itch may be best reported from patients themselves (Charman et al., 2003). The patient perspective is currently seen as an essential component when considering trial outcomes. This can be attributed to a backdrop of:

1) the changing landscape of healthcare focusing on the management of long-term conditions where the goals of healthcare intervention have shifted,

2) increased attention to involving patients in decision making requiring outcomes to be relevant and interpretable for them, and

3) the need to understand the range of benefits of treatment in relation to their cost with increasing demands on the financial resources of healthcare systems (Fitzpatrick et al., 1998).

The Food & Drug Administration (FDA) (2009) defines a patient-reported outcome (PRO) as "a measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else. A PRO can be measured by self-report or by interview provided that the interviewer records only the patient's response" (Food and Drug Administration, 2009). Given that eczema develops in young children, it is also important to consider the extent to which measurements incorporate the child's perspective. The FDA distinguishes between proxy-reported outcomes and observer-reported outcomes. Proxy-report is used to refer to a measurement based on a report by someone other than the patient reporting as if they are the patient, whereas an observer-report is used to refer to a measurement based on a clinician or caregiver's observation, where they may also interpret or give an opinion based on their observation (Food and Drug Administration, 2009).

Long & Dixon (1996) propose two spectrums that reflect the patient-centeredness in outcome measures that can vary by both degree of patient perspectives being incorporated in the content of the measure and the method used to gather information from patients. Regarding incorporation of patient perspectives, one end of the spectrum is professionally defined in that patient perspectives are not incorporated into the outcome measure, and the other end of the spectrum is where both content and form of each patient's perspective is preserved (Long and Dixon, 1996). Regarding methods used, one end of the spectrum is assessment by health care professionals with no information provided directly from patients, whilst the

other end of the spectrum is where information is supplied directly from the patient (Long and Dixon, 1996).

1.2.1 Patient-reported outcome measures

This thesis uses the term 'patient-reported outcome measures' to refer to standardised questionnaires that ask people with eczema or their caregivers about their perspective on an aspect of their or their child's health. Standardised questionnaires provide each individual with the same questions in the same format and answers are restricted to a range of options (Collins, 2003). The aim of this process is to create data which can be quantified and analysed statistically. The standardisation of a measure aims to ensure that the differences observed are the result of real changes and not artefacts of the way the data has been collected (Streiner et al., 2015).

A common criticism of patient-reported outcome measures is that they lack objectivity. Objective measures can be defined as measures which involve no human judgement in the collection and processing of information (McDowell, 2006).In contrast, subjective measures can be defined as measures that involve a person (healthcare professional, patient, caregiver) making a judgement that forms the indicator of health (McDowell, 2006). However, the objectivity argument is often only targeted self and caregiver report. There is a false assumption that clinician-reported measurement is objective. However, that is not the case as the clinician is still required to make a judgement.

Furthermore, measures that are considered to be objective may still require subjective interpretation. For example, taking a medication may appear to be an objective measure, but if it requires an individual to self-report medication taken it requires subjective interpretation (McDowell, 2006). Furthermore, the inclusion of subjective concepts can be viewed as a requirement to capture aspects of health and illness that are important and relevant to individuals and reducing outcomes to only those that can be objectively measured could omit important outcomes. One

situation where objectivity is a key concern is when there is a difficulty with blinding participants within a clinical trial. If a patient or caregiver cannot be blinded to their allocation, usually because they are aware that they are receiving an active treatment, this could influence their judgement and introduces a risk of bias.

1.2.1.1 Cognitive theories of questionnaires

There are assumptions made within the scientific model of standardised questionnaires (Collins, 2003). It is assumed that all respondents understand the questions in a consistent way, that the questions are asking for information that respondents have and can retrieve, and that the wording of the questions themselves provide respondents with all the necessary information they require to answer the question as the researcher desires (Collins, 2003). A synergy in research development across the fields of psychology and survey methodology has provided cognitive theories to test the assumptions of standardised questionnaires, which has led to the development of methods of cognitive testing (Willis, 2015).

There are four processes that forms the basis of most models: comprehension, retrieval, judgement and response (Tourangeau et al., 2000). Whilst the original model was a staged-approach, theory has progressed to suggest that interactions can occur between any stages of the process (Collins, 2003). It has been evidenced that there are occasions when individuals will not perceive there to be sufficient motivation to engage in the cognitive effort required to answer the questions. Therefore, individuals may engage in a process called 'satisficing', thus not engaging in the cognitive process thoroughly or missing out a cognitive process (Krosnick, 1991). A limitation of cognitive models has been that they disregard any non-cognitive processes, such as the sociocultural context and the respondent's world or the influence of an interviewer (if one is present) and how this can influence a respondent's relationship to the questions, and subsequently models have been developed that attempt to incorporate more comprehensive factors (Jobe and Herrmann, 1996).

1.2.2 Reflective and formative models

Constructs captured by self-report can be single-item or multi-item instruments. When a construct is a complex phenomenon, it is often measured with a multi-item instrument and this requires careful consideration about how the scores generated by items represent the construct to be measured (Edwards and Bagozzi, 2000). A conceptual framework is a model that represents the relationship between the items and the construct to be measured when using multi-item instruments; and broadly can be separated into reflective and formative models (De Vet et al., 2011).

In reflective models, each observed variable is a way of measuring the latent variable directly (*known as effect indicators*) (De Vet et al., 2011). Well-known measurement theories of classical test theory (CTT) and item response theory (IRT) are based on reflective models. CTT has been the predominant measurement theory over the last century and states that any observation is composed of two components; a true score and an error associated with the observation (Streiner et al., 2015). Recognising limitations to CTT, modelling approaches known as IRT (and closely related Rasch analysis) were developed. The items must be able to be placed on a unidimensional continuum representing differing 'levels' of a trait or latent construct. A second related assumption is that if the trait or latent construct was to be accounted for (conditioned out) of the model, there would be no correlation between any of the items, an assumption termed local independence (Streiner et al., 2015). Despite the potential benefits that IRT can bring to measurement, CTT has often prevailed due to its simplicity and applicability (Streiner et al., 2015).

An alternative to a reflective model is a formative model, where each observed variable is capturing a unique concept, which when combined together, form the latent variable *(known as causal indicators)* (De Vet et al., 2011). Given the predominance of CTT and IRT methodology in the development and validation of health-related measurement, it is often implicitly assumed that instruments have

reflective models. Distinguishing whether an instrument is based on a reflective or formative model is important for three main reasons:

- Firstly, utilisation of an incorrect measurement model can undermine the content validity of the construct and misrepresent the relationship between constructs (Coltman et al., 2008).
- Secondly, the best practice approaches for selecting items for inclusion in the instrument differs (De Vet et al., 2011).
- Thirdly, methods used to assess the measurement properties relating to the internal structure of a measure are only relevant for reflective models (Prinsen et al., 2018, Streiner, 2003).

1.2.3 Assessing the measurement properties of measurement instruments

Due to a lack of consistency in terminology and methodology used in evaluating measurement instruments, an international multidisciplinary team formed the COnsensus-based Standards for the selection of health Measurement Instrument (COSMIN) initiative to develop a consensus-based taxonomy to define domains, measurement properties that fit within those domains, and aspects of those measurement properties considered to be relevant for any measurement instrument used in any application (Mokkink et al., 2010b). Throughout this thesis, terms will be used consistently with this taxonomy. Within this section, evaluation of a measurement instrument in terms of the reliability, validity, responsiveness, and interpretability will be considered in turn. However, there are other criteria that should be considered when designing or selecting an instrument for use in clinical trials and core outcome sets, which will be discussed further in section 1.3.4.

1.2.3.1 Reliability: Are the instrument's scores reproducible? Are items in each dimension internally consistent?

Reliability is defined as the proportion of the total variance in the measurements which is due to 'true' differences between patients (Mokkink et al., 2010b). The

concept of 'true' in this context refers to the CTT notion that each observed score is composed of two components, a true score and error associated with that observation and true is the average score that would be obtained if the measurement was made an infinite number of times (Mokkink et al., 2010b).

Measurement error is defined as the systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured (Mokkink et al., 2010b). Measurements will almost always contain some measurement error. Measurement error and reliability are two distinct but related concepts. For example, if two measures of itch have the same level of reliability, but variation (standard deviation) between the individual's scores in the first measure of itch is low, whilst the variation between the individual's scores in another measure of itch are high, it is harder to distinguish between individuals in the first measure, and even if it has a small measurement error compared to the second measurement instrument this can reduce the reliability. This highlights how reliability is not an inherent property of an instrument, but a characteristic of the instrument scores when used in a population in a particular situation (Streiner et al., 2015).

The third measurement property within the domain of reliability, internal consistency, is defined as the degree of interrelatedness amongst the items (Mokkink et al., 2010b). Assessment of internal consistency is designed to test the CTT assumption that a pool of items are effect indicators that are tapping into the same latent construct (De Vet et al., 2011). However, this is not applicable to formative models where each item is not expected to be interrelated (Prinsen et al., 2018, Streiner, 2003).

1.2.3.2 Validity: does the instrument measure what it purports to measure?

Reliability is necessary but not sufficient in establishing the usefulness of a measure (Streiner et al., 2015). It may be possible to reproduce the scores of an instrument under different conditions, but this does not ensure that what is being measured is the desired construct. Validity, as an overall domain, is defined as the degree to

which an outcome measure measures the construct it purports to measure (Mokkink et al., 2010b).

Content validity is defined as the degree to which the content of an instrument is an adequate reflection of the construct to be measured (Mokkink et al., 2010b). To be able to assess content validity, it needs to be clearly documented what construct the instrument is aiming to measure (Patrick et al., 2011a).

Construct validity is the degree to which the scores of an instrument are consistent with hypotheses based on the assumption that the instrument validly measures the construct to be measured (Mokkink et al., 2010b). One element of construct validity is testing hypotheses of how the instrument will perform in relation to other measures (De Vet et al., 2011). Another aspect of construct validity is structural validity, which is concerned with the degree to which the scores of the instrument are an adequate reflection of the dimensionality of the construct to be measured (Mokkink et al., 2010b). This measurement property has been deemed only appropriate to assess when the measurement model of the instrument is considered to be reflective (Prinsen et al., 2018). Another element of construct validity is cross-cultural validity, which is defined as the degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version of the instrument (Mokkink et al., 2010b).

1.2.3.3 Responsiveness: Is the instrument sensitive to changes that are of importance to respondents?

Responsiveness is defined as the ability of an outcome measure to detect change over time in the construct to be measured (Mokkink et al., 2010b). Akin to construct validity, it is assessed via testing of hypotheses about whether scores of individuals that are expected to change do in fact result in a change. Within the context of a clinical trial, responsiveness of an outcome measure is essential as the primary

purpose of trial design is to be able to detect real change in health outcomes as a result of the intervention (Food and Drug Administration, 2009).

1.2.3.4 Interpretability: Can the scores on the instrument be meaningfully interpreted?

Whilst interpretability was not considered to be a measurement property within the COSMIN taxonomy, it is an important domain when evaluating an instrument. It is defined in the COSMIN taxonomy as the degree to which one can assign qualitative meaning – that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in score (Mokkink et al., 2010b).

1.3 Outcome measures in eczema clinical trials: introducing the HOME initiative

1.3.1 The need for evidence synthesis

"It is surely a great criticism of our profession that we have not organised a critical summary, by speciality and subspecialty, adapted periodically, of all randomised controlled trials." (Archie Cochrane, 1979)

The quote above sets out a vision for medical research by a pioneer in evidence-based medicine. Systematic reviews attempt to collate all empirical evidence with pre-specified criteria and use explicit and systematic methods to collate evidence with the aim of minimising bias (Green and Higgins, 2005). As the number of eczema trials has been increasing exponentially, there is an increased need to synthesise research to be able to directly compare the results of these trials and make meaningful conclusions (Flohr, 2011). When possible, authors of systematic reviews will conduct a meta-analysis to summarise the results of independent studies (Glass,

1976). Meta-analyses apply statistical methods to compare and pool the results of more than two studies (Glass, 1976).

However, the great vision of using systematic reviews and meta-analysis to improve the evidence base available to make healthcare and policy decisions often falls short due to inconsistency in outcome measures. It has been identified that diverse outcome measures are used across trials in multiple areas of health research. An early acknowledgement of the problem was articulated in a comprehensive review of randomised controlled trials in schizophrenia in 1998. In 2000 trials, over 600 interventions had been assessed and 640 different outcome measures had been used across the trials (Thornley and Adams, 1998). Systematic reviews of outcome measures in eczema clinical trials illustrate a similar situation regarding an inconsistency of outcome measures used across trials (Charman et al., 2003, Schmitt et al., 2007, Futamura et al., 2016, Gerbens et al., 2016).

1.3.2 Why inconsistent outcome measures are a problem

Whilst it is important that researchers can measure outcomes that they think are important to the research question the trial is asking, evidence suggests that the diversity in outcomes used comes from lack of careful planning of outcomes used (since many have often been instruments with lack of evidence of clear development or good measurement properties) and a lack of dialogue taking place across research teams. When trying to compare results in a Cochrane review of randomised controlled trials comparing chlorpromazine versus placebo, the authors who had previously conducted the comprehensive review of outcome measures across schizophrenia trials, reflected that the 'rating scales' used in trials were often not validated and varied widely in quality (Thornley et al., 2003). The same has been found in eczema clinical trials, with the systematic review on outcome measures included in eczema randomised controlled trials conducted in 2003 found only 27% of included trials used a published severity scale, whilst 14% used modified versions

of published scales and 59% used unnamed scales with no data on validity or reliability (Charman et al., 2003).

A key issue with such a plethora of outcomes being used is that when the individuals are trying to make sense of the evidence on a particular area or are trying to synthesis the findings from the evidence in a systematic review, their ability to make sense of the data is hindered by the diversity in outcome measures. This situation can be described as akin to comparing apples and pears. Even if outcome measures used in trials were well validated and well reported, it can be difficult to make any clinically meaningful sense if different outcome measures are used. To allow for direct comparison of study results via meta-analysis, uniform, valid and reliable outcome measures are needed (Flohr, 2011).

In addition to the difficulties synthesising research, a lack of standardisation of outcomes also leaves open the potential for bias (Clarke, 2007). Williamson et al, (2005) drew attention to a form of bias occurring within studies termed outcome selection bias. An example of this is where researchers may measure multiple outcomes in a trial, but selectively report only the outcomes that were statistically significant (i.e. supported the studies hypotheses). In their seminal paper highlighting the different areas of avoidable research waste, Chalmers and Glasziou (2009) outline bias or unusable reporting as one of the four key areas of research waste. They suggest that 50% of planned study outcomes are not reported and that the choice of primary outcome is often changed between trial protocol and trial reports (Chalmers and Glasziou, 2009). There is evidence that trials are more likely to report outcomes with statistically significant results than outcomes with non-significant results, which is likely to alter clinical decisions made from the reported results (Williamson et al., 2005). Within published eczema trials between January 2007 and January 2011, only 18 trials (17%) registered their trials prospectively and nominated a primary outcome (Nankervis et al., 2012). Even amongst the properly registered trials the descriptions of outcomes were often unclear so the authors of

the systematic review could only be confident that 5 (5%) of the trials did not have outcome reporting bias (Nankervis et al., 2012).

1.3.3 Core outcome sets

A core outcome set is a consensus-derived minimum set of outcomes that should be included in research studies or clinical practice (Schmitt et al., 2015). Core outcome sets have been proposed as a solution to variation in outcome measures between clinical trials. The core outcomes do not need to be the primary outcome, as this should be determined by what outcome is the most important to assess in that individual trial (Schmitt et al., 2015). There is widespread recognition for core outcome sets, and the Core Outcome Measures in Effectiveness Trials (COMET) initiative provides guidance on methodology of this approach, a community for core outcome set developers and a database to help disseminate core outcome sets (Williamson et al., 2011).

1.3.3.1 Stakeholder/public and patient involvement

Stakeholders are described for the purpose of this thesis as individuals who have a personal or professional interest in the benefit of a research project. Subsumed within the concept of stakeholder involvement is the concept Patient and Public Involvement (PPI), which focuses particularly on the involvement of patients or members of the public in conducting research.

Traditionally, outcome measures research has had relatively little input from patients and the public, with researchers and clinical experts typically deciding what should be measured and designing the content of patient-reported outcome measures (Staniszewska et al., 2012, Trujols et al., 2013). However, recent years have seen a growing emphasis in PPI across all areas of health research and there has been considerable attention paid to the role of PPI in developing core outcome sets. An often cited example of PPI impact in core outcome set development is how patient inclusion in OMERACT (a core outcome set initiative in rheumatology)

meetings identified the previously neglected outcome of fatigue (Kirwan et al., 2007). Whilst many of the challenges faced in core outcome set development will reflect PPI challenges in other areas of health research, distinct issues identified by core outcome set developers include the methodology and the objectives of a core outcome set perhaps feeling removed from the patient's immediate concerns, engaging meaningfully with patients across multiple nations, and bringing multiple key stakeholder groups across multiple countries together (Young and Bagley, 2016).

1.3.4 The Harmonising Outcome Measures in Eczema (HOME) initiative

The HOME initiative (www.homeforeczema.org) is an international group working together to agree a core outcome set for eczema clinical trials (Schmitt and Williams, 2010). Prior to the body of work presented in this thesis, there had been four international face-to-face consensus meetings involving multiple stakeholders including healthcare professionals, methodologists, pharmaceutical industry representatives and patients (Chalmers et al., 2014, Chalmers et al., 2016, Schmitt and Williams, 2010, Schmitt et al., 2012).

The HOME Roadmap is a framework to provide a standardisation of developing core outcome sets in dermatology (Schmitt et al., 2015). The first step of the HOME roadmap is to define the scope and the applicability of the core outcome set by considering the population, intervention, setting, geographical location and stakeholders for the core outcome set (Schmitt et al., 2015). For the HOME initiative there is international stakeholder involvement and intent for the core outcome set to be used globally. It aims to be relevant for both children and adults with eczema. The group includes clinicians, methodologists, representatives from the pharmaceutical industry and patients (Schmitt and Williams, 2010).

The second step of the HOME roadmap requires a consensus process with all stakeholder representatives to agree on the domains to be included in the core outcome set. Prior to the HOME I consensus meeting, there was an international modified e-Delphi study that found consensus for inclusion of the domains

symptoms, physician-assessed clinical signs and long-term control of flares (Schmitt and Williams, 2010).The HOME II consensus meeting in Amsterdam in June 2011 refined the domains identified in the Delphi process and consensus voting led to four domains being recommended: clinician-reported signs, patient-reported symptoms, long-term control of flares and quality of life (Schmitt et al., 2012). It was later decided that the domain long-term control (of flares) was to be applied to trials more than 3 months in duration only (Chalmers et al., 2014).

The third step of the HOME roadmap aims to agree on the core outcome instruments to be recommended to measure the core outcome domains selected in step 2. The process of step 3 is outlined in Figure 1-2. The processes of step 3 were used by the HOME members to recommend Eczema Area and Severity Index (EASI) for measuring signs at the HOME III meeting and the Patient-Oriented Eczema Measure (POEM) for measuring symptoms at the HOME IV meeting (Chalmers et al., 2014, Chalmers et al., 2016). However, the group recommended that evidence gaps identified for measurement properties of the POEM be subsequently assessed (Chalmers et al., 2016).

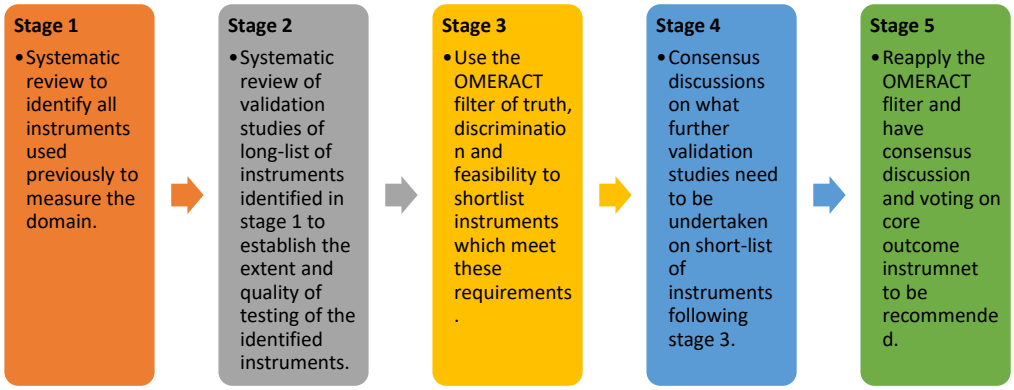


Figure 1-2 Step 3 of the HOME Roadmap

Step 4 of the HOME roadmap involves dissemination and revision of the core outcome set. A core outcome set must be “forever preliminary”, therefore new studies can serve to refine a core outcome set (Schmitt and Williams, 2010).

Since the development of the HOME roadmap, a consensus process has led to the co-production of guidance by two initiatives called COnsensus-based Standards for the selection of health Measurement Instrument/Core Outcome Measures in Effectiveness Trials (COSMIN/COMET) to aid the selection of outcome measurement instruments for a core outcome set (Prinsen et al., 2016). Two notable recommendations beyond the HOME roadmap are the suggestion of minimal validation requirements of content validity and internal consistency *or* test-retest reliability *or* inter-rater reliability for efficiency and recommending only one instrument per domain (Schmitt et al., 2015, Prinsen et al., 2016). Whilst the HOME roadmap does state that "ideally, one best instrument should be defined for each core outcome domain", the COMET/COSMIN guidance uses wording that puts more emphasis on this requirement for only one instrument (Schmitt et al., 2015).

To summarise the recommendations of the HOME initiative prior to the body of work within this thesis, the four core outcome domains recommended for collection in all eczema trials are clinician-reported signs, patient-reported symptoms, long-term control (of flares), and quality of life. It was recommended that clinician-reported signs be collected using the EASI and patient-reported symptoms are collected using the POEM, although further training materials and research into the advantages of dichotomising EASI scores and further assessment of POEM measurement properties was recommended (Schmitt et al., 2014, Spuls et al., 2017). Methods of data collection for the outcome domains long-term control of eczema (initially long-term control of flares, but subsequently renamed) and quality of life had not yet been recommended.

1.4 Thesis aims and objectives

The overall aim for this thesis is to inform the HOME initiative consensus-based discussions by providing evidence that contributes to the development of the core outcome set for eczema clinical trials. The thesis achieves this overall aim via four key objectives.

Objectives 1 & 2 contribute to the 'long-term control of eczema' domain:

1) Use a qualitative approach to understand key stakeholders' perspectives (people with eczema, caregivers and healthcare professionals) regarding the domain 'long-term control of eczema' (addressed in Chapter 3 and Chapter 4).
2) Use a mixed-methods approach to develop an instrument to measure eczema control that is suitable for use in the core outcome set (addressed in Chapter 5).

Objectives 3 & 4 contribute to the 'patient-reported symptoms' domain:

3) To explore the measurement model of the POEM to guide future investigations into the measurement properties of the POEM (addressed in Chapter 6).
4) To use methods of calculating the minimally important change, the standardised effect size and the smallest detectable change to advance the interpretability of POEM scores (addressed in Chapter 7).

Section B

Informing the HOME initiative domain of long-term control of eczema

This section of the thesis includes four chapters (Chapter 2, 3, 4 and 5), that contribute to research efforts to determine how to measure the HOME initiative's domain of 'long-term control of eczema' in clinical trials.

Due to the complex nature of the construct of 'long-term control of eczema', Chapter 2 serves to orient the reader by considering different conceptualisations of long-term control of eczema, what has informed conceptualisation within this thesis, and what the implications of this conceptualisation are for measurement of control.

Chapter 3 and Chapter 4 both use explorative, qualitative approaches to understanding what the domain long-term control of eczema means to key stakeholders. Chapter 3 presents a study using online focus groups with adults with eczema and caregivers of children with eczema in the UK, to explore perceptions of 'eczema control'. This study tested the use of online focus groups to provide methodological guidance for the subsequent online focus groups in the international qualitative study presented in Chapter 4.

The international qualitative study in Chapter 4 involved two concurrent phases, one that used online focus groups in six countries to understand 'eczema control' from the perspective of people with eczema/caregivers, and one that used an online survey of the HOME membership to gather clinician/researcher perspectives. The findings from these chapters informed subsequent HOME initiative decisions regarding the definition of the long-term control of eczema domain, which led to the development of a patient-reported outcome measure of eczema control, called Recap of atopic eczema (RECAP), which is presented in Chapter 5.

Chapter 2

Defining ‘eczema control’:

Conceptualisation and the implications for measurement

2.1 Chapter aims

This chapter aims to map how the conceptualisation of ‘eczema control’ developed throughout the studies presented in the rest of this thesis section. It is crucial to recognise that there are multiple ways of conceptualising the construct eczema control, and that the context of informing an agreed working definition for a core outcome domain for the HOME initiative has shaped decisions on how to define ‘eczema control’ within this thesis. Figure 2-1 presents the key developments in conceptualisation that have informed the work within this thesis, and the remainder of this chapter will describe these key developments.

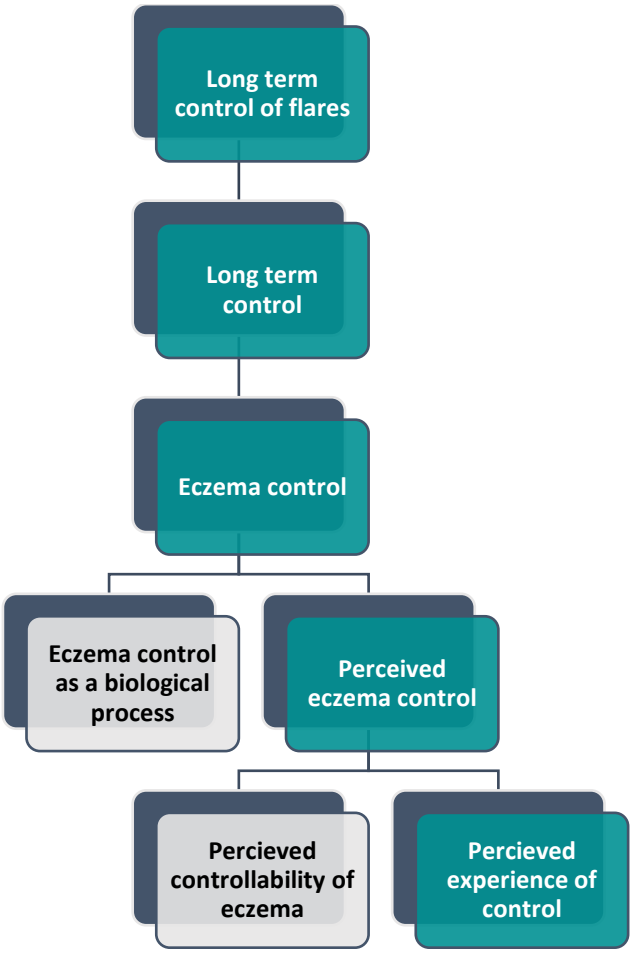


Figure 2-1 Conceptualisation of the core outcome domain

2.2 Moving from 'long-term control of flares' to 'eczema control'

The HOME initiative initially agreed 'long-term control of flares' as a core outcome domain (Schmitt et al., 2012). This process was informed by a three stage e-Delphi, which is an online consensus activity that uses a structured iterative process where participants assess an issue and in subsequent rounds are given feedback on their responses comparatively to other responses, and offered the opportunity to modify their response (Schmitt et al., 2011). Consensus was reached in the e-Delphi that 'long-term control of flares' should be included as a core outcome domain, and this was ratified in a vote at the HOME II meeting (Schmitt et al., 2012).

The challenge of gaining consensus on a working definition for the domain has been a recurring challenge that has hindered the progress of the HOME initiative in selecting a measurement instrument that can assess 'long-term control of flares'. The HOME initiative initially began with assessing the evidence base and discussing the definition of 'flares' (Schmitt et al., 2012, Chalmers et al., 2014, Langan et al., 2014). Following challenges reaching consensus with the concept 'flares', another concept of 'well-controlled weeks', that was adapted from the asthma literature, was also discussed (Langan et al., 2017, Chalmers et al., 2014). However, in discussions at the HOME III meeting, whilst methods of measurement of both 'flares' and 'well-controlled weeks' were proposed that were potentially useful in some contexts, it was acknowledged that as these were currently defined, they did not meet the needs of being applicable to all trials for a core outcome set (Chalmers et al., 2014). Reasons cited for concern about the applicability of these measures for all trials include intensive data collection (daily), and complexities in analysing and interpreting the data (Chalmers et al., 2014).

At the HOME III meeting, views of the attendees were also gathered, which demonstrated that there was wide variation in individuals' understanding of the domain (Chalmers et al., 2014). The difficulties in reaching consensus, it transpires, were at least partially because at time of agreeing consensus for the domain during the original e-Delphi and the HOME II meeting, stakeholders were agreeing on a 'name' for the domain, but held varied interpretations on what that construct entailed, and hence how it should be measured. The core outcome set methodology requires consensus to be able to progress to measurement instrument selection, since instruments will differ depending on the conceptualisation of the domain.

The disparity in interpretations across stakeholders may be in part because the domain name contains multiple elements that each require their own definition: 'long-term', 'control' and 'flares'. As a result of the difficulties to adequately define 'flares', the HOME initiative began to focus on defining 'long-term control', as

presented in Figure 2-1. However, during online focus groups with patients and parents presented in Chapter 3, participants made a further separation within the domain, which was 'eczema control' and 'long-term eczema control'. The approach taken within this thesis was to conceptualise 'eczema control', as this was the central concept discussed by participants, with an awareness that the HOME initiative would then still need to consider how 'eczema control' could be measured in the 'long-term'. This change in conceptualisation is presented in Figure 2-1.

2.3 'Perceptions' of eczema control

To overcome the barrier of a lack of consensus over the working definition for the domain, it was agreed at the HOME IV meeting in Malmö 2015 that international studies to gather patient and healthcare professional perspectives would be a helpful next step (Chalmers et al., 2016). As can be seen in the studies presented in Chapters 3 and 4, this marked the conceptualisation of 'eczema control' as 'perceived eczema control' as opposed to defining control as changes in the underlying biological processes of the disease, which is presented in Figure 2-1. Participants across all stakeholder groups tended to suggest that understanding the individual patient's perception of their eczema, and the impact that it has on them, was fundamental. Disease activity was discussed in the qualitative studies, but it was suggested that what was important was what individual noticed about changes in their eczema symptoms, rather than the underlying disease processes.

2.4 Something 'global'?

When interpreting the results from the qualitative studies presented in Chapter 3 and 4, it was not assumed that the various 'indicators of control' that were suggested would need to be combined into one measure of 'eczema control', but that they were potentially conflicting ways to view 'eczema control'. The multiple experiences that could form an individual's perception of 'control' was potentially a

barrier to reaching further consensus as there were so many options presented for conceptualising this domain.

The subsequent discussions at the HOME V meeting were trying to assess, out of all the different ways of viewing control that were presented from the qualitative studies, could the group reach consensus on what might be the best way to define 'eczema control' for the core outcome domain. Small groups agreed their three most important aspects relating to the long-term control of eczema domain, and subsequent whole group voting following discussions indicated that disease severity (signs and symptoms), quality of life and intensity of itch were considered the essential elements for measuring long-term control of eczema (Chalmers et al., 2018). However, when voting if the existing domains of signs, symptoms (including intensity of itch), and quality of life were sufficient, consensus was not reached (Chalmers et al., 2018). Further discussion suggested that a 'global assessment' of control may be required, and the HOME initiative agreed consensus to define the domain as signs, symptoms, quality of life and a patient global measure (91% agreed, 4% unsure, 4% disagreed) (Chalmers et al., 2018). This was the starting point for developing a new patient and parent reported outcome measure. It was not decided whether this patient global measure should be a single item measure or a multi-item measure. From the initial qualitative studies, and the HOME V discussions, it was proposed that a multi-item 'global' measure, which considered a range of experiences relevant for assessing control, when combined, would give a measure that provided a 'holistic' or 'global' assessment of perceived eczema control.

2.4.1 Experience of control or beliefs about the controllability of eczema?

Reflecting on the perspectives of participants that were presented in Chapter 3, 4 and 5, it was clear that there was further work to do in order to develop the RECAP instrument in regard to untangling what perceptions of eczema control were relevant to the construct of interest for this particular outcome measure. The RECAP development team made a key distinction between perceptions about 1) *the*

controllability of eczema and 2) the *level of control that has been attained*. The alternative approaches to defining the meaning of 'control' by Jan Walker were helpful for untangling this distinction in how 'control' was being understood by stakeholders (Walker, 2001).

1) *Controllability of eczema*

Theorising on perceptions about the controllability of eczema, The Common-Sense Model of Illness Representations proposes that individuals hold beliefs about their illness, one of which is the perceived curability/controllability of their illness (both either in their personal ability to control the illness or the their treatment's ability to control the illness) (Meyer et al., 1985, Diefenbach and Leventhal, 1996). Perceptions of higher levels of curability/controllability have been positively associated with adaptive outcomes across a range of illnesses and negatively related to disease state and distress (Hagger and Orbell, 2003). Applying this definition to examples relating to perceptions about the controllability of eczema, this may refer to the perceptions such as "if I avoid allergens I will have better control over my eczema", or "My treatment will always be able to get rid of any flares I have in the future".

2) *Level of control attained*

Theorising of perceptions about the attainment of control, this is about an outcome and the perception, based on an individual's experience, of whether that outcome has been achieved or not. This conceptualisation of perceptions of control appears to be mirrored in the way diseases such as asthma and urticaria have measured control (Weller et al., 2014, Nathan et al., 2004). Applying this definition to examples of perceptions of eczema control, this may refer to perceptions such as "My eczema has not impacted my life", or "I have not noticed any symptoms of my eczema". As presented in Figure 2-1, this is the way 'eczema control' has been defined for the development of RECAP.

2.5 Implications for measurement

2.5.1 Method of measurement

The choice to make the measurement instrument a patient reported outcome measure was, in part, dictated by the HOME V consensus vote that the domain needed to contain a patient global assessment. However, the conceptualisation of control as 'the perceived experience of control' also requires that it is reported by the patients (or their caregivers when they are too young to report themselves), which was a key finding from the studies presented in Chapters 3 and 4.

2.5.2 A formative measurement model

How eczema control is conceptualised has implications for the most appropriate measurement model to use in developing RECAP (presented in Chapter 5). The study team engaged in multiple discussions about whether the construct of interest for RECAP was best considered a reflective or a formative model. It was considered that each item was tapping into a different characteristic, and contributing part of the construct, and when considered together they form the whole construct. Therefore, it was decided that a formative model was most appropriate. One key tension in the development of formative models is that important items should be included, as otherwise the instrument is not measuring the construct of interest as comprehensively as possible, but that less important items are removed, and redundancy between items is reduced, so that each item is adding a unique contribution to the overall model. Throughout the development of RECAP in Chapter 5, the study team had to make some challenging decisions about which items should remain in the model.

2.6 Summary of Chapter 2

This chapter served to provide an overview of key developments in the conceptualisation of eczema control through interpretations formed throughout this

work and informed by the findings from studies presented in Chapter 3, 4 and 5 as well as decisions made at HOME meetings. It has also served to contextualise the body of work, as the position inherited from the history of the long-term control domain for the HOME initiative guided subsequent decisions about how to work towards a definition that could potentially meet the needs of the core outcome set.

Chapter 3

"When it goes back to my normal I suppose": a qualitative study to explore perceptions of 'eczema control' among people with eczema and caregivers of children with eczema in the UK

3.1 Introduction

The HOME initiative has agreed long-term control of eczema (originally termed 'long-term control of flares') is a domain that should be measured in all eczema clinical trials over 3 months in duration (Schmitt et al., 2012). For an outcome measure of long-term control of eczema to be recommended, there needs to be consensus on how to operationally define long-term control of eczema, and a lack of working definition for the domain has hindered progress with following the HOME roadmap to provide an evidence-base for consensus voting (Barbarot et al., 2016). It was agreed at the HOME IV meeting in Malmö 2015 that the next steps for the domain long-term control of eczema was to conduct international studies to gather patient and healthcare professional descriptions of 'eczema control' to inform a working definition for the domain (Chalmers et al., 2016).

3.1.1 A history of the long-term control of eczema domain

The HOME initiative initially included 'long-term control of flares' as the core outcome domain (Schmitt et al., 2012). Flares have been defined heterogeneously in the literature (Langan et al., 2006b). How researchers define flares and how they subsequently measure them are intrinsically linked. A systematic review of the definitions of eczema flares, that was originally conducted in 2006 but later updated in 2014, found that flares are measured in a variety of ways in the literature, which could be categorised into an arbitrary cut off on a measurement instrument, a composite score from multiple measures or a behavioural outcome (Langan et al.,

2014, Langan et al., 2006b). Very few had measured flares using a patient-reported outcome measure (Langan et al., 2014, Langan et al., 2006b). The authors did not find any method of capturing flares that contained both characteristics of being feasible to collect and being recorded at the time flare symptoms are experienced (Langan et al., 2014).

Measurement properties have only been assessed for a limited number of methods for capturing long-term control of flares. Two studies used the definition "an episode requiring escalation of treatment or seeking additional medical advice" to measure flares and escalation therapy was defined on entry into the study based on discussion with patients or their guardians, and patients daily rated yes or no for whether therapy was escalated (Langan et al., 2009, Thomas et al., 2015). A validation study found this measure had good face validity and construct validity (Thomas et al., 2015). However, concerns around resource intensive data management and difficulties pooling data may mean this measure is not suitable for all trials (Thomas et al., 2015).

Another method where measurement properties have been assessed is 'well controlled weeks'. Well controlled weeks are defined as having 2 days or fewer with both 1) symptoms greater than a pre-specified level, and 2) escalation of treatment required. This method was found to have good feasibility and construct validity, but there were concerns raised by the authors regarding ceiling effects in moderate to severe disease (Langan et al., 2017).

Flares represent one aspect of long-term eczema control and given the challenges that have occurred with defining flares and finding methods to capture flares, the HOME initiative undertook a systematic review in line with the first step of stage 2 in the HOME roadmap that looked at the domain of long-term control of eczema more broadly than only defining flares (the HOME roadmap is outlined in section 1.3.4). A systematic review was conducted to capture the ways that long-term control of eczema has been measured in published randomised controlled trials between 2000

and 2013 (Barbarot et al., 2016). This systematic review found that the most common method, in 91% of randomised controlled trials, was repeated measurement of eczema outcomes such as clinical signs, quality of life and itch (Barbarot et al., 2016). Most frequently used measures were clinician-reported outcomes, and they were most often reported on a monthly basis (40%) (Barbarot et al., 2016). The frequency of use suggests that repeated measures is acceptable to trial designers, however 39.7% of the included trials used 'inappropriate' statistical techniques such as repeated significance testing at multiple time points and 17.3% used 'inefficient' techniques such as not using data from all time points collected (Barbarot et al., 2016).

The second most commonly used measurement (27%) was use of regular eczema medications, often topical steroids but in a minority of studies calcineurin inhibitors or antihistamines and antibiotics (Barbarot et al., 2016). Medication use recorded quantity, potency and frequency of application. An often-linked measurement that was reported separately here is flares (25% of the trials used this outcome) (Barbarot et al., 2016). Using categorisation of flare outcomes developed by Langen et al. (2014) it was found that use of an arbitrary cut off was the most commonly used flare outcome (35%). Behavioural measures (23%) and composite measures (35%) were used too (Barbarot et al., 2016, Langan et al., 2014). Time to first flare was the most common analysis method of flare outcomes (Barbarot et al., 2016).

3.1.2 The need for qualitative research on an international scale

Despite a lack of clear definition, and hence multiple forms of measurement, it is apparent that long-term control of eczema is an important outcome for people with eczema. Seventy-five percent of people with eczema in a survey sample stated that being able to effectively control their eczema would be the single most important improvement to their quality of life (Zuberbier et al., 2006). Patients are key stakeholders and their involvement in developing core outcome sets may lead to outcomes not previously identified by other stakeholders (Williamson et al., 2012). It

is important that the standardised definition and core outcome measurement instrument of long-term control of eczema captures the experiences of people with eczema and caregivers of children with eczema as they are best placed to understand whether the eczema is controlled (Langan et al., 2014).

Qualitative research is increasingly being used in the development of core outcome sets to enable meaningful representation of all stakeholders (Keeley et al., 2016). Qualitative methods may be used to inform core outcome set development in numerous ways. Some tasks where qualitative methods can be beneficial include identifying what outcomes are important to stakeholders, facilitating understanding of why different outcomes are important to different stakeholders, determining the scope of outcomes, identifying appropriate language to use in a Delphi survey, to compare data from different stakeholders and other sources of outcome data such as systematic reviews (Keeley et al., 2016). The appropriate method of data collection will depend on the purpose of the research (Keeley et al., 2016).

Two commonly used methods of data collection in qualitative research are one-to-one interviews or focus groups. The benefit of focus groups is the interaction between participants that allows variation between participants to be directly addressed, however the structure is less conducive to following the individual journey for each individual patient (Keeley et al., 2016).

Online methods are becoming increasingly popular in health research due to their ability to reduce costs and the widespread availability of internet access. Although the study presented within this chapter focuses on understanding perspectives of long-term control of eczema for people with eczema[1] and caregivers[2] in the UK, this study was conducted as part of a larger international study to gain international,

[1] The term 'patients' refers to people who are seeking or receiving healthcare (Greenhalgh, 2017). As participants were recruited through a variety of different platforms and not necessarily healthcare settings, participants in studies presented in this thesis will be referred to as 'people with eczema'.

[2] The term caregiver will be used in studies presented in this thesis to be inclusive of all parents, carers and guardians who provided their perspective of their child's eczema.

multiple stakeholder input on their view of the domain long-term control of eczema (presented in Chapter 4). Qualitative research across countries is a resource intensive task, therefore the use of online focus groups was utilised to facilitate the process and allow for wider inclusion.

3.2 Study Objectives

To engage people with eczema and caregivers to:

1. determine what long-term control of eczema means to them,
2. explore what aspects of long-term control of eczema are most important to them, and
3. explore what methods of measuring long-term control of eczema are feasible and acceptable.

3.3 Methods and Materials

All participants provided online consent prior to participation. The protocol is on the Centre of Evidence Based Dermatology's website (http://www.nottingham.ac.uk/research/groups/cebd/resources/protocol-registration.aspx). This study was approved by the University of Nottingham's Medical and Health Sciences Research Ethics Committee (Ref: F14062016 SoM ROD).

3.3.1 Participant selection

Inclusion criteria were adults aged ≥16 years or caregivers of children with eczema who self-reported diagnosis of eczema by a doctor. Participants were recruited through the National Eczema Society Facebook page, Twitter posts, the University of Nottingham's website and email invitations to people who had consented to being

contacted for eczema research. If interested, they could phone or email the research team to express an interest and were sent an online form to complete. Based on the survey responses, purposive sampling was used to maximise variation in participants relating to characteristics of self-reported eczema severity, ethnicity, age, sex, previous participation in clinical trials, disease duration and previous experience of healthcare services (Patton, 1990).

3.3.2 Procedure

Six semi-structured, synchronous, online focus groups were conducted between August and October 2016 using a text-based online chatroom. Predominantly, online focus groups have been asynchronous in nature (not occurring in real time) (Gaiser, 1997). While asynchronous methods may allow for wider participation, they have been criticised as not being sufficiently responsive to be considered a 'focus group' (Bloor et al., 2001). Previous studies using synchronous online focus groups have found that the immediacy and dynamism of face-to-face conversations was mimicked (Stewart and Williams, 2005, Fox et al., 2007).

Table 3-1 Online Focus Group Timetable

Date	Time	Participants
Tuesday 23rd Aug	6:30pm – 8:00pm	People with eczema
Wednesday 24th Aug	6:30pm – 8:00pm	Caregivers
Tuesday 20th Sept	6:30pm – 8:00pm	Caregivers
Wednesday 21st Sept	6:30pm – 8:00pm	People with eczema
Saturday 8th Oct	10:00am – 11:30am	Caregivers
Saturday 8th Oct	12:00am- 1:30pm	People with eczema

Groups took place on a weekday evening or a weekend in the daytime (Table 3-1). Participants were allocated to one focus group only. Groups included either adults (16 years or over) who had eczema, or caregivers of children with eczema.

The recruitment procedure is outlined in Figure 3-1. Participants joined a group hosted on www.chatstep.com from a location of their choice. Participants entered a

nickname of their choice and the room name and password that were provided by email. Between two and four researchers were present at the discussion to do tasks such as time-keeping, ensuring all participants remained included, responding to technical issues, reading responses and typing responses. LMH and one other researcher (JRC, KST or SG) logged into each group to facilitate. All of the research team were introduced simply as 'researchers', although FC is a registered nurse and MS is a general practitioner. KST had previous experience facilitating online focus groups. Four participants knew some of the researchers prior to the discussions due to involvement in other research activities. LMH established a relationship with the participants prior to the discussion via email and sometimes telephone contact. Sessions lasted approximately 70–80 minutes.

Focus groups followed a common topic guide (final topic guide in Appendix A). The questions were revised iteratively based on what was perceived as facilitating engagement and relevant responses during the focus groups (e.g., stopped asking participants to think about going for a check-up with their doctor as this raised health service issues that were not the subject matter of the study). Questions were framed to elicit everyday experiences of people living with eczema.

The transcript was downloaded after each session. LMH made field notes during and after discussions to reflect on her thoughts and feelings relating to the data (Ward et al., 2013). Participants were sent a debrief email and asked to provide feedback on their experience of taking part shortly after the group. In April 2017, participants were sent a summary of the results and asked for their feedback on the results.

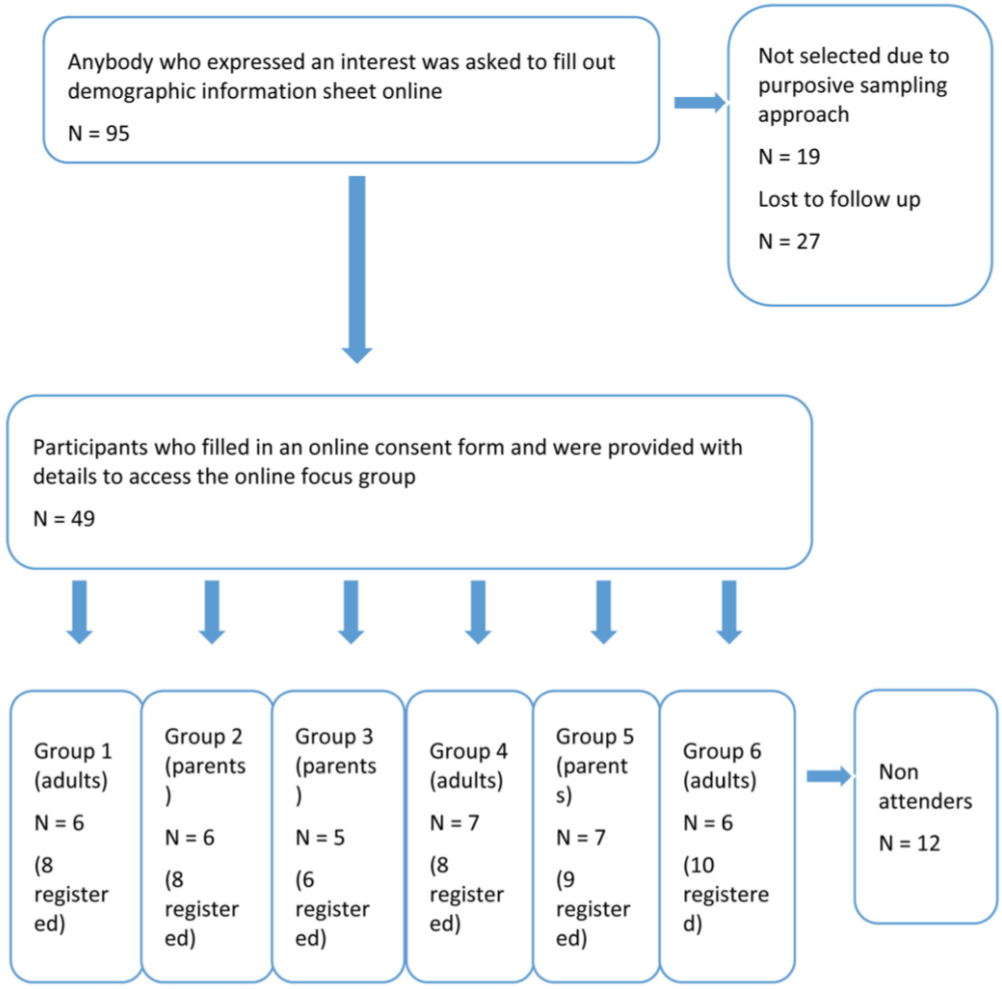


Figure 3-1 Recruitment procedure for UK online focus groups

Note. © Howells *et al.* 2017

3.3.3 Stakeholder / Patient & Public Involvement

The topic guide was developed with input from two caregivers of children with eczema and two adults with eczema, three of whom then participated in a group. It had been planned to conduct a pilot discussion group, but it was not possible to find a time that all could take part, so individuals were sent a copy of the topic guide and comments were returned via email or discussed on the phone. These suggestions were considered, and the topic guide was altered on the basis of this feedback. An

online survey was distributed by the CEBD on twitter to get stakeholder input on how to describe the concept of 'long-term control' by asking stakeholders to rank different terms. Sixty-one stakeholders filled out the survey (eczema patient representatives=32, healthcare professional=5, researcher=5, caregiver=19). The online survey of stakeholders preferred terms indicated that the most common terms that stakeholders would use to describe whether a treatment was working well was 'reducing flares' 35/61 and long-term control 31/61. These words were also the most highly ranked when choosing between terms.

However, following a short presentation of these findings at a CEBD Patient Panel Day in September 2017, patients warned against using the word control with the reasons that it could give an implication that you are doing something wrong, with eczema you don't always feel in control, and it was seen by some as a process rather than an outcome. Reducing flares was also discussed as potentially being too emotive similar to control. Therefore, panel members suggested using the term 'long-term management' or 'flare management'. The study materials used the term 'long-term management' on the basis of these discussions, and during the discussion the facilitators were careful not to introduce language about 'control' or 'flares', but let the language be guided by participants. The language participants began to use was subsequently adopted by the facilitators and often included 'flares' or 'control'.

3.3.4 Analysis

Framework analysis was used to analyse transcripts (Ritchie and Spencer, 2002). Themes were identified at the semantic level, and analysis was conducted in an essentialist/realist framework (Braun and Clarke, 2006). LMH conducted the analysis, but other team members were involved in refining, interpreting and mapping themes. The thematic framework was driven by familiarisation with the transcripts; however, some pre-existing concepts from the literature were considered when developing the framework. Themes were indexed using NVivo V.11. The framework

was iteratively adapted to best fit the data (i.e. as new themes were identified these could be incorporated into the framework). Framework matrices were developed and used for interpreting and mapping the data. No new themes were identified in the final two groups so the researchers were confident data saturation had occurred to the point that additional groups would not have modified the thematic framework and concluded no further UK online focus groups were required (Fusch and Ness, 2015).

3.4 Results

3.4.1 Participant characteristics

Table 3-2 summarises the participants' characteristics (n=37). The majority of participants had been seen in secondary care (had been referred to a hospital specialist) for their eczema (n=34/36, 94%). A minority had taken part in clinical trials (n=4/37, 11%).

Table 3-2 UK online focus group participant characteristics

	Adults with eczema	Caregivers of children with eczema	Children with eczema (reported by caregivers who participated)
N of participants	19	18	-
Sex, n (%)			
Male	4 (21)	-	9 (50)
Female	15 (79)	18 (100)	9 (50)
Age in years, n (%)			
≤5	-	-	12 (67)
6-11	-	-	5 (27)
12-15	-	-	-
16-25	6 (31.67)	1 (6)	1 (6)
26-40	6 (31.67)	15 (83)	-
>40	6 (31.67)	2 (11)	-
unknown	1 (5)	-	-
Ethnicity, n (%)			
White	16 (84)	-	17 (94)
Asian	1 (5)	-	-
Black	-	-	-
Mixed	2 (11)	-	1 (6)
Years since eczema diagnosis			
≤ 5	1 (5)	-	12 (66)
6-11	2 (11)	-	4 (22)
12-15	-	-	-
16-25	5 (26.34)	-	1 (6)
26-40	5 (26.34)	-	-
>40	5 (26.34)	-	-
unknown	1 (5)	-	1 (6)
Self or caregiver reported current disease severity, n (%)			
clear	2 (11)	-	-
Almost clear	2 (11)	-	2 (11)
Mild	4 (21)	-	4 (22)
Moderate	9 (47)	-	7 (40)
Severe	1 (5)	-	5 (27)
Very severe	1 (5)	-	-

3.4.2 Feasibility and acceptability of online focus groups

Out of participants that had provided consent and signed up to a group, the attendance rate was 76% (37/49). Eleven of 37 participants provided feedback on

their experience of taking part in the online focus group. Participants liked the convenience of the online method.

> *Preferred online discussion, much more convenient, no need for childcare. (Participant 29, group 5, caregiver to son, aged 9 months)*

However, some caregivers were distracted or unable to attend the evening groups (18:30–20:00) due to their children's bedtime coinciding with the discussion.

> *I have a screaming tired 3 year old so let's see if we can do bedtime simultaneously with all her zinc wraps etc. ha! Prob not wisest choice of times oops! (Participant 12, group 2, caregiver to daughter, aged 3 years)*

Participants liked the group size, privacy and moderation by the researchers. The method provided anonymity and a non-intimidating platform for all voices to be heard.

> *I thought the discussion group was well moderated and focused, and gave everyone ample opportunities to contribute - and allowed time to think and send responses. (Participant 1, group 1, male, aged 33 years)*

The chatroom indicated when individuals were typing, so the facilitators tried to ensure everyone had responded before advancing discussion. However, typing abilities differed and some said the conversation moved too slowly, while others found the conversation pace to be fast.

> *I think the only negative was the conversation became a bit disjointed at times, as it took people time to type. I don't however know how you could address this. (Participant 28, group 5, caregiver to daughter, aged 5 years)*

3.4.3 Main themes on long-term control of eczema

Figure 3-2 shows how themes 1–3 relate to experiences and understanding of eczema 'control' amongst people with eczema and caregivers, which informs how it should be measured (theme 4).

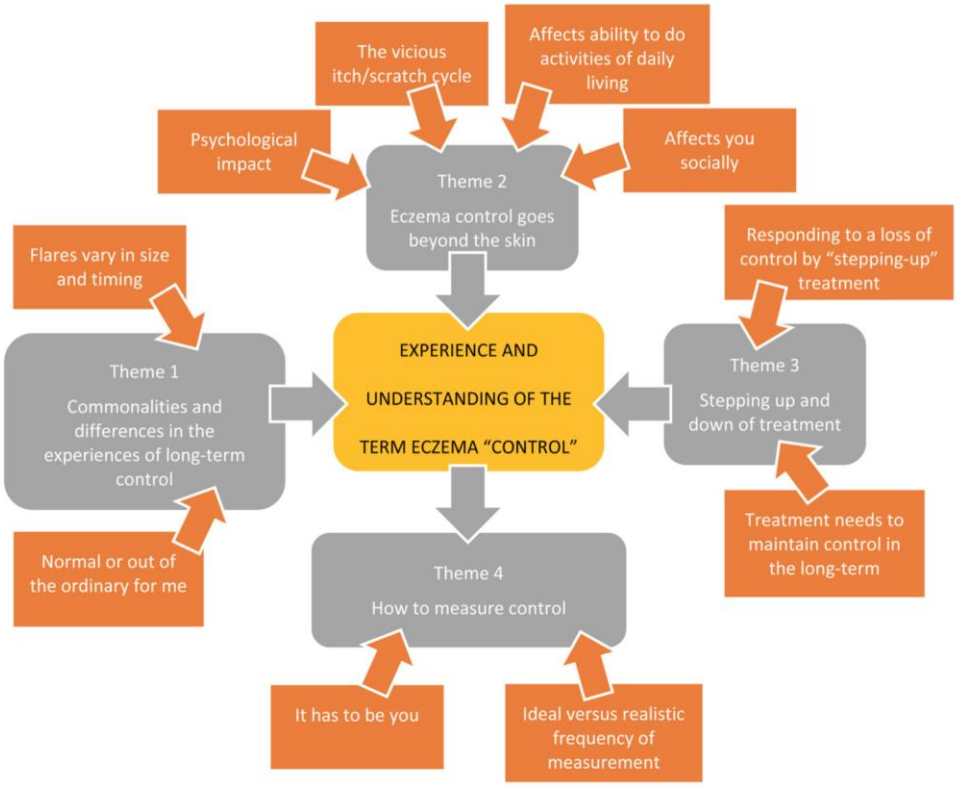


Figure 3-2 Mapping the thematic framework

Note. © Howells *et al.* 2017

Theme 1: commonalities and differences in the experiences of control

1a) Normal or out of the ordinary for me

Increased itching, increased redness and less sleep were often cited as indicators of a treatment not working. However, some participants expressed how their

symptoms, or their child's symptoms, of uncontrolled eczema have altered over time.

I have found some treatments help but he then presents with different types of eczema. Sometimes dry and cracked, other times wet and blistery, some times is spotty and other times is patchy. (Participant 9, group 2, caregiver to son, aged 7 years)

Individuals had different thresholds of acceptable level of control. For some, control was only achieved when there were no symptoms, while many viewed control as a reduction in their symptoms to what was 'normal' for them. For example, one participant described their normal as 'just mild pigmentation marks'. The variation in 'normal' for each individual has important implications for measuring control.

Establishing a baseline of what's 'normal' for the person and how this is deviated from – because it is different for everyone! (Participant 2, group 1, female, aged 23 years)

1b) Flares vary in size and timing

Variations in the severity of flares were apparent in descriptors of size such as 'small flares' and 'big flares'. The length of flares also varied both between and within individuals. Some described flares lasting hours, a day or a few days, while others described flares that lasted months.

If I catch a minor flare quickly enough I can sometimes control the irritation before it get completely out of control to what I call a meltdown. (Participant 5, group 1, female, aged 20 years)

The speed of onset of a flare was varied between individuals and changed for some individuals over time. Flares were often described as sudden in onset with limited early signs, while others described how the build-up could be gradual. Some caregivers said their child will know before them that a flare is coming.

The end of a flare was characterised by better sleep, itching gone or not bleeding on their bedsheets. However, a few commented how it was hard to determine the end of a flare as it never fully went away.

With the eczema being so severe a flare never truely ends, it just is better. (Participant 7, Group 2, caregiver to son, aged 3 years)

Theme 2: eczema control goes beyond the skin

2a) Psychological impact

Periods of uncontrolled eczema were characterised by low mood for adults, children with eczema and caregivers. Caregivers used emotive language to describe how they felt about periods of uncontrolled eczema but reported the psychological impact on their child in broad terms of how it affected the child's general demeanour. The level of control was said to result in either a 'happy child' or a grumpy/distressed child.

Total change in mood. She is happier, carefree and is not frowning. She can play more as not scratching all the time. She sleeps and is not grumpy, less tantrums as rested, able to concentrate and have patience due to sleeping well. (Participant 13, group 3, caregiver to daughter, aged 3 years)

Some groups discussed how flares were accompanied by apprehension due to lack of certainty of how bad the flare would be and how long it would last. A few participants also recalled that even when they were better they were apprehensive of a flare returning.

…fear that it will come back after it's got better is always there. (Participant 34, group 6, female, age unknown)

2b) The vicious itch/scratch cycle

Itchiness to the point that it was difficult not to scratch equated to uncontrolled eczema for many participants.

I couldn't even go to the toilet as he would be scratching, face weeping, awful. (Participant 29, group 5, caregiver to son, aged 9 months)

Scratching was said to often make the eczema worse, often leading to more itchiness, broken skin and bleeding, described as a 'vicious cycle' that needed to be broken to regain control.

It feels good to itch – a temporary relief! Although you know it is making it worse! (Participant 36, group 6, female, aged 41 years)

2c) Affects ability to do activities of daily living

Uncontrolled eczema could impact almost every aspect of people's lives. Sleep disturbance had an impact on concentration the next day. Itchiness and needing to scratch also impaired concentration; however, it was clear that tiredness exacerbated this problem.

I used to go into work on one hour's sleep…and then I'd spend my time staring into space, itching, feeling sorry for myself and unable to snap out of it. (Participant 35, group 6, female, aged 25 years)

Uncontrolled eczema and the sleep loss it causes affected the child's behaviour as well as their concentration. For example, the child was described as 'grumpy' and having 'tantrums'.

Another challenge was the time consumed during a flare period because of what they needed to do differently to normal, for example, applying treatments, although some participants suggested the burden was not limited to flare periods as they engaged in activities to prevent flaring that were also very time consuming.

Adults often felt restricted in clothing due to irritation that certain clothes caused, wanting to cover eczema with clothing, visibility of blood on light fabrics and flaky skin on dark fabrics. Some caregivers altered their child's clothing, which related to irritation rather than concerns about the visibility of the eczema.

I guess what you might be doing differently because of your eczema. Like clothes… covering up eczema, not black because of dry skin, not too light because I might itch and bleed. (Participant 34, group 6, female, age unknown)

Washing and exercise were problems for many adults due to being unable to face the sensations when cracked or dry skin came into contact with water or sweat. Caregivers expressed concern about their child's ability to learn, play and interact.

Several adults and caregivers said impairment to movement was one of the biggest problems faced. Eczema around joint areas, feet or hands impaired their ability to move and do things. Movement was often described as painful.

2d) Affects you socially

Some adults with uncontrolled eczema reported 'embarrassment' and 'social anxiety'. One group suggested this effect was increased when the eczema was more visible.

Embarrassing? I often tend to avoid eye contact, as I'm embarrassed with what people might see…and think. I've lost count of the times that I've been asked things like 'Who beat you up?', 'Dud you cut yourself shaving', and 'Ugh, is that contagious?'. (Participant 18, group 4, male, aged 43 years)

Both adults and caregivers expressed feelings of isolation due to the lack of shared experience and understanding from others, rather than from direct social exclusion (although this was apparent for one participant). Some caregivers expressed concern

for their and their child's physical isolation due to preventing exacerbations and treating the eczema reducing socialisation opportunities.

How they interact, their developmental markers, how much socialisation they get, for us when his skin is bad it's the socialising and getting out of the house that suffers. (Participant 27, group 5, caregiver of a male child, aged 16 months)

Theme 3: stepping up and down of treatment

3a) Responding to a loss of control by 'stepping-up' treatment

Most participants altered treatment depending on the level of eczema control. Many ways of measuring control relating to treatment changes were mentioned, including number of treatment prescriptions, use of extra treatments, use of higher strength treatments, time spent using treatments, frequency of treatment and amount of treatment used.

While most discussed visiting the doctor to seek treatment if the eczema was not controlled, stepping-up treatment was complicated by various factors. Not all participants visited a healthcare professional to step-up treatment. Many would use stronger creams they had at home or increase the amount used of the treatment they are already prescribed. Participants often wanted to step-up treatment promptly to prevent further exacerbation of the eczema, but a few felt their doctor did not understand this urgency and would sometimes delay or avoid giving certain prescriptions.

However, some participants were reluctant to change their treatment in response to worsening of the eczema. Some were fearful that if they use the strongest treatment available, they may have nothing left to use if it gets worse. There were beliefs that a treatment would eventually stop working if you used it too much. Some did not change their treatment as they wanted to try non-pharmacological solutions such as

exercise, diet and silk sleeping suits for children. However, sometimes these methods were used in combination with pharmacological treatments.

> *With that in mind we have been trying a lot of other things and we have been reducing the creams down as much as possible so they will be at maximum effectiveness for flare ups, It's a scary prospect isn't it - having nothing to use as an alternative! (Participant 25, group 5, caregiver of daughter, aged 11 months)*

3b) Treatment needs to maintain control in the long term

It was discussed that control has been gained or a flare has come to an end when you could return to maintenance treatment routines and the eczema symptoms remained controlled.

> *I usually think a flare has ended if I've not had to put steroid cream on that day (Participant 1, group 1, male, aged 33 years)*

Alternatively, if the worsened eczema returned once stepped-up treatment was ended, this was not considered to be long-term control. Some participants described a continuous cycle of stepping-up and stepping-down treatment where the eczema did not remain improved when returning to maintenance treatments. However, it was thought some participants were searching for a 'cure' rather than 'control'.

> *Sometimes we are advised to use steroids for x long… then as soon as stops everything flares up again. Not a long term solution really. (Participant 29, group 5, caregiver to son, aged 9 months)*

Theme 4: how to measure control

Suggested features of long-term control of eczema to be measured were diverse. They can be categorised as observable signs or symptoms (e.g., redness), unobservable symptoms (e.g., itch), treatment used (e.g., escalation of treatment),

scratching (e.g., number of times scratched) the effects on the individual's life (e.g., ability to do everyday activities) and the psychosocial impact (e.g., general mood).

4a) It has to be you

Participants unanimously thought measures should include self-report. Measurement by a doctor was suggested to be too infrequent to capture the fluctuations in eczema.

> *Again it depends on each child, but certainly more frequently than the usual 3 months between consultants visits. We can be fine in the morning as horrendous by bedtime. (Participant 30, group 5, caregiver of male child, aged 11 years)*

Self-report was also preferred because doctors assess the physical aspects of eczema and not how it affects the individual beyond this. The broad array of ways it can affect individuals is illustrated in theme 2.

> *Has to be you. SO subjective a topic, and nurses' doctors can only observe so much - and not the effects it has personally! (Participant 2, group 1, female, aged 23 years)*

Some participants suggested they need to be able to make comparisons with previous levels of control for self-report to be meaningful. This links to theme 1, where it was discussed how everyone has different experiences of 'normal'.

> *It's usually 'compared to what', or on a scale of 1–10, where no number has a real meaning. (Participant 22, group 4, female, aged 31 years)*

Other concerns with self-report included forgetting how the eczema has been if measurements are far apart (i.e., looking back over the last month would be difficult) or 'kidding' themselves their eczema is better. Some participants suggested a measure should be quantifiable and percentage of body involvement was suggested.

Others thought percentage of body affected would not capture the severity of an area affected or if they felt more burdened by eczema on certain body parts.

Although participants highlighted the importance of looking beyond the visual effects of eczema, photographs of the skin were frequently reported as a way to show doctors the eczema fluctuations between visits. Both photographs and diaries were used for self-reflection on eczema changes. Numerous participants said reflecting on how the eczema used to be worse helped them cope. A few were concerned photographs may not show the eczema properly.

4b) Ideal versus realistic frequency of measurement

Participants varied hugely in how often their eczema or their child's eczema needed to be measured from 'how about 24 hours watch!!!' to every few months. For many participants at least daily was deemed necessary to capture a full picture of the disease. However, many felt that the frequency necessary would not be easily achieved for them due to the burden of recording eczema activity. Reasons for this included being too busy, forgetting or not being disciplined enough.

but to be honest don't have time when I am a busy working mother, sometimes it only gets a glance. (Participant 4, group 1, female, aged 50 years)

Analysing group interactions showed caregivers wanting to appear willing to help. Therefore, caregivers were sometimes hesitant to share their reservations with measuring the eczema too frequently, but when prompted with suggestions from previous groups, they did share this concern.

While a few were prepared to measure controlled eczema, many were prepared to measure at frequent intervals during a flare, but not when eczema was controlled. They did not want to be reminded of the eczema, whereas during a flare they would measure it because they cannot help thinking about it.

It's all you think about when it's bad so you'd be prepared to do anything to make it better, like recording its current state, If it was in a bad condition, probably 30 min a day but when it's improved I would rather not think about it, so maybe 10 min. (Participant 33, group 6, female, aged 17 years)

Some suggested frequency should change depending on the stage of treatment. It was thought it would be more necessary to understand fluctuations in periods of uncertainty (e.g., starting a new treatment and ending a treatment).

To overcome the problem of the ideal measurement not being realistic, a few suggestions were made that frequency could be flexible to the individual's needs or measuring tools could be designed so you could skip questions if not needed.

3.4.4 Participant feedback on results

All participants were given the opportunity to feedback on the results, but only four participants provided feedback. All found the analysis to be accurate and insightful. One participant reiterated their belief that an assessment of 'baseline' and 'flaring' should include a holistic appraisal of lifestyle, including mental well-being.

3.5 Discussion

3.5.1 Main findings

This study aimed to find out what long-term control of eczema means to people with eczema and caregivers and what was important to them. Conducting online focus groups allowed us to explore the experiences and understanding of eczema 'control' among people with eczema and caregivers of children with eczema across the UK. Experiences such as symptoms of the eczema worsening or flaring, the impact socially, psychologically and ability to carry out daily activities, scratching, and changes in the treatment and management of the condition were all prominent

features of uncontrolled eczema (Figure 3-2). While some of these experiences have been reported elsewhere, this study specifically helps us understand how these experiences relate to the concept of 'control' (Chamlin et al., 2004, Gore et al., 2005, Santer et al., 2015a, Noerreslet et al., 2009, Santer et al., 2012, Santer et al., 2013). Dissatisfaction with treatment that did not lead to controlled eczema once the treatment was stopped mirrors interviews with caregivers that found dissatisfaction with the trial-and-error approach to eczema treatment in primary care (Santer et al., 2012).

The language used to discuss long-term control of eczema can vary. 'Long-term control' is a combination of two concepts: the timeframe and disease activity. There has been international consensus that long-term control of eczema should be measured in clinical trials of eczema treatments that are 3 months or longer in duration (Schmitt et al., 2012, Chalmers et al., 2014). The trial context was not discussed with participants in these focus groups so that the focus was on participants' individual and everyday experiences. Participants spontaneously used the term 'control' during discussions. Participants also frequently used the terms 'flare'/ 'flare-up', but what experiences constituted a flare was highly variable. Some people had 'chronic' flares lasting for months, whereas some lasted hours or days.

The most common way of measuring long-term control of eczema in published randomised controlled trials has been repeated measures of clinician-reported signs, usually on a monthly basis (Barbarot et al., 2016). Participants talked about how they would measure control as repeated measures of various factors such as quality of life, itch or mood. They generally preferred self-reported measures due to not all effects being observable by a doctor and enabling frequent measurement to capture fluctuations in eczema more fully. There are some patient-reported outcome measures that have been previously developed for use in eczema clinical practice and research, but none sufficiently capture all aspects of eczema control from a patient perspective (Gerbens et al., 2016, Heinl et al., 2016, Barbarot et al., 2016).

This is perhaps because previous patient-reported outcome measures have not been developed to measure this specific construct.

Published randomised controlled trials have measured flares using number of flares, time to first flare and, to a lesser extent, duration of a flare or remission period (Barbarot et al., 2016). The size and length of flares could differ substantially and have varying effects on the individual, which highlights how measures only capturing frequency or length of flares would not capture the impact each flare has on an individual.

Treatment escalation has been used to measure flares. Two studies have defined a flare as being: 'an episode requiring escalation of treatment or seeking additional medical advice' (Langan et al., 2009, Thomas et al., 2015). A similar definition has been proposed by the European Task Force on Atopic Dermatitis: 'acute, clinically significant worsening of signs and symptoms of atopic dermatitis requiring therapeutic intervention' (Wollenberg et al., 2016). While participants did escalate treatment in response to a flare, there were a number of complexities that could make this difficult to implement as a measure in all trials. The Necessity-Concerns Framework suggests people hold beliefs about how necessary a treatment is to maintain their health and concerns about the treatment having adverse effects, both of which can influence treatment adherence (Horne et al., 1999). Beliefs expressed by some participants in this study suggest they are concerned about their medication as they were reluctant to step-up treatment, whilst others were concerned about stepping-up treatment for fear that they would be left with no alternatives once they had stepped up the treatment. Therefore, lack of treatment adherence may present a problem for using treatment as an indicator of eczema 'control'. Alongside changes to treatment, a behavioural response to loss of control many participants cited and suggested measuring was increased scratching, which is one method that has been considered as a way of measuring long-term control (Langan et al., 2006b, Thomas et al., 2002).

3.5.2 Strengths and limitations

To our knowledge, this is the first qualitative study to consider long-term control of eczema from the perspective of people with eczema and caregivers. The online focus groups reduced barriers of geographical location and time constraints and removed costs for travel, venue and transcription. Previously, online focus groups have been cited as enabling participation by people with visible skin conditions who lack the confidence to attend a face-to-face focus group (Fox et al., 2007). This is supported by feedback from participants that they liked the anonymous and inclusive nature of the group.

Non-response bias is where respondents differ in a meaningful way from non-respondents. There was a predominance of people of white British ethnicity and of female adults with eczema, despite targeted efforts to recruit male and ethnic minority participants. This is a trend seen across a variety of research studies. Out of those who filled out their characteristics and were subsequently invited to an interview, there was no discernible difference in the characteristics collected of people who participated and those who were lost to follow up and did not participate.

Only two participants reported never having seen a hospital doctor for their eczema, therefore the sample under-represents experiences for participants treated in primary care, which is the majority of eczema patients in the UK (Emerson et al., 1998). It is likely that those who have more severe eczema were likely to have seen the advertisements for the study as they will perhaps be more likely to seek out eczema related online content or have taken part in eczema research previously. Despite not being typical of a primary care population, there were 8 (44%) of adults and 6 (33%) carers of children who reported disease was currently clear to mild, so this indicate that the results may also apply to milder disease.

Online methods present distinct challenges for qualitative research. Language needs to be clear as there is no voice intonation or non-verbal cues. Prior preparation of questions allowed precision in phrasing. Participants valued that typing allowed them to provide considered answers. Threading, where multiple strands of conversation occur in parallel, is typical of online focus groups (Moore et al., 2015, Stewart and Williams, 2005, Fox et al., 2007). Addressing an individual's comment while simultaneously inviting the whole group to comment on that specific point helped to maintain a coherent discussion and allow further exploration of the topic.

Qualitative research has the potential to develop both breadth and depth of knowledge (Keeley et al., 2016). Breadth is required to understand the scope of the experience and depth enables understanding the detailed complexities of the experience (Keeley et al., 2016). The method tended to produce brief responses but prompting was successful at eliciting more detailed responses. Reading responses was difficult when the discussion moved quickly, so having multiple researchers available to engage in different tasks was beneficial. Eleven participants provided feedback on the methodology; however, only four provided feedback on results. However, all feedback suggested they found the results to be an accurate representation of the discussion.

3.5.3 Implications for research and clinical practice

This study investigates an aspect of eczema that has been under-researched from the perspective of people with eczema and their caregivers. The box below (Figure 3-3) highlights the important ways this study may inform decisions on the best way to measure long-term control in people with eczema.

This study is the first stage of an international qualitative research project that will aim to understand long-term control of eczema for people with eczema and caregivers in different countries to ensure an international perspective. The methodology is globally accessible and facilitates composite analysis to identify differences across countries. Since online focus groups are a relatively novel method,

the lessons learnt from this study will provide guidance for international collaborators. It was also acknowledged that there may be limitations in the depth of understanding that can be obtained using these relatively novel online methods, therefore, later work aimed to see if these findings were reproduced in a face to face focus group in a UK setting (Chapter 5).

From these findings, it is recommended that the following seven points should be taken into consideration when deciding how to measure long-term control of eczema:

1. The need to scratch uncontrollably, the psychological impact, the social impact, symptoms (including itch, pain, sleep), impaired movement, the ability to do everyday activities and treatment used were all indicators of level of control.
2. Understanding the baseline of what is normal for an individual was considered important for understanding eczema control.
3. Not everyone with eczema experiences "flares" or finds it easy to notice changes in their eczema.
4. The behaviour of stepping up treatment for a flare was common but was complicated by factors such as difficulty getting a prescription and concerns about stepping-up treatment when already using maximum treatment available.
5. For some participants taking treatments that are recommended for short term use, lack of control referred to "rebounds" after the treatment ends. Long-term control of eczema measurement should take into account control both during and after treatment.
6. The patient/caregiver perspective was considered important to fully capture eczema control. The ability of caregivers to report eczema control in young children was not questioned by caregivers.
7. The acceptance of measurement frequency varied between participants and for an individual over time depending on lifestyle and commitments, the treatment stage and level of eczema control.

Figure 3-3 Box with seven key recommendations for measuring long-term control of eczema

3.6 Summary of Chapter 3

This qualitative study has shown the complexity of the experience of long-term control of eczema for people with eczema and caregivers in the UK. 'Eczema control' is multi-faceted and can have a variety of meanings for people with eczema and caregivers of children with eczema, which has important implications for how long-term control of eczema may be measured. Overall, this pragmatic online research method was embraced by participants and enabled qualitative research to be conducted effectively with limited resources. This suggested the methods could be used in other countries to build on the international input informing the HOME initiative. Chapter 4 will present how this study was scaled up to meet the aims of the HOME initiative.

Chapter 4

Defining and measuring "eczema control": An international qualitative study to explore the views of those living with and treating eczema

4.1 Introduction

Content validity is essential for an instrument to be selected for inclusion in a core outcome set and is defined as "the degree to which the content of a health-related patient-reported outcome instrument is an adequate reflection of the construct to be measured" (Mokkink et al., 2009, Brod et al., 2009). A key initial step in establishing content validity is defining the construct to be measured. A construct is a mental abstraction that provides a common language that has a shared meaning to help us communicate in a clear and precise manner. A conceptual framework of the construct is then developed to allow judgement of whether an instrument adequately reflects the concepts relevant to the construct of interest (Brod et al., 2009). Qualitative research helps to build a definition of a construct that authentically and comprehensively reflects patient experiences (Apfelbacher and Nelson, 2017).

It was agreed at the HOME IV meeting that the next steps for the domain 'long-term control of eczema' was to conduct international studies to gather patient and healthcare professional descriptions of 'eczema control' (Chalmers et al., 2016). Chapter 3 presented a UK based study that began to fill this research gap by gathering the perspectives of people with eczema and caregivers. However, given the scope of the HOME core outcome set, it was considered important that exploration of a HOME domain should be as globally inclusive as possible and consider the views of multiple stakeholders. Therefore, this chapter builds on this

body of work to consider the perspectives of people with eczema, caregivers, clinicians and researchers across multiple nations.

4.2 Aims and Objectives

This study aims to facilitate evidence-based discussions within HOME about how to define and measure “long term control of eczema” and enable assessment of the content validity of existing outcome measurement instruments.

The study objectives were:

(1) to understand what long-term control of eczema means to people living with eczema (including adults with eczema and caregivers of children with eczema) and people treating eczema (clinicians/researchers).

(2) to explore the potential feasibility and acceptability of different ways of measuring long-term control of eczema.

4.3 Methods

4.3.1 Study Design

To obtain the perspectives of people with eczema/caregivers and clinicians/researchers treating eczema, online focus groups and an online survey were carried out respectively. This article presents an international collaboration of multiple data sources. The online focus groups received ethical review from the following institutions: the University of Nottingham, Faculty of Medicine & Health Sciences Research Ethics Committee (F14062016 SoM ROD), METc Groningen (METc 2016/664), Osaka Habikino Medical Center (831 on 30th March 2017), Uppsala University (2017/106). The following ethical review bodies declared this research

exempt from requiring ethical review: Nantes University Hospital Ethics Committee, Lurie Children's Hospital Institutional Review Board (2017-1033). The online survey was completed by the HOME membership and ethics committee review was not required.

4.3.2 Online focus groups with people with eczema and caregivers

4.3.2.1 Participant selection

People with eczema and caregivers were invited to take part in the study. The sampling strategy aimed to purposefully include a diversity of participants regarding age, sex, ethnicity, eczema severity and disease duration, but the practicality of who could be conveniently accessed to take part in the study was also a factor (Ritchie et al., 2013). Recruitment was via social media or approaching patients in clinics and varied by country (Appendix B provides more details per country).

4.3.2.2 Procedure and materials

Online focus groups were conducted via text-based chatroom websites, where facilitators and participants type responses in real-time, between August 2016 and June 2017 in the UK, the Netherlands, France, Sweden, USA and Japan. Apart from the UK where six focus groups were conducted, each local team conducted two focus groups; one with adults with eczema and one with caregivers of children with eczema. The countries were chosen on a basis of where HOME members collaborating on this project were based. Local teams carried out focus groups in their own country and native language. The table in Appendix B documents the key aspects of the procedure for each country.

The study was centrally co-ordinated from the UK. The UK team supported collaborators in the conduct of their online focus groups by providing methodological guidance both via documents and teleconferences. The methodology was informed by the format of the initial UK-based focus groups. Full

details of the procedure of the UK focus groups has been described in section 3.3.2. The methodological guidance focused on key issues to consider when running the online focus groups i.e. terminology, recruitment, setting up the chatroom, strategy for asking questions and responding to answers. The collaborators were also sent ethics application, protocol, topic guide and participant information sheets that were used for the UK groups. Teleconferences and email communication were then used to answer any queries about the procedure and cover any challenges that collaborators were facing (e.g. applying for ethical approval). Each focus group followed a common semi-structured topic guide. The topic guide from Chapter 3 had been used and adapted throughout the UK focus groups, therefore the final version used in the UK focus groups was adopted for subsequent focus groups (Appendix A). However, where required the individual groups translated the topic guide and adapted it to meet their needs.

4.3.2.3 Analysis

Each local team used a thematic framework to map findings to the sub-themes that were based on the findings from the UK focus groups (section 3.4.3) and also highlight findings that did not fit within the framework (Ritchie and Spencer, 2002). This was done in the native language of the participants. The detailed UK findings were not shared with the other groups until they had analysed their findings. A summary of findings and key quotes were produced in English by each local team and provided to the UK team. The UK team (LMH and JRC) collated and compared the data from all countries in a thematic framework. Any areas of uncertainty were discussed with the researchers who collected the data.

4.3.3 Online survey of HOME membership

4.3.3.1 Participant selection

All members of HOME were invited to participate. HOME is an international, multi-stakeholder group including clinicians (mainly dermatologists), methodologists,

patients, patient representatives and members of the pharmaceutical industry (http://www.homeforeczema.org/).

4.3.3.2 Materials and procedures

This online survey took place during September and October 2016 using SurveyMonkey Inc. software. Using mainly open questions with free text responses, participants were asked what they considered to be long-term control of eczema and their views on different ways of measuring long-term control (Appendix C provides survey content) identified from a systematic review (Barbarot et al., 2016).

4.3.3.3 Analysis

Thematic analysis was used to analyse textual data (Braun and Clarke, 2006). JRC compiled the themes and IN and LMH reviewed the themes, which were then discussed and revised.

4.3.3.4 Combining the focus group and survey findings

LMH and JRC combined the themes from the online focus groups with people with eczema/caregivers and the clinician survey into an overarching thematic framework. The UK online focus groups presented in Chapter 3 were included in this analysis. These themes were then shared with all authors so that the results could be compared with original data sources and local teams could confirm that the results appropriately reflect their findings.

4.3.4 Stakeholder / Patient & Public Involvement

Since this study was in part an extension to the study presented in Chapter 3, the PPI presented in section 3.3.3 is relevant to this study too. However, in some respects the whole project presented within this chapter can be seen as engagement to get wider stakeholder input on the domain of 'long-term control of eczema' to inform the HOME initiative. This stakeholder input was vital to lead to the outcomes of a

long-list of concepts that could be discussed for consideration at the HOME V meeting and to provide feedback on the feasibility and acceptability of different approaches to measurement so that this could also be presented at the HOME V meeting.

4.4 Results

4.4.1 Participant Characteristics

Sixteen online focus groups took place including 97 people with eczema/caregivers across six countries. Table 4-1 provides participants' characteristics.

Table 4-1 International online focus groups participant characteristics

	UK**	The Netherlands	France	Sweden	USA	Japan	All countries
Focus groups, N	6	2	2	2	2	2	16
Participants, N	37	15	9	15	8	13	97
People with eczema, N	19	7	5	7	4	10	52
Caregivers, N	18	8	4	8	4	3	45
Self or caregiver reported sex, n							
Male	13	6	4	4	0	8	35
Female	24	9	5	4	8	5	55
Unknown	0	0	0	7	0	0	7
Self or caregiver reported age in years, n							
≤5*	12	5	0	4	0	2	23
6-11*	5	1	4	2	3	1	16
12-15*	0	2	0	2	1	0	5
16-25	7	0	0	2	3	4	16
26-40	6	1	2	2	0	0	11
>40	6	6	3	2	1	6	24
Unknown	1	0	0	1	0	0	2
Self or caregiver reported ethnicity, n							
White	33	14	9	15	4	0	75
Asian	1	0	0	0	2	13	16
Black	0	0	0	0	1	0	1
Mixed	3	1	0	0	1	0	5
Unknown	0	0	0	0	0	0	0
Self or caregiver reported years since eczema diagnosis, n							
≤5	13	5	1	5	0	3	27
6-11	6	1	3	2	3	0	15
12-15	0	2	0	1	0	2	5
16-25	6	0	2	3	4	4	19
26-40	5	2	1	3	0	2	13
>40	5	5	1	1	1	2	15
Unknown	2	0	1	0	0	0	3
Self or caregiver reported current disease severity, n							
Clear	2	0	1	0	0	0	3
Almost clear	4	2	2	3	0	2	13
Mild	8	4	5	7	1	4	29
Moderate	16	5	1	3	3	4	32
Severe	6	4	0	1	4	3	18
Very severe	1	0	0	0	0	0	1
Unknown	0	0	0	1	0	0	1

*Participants were caregivers for all aged under 16 years.

**UK online focus groups were also presented in Chapter 3.

Sixty-two out of 251 HOME members from sixteen countries completed the online survey, a response rate of 25%. Most (81%) were clinicians, plus 5 methodologists/non-clinical researchers and 7 patient representatives. Since this survey principally represents the views of clinicians, this group is hereafter referred to as clinicians for brevity. Sixteen countries across six continents were represented (Figure 4-1).

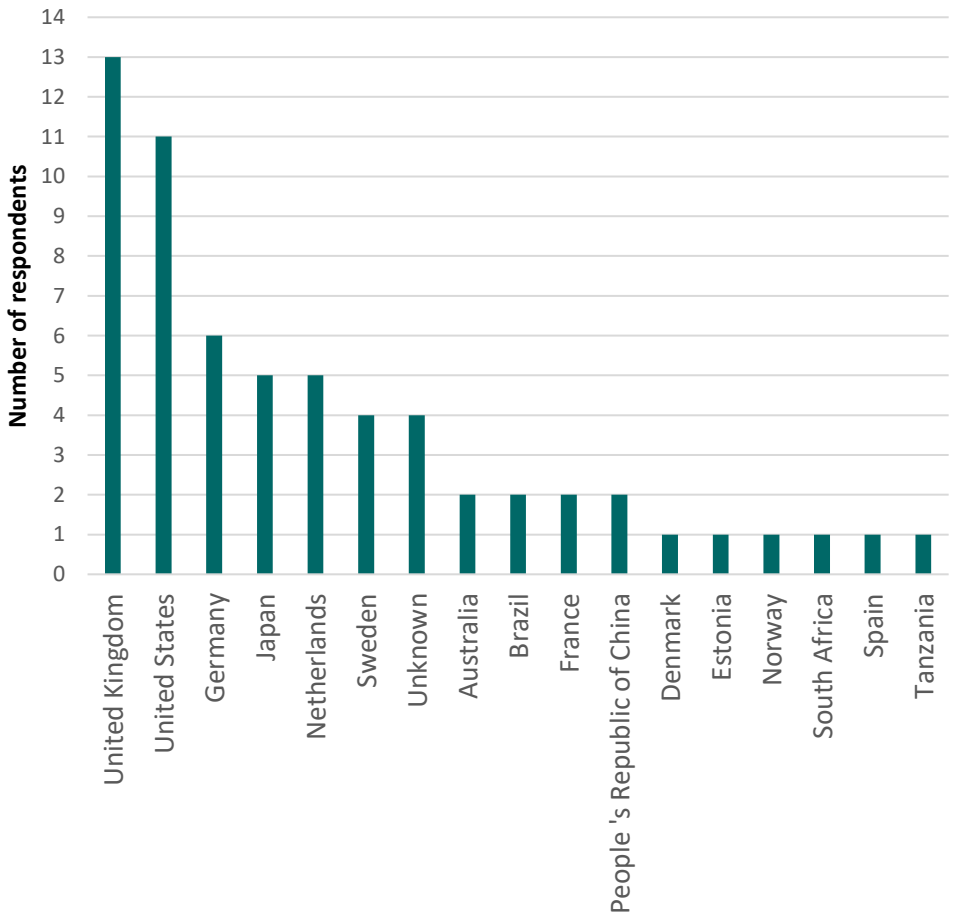


Figure 4-1 Number of online survey respondents by country

4.4.2 Overarching themes

The results can be structured into two overarching themes; i) long-term control as a multifaceted construct and ii) long-term control is complex to measure. Illustrative quotes that link themes to the data are presented in Table 4-2 and Table 4-3.

1) Long-term control as a multifaceted construct

Most participants described multiple related aspects of eczema when thinking about 'what is long-term control of eczema?' which divide into four key sub-themes (Figure 4-2).

1a) Long-term control as disease activity

People with eczema/caregivers and clinicians frequently described long-term control as a reduction in disease activity. Signs and symptoms such as the level of itch, pain, and redness were often mentioned, although people with more severe eczema were also concerned about complications such as infection and bleeding. Clinicians used terms such as reduced intensity, minimal signs and/or symptoms or minimal disease activity. Differences in some symptoms by country were noted; for example, only participants in France discussed "smoothness of skin" as an indicator of control.

Disease activity in relation to flares of activity was also described. A reduction in the intensity, number, frequency and duration of flares were all parameters suggested as indicators of disease control. However, the concept of "flares" was not seen as universally useful, as a flare was difficult to define and did not always correlate with disease activity. For example, those with a continually high level of disease activity may not experience flares, despite having uncontrolled eczema.

1b) The experience of long-term control goes beyond the skin

A positive impact on daily activities was considered important to people with eczema/caregivers and clinicians. People with eczema/caregivers reported a wide

range of daily activities that were affected by eczema, but these varied among countries. For example, in the UK, washing, exercise, and clothing choice were discussed, whereas in Japan it was disturbance of concentration while reading books and watching TV. Scratching, sweating, pain, sleep disturbance and lack of ability to concentrate were amongst the ways that people with eczema and caregivers described eczema having an impact on their daily life.

The emotional impact of uncontrolled eczema was raised by people with eczema/caregivers, who often reported high levels of distress when eczema was uncontrolled. Feeling frustrated, miserable and "stressed" when eczema was uncontrolled, apprehensive about the return of flares, and the social impact on people with eczema and their families were all mentioned. Feeling embarrassed or receiving comments from others were discussed, particularly when the eczema was on a visible area such as the face. Some people with eczema/caregivers felt that eczema controlled their lives and prevented social activities, such as visiting friends and family and school attendance. Clinicians also described the impact of uncontrolled eczema on patient's quality of life.

1c) Long-term control linked to treatment and management decisions

Long-term control was also linked to treatment use by people with eczema/caregivers and clinicians in all countries. Reducing treatment or returning to maintenance treatment were indicators of re-gaining control. Using only maintenance treatment and ability to self-manage were indicators of ongoing control. Seeking help from a doctor, stepping up treatment, or increasing the amount or frequency of treatment were all described as indicators that the disease was uncontrolled.

1d) Control is an individual experience

Control of eczema was largely considered to be an individual experience. Individual people with eczema/caregivers reported different aspects of the disease as being

representative of a lack of control, such as specific symptoms, the need to increase treatment, or the impact on particular aspects of life. The level at which disease activity or impact represents control varied between individual people with eczema/caregivers. Some expressed that feeling the eczema has “completely receded” would represent control, whereas for others, a reduced and acceptable level of disease activity was considered controlled. Clinicians often linked control to being what is acceptable to the individual patient such as “can live with” or that is “acceptable to the patient”.

Table 4-2 Illustrative quotes for Theme 1

Theme 1: Long-term control is a multifaceted concept		
Sub themes	**Illustrative quotes (people with eczema/caregivers)**	**Illustrative quotes (clinicians/researchers)**
1a) Disease activity		
Improved signs and symptoms	I feel long-term control means less pain, less itch, less scratch – adult, Japan	Achievement and maintenance of a low level of symptoms and signs of AD over time – clinician, Germany
Flares	My skin really hates me. Usually when it has completely receded, I am still waiting for this eczema to die down completely. – adult, UK	A treatment plan that prevents flares for a longer period of time (years) – Clinician, Denmark
1b) Beyond the skin		
Emotional impact	Her whole demeanour changes too - weepy, fiery temper, generally sad. – caregiver, UK In case of big flare, my skin gets worse and worse spontaneously even though I do not scratch. I feel quite depressed in this process – adult, Japan	
Social impact	Children at school can be quite unkind when it's sore looking – caregiver, UK How they interact, their developmental markers, how much socialisation they get, for us when his skin is bad it's the socialising and getting out of the house that suffers – caregiver, UK	Re-establishment of normal sleep patterns Re-establishment of normal social activities and ADLs [activities of daily living]. Re-establishment of normal family dynamics – clinician, South Africa
Family impact	"center of our lives" - caregiver, USA	
Effect on my day (including the impact of scratching and loss of concentration)	I would say how much I feel my eczema is bothering me in everyday life, what you might be doing differently because of your eczema – adult, UK and then I'd spend my time staring into space, itching, feeling sorry for myself and unable to snap out of it – adult, UK Permanent sensation of scratching non-stop - adult, France	Living with eczema of a tolerable level without flares that start to interfere with work and play – clinician, UK Increase in symptoms, itching, scratching behaviour that may impact on daily activities / sleep and quality of life. – clinician, UK
Sleep disturbances	It's important to me to create more rest in the long-term and to be able to sleep well. Then I have more	Disease improvement that no longer affects sleep and daily function that lasts for more than 6 months – researcher, US

	energy to do my daily things – adult, The Netherlands How much trouble I have, for example, I get a problem with sleep when the eczema is at its worst – adult, Sweden	
1c) changes in treatment and management		
Using only maintenance treatment	can return to normal maintenance routines – adult, UK	No itch; controlled by emollient only. – clinician and researcher, Japan
Stepping down treatment	…I realize long-term control by frequency of TCS application. While maintaining by tapering frequency of TCS, I recognize long-term control by no flare despite less frequent TCS. – adult, Japan	The end of using the rescue medication, because overall disease severity is back to pre-flare levels. – clinician, Germany
Self-management of the eczema	I guess so. We see the doctor more during a bad patch, but when I can manage it at home we see him less. I only go to him if I need him when it's beyond me level of helping. – caregiver, UK	Self-management of control of eczema flares – researcher, Germany
1d) An individual experience		
A level of eczema acceptable to me	I think everyone has his own definition of 'control'. For me the bar is set fairly low. I will have itch every day and I accept it. As long as it's not constantly there. – Adult, Netherlands	Over time I have been taught by patients that individually, patient satisfaction with their skin and the treatments used varies enormously and what for me as a clinician is poor control is entirely acceptable for a patient who has reached a level of control that they feel manageable and which does not interfere with their life style psychosocial health or general health (read adverse effects especially). How does one integrate measures of this sort into trials and give them meaning? – clinician and researcher, South Africa

Note. Some quotes have been translated into English from another language. Spelling and grammar have been edited for ease of reading.

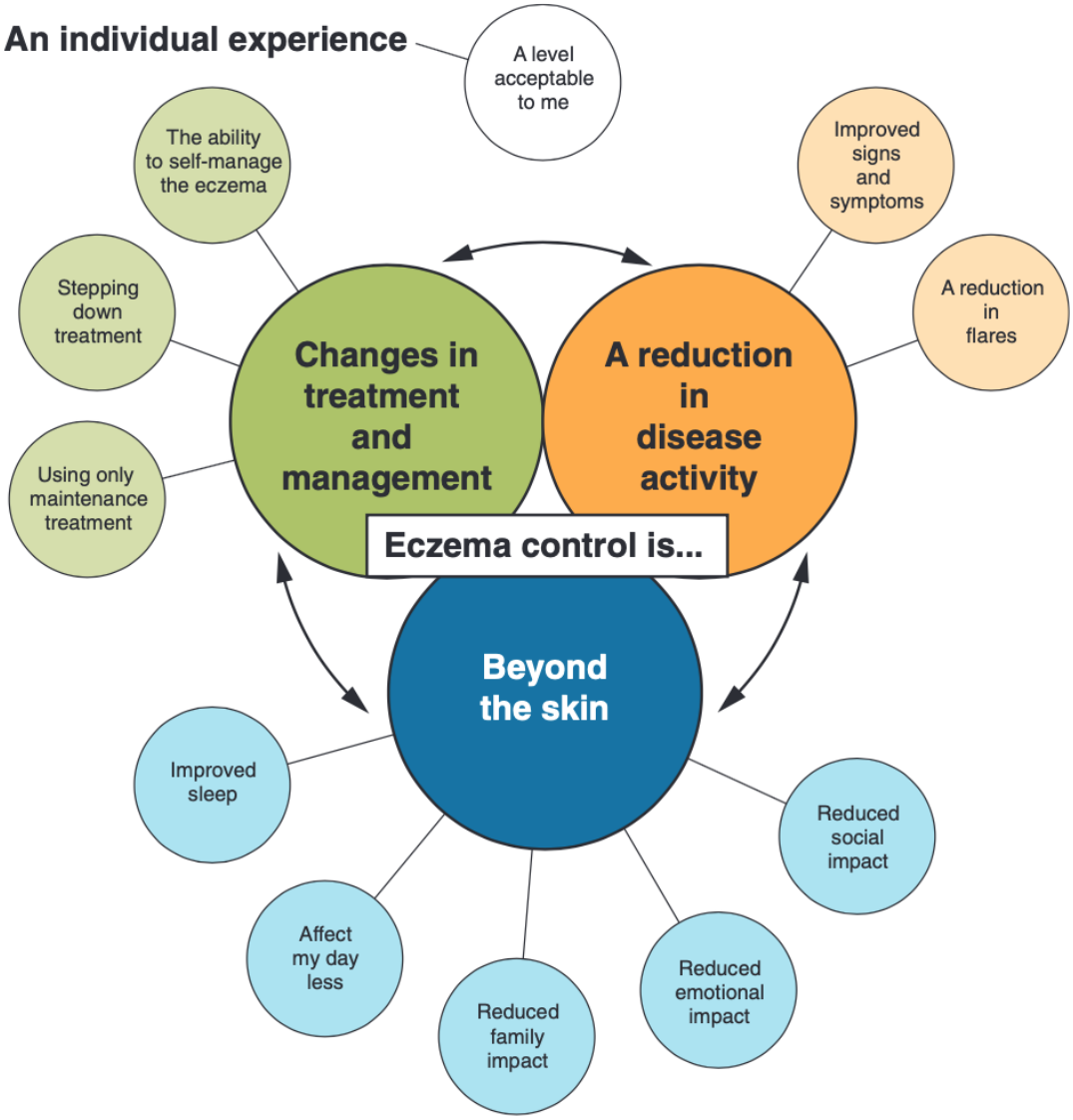


Figure 4-2 Graphic illustration of eczema control.

Note. © Howells *et al.* 2019

2) Long-term control is complex to measure

People with eczema/caregivers and clinicians also discussed what they felt was important when measuring long-term control, and the different views reflect the complexity involved in measuring this multi-dimensional construct.

2a) Who should measure long-term control? People with eczema/caregivers vs. independent observers

The majority of people with eczema/caregivers felt strongly that they were best placed to understand and measure their own (or their child's) eczema and could measure disease activity between visits to their doctor. Some people with eczema/caregivers felt clinicians should also measure the eczema, as they have expertise, experience and perceived them to be less subjective. Clinicians acknowledged the importance of patient-reported outcomes as an important way of capturing aspects of the disease that cannot be assessed by a clinician, but for measuring long-term control in clinical trials they were also concerned about potential for bias, reproducibility and scientific acceptability, and potential discrepancies between patient reported and clinician reported outcomes. These findings suggest stakeholders see benefits of both patient and clinician reporting as it was felt important to both measure aspects of control that only people with eczema/caregivers can assess that captures their everyday experience as well as having a method of measuring long-term control that can be assessed by independent observers.

2b) The burden and feasibility of measuring long-term control

Many people with eczema/caregivers raised concerns about the time and effort required to measure long-term control comprehensively, whereas some, particularly caregivers of children with eczema, were prepared to go to great lengths to ensure that the level of control was captured accurately and frequently. Clinicians also highlighted the potential burden of frequent measurement on people with eczema/caregivers resulting in the generation of large amounts of potentially redundant data. In the USA, the Netherlands and the UK, people with eczema/caregivers discussed the role that technology such as smartphone applications and photographs could help them measure their eczema more frequently and for long periods of time.

Table 4-3 Illustrative quotes for Theme 2

Theme 2: Long-term control is complex to measure		
Sub themes	**Illustrative quotes from people with eczema/caregivers**	**Illustrative quotes from clinicians/researchers**
2a) Who measures long-term control	You know your own body and eczema best, a doctor has expertise and experience. – adult, The Netherlands Has to be you. SO subjective a topic, and nurses' [and] doctors can only observe so much - and not the effects it has personally! - adult, UK	Recording of observed signs in parallel with patient-reported symptoms and QoL will add information. Taken together, these measures would probably be judged by many as more robust and valid. – clinician and researcher, Sweden These measures might be more standardized across a population. – clinician and researcher, United States
2b) The burden and feasibility of measuring long-term control	Again it depends on each child, but certainly more frequently than the usual 3 months between consultant visits; we can be fine in the morning as horrendous by bedtime. – caregiver, UK I do not want to observe every bad aspect of my skin, as it makes me depressed. I never want to take a picture of worsening skin. I think we should look at the better aspects. – adult, Japan The time interval might depend on individuals needs and severity of eczema. – adult, Sweden	Motivation of patient to describe regular frequent diary is needed. Reliability that the patient regularly and surely describe each outcome is doubtful. – clinician, Japan [Discussing concept of well-controlled weeks] Difficult to define/assess, Might be difficult to assess in a standardised way and might result in difficulties to merge data/compare trials - clinician and researcher, Sweden

Note. Some quotes have been translated into English from another language. Spelling and grammar has been edited for some quotes for ease of reading.

4.5 Discussion

4.5.1 Main findings

This study suggests eczema control is a multifaceted construct involving changes in the signs and symptoms of eczema, psychological, social and physical functioning, and the treatment and management of the condition. Indicators of control and the acceptable level of control can vary between individuals. Both patient-reported and clinician-reported outcomes were considered important when measuring eczema control.

There is no previously agreed definition of an eczema flare, although many definitions have been proposed and used in trials, often with little validation (Langan et al., 2006b, Langan et al., 2014). However, this current study suggests that eczema control is viewed by people with eczema/caregivers and clinicians as a broader concept than these previous definitions.

4.5.2 Implications for measuring long-term control

1) No measurements in clinical trials to date cover all important concepts

Although clinical trials have captured individual sub-domains identified in this study as being important to people with eczema/caregivers and clinicians, the multiple aspects important to long-term control have not routinely been captured as a unified construct using a single instrument (Barbarot et al., 2016).

2) Variation in experience will be a challenge for measurement

The variability of experiences, both for an individual over time and between individuals, present numerous challenges to the measurement of long-term control of eczema. There was variability in the details of what people responded as being important to them. For example, a lot of participants gave different activities specific that indicated to them that eczema was uncontrolled. For example, for some it was to do with washing, for others it was watching TV or reading, for others it was to do with getting dressed, whilst for others it was whether their child could play. Therefore, if designing a measurement tool, it would be hard to capture aspects that were relevant to all if the items were too specific.

The variability between people with eczema regarding the level of control considered to be acceptable, highlighted by both people with eczema/caregivers and clinicians, presents a challenge for measuring eczema control. What constitutes control for an individual with eczema may be driven by their expectations of the disease course, their treatment, and the degree to which they have accepted having

the disease and the lack of a cure. It is possible that expectations about the level of disease control in eczema may be altered in the future by advances in treatments. It is important to consider how the expectations of the person may impact measures of eczema control. For a self-reported instrument, there may be changing standards depending on level of adjustment to the condition.

3) Patient/caregiver and clinician reported outcomes were favoured

Both patient/caregiver and clinician reported outcomes were considered important in the measurement of eczema control. There is an increasing acceptance within the medical community of patient-reported outcome measures and a patient-centred approach to healthcare and a systematic review showed that patient-reported symptoms were reported in 78% of eczema clinical trials (Gerbens et al., 2016). However, with both patient/caregiver and clinician reported measures, consideration of the burden and feasibility of measurement is needed.

4.5.3 Strengths and Limitations

The learning from the initial UK-based online focus groups was shared with other authors running the subsequent online focus groups in other countries to elicit more detailed responses and prompt a coherent discussion amongst participants. The analysis was initially conducted by the local teams. This was chosen to ensure the same people who collected the data also analysed the data, due to the importance of 'familiarisation with data' as an early stage of the data analysis process (Ritchie and Spencer, 2002). It also overcame barriers relating to data sharing between institutions and language translations which would have been challenging to conduct with the research budget available and to be conducted in time for presenting the findings at the HOME V consensus meeting. A limitation of this is that the overall analysis and interpretation of the results was based on summaries and key quotes and so the co-ordinating researchers were a step removed from the data and could have potentially missed some nuances. However, at this stage of the project, the purpose was to guide a broader understanding of how to conceptualise the domain

'long-term control of eczema', and therefore this was perceived by the research team as a useful way of gathering international perspectives for this purpose.

A convenient online survey using mainly open text responses was used to gain the clinician perspective. This method was considered appropriate because HOME members are familiar with the concepts being discussed and responses were generally clear and relevant.

Given that only HOME members were invited to take part in the survey, the sampling method only targeted clinicians who are interested in eczema research. This study was designed to inform measurement; therefore, sampling targeting those interested in eczema research may produce valuable insight. However, although the HOME initiative is an international collaboration with people participating from diverse locations, it is still predominantly made up of Western countries, therefore experiences from certain cultures may be missed. Out of the members emailed, only 25% of the HOME membership took part in the survey, and it is possible that views may not have been representative of the entire HOME membership. It is likely that those who did not respond were less engaged in the subject and consensus process. The findings from this study informed decisions at the HOME V meeting in Nantes, France, therefore allowing HOME members who did not participate in the survey further opportunity to input into HOME consensus decisions (Chalmers et al., 2018).

Considering how the characteristics of the online focus groups sample relate to the wider population, efforts were taken to include focus groups from different geographical locations. However, despite the inclusion of Japanese participants, the availability of resources resulted in a predominance of participants in Western countries. In particular, only one participant reported Black ethnicity. Given the potential differences in how eczema presents in different skin types and the potential differences in how eczema might impact people from different cultures,

this is a meaningful way that respondents may differ from non-respondents, and this bias should be noted as a limitation when interpreting the results of this study.

4.5.4 Implications and future directions

The studies presented in Chapter 3 and 4 provide information about different stakeholder perspectives to inform voting on a consensus-based definition of long-term control of eczema by the HOME initiative. This study directly informed discussions about the content validity and feasibility of different methods of measuring long-term control at the HOME V consensus meeting in Nantes 2017, allowing the HOME group to move towards consensus on standardising the measurement of this domain (Chalmers et al., 2018).

Figure 4-3 Small group discussion at the HOME V meeting

The results of this study led to the decision to develop a patient-reported outcome measure of eczema control, which is presented in Chapter 5. Although this study was

aimed to inform measurement of eczema control in clinical trials, it may also be appropriate to inform measurement in routine clinical settings.

4.6 Summary of Chapter 4

People with eczema, caregivers and clinicians across multiple countries view long-term control of eczema as a multi-faceted construct involving changes in disease activity, the treatment and management of the condition, and psychological, social and physical functioning. This online approach to an international qualitative study is an example of how core outcome set developers with limited resources can engage with multiple stakeholder groups on an international basis to inform consensus meeting discussions. Chapter 5 will present a study developing a new instrument to measure eczema control that used the results of this study as the foundation for the conceptual framework of the new instrument.

Chapter 5

A patient-centred approach to developing an instrument to measure the experience of eczema control

5.1 Introduction

The HOME initiative recommends long-term control of eczema as a core outcome domain that should be measured in every clinical trial over 3 months in duration. The HOME V meeting in Nantes, France June 2017 resulted in consensus voting regarding the core outcome domain of long-term control of eczema (Chalmers et al., 2018). The meeting used consensus processes to further define the construct of interest for inclusion in the core outcome set.

Based on the results of the study that was presented in Chapter 4, alongside the results of a systematic review of how long-term control has been measured in previous clinical trials for eczema, discussions took place to define long-term control and consider which factors and items are important to consider when measuring long-term control (Barbarot et al., 2016, Chalmers et al., 2018). Small groups agreed their three most important aspects relating to the long-term control of eczema domain, and subsequent whole group voting following discussions indicated that disease severity (signs and symptoms), quality of life and intensity of itch were considered essential for measuring long-term control of eczema (Chalmers et al., 2018). However, when voting if the existing domains of signs, symptoms (including intensity of itch), and quality of life were sufficient, consensus was not reached (Chalmers et al., 2018).

Further discussions resulted in a vote on whether long-term control of eczema (in trials of at least 3 months) should be defined by signs, symptoms, quality of life and a patient global measure where consensus was reached (91% agreed, 4% unsure, 4% disagreed) (Chalmers et al., 2018). Following this meeting, the remaining challenge

was exploring what approach to a “patient global measure” would be most appropriate to use in eczema clinical trials and suitable for the core outcome set.

Figure 5-1 Voting session at HOME V meeting

One approach could be a single global item developed to directly assess an individual’s experience of eczema control. Single global items are appealing for their simple, pragmatic design. It has been cited that they are easier to administer, less burdensome to participants and this could also result in higher completion rates than multi-item instruments (de Boer et al., 2004).

However, for complex constructs, such as the experience of eczema control, the wording of single global items can be more open to varied interpretation between individuals and may be a difficult concept for people to have an appropriate schematic framework available to them (Fayers and Machin, 2013). With a single item, respondents are required to think of all the information relevant to them relating to the construct, ignore anything that is not relevant to them and weigh up their experience in a single score (de Boer et al., 2004). Single global items have also

been demonstrated as being less sensitive to change, which is an important criterion for the measure to meet if it is to be useful in demonstrating the effectiveness of intervention (Bowling, 2005). Therefore, when designing an instrument to capture a complex construct such as eczema control, it is important to consider the length of the instrument and burden on participants weighed against gaining precise information that can more easily detect small but clinical changes (Bowling, 2005).

The findings in Chapter 3 and Chapter 4 of this thesis and HOME membership consensus voting suggests that the experience of eczema control is made up of multiple aspects that together create the experience of eczema control (Howells et al., 2017, Howells et al., 2018b). Measuring such a complex construct over time can be challenging and is something that has been considered by others wishing to capture control in chronic diseases such as asthma and urticaria (Weller et al., 2014, Nathan et al., 2004).

5.2 Aims and objectives

This study aimed to develop a new outcome measurement instrument to capture the experience of eczema control in clinical trials and clinical practice.

1. to develop an instrument to capture eczema control that is suitable for use in both adults and children with atopic eczema;
2. to conduct preliminary validation of the new instrument.

5.3 Methods

5.3.1 Design

This mixed-methods study progressed through five stages of instrument development as summarised in Figure 5-2. Methodological guidance for instrument development were followed for study design (Patrick et al., 2011b, Patrick et al.,

2011a, De Vet et al., 2011, Streiner et al., 2015, Food and Drug Administration, 2009). The instrument was co-designed by an international expert panel consisting of three dermatologists, a dermatology nurse, a general practitioner, two people with eczema, two caregivers of children with eczema, four methodologists, and a psychologist. Five countries were represented on the expert panel. This project has been approved by the University of Nottingham's Faculty of Medicine & Health Sciences Research Ethics Committee (Refs: 18-1805 and F14062016 SoM ROD). The study protocol (https://www.nottingham.ac.uk/research/groups/cebd/documents/methodological-resources/protocol-index-for-eczema-control-v1.pdf) and study data analysis plan (https://www.nottingham.ac.uk/research/groups/cebd/documents/researchdocs/data-analysis-version-1.2-11.2.18.pdf) were prospectively registered on the CEBD's protocol registration portal.

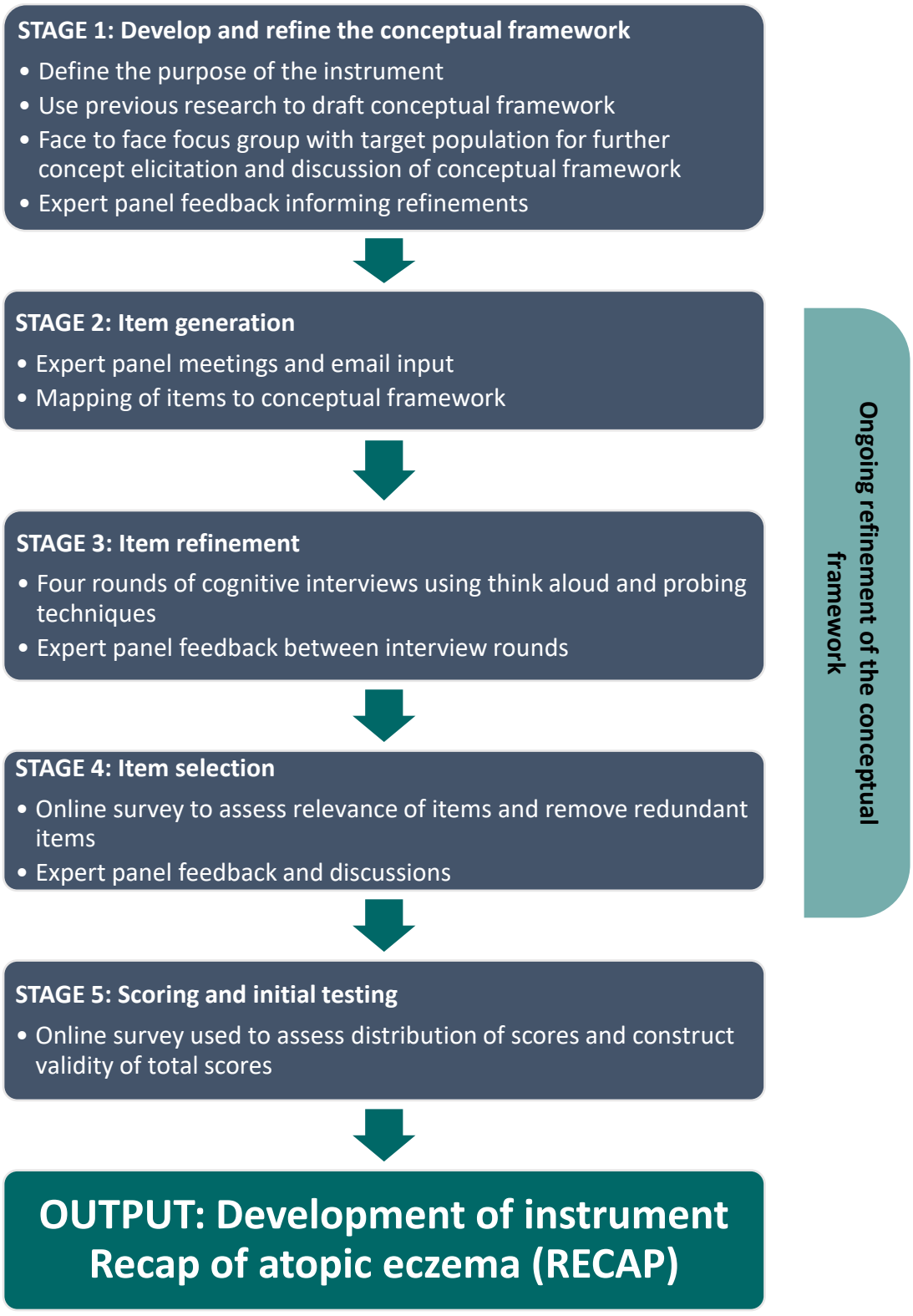


Figure 5-2 Study design for developing RECAP

5.3.2 Stages of Instrument Development

5.3.2.1 STAGE 1: Developing and refining the conceptual framework

The box below outlines the intended purpose of the instrument (Figure 5-3). This definition was developed as an interpretation of eczema control as an 'experience' perceived by individuals. This was developed from the qualitative studies presented in Chapter 3 and 4, and as a result of the HOME discussions that called for a global patient assessment of eczema control. Chapter 2 explains in more detail how this construct definition developed throughout the research process.

Intended purpose:	**Decision:**
Intended construct of interest:	The experience of eczema control. This was defined in this study as "the extent to which the various manifestations of eczema and the impact that these have for an individual are removed or meaningfully reduced."
Intended target population:	Individuals with eczema of all ages. However, for younger children who do not have the cognitive abilities to answer the questionnaire alone, it is intended that the information will be provided by caregivers or with the assistance of caregivers. The questionnaire is not intended to be exclusive to use in a single disease severity, disease duration, sex or ethnicity.
Intended context for use:	Primarily designed for use in clinical trials assessing any type of intervention in people with eczema. As a secondary aim, it was also anticipated that the instrument should be appropriate for use in clinical settings.

Figure 5-3 Box to outline intended purpose of the instrument

The conceptual framework was drafted by LMH, JRC, CA and KST through synthesising findings from an international qualitative study (Howells et al., 2018b, Howells et al., 2017), an international quantitative study (von Kobyletzki et al., 2016), and a systematic literature review (Barbarot et al., 2016) relating to the construct of interest. A final face-to-face focus group involving people with eczema and caregivers of children with eczema was conducted to confirm the conceptual framework and to ensure that key items had not been overlooked. This focus group was moderated by author C.A., who has experience moderating groups and training in qualitative research. The findings from the focus group were discussed by the expert panel who used this conceptual framework as a starting point to begin item design.

5.3.2.2 STAGE 2: Item generation

Guided by the conceptual framework and guidance on designing questionnaire items, the expert panel members submitted ideas for items to include in the instrument. The items were then categorised and discussed by the panel. The items were either discarded, kept or amended to produce an initial working list of items.

5.3.2.3 STAGE 3: Item refinement

Cognitive interviews, which used a range of think-aloud and probing techniques, were used to improve the comprehension, comprehensibility and relevance of the questionnaire. The target population were adults (16+ years) with eczema or caregivers of children with eczema living in the UK. Children under 16 years could take part if their caregiver was present. Participants were recruited using existing mailing lists of participants interested in eczema-related research at the Centre of Evidence Based Dermatology (CEBD) and social media. All participants had to be proficient in the English language. There was no exclusion of participants based on age, eczema severity, sex and ethnicity and purposeful sampling aimed to achieve a diverse range of participants based on these characteristics.

Cognitive interviews lasting approximately one hour each took place either face to face, on the telephone, or video call depending on participant preferences. If the interview was conducted remotely, participants were asked to complete an online consent form prior to the interview and verbal consent was confirmed at the start of the interview. They were also sent an electronic copy of the items to access during the interview. If the interview was face to face, the participants were asked to fill out a paper consent form at the start of the interview and provided with a paper copy of the items at the start of the interview.

All interviews were conducted by LMH, who has experience and training in qualitative research and who received additional training in cognitive interviews from a methodological advisor (PL). Interviews followed a semi-structured interview guide (Appendix D). The interview guide was developed considering cognitive theories of questionnaires (Tourangeau et al., 2000, Collins, 2003). Interviews began with an explanation of the purpose of the interview and an opportunity for participants to share their experience of eczema and demographic information. The participants were then asked to answer the items using think-aloud methods (Willis, 1999). The interviewer planned breaks to summarise and use pre-planned probes to encourage elaboration by participants. Once the think-aloud process was applied to all items, the interviewer then probed the participants about the items as a global set. Interviews took place in rounds, with the expert panel refining the content in between rounds, and subsequent rounds assessing if the changes had addressed the initial problems. It was planned that rounds would be continued until no further refinements were required. Participants were debriefed at the end of the interview, where the purpose of the interview was reiterated in more detail and the participants were given an opportunity to ask any questions.

5.3.2.4 STAGE 4: Item selection

An online survey was used to conduct an impact analysis, which uses information about frequency of occurrence and the importance of the experiences to assess

relevance of each experience. This approach was adapted from the impact analysis used in the development of the Urticaria Control Test (Weller et al., 2014). The online survey was also used to conduct a multivariable regression analysis to reduce the items to ensure each item asked is carefully chosen to add unique information about the experience of eczema control.

The target group were adults (16+ years) with eczema or caregivers of children with eczema living in the UK. Children under 16 years could take part if their caregiver was present. Participants were recruited using existing mailing lists of participants interested in eczema-related research at the CEBD, social media, and posters in various public settings (shops, cafes, libraries, universities, healthcare centres) were used with permission. All participants had to be proficient in the English language. There was no exclusion of participants based on age, eczema severity, sex and ethnicity.

Variables included in the survey were age, sex, ethnicity, 'bother caused by the eczema' global eczema severity, the Patient-Oriented Eczema Measure, the items still being considered for inclusion following the item refinement stage, frequency of occurrence of experiences over the past year, importance of experience. A full copy of the survey can be found in Appendix C.

5.3.2.5 STAGE 5: Scoring and initial testing

The online survey described in stage 4 was also used to collect data on the final items chosen for inclusion so that these could be scored and tested. Overall scores for the final items were generated using scoring rules determined by the expert panel. These scores were then tested for distribution of scores and construct validity.

5.3.3 Analysis

5.3.3.1 Focus group data (stage 1)

Experiences of eczema control were mapped onto the theoretical framework developed in Chapter 3. LMH and PL independently coded the data and met to discuss any discrepancies in coding. A qualitative descriptive approach was used to analyse the participant's responses to the conceptual framework (Sandelowski, 2000, Sandelowski, 2010, Neergaard et al., 2009).

5.3.3.2 Cognitive interview data (stage 3)

Data were analysed using a problem-focused approach using a hybrid model that used a top-down framework that coded for both problems relating to cognitive processes and question features (Table 5-1). Top-down frameworks for coding differ from inductive coding as they focus on the areas of particular interest to the researchers, and are suited for detecting problems with items that need further refinement (Willis, 2015). However, this framework can be refined using inductive coding (inductive codes added are highlighted in green). The framework was developed by LMH, but drew on theories of cognitive processing (Tourangeau et al., 2000, Collins, 2003), and drew on frameworks that have been developed in previous research (Willis, 2015). All data were analysed by LMH, with secondary coding on transcripts selected to cover a range of codes and areas that had been challenging to code by JRC or AVS. Discrepancies in coding were discussed and resolved via discussion. SG, KST, LMH and JRC were all involved in ongoing discussions about the coding and between each round of interviews, any problems that were identified and potential solutions were fed back to the expert panel for their further ideas and input. Interview rounds were continued until the authors were satisfied that there were no major problems that could be identified or resolved via further cognitive interviewing and the final round tested items that were ready for the online survey.

Table 5-1 Coding framework used for analysis of cognitive interviews (inductive coding in green)

Code	Label	Elaboration
1	**Comprehension**	Item has ambiguous meaning, lack of clarity in wording, uses obscure or difficult language
2	**Intended construct**	Raised a concern about if participant is responding in a way that is capturing the intended construct
2.1	**Beliefs about their eczema and/or treatments affecting response**	i.e. related to eczema but not the concept we are trying to capture
2.2	**None eczema related issues affecting response**	i.e. not related to the eczema, other diseases, other reasons
3	**Knowledge**	Participant lacked the information needed to answer the question
4	**Applicability**	Item was not relevant or applicable to the participant, question had made assumptions
5	**Sensitivity / Desirability**	Item raised concerns or wording was too sensitive, desirability bias likely to occur
6	**Memory retrieval**	Participant had difficulty recalling information required, high level of detail required, recall period too long, felt they had a shortage of cues
7	**Calculating response**	Participant had to make a complex estimation to decide upon a judgement or evaluation, had to use heuristics to provide answer
8	**Assigning response options**	Response options were undefined or vague, used inappropriate units, unclear what they referred to, overlapping categories, missing categories
8.1	**Distinguishing between response option types**	i.e. not clearly distinguishing frequency response options from intensity response options
9	**Other concerns raised**	Problems identified that do not fit within the above codes
9.1	**Aim of the questions**	Uncertainty about the aims of the questions
9.2	**Uncertainties when making comparisons**	Uncertainties about what experience to compare current experience to

5.3.3.3 Online survey data (stages 4 and 5)

Data were analysed using STATA 15.

5.3.3.3.1 Impact analysis

For each item concept, the proportion of individuals who had experienced that in the past year was multiplied by the mean score of the importance rating (1 = not important, 5=extremely important) to give an impact score ranging from 0 to 5. Whole sample analysis and sub-group analysis by age of person with eczema (0-4

years, 5-15 years, and 16+ years) were decided a priori. It was predefined that an impact score of less than 2 in any group analysed indicated that an item should not be considered for inclusion in the instrument.

5.3.3.3.2 Multivariable regression analysis

The potential items were entered as independent variables into multivariable linear regression models. This approach was chosen as formative measurement models are built on the underlying assumptions of regression models (Peterson et al., 2017). The aim of this process was to ensure that each item is adding a distinct concept to the overall construct and explore removing items that overlap conceptually with items already included. The dependent variable was 'bother caused by eczema' (0-10 points). This dependent variable was chosen because there was no 'gold standard' measure of 'eczema control' that could be used. 'Bother caused by eczema' was agreed by the expert panel to be the most closely aligned measure available that was in line with the concept of 'eczema control' as defined for the development of RECAP.

Sample size was calculated as at least 10 cases per independent variable (Harrell, 2015). The backward elimination variable selection technique was used to determine which items remained in the model, which is recommended when using automatic selection procedures as it evaluates each predictor after accounting for other variables, in contrast with univariable screening (Royston and Sauerbrei, 2008). The stopping criteria for this process was p=0.157, which is recommended for sample sizes with between 10 and 25 events per parameter as a proxy for Akaike's information criteria (AIC) (Sauerbrei, 1999, Heinze et al., 2018). AIC is a relative measure of information using the likelihood function and includes a penalty for the number of parameter estimates.

Multicollinearity is when independent variables are highly correlated, which can result in unreliable regression coefficients and testing the significance of regression

coefficients can be misleading. Table 5-2 shows that correlations range from -0.31 to 0.49, which indicated multicollinearity was not likely to be a problem. Following regression analyses ran, the variance inflation factor (VIF) was checked, and it was 2.33 for model 1 and 2.22 for model 2. It has been suggested as a rule of thumb a VIF higher than 10 is a concern (Armitage et al., 2008).

Table 5-2 Assessment of correlations between independent variables

	Acceptability of the eczema	Itchy skin	Sleep disturbance	Unable to stop scratching	Getting in the way of day to day activities	Affecting how been feeling	Stopped from doing something wanted or needed to do	Having flares	Having any symptoms	recap13painful_sore	Intensely itchy skin
Itchy skin	0.29	1									
Sleep disturbance	-0.05	0.02	1								
Unable to stop scratching	-0.17	0.14	0.05	1							
Getting in the way of day to day activities	-0.14	-0.16	0.04	0.09	1						
Affecting how been feeling	-0.17	0.08	0.07	0.35	0.003	1					
Stopped from doing something wanted or needed to do	-0.24	-0.24	0.03	0.06	0.06	-0.05	1				
Having flares	-0.20	0.14	0.09	0.24	0.08	0.22	0.06	1			
Having any symptoms	0.27	0.56	0.04	0.12	-0.10	0.01	-0.14	0.11	1		
Painful/sore	-0.09	0.11	0.15	0.32	0.07	0.24	-0.06	0.34	0.16	1	
Intensely itchy skin	-0.19	0.08	0.09	0.52	0.08	0.39	-0.009	0.23	0.15	0.49	1
Global	-0.31	-0.09	-0.02	0.16	0.25	0.14	0.16	0.15	-0.13	-0.02	0.06

Linear regression analysis cannot fully capture the extent of a curvilinear relationship between a dependent and independent variable (Altman, 1990). The linearity of the relationship between the dependent variable and independent variable was assessed by looking at scatterplots between the dependent variable and each independent variable. Residuals are the difference between the observed value of the dependent variable and the predicted value in the model. Ideally, in a linear regression model, residuals should be normality distributed (Altman, 1990). The histograms of the residuals for each independent variable were assessed. The assumption is violated if the data does not show a 'bell curve'. Ideally, the residuals should also show homoscedasticity. This was assessed by looking at a scatterplot of the residuals across each point on the independent variable. Homoscedasticity is violated if there is a clear pattern in the distribution of the data (Altman, 1990). The assessments of linearity, normality of residuals and homoscedasticity of residuals were all interpreted as indicating it was appropriate to conduct the planned analyses.

5.3.3.3.3 Scoring

The expert panel agreed scoring rules resulted in all RECAP items being scored from 0-4 and weighed equally and added together (total scores ranging from 0-28), with a higher score indicating less eczema control.

5.3.3.3.4 Distribution of scores

Assessment of histograms. A floor or ceiling effect was defined prior to data collection as more than 15% of participants achieving the highest or lowest possible score (Terwee et al., 2007).

5.3.3.3.5 Construct validity

Pearson's correlation coefficients were used to assess the relationship between POEM and the newly developed instrument (and assumptions for Pearson's correlation were met). It was hypothesised that correlations would be at least 0.3 (moderately correlated), given that the POEM measures patient-reported symptoms, which is a construct that is considered one element of the experience of eczema control. Mean RECAP scores were calculated according to known groups on a single global eczema severity item and scores based on the POEM severity categories (Charman et al., 2013). It was hypothesised that those categorised with more severe eczema would have higher mean scores on RECAP than those with lower severity categories on the corresponding instruments.

5.3.4 Stakeholder / Patient & Public Involvement

This study was co-designed with a range of stakeholders. The expert panel members were purposefully made up of people with diverse experiences and skills, and four members of the public were part of the expert panel and they took part in one face-to-face meeting where 4 participants (one PPI member) used videoconference to join remotely, multiple teleconferences and email exchanges. The expert panel members were all co-applicants and given the opportunity to comment on the ethics application, the study protocol and the study materials. PPI was particularly important regarding the wording and design of materials (including information sheets, consent forms, advertisement/posters, and the online survey). The expert panel fed back on the conceptual framework and collaborated on the design of the items, refinement of items and reduction of items. All PPI members had involvement in trials or outcome research previously.

NR and TB also piloted the cognitive interview and provided feedback on the process and provided a training opportunity so that the interviewer (LMH) could reflect on the process and make changes to the interview guide before beginning the interviews.

The CEBD Patient Panel day in September 2018 (a patient and public involvement day at the University of Nottingham) provided input during item refinement process with some key, targeted queries to aid finding solutions to a problem exposed in the cognitive interviews.

5.4 Results

5.4.1 Participant characteristics

5.4.1.1 Focus group (stage 1)

Six people took part in a face-to-face focus group. Three participants were adults with eczema, two participants were adults who had eczema themselves as well as experiences of caring for their children with eczema and one participant was a caregiver of a child with eczema. Eczema severity ranged from mild to very severe. Four participants were female and two were male.

5.4.1.2 Cognitive interviews (stage 3)

Thirteen people took part in a think-aloud interview. Eight answered the questions as adults with eczema and five as caregivers of children with eczema. Eczema severity ranged from mild to very severe. Age of adults with eczema ranged from 37 to 64 years. All adults taking part were female and all reported onset of eczema as young children. Self-reported ethnicity was White British (n=5), While Scottish (n=2) and Sikh (n=1). Age of children of the caregivers taking part ranged from 2 years 8 months to 14 years. Three of the children were male and two were female. Onset of eczema was reported from 8 weeks old to 2 years and 6 months old. Ethnicity was described as White British (n=4) and Welsh/Maltese (n=1).

5.4.1.3 Online survey (stage 4 and 5)

Table 5-3 provides the participant characteristics. A total of 337 entered the online survey, but n=2 did not meet the inclusion criteria, n=5 did not fill out any variables and n=6 did not provide information beyond demographic variables. Only 324 participants completed any of the key variables, so data from these participants are reported here.

Table 5-3 Online survey participant characteristics

	N	%	Mean (SD)	Range
Age	324	-	22.71	0-66
Under 5 years	62	19.14	-	-
5-15 years	77	23.77	-	-
16+ years	185	57.10	-	-
Sex	322	-	-	-
Male	110	34.16	-	-
Female	211	65.52	-	-
Non-binary	2	0.62	-	-
Rather not say	1	0.31	-	-
Ethnicity	321	-	-	-
White	300	93.46	-	-
Bangladeshi	1	0.31	-	-
Black Caribbean	2	0.62	-	-
Chinese	6	1.87	-	-
Indian	6	1.87	-	-
Mixed Race	5	1.56	-	-
Other Asian (non-Chinese)	1	0.31	-	-
Sikh	1	0.31	-	-
Total POEM score	263	-	15.12 (7.37)	0-28
POEM severity banding	263	-	-	-
Clear-Almost clear	13	4.94	-	-
Mild	33	12.55	-	-
Moderate	95	36.12	-	-
Severe	96	36.50	-	-
Very severe	26	9.89	-	-
Global severity	266	-	-	-
Clear	6	2.26	-	-
Almost clear	34	12.78	-	-
Mild	65	24.44	-	-
Moderate	121	45.49	-	-
Severe	40	15.04	-	-
Bother score	324	-	5.65 (2.56)	0-10

5.4.2 Key stages of instrument development

5.4.2.1 STAGE 1: Developing and refining the conceptual framework

Figure 5-4 shows the initial conceptual framework that was presented to members of the focus group (although more detail was included relating to each concept). Analysis of the discussions confirmed that the framework represented an accurate model of 'eczema control' and that the conceptual framework was comprehensive. The focus group analysis confirmed the conceptualisation of 'eczema control', indicating that data saturation for concept elicitation was reached. Nevertheless, some minor refinements were suggested, as summarised in Table 5-4. The final conceptual framework for the RECAP instrument, after refinement based on all stages of instrument design is presented in **Error! Reference source not found.**. The c onceptual framework suggests that a formative model is the best approach to developing the measurement model for this construct of interest as multiple unique factors are relevant to the experience, which when combined together, form the latent variable (De Vet et al., 2011).

Figure 5-4 Initial conceptual framework

Table 5-4 Refinements to the conceptual framework

Refinement of conceptual framework	Reasons why
Addition of concept predictability of eczema	Participants at the focus group expressed a concern that the predictability of eczema, which related to eczema control in their perception, was not included in the conceptual framework.
Removal of item on predictability	The cognitive interviews suggested that an item asking directly about the predictability of the eczema was not interpreted in line with the construct of interest. It may be that this concept is a related but distinct outcome to be measured.
Removal of concept impact on family	The expert panel meeting led to discussions about designing items on the impact on family and it was felt strongly amongst stakeholders including patients that this concept was not universal to all. It was also suggested to be a related but distinct construct.
Removal of treatment and management concepts	The cognitive interviews revealed issues regarding the applicability and relevance of treatment items. There were multiple reasons why this was the case, which are outlined below. 1) The intended purpose of this measurement instrument to be used in all trials. The changes in treatment and management will differ across different severities of eczema where different types of treatments are used. 2) There may also be trials were the treatment used is being tightly regulated or out of the individual's decision to change treatment and management options. 3) The cognitive interviews revealed there were issues with developing treatment and management related questions that were applicable across all of the target population. 4) The expert panel discussed the issues above and discussed the inclusion of these concepts within the framework. Some members were conflicted about removing these concepts as they understood the operational difficulties of including them, but they felt they were important as they had been recognised as an indicator of control by stakeholders, whilst others felt they were not necessary to measure the construct of interest. It was agreed they would be removed.
Removal of items on social impact	The online survey revealed that the items regarding social impacts were not applicable and relevant to young children. This finding was discussed amongst the expert panel who approved removing this concept.
Change in way overall individual perceptions and acceptability were included in the framework	The first expert panel meeting resulted in the acknowledgement of multiple global concepts that made a fourth area of interest. Therefore, it became a unique element to be included in the conceptual framework other than an overriding viewpoint.

5.4.2.2 STAGE 2: Item generation

Expert panel members were each asked to submit questions that could be used to capture the key elements of eczema control as outlined in the conceptual framework. This process resulted in an initial list of 154 ideas, although many of the ideas in this list were different suggested wording of how to capture one aspect within the conceptual framework and gave multiple alternate options to consider. This was subsequently reduced to 25 agreed items that were approved by the expert panel to be tested in the next phase of the development process.

5.4.2.3 STAGE 3: Item refinement

Table 5-5 illustrates the changes that took place between four rounds of cognitive interviews. The results are detailed below, but to summarise, the recall period was changed from 4 weeks to 1 week, the response options were changed from 4 to 5, the items were changed from statements to questions, wording was changed to provide clarity, and language was amended to reflect terms that felt more resonant to respondents and increase the confidence of respondents in their ability to answer the questions. By the end of the interviews 15 items remained for further testing. . However, only 14 items were included in subsequent analysis as the expert panel made the decision to remove the remaining treatment-related item, having reflected on the HOME V meeting decisions that treatment-related measures are not feasible for use in many clinical trials, the cognitive interview findings and the conceptual framework refinements (Table 5-4).

Changes to recall period

Participants found it difficult to remember the recall period throughout answering the questions, as well as finding 4 weeks a long time to recall their experience. Therefore, following round 1 the recall period was stated in each question and reduced to 1 week.

"So the first thing I have to think about is if it's four weeks where does that four weeks start from, and what's been happening in my life at that time, and can I remember what my skin's been doing in that last four weeks, and that's not necessarily an easy thing, it takes quite a lot of thought to try and work all of that out." (Elsa, person with eczema)

Changes to response options

Participants felt that the 4-point rating scales did not cover the full range of experience, expressing a desire to answer in a way that fell between two of the response options. Therefore, 5-point rating scales that aimed to more finely capture the full continuum of experience were presented in round 2 onwards.

"I suppose, this is one way it would be helpful to probable just have a rarely because I wouldn't go as far as to say some of the time, but there have been on one or two occasions, where he's just really, really scratching. Where eczema is… and we always kind of distract him from scratching himself. So, but I'd say that's probably been once or twice over the past four weeks and yes, he does seem distracted from what he's doing before because he suddenly wants to scratch and scratch and scratch. So, it's closest to none of the time, some of the time because it literally only… I don't know. Some in my mine, is I would expect it to be once or twice a week over the past four weeks, but yes. Can I write rarely on it?" (Natalie, caregiver)

Challenges with multiple possible interpretations of the response options presented as a 'proportion of the time' were addressed by changing to number of days.

"So, I guess, all of the time is every hour of every day. And most of the time would be many occasions through the day with some of

the time I think maybe just one or two things in a day. Then it's hard to know the severity. So, overnight he's been waking once or twice normally, so that's not all the time. But that is every night. So, that's hard to know." (Leslie, caregiver)

Whilst still using a 4-week recall period the number of days response options remained as descriptors rather than specifying a number of days, which still required difficult calculations for respondents, as they were trying to convert the "vague" descriptors within the response options into how they would interpret how many days in the last 4 weeks they felt that described. Since the version developed for round 3 had a recall period of one week, it was possible to alter the response options to be more specific within the description what range of days were included in each response category.

Change from statements to questions

Using statements for items did not match the 'number of days' response options. Participants in round 1 also felt the statements led to a desire to respond in a 'yes/no' format, whereas the developers wanted to measure each item as a range of experience. Therefore, the items were changed into a question format.

Removing and adding items

Entering the cognitive-interview phase, the expert panel had included more than one item that tapped onto each important concept within the conceptual model, to help determine which way of asking about a particular construct would be most appropriate in this setting. Some of the items were judged as not being easily interpreted or as clearly aligned with the intended construct as others, and therefore removed during the interview rounds, and in some cases additional items were added as new ways to approach capturing the concepts.

There was one area where the removal of items resulted in the refinement of the conceptual model. Four items were developed that were intended to capture

constructs relating to the treatment and management of the condition. However, in designing and testing the items it became clear that it was difficult to establish items that would be applicable across all populations and for all types of clinical trials. Therefore, all treatment questions were removed at this stage except for one which was adjusted into a global question about the impact of treatment on control to improve applicability.

Changes to language used in items

The term "control" (and its derivatives) was thought to be a term that may be interpreted differently by all. The developers initially felt rephrasing to give permission for respondents to use their own interpretation may be sufficient, given that the experience of control was conceptualised as an individual experience within the conceptual framework. However, further exploration made it apparent that there were multiple interpretations of the term and some were not in line with the construct of interest. Therefore, it was decided that the term 'control' should be avoided in all items.

> *Well, I'd say that to you, so control of eczema that you manage it, so that you've done all your creams, you've done all your lotions, you've taken whatever tablets you take, so you've followed your regime to manage your eczema or is it about say, managing it so that it's clear so – I'm not clear what you're asking so, that's why I say it depends on whether it's control and management or if its clearing it up. So, if it's clearing it up the answer is no days because it never goes. If it's control and management inline of what I've to do, then it's every day. (Amy, person with eczema)*

There were other changes in the language used designed to provide clarity, use language that felt more resonant to respondents and increase the confidence of respondents in their ability to use their knowledge to answer the question.

Table 5-5 Documentation of changes made following each round of cognitive interviews

Part of questionnaire	Round 1	Round 2	Round 3	Round 4
Recall period	4 weeks (in instructions only)	4 weeks (in instructions and questions)	1 week (in instructions and questions)	1 week (in instructions and questions)
Number of response options per item	4	5	5	5
Response option wording	None of the time / Some of the time / Most of the time / all of the time	No days / Hardly any days / Some days / Most days / Every day	No days / 1-2 days / 3-4 days / 5-6 days / Every day	No days / 1-2 days / 3-4 days / 5-6 days / Every day
	Not at all / A little / A lot / Completely	Not at all / A little bit / Quite a lot / A huge amount / Completely	Not at all / A little bit / Quite a lot / A huge amount / Completely	Not at all / A little bit / Quite a lot / A huge amount / Completely
			Not at all controlled / A little controlled / Quite controlled / Mostly controlled / Completely Controlled	Very good / Good / OK / Bad / Very bad
			Not at all acceptable / Not very acceptable / Quite acceptable / Mostly acceptable / Completely acceptable	Completely acceptable / Mostly acceptable / Quite acceptable / Not very acceptable / Not at all acceptable
Item wording that was refined	Overall, my eczema has been well controlled.	On how many days in the last 4 weeks would you describe your eczema as well controlled?	Over the last week, how would you describe your eczema? / Over the last week, on how many days would you describe your eczema as having been well controlled?	Over the last week, how has your eczema been?
	My level of eczema control has been acceptable to me.	On how many days in the last 4 weeks has your level of eczema control been acceptable to you?	Over the last week, how acceptable has your level of eczema control been to you? / Over the last week, on how many days has your level of eczema control been acceptable to you?	Over the last week, how acceptable has your eczema been to you?
	My skin has been itching	On how many days in the last 4 weeks has your	Over the last week, on how many days has your skin been	Over the last week, on how many days has

	because of my eczema.	skin been itching because of your eczema?	itchy because of your eczema?	your skin been itchy because of your eczema?
	My skin has felt painful/sore because of my eczema.	On how many days in the last 4 weeks has your skin felt painful/sore because of your eczema?	Over the last week, on how many days has your skin felt painful or sore because of your eczema?	Over the last week, on how many days has your skin felt painful or sore because of your eczema?
	I have been experiencing eczema symptoms.	On how many days in the last 4 weeks have you experienced at least one eczema symptom?	Over the last week, on how many days have you had any signs or symptoms of your eczema?	Over the last week, on how many days have you had any symptoms of your eczema?
	I have been experiencing eczema flares.	On how many days in the last 4 weeks have you been experiencing eczema flares?	Over the last week, on how many days have you experienced an eczema flare?	Over the last week, on how many days have you experienced an eczema flare?
	I have felt isolated because of my eczema.	On how many days in the last 4 weeks have you felt isolated because of your eczema?	Over the last week, on how many days have you felt isolated because of your eczema?	Over the last week, on how many days have you felt isolated because of your eczema?
	I have felt embarrassed because of my eczema.	On how many days in the last 4 weeks have you felt self-conscious or embarrassed because of your eczema?	Over the last week, on how many days have you felt self-conscious or embarrassed because of your eczema?	Over the last week, on how many days have you felt self-conscious or embarrassed because of your eczema?
	My eczema has affected how I have been feeling.	On how many days in the last 4 weeks has your eczema affected how you have been feeling?	Over the last week, on how many days has your eczema affected how you have been feeling?	Over the last week, on how many days has your eczema affected how you have been feeling?
	My eczema has been getting in the way of my everyday life.	In the last 4 weeks, how much has your eczema been getting in the way of your everyday life?	Over the last week, how much has your eczema been getting in the way of your day to day activities?	Over the last week, how much has your eczema been getting in the way of your day to day activities?
	My sleep has been disturbed because of my eczema.	In the last 4 weeks, how much has your sleep been disturbed because of your eczema?	Over the last week, how much has your sleep been disturbed because of your eczema?	Over the last week, how much has your sleep been disturbed because of your eczema?
	My eczema treatment has been enough to	Thinking about all the eczema treatments you	Thinking about all the eczema treatments you	Thinking about all the eczema treatments you

	control my eczema.	have used in the last 4 weeks, on how many days has your treatment been enough to control your eczema?	have used in the last week, on how many days has your treatment been enough to control your eczema?	have used in the last week, on how many days has your treatment been enough to manage your eczema?
Items that were added				Over the last week, on how many days were you unable to stop scratching?
			Over the last week, on how many days has your skin felt intensely itchy because of your eczema?	Over the last week, on how many days has your skin felt intensely itchy because of your eczema?
		On how many days in the last 4 weeks has your eczema stopped you doing something you wanted or needed to do?	Over the last week, on how many days has your eczema stopped you doing something you wanted or needed to do?	Over the last week, on how many days has your eczema stopped you doing something you wanted or needed to do?
Initial items that were removed before the final round	My life has been impacted by my eczema.	In the last 4 weeks, how much has your life been impacted by your eczema?	Over the last week, how much has your life been affected by your eczema?	
	I have been distracted because of my eczema.	On how many days in the last 4 weeks have you been distracted because of your eczema?	Over the last week, on how many days have you been distracted because of your eczema?	
	Overall, my eczema has been controlled.	On how many days in the last 4 weeks would you describe your eczema as under control?		
	I have modified my everyday life because of my eczema.	In the last 4 weeks, how much have you modified your everyday life because of the eczema?		
	I have been able to control my eczema.			
	I have been affected by my eczema symptoms.			

	I have been affected by eczema flares.			
	My eczema has felt unpredictable.			
	My eczema has affected my appearance.			
	My eczema has been on my mind.			
	I feel able to manage my eczema using the treatment I have.			
	The amount of time I have spent treating my eczema troubles me.			
	My eczema treatment has not been enough to control my eczema.			

Note. The self-reported version is reported here for brevity, but the same changes were made for the caregiver-reported version

5.4.2.4 STAGE 4: Item selection

5.4.2.4.1 Impact analysis

Data on frequency, importance and impact scores from the online survey are presented in Table 5-6.Feeling isolated scored less than 2 across all groups. Feeling self-conscious scored less than 2 in the age group 0-4 years. It was pre-defined that any item with an impact score of less than 2 for any of our target groups was considered not relevant and therefore excluded from the subsequent regression analysis. Items on the 'acceptability', 'overall individual perception', and 'treatment been enough' were not included in the impact analysis as it did not make conceptual

sense to ask whether these experiences had occurred in the last year to assess their relevance.

Table 5-6 Results of impact analysis

	Frequency (proportion)				Importance (Mean score)				Impact score (Frequency x Importance)			
Age group	All	0-4	5-15	16+	All	0-4	5-15	16+	All	0-4	5-15	16+
Itchy skin	1	1	1	1	4.77	4.89	4.82	4.7	4.77	4.89	4.82	4.70
Flare	0.9963	1	1	0.9935	4.6	4.81	4.64	4.5	4.58	4.81	4.64	4.47
Had any symptoms	0.9963	1	1	0.9935	4.57	4.63	4.54	4.55	4.55	4.63	4.54	4.52
Skin painful or sore	0.9925	1	0.965	1	4.63	4.74	4.73	4.56	4.60	4.74	4.57	4.56
Intensely itchy skin	0.9736	0.9811	0.9649	0.9742	4.55	4.74	4.66	4.45	4.43	4.65	4.50	4.34
Unable to stop scratching	0.9586	0.9444	0.9474	0.9677	4.58	4.7	4.79	4.46	4.39	4.44	4.54	4.32
Eczema affecting how been feeling	0.937	0.9259	0.9298	0.9419	4.4	4.5	4.59	4.29	4.12	4.17	4.27	4.04
Disturbed sleep	0.9023	0.9259	0.9483	0.8766	4.24	4.35	4.81	3.99	3.83	4.03	4.56	3.50
Eczema getting in the way of day to day activities	0.8647	0.8148	0.8966	0.8701	4.21	4.22	4.47	4.12	3.64	3.44	4.01	3.58
Stopped from doing something wanted or needed to do	0.7895	0.7222	0.8621	0.7806	4.14	4.17	4.38	4.04	3.27	3.01	3.78	3.15
Feeling self-conscious or embarrassed	0.7857	0.2778	0.8596	0.9355	4.3	3.61	4.59	4.39	3.38	1.00 *	3.95	4.11
Feeling isolated	0.4906	0.1852	0.569	0.5677	3.5	3.54	4.28	3.15	1.72 *	0.66 *	2.44	1.79 *

Note. *An impact score of <2 was defined a priori as indicating an experience was not relevant to include in the multivariable linear regression analysis. **Items on the 'acceptability, 'overall individual perception', and 'treatment been enough' were not considered appropriate for inclusion in the impact analysis

5.4.2.4.2 Multivariable linear regression analysis

Two models were developed using multivariable linear regression analysis. The expert panel decided which items should remain for consideration in the final set, using the evidence from all previous stages of development. It was determined that two models would be conducted to be chosen between.

- **Model 1 (the model chosen for the items in the instrument by the expert panel)**

The first model contained all 12 items that were still under consideration for inclusion in the final set of items as predictor variables. The bother scale was used as the outcome variable.

Five predictor variables were removed from the model following a backward elimination item reduction technique with a stopping criterion of $p = 0.157$. These included items 'being unable to stop scratching' ($p = 0.809$), 'stopped from doing something wanted or needed to do' ($p = 0.438$), 'having flares' ($p = 0.314$), 'having any symptoms' ($p = 0.809$) and 'painful or sore skin' ($p = 0.612$).

The results of the regression indicated that the seven remaining predictor variables explained 71.1% of the variance in bother, $R^2 = 0.718$, adjusted $R^2 = 0.711$, $F(7, 256) = 93.19$, $p < 0.001$. Table 5-7 shows the predictor variables that remained in the model.

Table 5-7 Model 1 Final Output

Predictor Variables	β	P-value
Acceptability of eczema	0.30	0.017
Itchy skin	0.19	0.053
Sleep disturbance	0.14	0.127
Getting in the way of day to day activities	0.32	0.01
Affecting how been feeling	0.13	0.102
Intensely itchy skin	0.22	0.009
Global	0.92	> 0.001

- **Model 2 (an alternative model that was not chosen by the expert panel)**

The second model contained 10 items entered into model 1 as predictor variables, with the exclusion of 'acceptability of eczema' and 'global' due to expert panel concerns that the more global nature of these items may be problematic for inclusion in the regression model with more specific items. The bother scale was used as the outcome variable.

Two predictor variables were removed from the model following a backward elimination item reduction technique with a stopping criterion of $p = 0.157$. These included items 'being unable to stop scratching' ($p = 0.808$) and 'painful or sore skin' ($p = 0.273$).

The results of the regression indicated that the seven remaining predictor variables explained 61.5% of the variance in bother, $R^2 = 0.627$, adjusted $R^2 = 0.615$, $F(8, 256) = 53.74$, $p < 0.001$. Table 5-8 shows the predictor variables that remained in the model.

Table 5-8 Model 2 Final Output

Dependent Variables	β	P-value
Itchy skin	0.32	0.007
Sleep disturbance	0.26	0.016
Getting in the way of day to day activities	0.64	>0.001
Affecting how been feeling	0.24	0.013
Stopped from doing something wanted or needed to do	-0.21	0.068
Having flares	0.41	>0.001
Having any symptoms	0.20	0.069
Intensely itchy skin	0.21	0.038

5.4.2.4.3 Expert panel review

The expert panel agreed model 1 as the final set of items because it explained a larger proportion of the variance than model 2. Model 2 was run due to concerns that the global item and acceptability item may closely link with bother concept and remove all other important concepts from the model, however the items removed at

this stage from model 1 were reviewed and it was not felt that excluding them would have any major impact on the conceptual framework and that the items included were comprehensive. The items that were removed from model 1 represent concepts from the original conceptual model that are already covered in another item in model 1, therefore reducing redundancy in items. The final conceptual framework is presented in Figure 5-5.The self-report version of RECAP can be found in Figure 5-6 and the caregiver-report version of RECAP can be found in Figure 5-7.

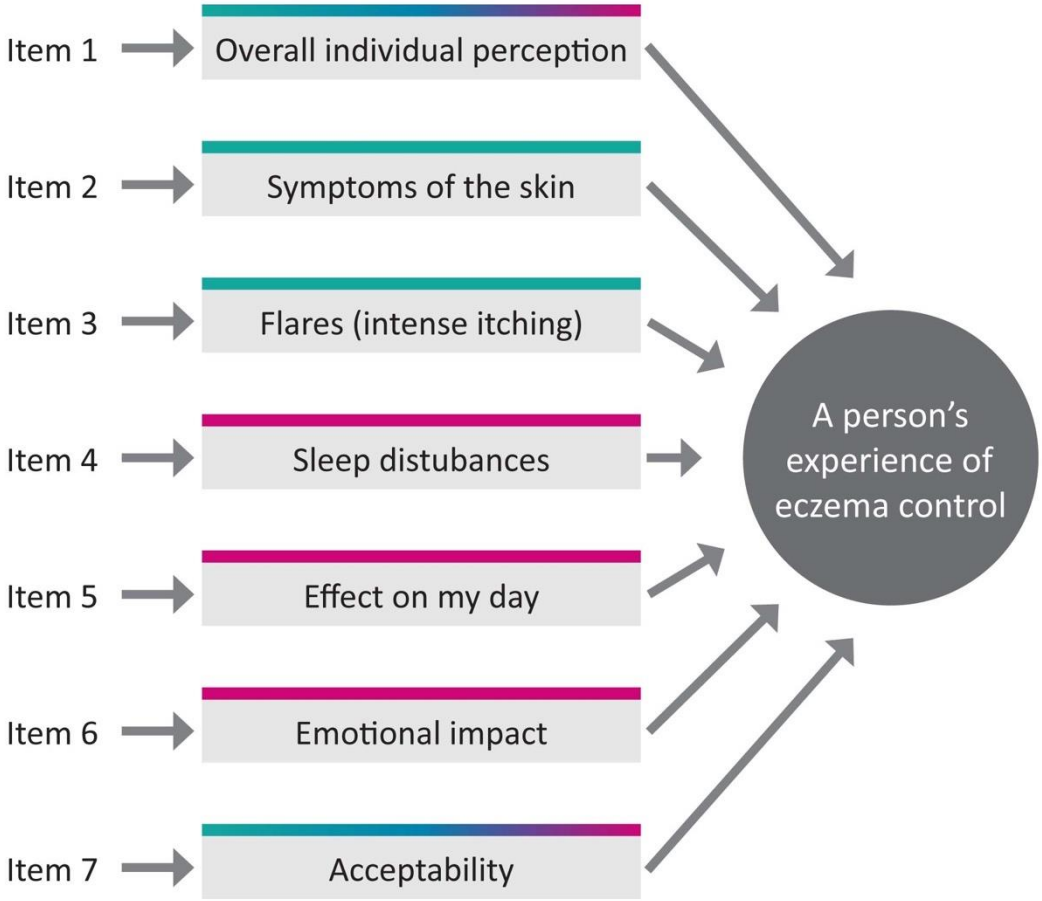


Figure 5-5 Final conceptual framework

Recap of atopic eczema (RECAP)

The questions below provide a snapshot of how your eczema has been over the last week from your point of view. Please only select one response for each question. Try and respond to every question, but if you are unable to respond then leave it blank.

1. Over the last week, **how has your eczema been?**

Very good	Good	Ok	Bad	Very Bad

2. Over the last week, on how many days has your **skin been itchy** because of your eczema?

No days	1-2 days	3-4 days	5-6 days	Every day

3. Over the last week, on how many days has your **skin been intensely itchy** because of your eczema?

No days	1-2 days	3-4 days	5-6 days	Every day

4. Over the last week, how much has your **sleep been disturbed** because of your eczema?

Not at all	A little bit	Quite a bit	A huge amount	Completely

5. Over the last week, how much has your eczema been **getting in the way of day to day activities**?

Not at all	A little bit	Quite a bit	A huge amount	Completely

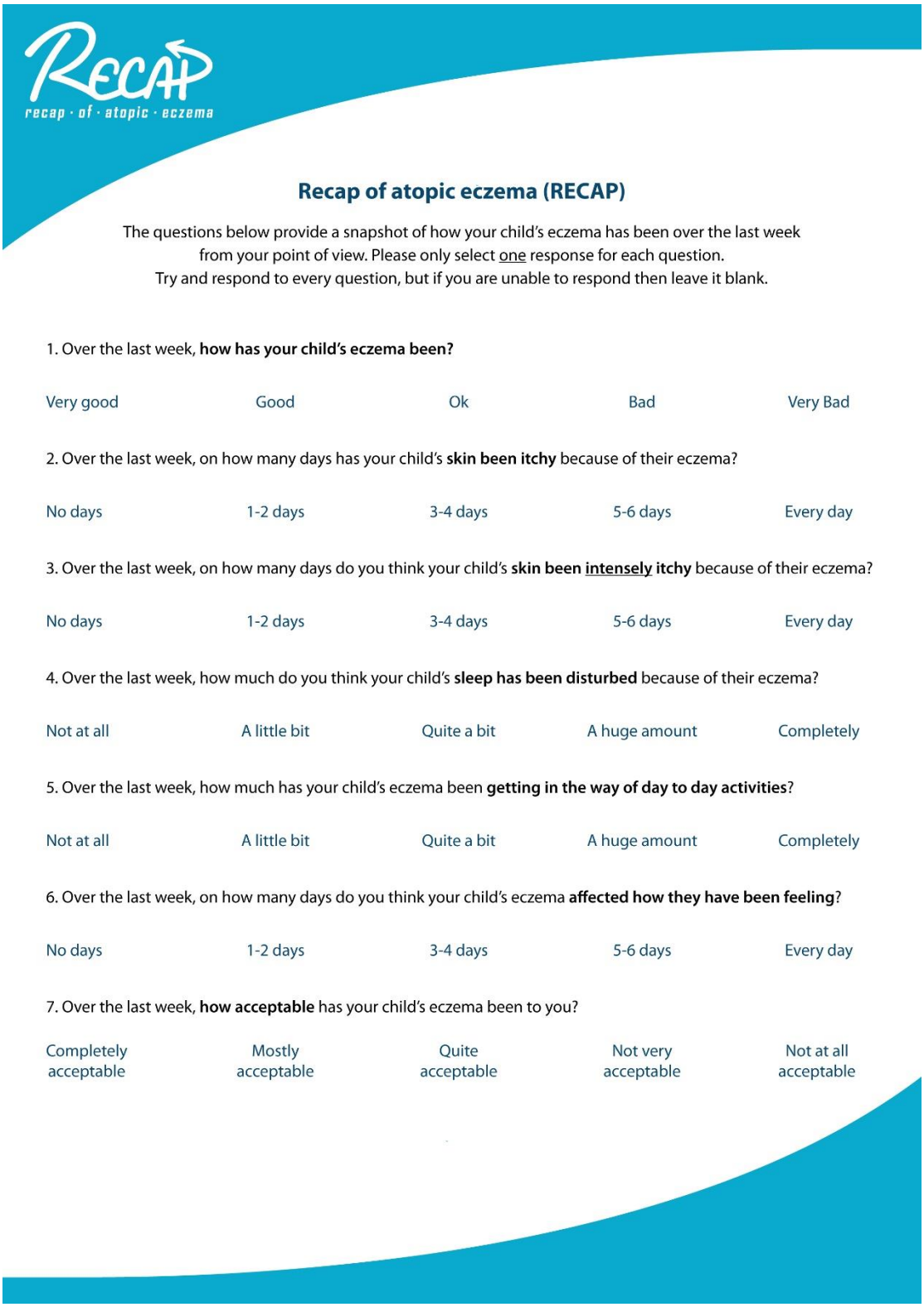
6. Over the last week, on how many days has your eczema **affected how you have been feeling**?

No days	1-2 days	3-4 days	5-6 days	Every day

7. Over the last week, **how acceptable** has your eczema been to you?

Completely acceptable	Mostly acceptable	Quite acceptable	Not very acceptable	Not at all acceptable

Figure 5-6 Self-report version of Recap of atopic eczema (RECAP)



Recap of atopic eczema (RECAP)

The questions below provide a snapshot of how your child's eczema has been over the last week from your point of view. Please only select one response for each question.
Try and respond to every question, but if you are unable to respond then leave it blank.

1. Over the last week, **how has your child's eczema been?**

Very good | Good | Ok | Bad | Very Bad

2. Over the last week, on how many days has your child's **skin been itchy** because of their eczema?

No days | 1-2 days | 3-4 days | 5-6 days | Every day

3. Over the last week, on how many days do you think your child's **skin been intensely itchy** because of their eczema?

No days | 1-2 days | 3-4 days | 5-6 days | Every day

4. Over the last week, how much do you think your child's **sleep has been disturbed** because of their eczema?

Not at all | A little bit | Quite a bit | A huge amount | Completely

5. Over the last week, how much has your child's eczema been **getting in the way of day to day activities**?

Not at all | A little bit | Quite a bit | A huge amount | Completely

6. Over the last week, on how many days do you think your child's eczema **affected how they have been feeling**?

No days | 1-2 days | 3-4 days | 5-6 days | Every day

7. Over the last week, **how acceptable** has your child's eczema been to you?

Completely acceptable | Mostly acceptable | Quite acceptable | Not very acceptable | Not at all acceptable

Figure 5-7 Caregiver-report version of Recap of atopic eczema (RECAP)

5.4.2.5 STAGE 5: Scoring and initial testing

5.4.2.5.1 Scoring

Expert panel discussions resulted in the following scoring rules:

Each of the seven questions carries equal weight and is scored from 0 to 4:

Very good	**=0**	**No days**	**=0**	**Not at all**	**=0**	**Completely acceptable**	**=0**
Good	**=1**	**1-2 days**	**=1**	**A little bit**	**=1**	**Mostly acceptable**	**=1**
OK	**=2**	**3-4 days**	**=2**	**Quite a lot**	**=2**	**Quite acceptable**	**=2**
Bad	**=3**	**5-6 days**	**=3**	**A huge amount**	**=3**	**Not very acceptable**	**=3**
Very bad	**=4**	**Every day**	**=4**	**Completely**	**=4**	**Not at all acceptable**	**=4**

Note:

- If one question is left unanswered this is scored 0 and the scores are summed and expressed as usual out of a maximum of 28.
- If two or more questions are left unanswered the questionnaire is not scored.
- If two or more response options are selected, the response option with the highest score should be recorded.

The scoring rules match those of the POEM, which was chosen in part to increase consistency across POEM and RECAP instruments and reduce the likelihood of errors in scoring. The assumptions made for these scoring rules are outlined below:

1. If one question is left unanswered, it is assumed that the item is not of relevance to the individual's control, and it is considered most appropriate to give 'no relevance' a score of zero.
2. If two or more questions are left unanswered, it is assumed that there is a potential problem with filling out the questionnaire, and it is regarded as too much missing data within the items to generate a total score.
3. If two more responses are selected in one item, it is considered preferable to try and provide a score and reduce missing data, and the response option with the highest score is chosen as it is assumed this will reduce the likelihood of underestimating the impact of eczema on the individual.

5.4.2.5.2 Measurement properties

Figure 5-9 illustrates the number of participants who scored each of the available scores on the final instrument. The results show a normal distribution of scores and no floor or ceiling effects are present. A floor or ceiling effect was defined prior to data collection as more than 15% of participants achieving the lowest or highest possible score respectively, which has been previously suggested for the development of patient-reported measurement instruments (Terwee et al., 2007). Five participants (1.5%) scored the lowest possible score and two participants (0.62%) scored the highest possible score.

The scores for the final instrument were significantly positively correlated with POEM scores, $r(258) = 0.83$, $p < 0.001$, which is in line with the hypothesis about convergence validity (construct validity). The scatterplot in Figure 5-8 presents the correlation between the POEM and RECAP scores.

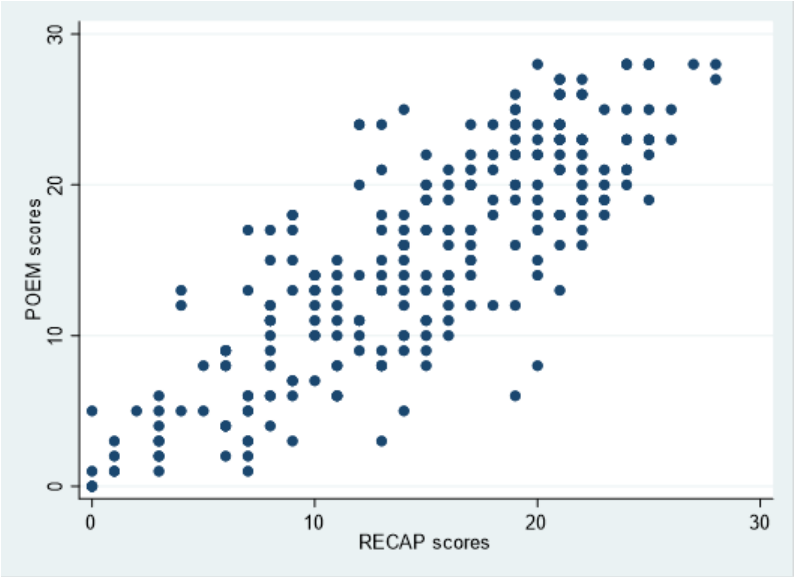


Figure 5-8 Scatterplot of correlation between POEM and RECAP scores

Table 5-9 illustrates how for each increase in severity banding according to established POEM severity bandings and a single item global severity measure corresponded with a larger mean RECAP score for those scoring within that severity

category (Charman et al., 2013), which is in line with hypotheses about discriminative validity (construct validity).

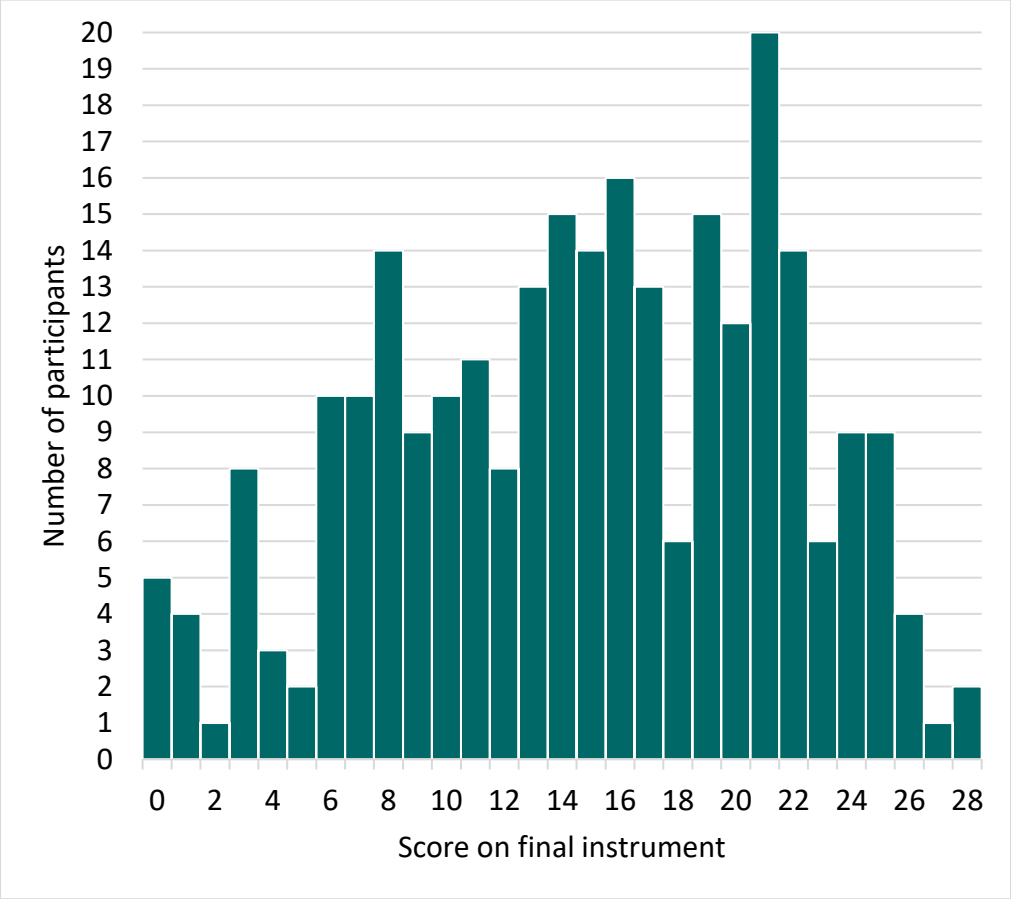


Figure 5-9 Distribution of scores on final instrument

Table 5-9 Mean RECAP scores by severity categories

Severity Categories	Mean (SD)	N	Range
POEM severity banding			
Clear – almost clear (0-2)	2.46 (2.67)	13	0-7
Mild (3-7)	7.15 (3.99)	33	0-19
Moderate (8-16)	12.64 (4.13)	94	4-22
Severe (17-24)	18.72 (4.25)	94	7-26
Very severe (25-28)	22.69 (3.18)	26	14-28
Global severity response option			
Clear	0.67 (1.21)	6	0-3
Almost Clear	6.88 (4.94)	34	0-21
Mild	10.86 (4.10)	64	3-20
Moderate	17.24 (4.50)	120	6-26
Severe	22.15 (3.23)	39	14-28

5.5 Discussion

5.5.1 Main findings

This study has developed a patient-reported outcome measure, called Recap of atopic eczema (RECAP) which is a patient-centred, easy to use measure of 'an individual's experience of eczema control' intended for use in clinical trials and routine care. The development process was designed to maximise the comprehensiveness, comprehensibility and relevance of the items to people with eczema/caregivers (Mokkink et al., 2017).

The study demonstrated that asking directly about an individual's 'experience of eczema control', the construct of interest, resulted in variations in interpretation between individuals and individuals reported having difficulty understanding the meaning of the question. The word 'control' was varied in its use in participants' everyday language. Therefore, the construct of interest was seen to be best approached via considering multiple aspects relating to the experience of eczema control. This multi-item instrument was built on a conceptual framework that suggested a formative model was the most appropriate underlying measurement model. This is where each observed variable is capturing a unique concept, which when combined together, form the latent variable *(causal indicators)* (De Vet et al., 2011).

Given the nature of the construct of interest, it was considered important to ask about a period of time rather than a single time point. The FDA guidance states a preference for items with short recall periods or items that ask patients to describe their current or recent state (Food and Drug Administration, 2009). The FDA guidance has been critiqued for ignoring the complex interplay of factors that determine which is the most appropriate recall period for a given instrument in a particular context, which is why it was initially assumed that a long-time period may be appropriate for the construct of interest in this population (Norquist et al., 2012).

However, the 4-week period initially used in this study was found to affect the ability of patients to calculate a response due to difficulties with recall and averaging out experience if it had varied greatly over that period of time.

The response options were rating scales developed using the principles of Kossnick and Presser (2010), who have presented four conditions that must be met for a rating scale to work effectively, which can be summarised as: cover the entire continuum of the concept, appear equally spaced conceptually, avoid responses where meanings overlap and use words so that each respondent has a stable and precise understanding of each point. The ongoing feedback from responders helped align the response options with these principles.

5.5.2 Implications of altering the conceptual model

The conceptual model initially proposed was refined throughout the development process of RECAP, which reflects the iterative nature of instrument development. However, it is important to reflect on why these refinements were considered appropriate when designing the instrument and what might be the potential implications of these refinements.

Firstly, the treatment and management related concepts, including stepping down treatment, using maintenance treatment only and self-management of the condition were not included in the final model. This was largely due to problems identified during testing items in the cognitive interview phase of the study. These issues have been outlined in Table 5-4. There was also some evidence from the cognitive interviews that these items were not considered by participants to be essential to include, hence supporting their removal.

Secondly, in the initial conceptual model, “an acceptable level of control is an individual experience” was an overarching concept that was considered important, but it was not initially clear how this fit within the design of the instrument. Through expert panel discussions when interpreting the findings from the face to face focus

group and designing items, it was acknowledged that items about the 'acceptability of eczema' to an individual and the individual's personal overall perception of 'how the eczema had been' were unique perceptions about the experience of eczema control that could be included as items in the measure.

Thirdly, guided by the definition of the construct of interest to be 'the extent to which the various manifestations of eczema and the impact that these have for an individual are removed or meaningfully reduced', which was derived from the qualitative studies presented in Chapter 3 and 4 and HOME V meeting decisions, this brought into focus items that were designed that were actually tapping into 'eczema control' that was not in line with the construct of interest. For example, 'I have been able to manage my eczema' and 'My eczema has felt unpredictable' made individuals in the cognitive interviews think about, and answer the question in relation to, their beliefs about the controllability of eczema, rather than the construct of interest which was their perceived experiences over the last week. This conceptual distinction has been discussed further in Chapter 2.

Finally, the final item set included in RECAP correlated highly with POEM scores. This could be an indication that although these instruments were designed with the intention of measuring different constructs conceptually, they are measuring closely related constructs. This is not surprising, as eczema symptoms experienced by individuals are closely related to the experience of eczema control, and some items to measure each construct overlap. However, further research will be required to further unpick the relationship between the constructs that RECAP and POEM are measuring.

5.5.3 Strengths and limitations

This study took a complex construct and used a patient-centred approach to develop a simple instrument that was acceptable to a range of stakeholders. The study followed guidance from leading organisations that can aid the development and

reporting of patient-reported outcome measures including the FDA, ISPOR and COSMIN (Food and Drug Administration, 2009, Patrick et al., 2011a, Patrick et al., 2011b, Mokkink et al., 2017).

The instrument was purposefully developed so that it could be applied across all age groups. The recruitment methods were also varied to try and reach different audiences. However, it is possible that there are potential biases in the types of people who would be willing to take part in an online survey voluntarily. For example, participants predominantly described their ethnicity as white. However, quite importantly, throughout all stages of the study it seemed we were successful in gaining experiences from people with a range of eczema severity. However, it could be argued that the severity of eczema reported by participants in the online survey does not match the severity distribution in the UK, where the majority of people have mild eczema (Emerson et al., 1998). This may be in part due to recruitment strategies meaning the advertisements were seen more often by people with more severe eczema. It is possible, therefore, that the items considered relevant to this sample are not the same items that would have been relevant to a sample with predominantly mild eczema. Whilst this severity distribution does not reflect the general population, it might be more likely to match the severity distribution seen in many eczema clinical trials, which are predominantly in moderate to severe disease (Nankervis et al., 2016).

The initial development phase involved testing of the instrument in a UK population and English language only due to resources available. However, involvement of stakeholders across different countries in the development team was utilised to try and anticipate any difficulties in adaption and translation that could be foreseen by the team.

Historically, children's experiences of health and illness have not been sought in research which could be due to concerns about consent, lack of evidence and feasibility (Hargreaves et al., 2018). However, there is a body of evidence that

suggests age-appropriate surveys could be used from 8 years onwards (Hargreaves et al., 2018). Although this study did not include as many younger people as was originally anticipated throughout the process, RECAP has been designed so that children can self-report if they are able and willing to answer the questions.

5.5.4 Practical implications

The development team aimed to create an instrument that could potentially be useful for inclusion in the HOME initiative's core outcome set domain of long-term control of eczema. The development (and subsequent validation study) of RECAP were presented at the HOME VII meeting in Tokyo, Japan, April 2019, where voting on instruments for the domain 'long-term control of eczema' took place. It was agreed that RECAP, alongside another instrument called the Atopic Dermatitis Control Test (ADCT), should be included as options for core outcome instrument(s). Reflections on RECAP as a HOME recommended core measurement instrument are presented in section 8.2.

5.5.5 Future research directions

A study by Bhanot et al. (manuscript in preparation) has already been conducted to assess the measurement properties of RECAP in a community-based UK sample. The study suggested RECAP had no floor or ceiling effects, good construct validity, good test-retest reliability and good responsiveness in this population. Further studies should assess the measurement properties of the RECAP measure to establish the contexts in which it is fit for purpose and to improve the usability and interpretability of the measure. It will be helpful if RECAP is utilised in clinical trials, other research studies or routine practice to collect data to aid understanding of how RECAP performs in different settings.

Since this measure was developed using a formative approach to model development, assessment of the measurement properties concerned with the internal structure of the measure (structural validity, internal consistency and

measurement invariance to establish cross-cultural validity) are not considered appropriate due to the formative model not being based on the same assumptions as reflective models.

It will also be important that the frequency of measurement of RECAP (and other measures of eczema control) is considered in further research and this is discussed at HOME meetings in relation to the core outcome set. Finally, further work could focus on delivering the RECAP in different formats (i.e. paper version, online version, mobile application) to improve the usability of the instrument in varied settings. Further work could also assess how well children of different ages are able to respond to the questions and perhaps develop aids (e.g. pictures) to improve ability for younger children to self-report.

5.6 Summary of Chapter 5

A process of developing a conceptual framework, item generation, item refinement, item selection and iterative refinement of the conceptual framework resulted in a newly-developed outcome measure to capture the experience of 'eczema control' over the past week. This instrument provides a potential solution for researchers, clinicians and patients desiring to capture this concept in clinical trials, other research designs and clinical practice.

5.7 Reflecting on recommendations from Chapter 3

Figure 3-3 provided seven key recommendations that should be taken into consideration when deciding how to measure long-term control of eczema on the basis of the UK online focus group findings. Although the development of RECAP was not specifically designed to meet these recommendations, they were considered in the development of RECAP, and it is helpful to reflect to what extent RECAP aligns with these recommendations. Where differences occur, it is also helpful to explore the reasons why. Each recommendation is presented and reflected on below.

1. **The need to scratch uncontrollably, the psychological impact, the social impact, symptoms (including itch, pain, sleep), impaired movement, the ability to do everyday activities and treatment used were all indicators of level of control.**

Although the indicators of 'eczema control' were slightly refined further in Chapter 4 to reflect findings from the international online focus groups and the HOME membership survey, the indicators of level of control elicited from the qualitative studies were the starting point for developing a conceptual framework and developing RECAP items. The final items did not contain items on the treatment used, with reasons outlined in Table 5-4.

2. **Understanding the baseline of what is normal for an individual was considered important for understanding eczema control.**

This is one of the major challenges for measures of long-term conditions. One way of understanding the baseline of normal would be to ask individuals to compare how their eczema is now to a previous state, known as a transition measure. However, this should not be applied uncritically, as there is ongoing theoretical discussions and empirical investigations into whether measures asking individuals to make such comparisons are valid due to issues with recall (Kamper et al., 2009, Streiner et al., 2015).

With regards to RECAP, which does not include a direct comparison within the questions asked, determining an appropriate baseline is a trial design issue. Usually, trials measure a baseline at the start of a trial to be used as the comparator point, but it is an ongoing research question requiring further discussion as to whether this is adequate in a long-term relapsing and remitting condition such as eczema.

3. **Not everyone with eczema experiences "flares" or finds it easy to notice changes in their eczema.**

This difficulty was considered during the development of RECAP, and therefore the initial items included multiple ways of approaching the concept of 'flares' that might be acceptable and relevant to most people. Asking about 'intensely itchy skin' was one of the ways this was approached, as intensely itchy skin is something participants said was an experience characteristic of when eczema is 'flaring', and this item was included in the final set.

4. **The behaviour of stepping up treatment for a flare was common but was complicated by factors such as difficulty getting a prescription and concerns about stepping-up treatment when already using maximum treatment available.**

The differences amongst participants in the online focus groups in how they responded to the level of control by stepping up and stepping down of treatment were mirrored in the cognitive interviews. Given the intended purpose of RECAP to be appropriate for use in all clinical trials, it was also a problem including items relating to this, as not all regimens in trials will allow for this behaviour in the treatment protocol.

5. **For some participants taking treatments that are recommended for short term use, lack of control referred to "rebounds" after the treatment ends. Long-term control of eczema measurement should take into account control both during and after treatment.**

RECAP can be used during or after treatment as it does not ask directly about treatment taken. Therefore, this is a trial design issue and can be implemented when using RECAP in a trial by measuring RECAP both during and after the treatment.

6. **The patient/caregiver perspective was considered important to fully capture eczema control. The ability of caregivers to report eczema control in young children was not questioned by caregivers.**

This was reflected in the decision to develop a patient-reported outcome measure and to provide two options of self and carer completed versions. Careful attention was considered as to what caregivers were able to respond about and how this was phrased so that they felt they may not be able to make judgements and provide responses.

7. **The acceptance of measurement frequency varied between participants and for an individual over time depending on lifestyle and commitments, the treatment stage and level of eczema control.**

The time frame of the measurement was given careful consideration. Given the desire to capture eczema control over longer time periods and minimize burden, it was initially attempted to ask about the last 4 weeks. However, this presented problems with recall and calibration of an answer over a period that might have been very changeable. The cognitive interviews suggest one week was an acceptable recall period. However, it remains a further decision during trial design whether RECAP is administered to participants as frequently as weekly or measured at less regular intervals. Trialists should consider what measurement of frequency is likely to be acceptable to participants taking part in that specific trial, which might depend on factors such as the severity of disease, the type of treatment being used and the duration of follow-up.

Section C

Informing the HOME initiative domain of patient-reported symptoms

This section moves the focus to the HOME initiative core outcome domain of patient-reported symptoms. This domain was further progressed in the HOME roadmap process than the long-term control of eczema domain, and the POEM had been recommended as the instrument that should be used in all eczema clinical trials. However, exploration of the measurement properties of the POEM that have not been fully addressed is required to ensure that it is fit for purpose in the core outcome set.

Chapter 6 explores the conceptual framework of the POEM. Chapter 7 was designed to improve the interpretability of the POEM scores and presents two studies that aim to provide meaning to POEM change scores and calculates the smallest detectable change and minimally important change of the POEM.

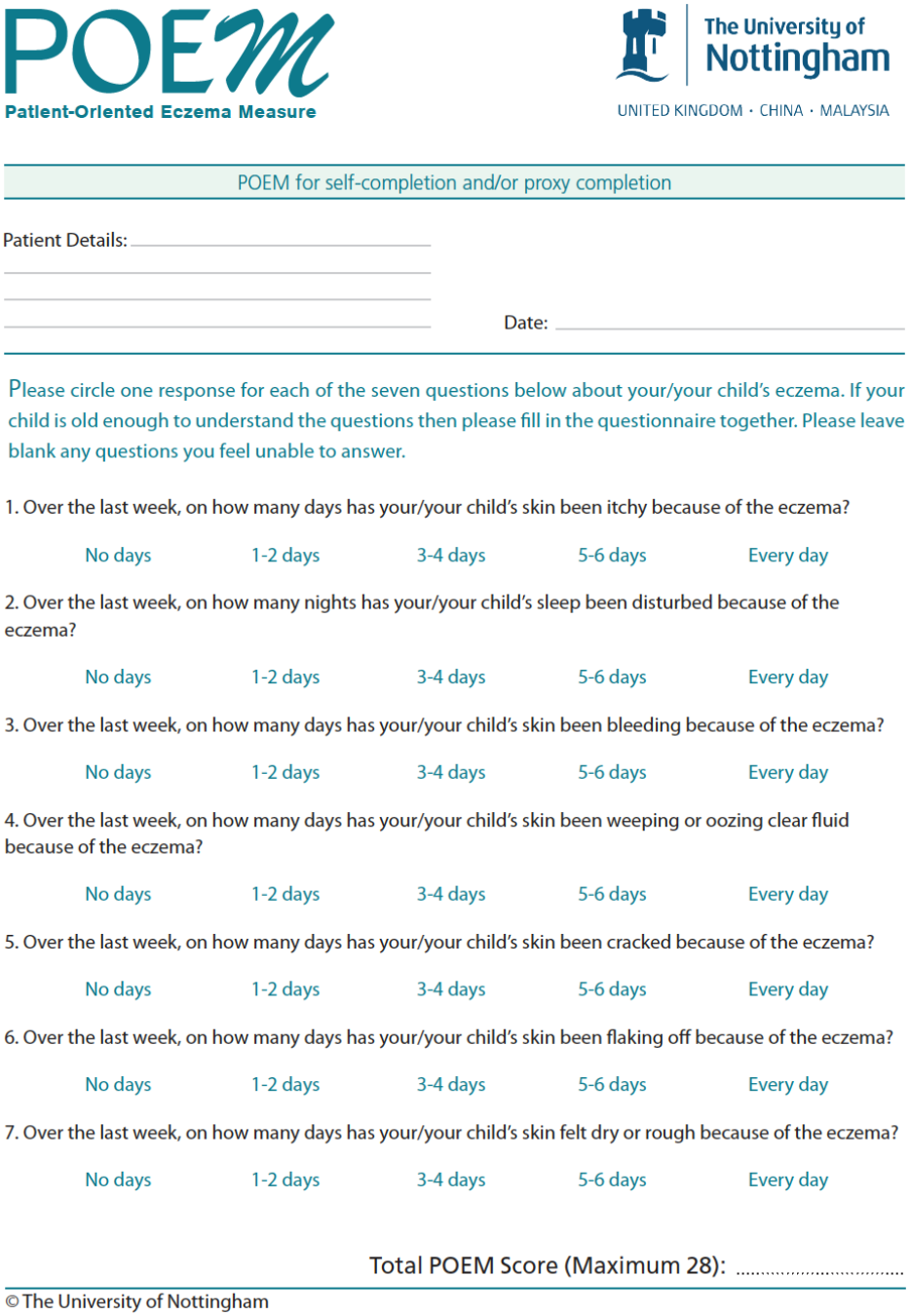
Chapter 6

Examining the measurement model of the Patient-Oriented Eczema Measure (POEM): rethinking assessment of the internal structure

6.1 Introduction

The POEM is a patient-reported outcome measure developed by Charman et al. (2004) that can be completed by individuals with eczema or caregivers of children with eczema (Figure 6-1). The POEM asks how frequently a symptom has occurred over the last week on a 5-point Likert scale from “No days” to “Every day” (Charman et al., 2004). The symptoms measured in the POEM are itch, sleep, bleeding, weeping/oozing, cracking, flaking, and dryness (Charman et al., 2004).

In accordance with the HOME roadmap, a range of instruments as candidates to be the core measurement instruments for the HOME domain ‘patient-reported symptoms’ were identified in a systematic review of instruments used to measure patient-reported symptoms in randomised controlled trials (Gerbens et al., 2016). This was followed by a systematic review of studies assessing the measurement properties of symptom measurement instruments for eczema published up until August 2015 (Gerbens et al., 2017). This review was guided by the original COSMIN checklist to determine the quality of each measuring instrument (Mokkink et al., 2010a). The COSMIN checklist is a consensus-based checklist to evaluate the methodological quality of studies on measurement properties.



POEM
Patient-Oriented Eczema Measure

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

POEM for self-completion and/or proxy completion

Patient Details: _____

_____

_____

_____ Date: _____

Please circle one response for each of the seven questions below about your/your child's eczema. If your child is old enough to understand the questions then please fill in the questionnaire together. Please leave blank any questions you feel unable to answer.

1. Over the last week, on how many days has your/your child's skin been itchy because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

2. Over the last week, on how many nights has your/your child's sleep been disturbed because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

3. Over the last week, on how many days has your/your child's skin been bleeding because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

4. Over the last week, on how many days has your/your child's skin been weeping or oozing clear fluid because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

5. Over the last week, on how many days has your/your child's skin been cracked because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

6. Over the last week, on how many days has your/your child's skin been flaking off because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

7. Over the last week, on how many days has your/your child's skin felt dry or rough because of the eczema?

No days 1-2 days 3-4 days 5-6 days Every day

Total POEM Score (Maximum 28): .....

© The University of Nottingham

Figure 6-1 Copy of the Patient-Oriented Eczema Measure (POEM)

Note. **© University of Nottingham**

In relation to the POEM, the review concluded that there was limited evidence for good internal consistency, and moderate evidence for good construct validity, good responsiveness and good content validity (Coutanceau and Stalder, 2014, Schram et al., 2012, Charman et al., 2004). Interpretation of POEM has been assessed in the form of the minimally important change and severity bandings (Schram et al., 2012, Gaunt et al., 2016, Charman et al., 2013, Howells et al., 2018a). Based on the available evidence, the POEM was recommended as the core outcome instrument for measuring symptoms in eczema trials at the HOME IV meeting (Chalmers et al., 2016). However, the group discussed that there were some validation gaps that should be explored (Chalmers et al., 2016). Structural validity and cross-cultural validity were identified as having an absence of evidence, and reliability and measurement error were considered to have unclear evidence (Chalmers et al., 2016).

The original aim was to design a study to address the structural validity evidence gap. Structural validity is the degree to which the scores of the instrument are an adequate reflection of the dimensionality of the construct to be measured (Mokkink et al., 2010b). It is suggested that for a developed instrument, confirmatory factor analysis of dimensions assumed by the instrument are tested (De Vet et al., 2011). However, to begin the task of assessing structural validity, the researcher is faced with considering in detail the measurement model underlying the instrument. By considering the measurement model of the POEM, questions were raised about the appropriateness of testing measurement properties relating to the internal structure (structural validity and internal consistency) (Prinsen et al., 2018).

Reflective models and formative models have been separated into two types of measurement models. In reflective models, each observed variable is a way of measuring the latent variable directly (effect indicators), whilst in formative models, each observed variable is capturing a unique concept, which when combined

together form the latent variable (De Vet et al., 2011). The concepts of reflective and formative models were introduced in section 1.2.2 of this thesis.

Mental experiments are structured ways of testing out ideas. Kuhn (1964) described the role of scientific thought experiments as "a way to disclose nature's failure to conform to a previously held set of expectations" as well as "suggest particular ways in which both expectation and theory must henceforth be revised". Mental experiments have been proposed as a mechanism for guiding conceptual decisions regarding the reflective or formative nature of a measurement model (De Vet et al., 2011).

6.2 Chapter aims

The aim of this chapter was to explore what is known about the measurement model of the POEM and use mental experiments to appraise whether the POEM should be considered as a reflective or formative model.

6.3 The measurement model of the POEM: moving from implicit to explicit

This section will first look at what is known about the POEM's construct of interest and approach to development to understand what is known about the measurement model of the POEM. Then it will turn to examine the key characteristics of reflective and formative models and use mental experiments to assess which the POEM best fits.

6.3.1 What is known about the POEM's measurement model?

Charman et al. (2004) described POEM as "a simple, valid, easily interpreted, and reproducible tool for assessing eczema and monitoring aspects of the disease that

are important to patients." Their use of the plural suggests that the POEM is intended to combine measurement of more than one aspect of disease. Further, the developers purport that the POEM "captures the fluctuating and chronic nature of eczema and provides a more comprehensive assessment of patient symptoms than that obtained by measuring itch and/or sleep disturbance alone" (Charman et al., 2004). Therefore, again alluding to the combining of more than one symptom.

The developers also distinguish between the construct measured by the total score, "eczema-related morbidity", versus using the scores of "individual variables can provide useful information on whether acute (e.g., weeping, bleeding) or chronic (e.g., itching, dryness) changes are predominant and can help target therapy accordingly" (Charman et al., 2004). In clinical practice, indicators that a specific symptom was problematic may help clinicians to further target their therapy. This suggests further evidence that the developers' view POEM as combining individual symptoms that together measure the construct "eczema-related morbidity". Furthermore, it provides insight that they did not view the symptoms as being likely to all change in equal measure at equal timing. The following section will articulate how these features suggest the POEM is based on a formative measurement model.

6.3.2 Mental experiments: is POEM based on a reflective or formative model?

Table 6-1 summarises features that can be differentiated between reflective and formative models (De Vet et al., 2011, Peterson et al., 2017). The following sub-sections will work through what is understood about each characteristic in relation to the POEM to help determine how the measurement model of POEM is most appropriately viewed.

Table 6-1 Differentiating features of reflective and formative models

Features	Reflective models	Formative models
Direction of causality	The items describe manifestations/reflections of the underlying construct. Imagine changing the latent variable and ask whether this is likely to change the value of the indicator(s). Intervention on indicators assessed by specific items do not cause changes in the underlying construct.	The items are characteristics that induce of form the underlying construct. Imagine a change in the indicator and ask whether this change is likely to change the value of the latent variable. Items are defining characteristics of the construct, such that intervention on indicators assessed by specific items cause changes in the construct.
Redundancy of items	Each item is a reflection of the latent construct; therefore, within each dimension of a questionnaire, all items are replications of the same construct. Items should be related to one another and the construct should have useful redundancy.	Each item contributes a part of the construct, and together the items form the whole construct. Items should be related to the construct with no redundancy.
Item covariation	Items that correspond to the same underlying latent construct are correlated and replaceable with one another. Items are expected to covary; as responses on one item go up, responses on the other also tend to increase.	Items do not necessarily correlate with each other, and thus are not interchangeable, one item cannot be replaced by another. Items are not expected to covary; as responses to one item increase, there is no expectation of an increase in response on others.
Effect of item removal on construct	Removal of one item does not alter the construct. Items that are measuring the same underlying latent construct can be removed without changing the nature of the construct being measured.	Removal of one item may alter the construct in a formative model. Missing an important item inevitably means the construct is not measured comprehensively.
Methods used to choose items	Interested in inter-item correlation, exploratory and confirmatory factor analysis used.	The assessment of importance and the elimination of less important items should take place during field-testing. Interested in maximal variation, underlying structure based on regression model.

Note. Information within table is based on sources De Vet et al., 2011 and Peterson et al., 2017

6.3.2.1 Direction of causality

Figure 6-2 represents the POEM as a reflective model and Figure 6-3 represents the POEM as a formative model. The direction of causality can be a challenging distinction to make, due to the often seemingly bi-directional relationship between items and constructs. POEM developers' description of the construct of interest suggests that each item relating to the individual symptom can be conceptualised as forming their overall perception of their eczema symptoms. If there is a change in an

item (e.g. frequency of itchy skin perceived) this is likely to change the value of the construct of interest (perceived eczema severity overall). An intervention on an item (e.g. to improve skin feeling dry or rough), would lead to a change in overall perceived eczema symptoms. This is in line with a formative model (Table 6-1).

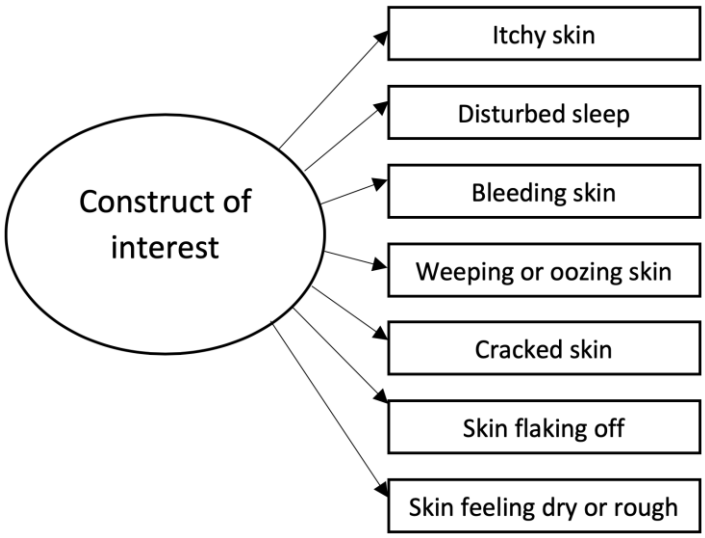


Figure 6-2 Representing POEM as a reflective model

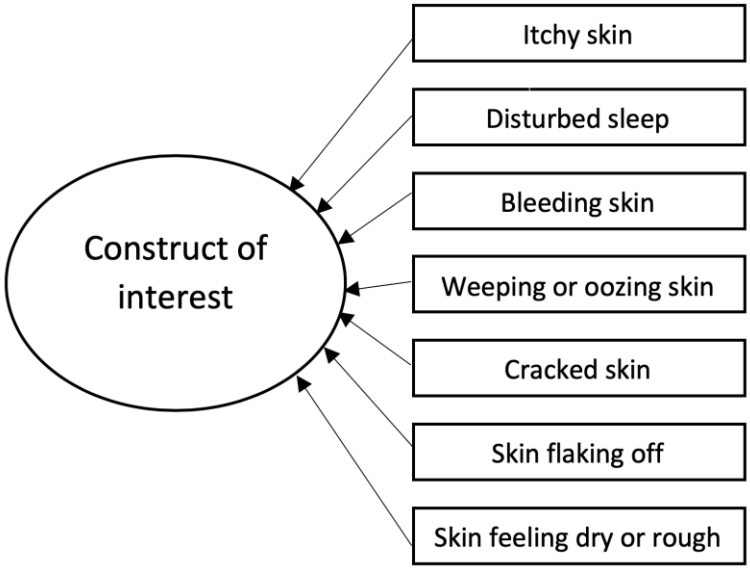


Figure 6-3 Representing POEM as a formative model

6.3.2.2 Redundancy of items and item covariation

Within the POEM, each item provides unique information about an additional concept and are not interchangeable. For example, how often an individual reported itch in the last week and how often an individual reported dry skin in the last week are clearly trying to measure unique aspects of that person's experience of eczema symptoms that week. Within the description of the development of the POEM, there is no indication that the developers aimed to have a group of items to measure closely related concepts. In fact, the developers aimed to create a manageable list of all different symptoms that were important to patients. It appears that the approach to redundancy of items in the POEM is in line with the development of a formative model.

The POEM is made up of symptoms that the developers predicted may occur at different time points (differentiated as more acute and more chronic symptoms). Therefore, whilst the developers clearly anticipated the items would be somewhat related since they conducted a test of internal consistency, conceptually it is predicted that correlations between the items may vary over time and the items were not designed to be replaceable with one another (Charman et al., 2004). Therefore, the emphasis on item correlation required for a reflective model may not apply for this instrument (Table 6-1).

6.3.2.3 Effect of item removal on construct

Removal of items in the POEM would lead to a variation in the concept being measured. For example, a patient with chronic dry uncomfortable lichenified eczema would score zero if the POEM only asked patients about weeping or bleeding, and vice versa, a patient with recurrent infected flare ups would score zero if the POEM only asked about dryness or flaking. Removing any item would alter the construct, which is suggestive of a formative model. If the model was reflective, there would be a level of redundancy that meant removal of items would not fundamentally alter the construct being measured (Table 6-1).

6.3.2.4 Methods used to choose items

Moving from the conceptual nature of models to the methods employed to develop measures built on certain models, the POEM appears to be developed in line with principles used for measurement based on formative models. Reflective models are developed using factor analysis to ensure effect indicators with unity in meaning are combined and effect indicators with distinct meaning are scored separately. POEM developers did not use factor analysis. POEM was developed by assessing the frequency and the importance of symptoms for patients (Charman et al., 2004). Whilst the development processes of formative measures are not as well established as the development processes of reflective measures, the emphasis on including the most important items is common and often recommended (De Vet et al., 2011).

6.4 Discussion

By using mental experimentation to consider how the properties of the POEM fit with conceptual differences between reflective and formative measures, as well as differences in methods of development depending of the nature of the construct, it has been expressed that the POEM may be most consistent with a formative model. Thereby, proposing the conceptual model of POEM is depicted in Figure 6-3.

6.4.1 Strengths and limitations

Ideally, the measurement model should be explicitly reported by the developers of the outcome measure, and current guidance encourages this as it is recommended that explicit conceptual frameworks are developed to guide the measurement model for instrument development (Food and Drug Administration, 2009, De Vet et al., 2011, Patrick et al., 2011a). Since the POEM was reported before these guidelines, it is helpful to retrospectively explicitly state the intended measurement model. The original publication by Charman et al. (2004) comments on the construct of interest

and how the items are related, so inferences about the measurement model can be made on this basis. To gain further insight into the developers' intentions, the developers (HW and CC) were asked for comment and they offered comments that helped formulate what is presented in this chapter.

Mental experiments are widely regarded as the method to determine the nature of the measurement model of an instrument (De Vet et al., 2011). There is ongoing debate regarding the superiority of reflective models (Edwards, 2011). However, it is ultimately detrimental to misspecify the measurement model, as this can undermine the validity of the measure and result in inappropriate testing of the measure (Coltman et al., 2008, Prinsen et al., 2018, Streiner, 2003).

6.4.2 Implications for future research

Assuming a formative model for the POEM has implications for the assessment of the measurement properties of structural validity and internal consistency, as these properties are built on the assumptions that the measurement is based on a reflective model. Therefore, it is deemed inappropriate to assess these properties (Prinsen et al., 2018, Streiner, 2003).

Gerbens et al. (2017) conducted a systematic review of measurement properties of patient-reported symptoms as part of the HOME roadmap and found one study by Charman et al. (2004) that assessed internal consistency of the POEM, which was graded as being of "fair" methodological quality using the original COSMIN checklist (Gerbens et al., 2017, Mokkink et al., 2010a). It is suggested that this evidence should not be included for consideration when reviewing the measurement properties of the POEM due to the formative nature of the underlying measurement model.

6.5 Summary of Chapter 6

To conclude, this chapter has argued that POEM may be best considered as being developed with the underlying principles of a formative model, therefore assessing measurement properties relating to the internal structure of a measure (structural validity and internal consistency) is inappropriate. Relating to the previous chapter, during the development of RECAP in Chapter 5 the instrument was also deemed to be a formative model, therefore these measurement properties will also not be relevant when assessing the measurement properties of RECAP. The next chapter will proceed with looking at the interpretability of changes in POEM scores, which is required to enhance the usability of the instrument.

Chapter 7

Interpretability of change in Patient-Oriented Eczema Measure (POEM) scores

7.1 Introduction

The definition of interpretability was introduced in section 1.2.3.4 as the degree to which one can assign qualitative meaning – that is, clinical or commonly understood connotations – to an instrument's quantitative scores or change in scores (Mokkink et al., 2010b). The scores produced by a multi-item instrument are initially not easily interpreted and there have been approaches designed to aid interpretability.

7.1.1 Interpreting single scores

One approach to improve the interpretability of an instrument's scores is to use an external reference of well-known groups to assign clinical connotations to the scores. This has been previously assessed for POEM. Charman et al. (2013) used two global questions about overall assessment of eczema severity as an external criterion to assign severity categories to POEM scores which were 0–2 (clear/almost clear); 3–7 (mild); 8–16 (moderate); 17–24 (severe); 25–28 (very severe).

7.1.2 Interpreting changes in scores

Another approach to improve the interpretability of an instrument is to assign meaning to the change in scores. A concept found to be useful for interpreting changes in scores is the minimally important change (MIC), which has been defined as 'the smallest change in score in the construct to be measured which patients perceive as important' (Mokkink et al., 2009). There is a variety of terminology used to explain the concept of MIC, the most common alternatives being the minimal clinically important difference (MCID) and minimally important difference (MID). MIC tends to be used to refer to longitudinal within-person changes in scores, and MID for cross-sectional between-person differences (de Vet et al., 2006b). For

consistency, I have used MIC throughout this thesis. There is an ongoing debate about whether the methods currently used are appropriate to estimate the MID (de Vet et al., 2006b).

Understanding the MIC of the POEM can support sample size calculations for clinical trials and interpretation of trial results (Wright et al., 2012). MIC estimates also allow clinicians to interpret a patient's change in the POEM score in the clinical setting and aid decisions regarding whether a treatment alteration is required (Wright et al., 2012). It has been recommended that researchers use multiple methods and multiple datasets to triangulate MIC results (Beaton et al., 2002, Revicki et al., 2008, Crosby et al., 2003). Although a multitude of MIC estimates could detract from the usefulness of a universal MIC threshold, it is important to explore how the MIC of the POEM may vary to ensure it is meaningful in the context used.

7.1.2.1 Distribution-based approaches

One approach to assigning meaning to change scores is the standardised effect size, which was developed to assign meaning when the units of measurement are not intuitive (Coe, 2002). In its simplest form, the standardised effect size is the standardised mean difference. The most commonly applied interpretation of the standardised effect size is that 0.2 is a small effect, 0.5 is a medium effect, and 0.8 is a large effect (Cohen, 1969, as cited in Coe, 2002).

Using the standardised effect size approach to interpret change has since often been referred to as a distribution-based approach to assessing the MIC (Beaton et al., 2002, Cook, 2008, Copay et al., 2007, Wright et al., 2012). It was suggested by Norman (2003) following a systematic review that an effect size of 0.5 typically corresponded closely to patients with chronic disease identifying MIC in instrument scores for health-related quality of life. Others have argued that it would be more appropriate to interpret the change in score corresponding to an effect size of 0.2 as the MIC as this is often seen as a small effect size (Beaton, 2003). However, it has

been argued that both of these methods do not sufficiently capture the concept of MIC because it does not assess the importance of that change, and so has been suggested as a method more akin to minimal detection rather than minimal importance (De Vet et al., 2011, Turner et al., 2010).

Nevertheless, using the 0.5 SD method to improve the interpretability of change in instrument scores has been widely adopted due to the ease of calculation and readily availability of data required for such a calculation. However, it has been articulated that the SD used in the calculation needs to also reflect the SD of the population of interest for the clinical trial for the results to be used appropriately (Cook et al., 2014, Cook et al., 2015). Therefore, when using this method, it is important to think about the population of interest and how the SD may change in participants with different characteristics such as age, sex, ethnicity and disease severity.

7.1.2.2 Anchor-based approaches

"Anchor-based" approaches to assigning qualitative meaning to changes in scores are similar to approaches to assign meaning to single scores, where a well-known reference point can be used to assess how this corresponds to change on the instruments scores (De Vet et al., 2011). A certain amount of change on an external criterion, which should ideally be a related and well interpretable outcome measure (the 'anchor'), is said to correspond to a MIC on the measuring instrument of interest (De Vet et al., 2011).

The MIC of the POEM was previously estimated using a patient-reported anchor using what is referred to within this thesis as the 'within-patient' score change method (section 7.3.1.3.2.2 contains details on this method). The MIC was reported as 3.4 points in a study using datasets from two trials in adults with severe eczema (Schram et al., 2012). Subsequently, a study using data from a trial in children with mild to moderate eczema aged between 1 month and 5 years from primary care used a combination of anchor-based and distribution-based methods to calculate

the MIC. Results ranged from 2.5 to 4.27, which led the authors to conclude that the results broadly concurred with an MIC of three points (Gaunt et al., 2016).

7.1.2.3 Smallest detectable change

The MIC of an instrument must be greater than the smallest detectable change to be useful. In other words, the smallest detectable change can be defined as a change beyond measurement error (Mokkink et al., 2010b). It is important to be able to interpret changes in scores in relation to both the MIC and the smallest detectable change as a change in a score may be detectable beyond measurement error but not clinically relevant, and it may also be that clinically relevant change has occurred, but the instrument may be insufficient for capturing minimally important change (De Vet et al., 2011). The smallest detectable change of the POEM has not been previously assessed (Gerbens et al., 2017).

7.2 Aims and objectives

The aim of this chapter is to aid interpretability of change in POEM scores to improve the designing and interpreting of eczema clinical trials using the POEM. To achieve these aims this chapter includes two studies that have used existing clinical trial datasets that have included POEM. As presented in Figure 7-1, study 1 used one dataset to assess change in POEM scores using a variety of methods, whilst study 2 merged multiple datasets to use a distribution-based approach to explore change in POEM scores across different patient populations.

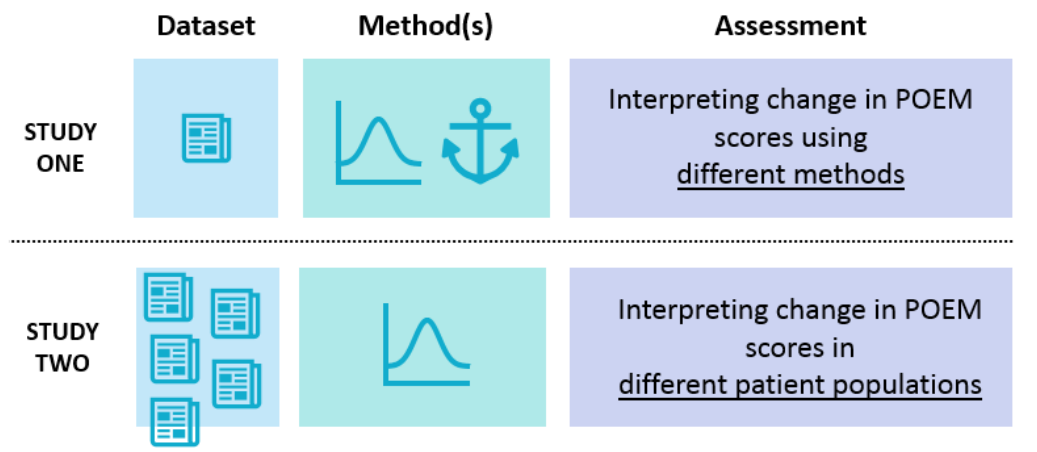


Figure 7-1 Schematic diagram of included studies

Note. Key to schematic diagram:

= number of datasets

= distribution-based methods used

= anchor-based methods used

Study 1 objectives:

- to calculate the smallest detectable change in the POEM,
- to estimate the MIC of the POEM by repeating methods used in previous studies and using additional methods;
- and to assess whether using a patient/parent or investigator static global assessment as an anchor influences the anchor-based MIC estimates.

Study 2 objectives:

- to assess how the POEM scores that correspond to a specified standardised effect size (calculated as 0.5 multiplied by SD at baseline) are influenced by age of the child, gender, ethnicity and disease severity.

7.3 Study 1

7.3.1 Methods and Materials

The study protocol was registered on the Centre of Evidence Based Dermatology's Protocol Registration on 5 January 2017 prior to data analysis: http://www.nottingham.ac.uk/research/groups/cebd/resources/protocol-registration.aspx. As this study made secondary use of an existing trial dataset, further ethics approval was not required, and this was confirmed by the University of Nottingham's Faculty of Medicine & Health Sciences Research Ethics Committee (Ref: 258-1712).

7.3.1.1 The CLOTHES trial

The CLOTHES trial is a parallel-group, randomized, controlled, observer-blind trial. Children aged 1–15 years were recruited from secondary care and the community and allocated to wear either silk garments plus standard care or standard care only. At study entry, one of the eligibility criteria was that participants had either a moderate (9–11) or severe (12–15) score on the Nottingham Eczema Severity Score (Thomas et al., 2017, Emerson et al., 2000).

7.3.1.2 Measures

POEM measures patient-reported frequency of itch, sleep disturbance, bleeding, weeping/oozing, cracking, and flaking and dryness/roughness over the past week. Each item is weighted equally and scored as 0 (no days), 1 (1–2 days), 2 (3–4 days), 3 (5–6 days) or 4 (every day). This analysis used POEM scores at baseline and at 6 months from the CLOTHES dataset. A Patient's/Parent's Global Assessment (PGA)

and Investigator's Global Assessment of Severity (IGA) were also used (these measures are described in Table 7-1).

Table 7-1 Measures used for anchors

Measure name	Question	Response options (tick one box)	Completed by:	Times collected to be used for anchors:
Static Patient/parent global assessment (PGA)	How is your / your child's eczema today?	Clear Almost clear Mild Moderate Severe Very Severe	Parent/legal guardian of child with eczema or child themselves if old enough (individual decision)	Baseline 6 months
Static Investigator global assessment (IGA)	How is the child's eczema today?	Clear Almost clear Mild Moderate Severe Very Severe	Research nurse (excluded measure when different nurse completed at different time points)	Baseline 6 months

7.3.1.3 Statistical analysis

The total POEM score was calculated as recommended by the developers by adding together the score from each item (Charman et al., 2004). If one item was missing the total score was still calculated, but total score was coded as missing if at least two items were missing. As only 9% of data collected at 6 months was missing, all analyses used complete case series. POEM scores from patients in both the treatment and control arm were not treated separately in this study. Except where stated otherwise, the statistical package Stata 14 was used to run analyses (StataCorp, College Station, TX, U.S.A.). No formal sample size was conducted, but the number of patients included in the analysis is likely to be sufficient. It has been recommended that validation studies contain at least 100 participants (De Vet et al., 2011).

7.3.1.3.1 Calculating the smallest detectable change

The smallest detectable change was calculated as $1.96 \times \sqrt{2} \times SEM$ (De Vet et al., 2011) The standard error of measurement (SEM) was calculated as:

$$SEM = SD(pooled) \times \sqrt{1 - ICC}\,.$$

The following calculation was used for SD (pooled):

$$SD(pooled) = \frac{\sqrt{SD(baseline)^2 + SD(6\ months)^2}}{2}$$

The intraclass correlation coefficient (ICC) for absolute agreement was derived from the test–retest reliability of the POEM, which was tested in the development of the POEM (Charman et al., 2004). This was considered acceptable as there is similar variability in the CLOTHES data, SD(baseline) = 5.36, and the dataset used initially to validate the POEM by Charman et al. (2004), SD(baseline) = 7.73.

7.3.1.3.2 Calculating the minimally important change (MIC)

7.3.1.3.2.1 Distribution-based methods

Two distribution-based methods were used to estimate the MIC. An effect size is measured by the difference between the score at baseline and follow-up, divided by the standard deviation of the baseline score (Copay et al., 2007). Multiplying 0.5 by the standard deviation of POEM scores at baseline was used as an estimate of the MIC (Norman et al., 2003). It has been suggested that the MIC should correspond to a smaller effect size, therefore 0.2 multiplied by the standard deviation of POEM scores at baseline was also used (Copay et al., 2007).

7.3.1.3.2.2 Anchor-based methods

Four anchor-based methods were used to estimate the MIC. The IGA and PGA scores were transformed into a change score to provide an anchor: score at time point 1 (baseline) minus score at time point 2 (6 months). This creates change scores on the

IGA and PGA that can range from –5 (worsened eczema severity) to 5 (improved eczema severity), that were used as anchors. The MIC was operationally defined as a positive one-point change on the PGA or the IGA anchor to indicate a minimal important improvement. This change indicates a change in severity banding on the PGA/IGA (e.g. from moderate to mild). This is also the cut-off used by Schram et al. (2012), enabling comparisons with previous MIC estimates.

For an anchor to be useful it must at least moderately correlate with the POEM change score ($r \geq 0.3$) (Revicki et al., 2008). This was assessed prior to the study with Pearson's r correlations between the POEM change scores and (i) the PGA change scores ($r = 0.55$) and (ii) the IGA change scores ($r = 0.46$), and both met this minimum criterion. The assumptions of Pearson's r correlations were met.

The first two anchor-based methods used mean change to analyse (i) 'within-patient' score change where the MIC is estimated as the mean change score of the smallest reported improvement (a positive one-point change in the anchor) and (ii) 'between-patient' score change where the MIC estimate is based on the relative change between the mean change score of the group with the smallest reported improvement on the anchor (a positive one-point change) and the mean change score of the group with no change on the anchor.

The third anchor-based method used the receiver operating characteristic (ROC) curve. The area under the curve of the ROC curve identifies the cut-off point on the POEM change scores that most optimally distinguishes between the anchor of IGA or PGA change scores ≤ 0 and IGA or PGA change scores ≥ 1 (Copay et al., 2007). The cut-off used to provide an MIC estimate will maximize the Youden J statistic: sensitivity – (1 – specificity) (Copay et al., 2007). The statistical package R (R Foundation, Vienna, Austria) was used for bootstrapping methods to allow us to calculate an MIC estimate with 95% confidence intervals (CIs) (Terluin et al., 2015).

The fourth anchor-based method used predictive modelling. This method uses logistic regression to predict whether a patient belongs to the improved (≥ 1) or not

improved group (≤ 0) on the IGA or PGA anchor using the change in POEM score as the predictor (Terluin et al., 2015). The MIC is estimated using the equation [ln(pre-odds) – C]/B, where C is the intercept and B is the regression coefficient for the change in POEM score from the logistic regression model. The pre-odds is calculated using the proportion improved based on the anchor divided by 1 minus the proportion improved based on the anchor (Terluin et al., 2015). The Microsoft Excel spreadsheet designed to aid confidence interval calculations of predictive modelling MIC estimates provided in supplementary materials by Terluin et al. (2015) was used. It has been suggested that if the proportion improved on the anchor does not equal 0.5, an adjusted MIC may need to be calculated (Terluin et al., 2017). As the proportion improved on the IGA anchor was 0.56 and the PGA anchor was 0.53, the adjusted MIC has not been reported here.

7.3.2 Results

A total of 300 children with eczema were randomized into the CLOTHES trial and completed the POEM at baseline; 174 (58%) were female and the majority were of white ethnic origin (n = 237, 79%). At 6 months, 273 participants (91%) had an assessment visit, therefore 273 patients were included in the anchor-based MIC methods that required this time point. Table 7-2 summarizes the age, disease severity and POEM scores of the sample. Figure 7-2 provides the distribution of POEM scores at baseline.

Table 7-2 Summary of participant characteristics and POEM scores

	N	Mean	SD	Min.	Max
Age (years)	300	5.06	3.63	1	15
Nottingham Eczema Severity Score (NESS)	300	13.13	1.62	9	15
POEM baseline	300	16.95	5.36	4	28
POEM 6 months	273	12.16	6.95	0	28
POEM change scores	273	4.78	7.14	-21	24

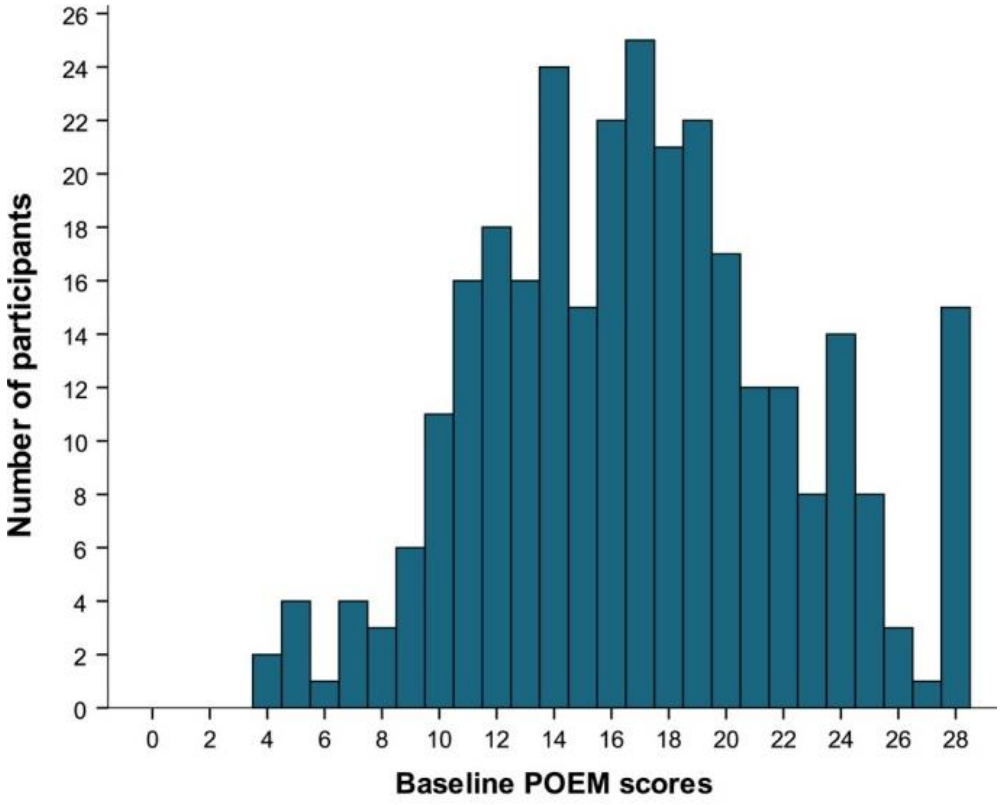


Figure 7-2 Distribution of POEM scores at baseline

7.3.2.1 Smallest detectable change

The intraclass correlation coefficient was 0.9847 and the pooled SD was 6.21, therefore SEM = 0.77.

Therefore, the smallest detectable change was calculated as:

$$SDC = 1.96 \times \sqrt{2} \times 0.77 = 2.134$$

The smallest detectable change in the CLOTHES dataset was 2.13 points on POEM.

7.3.2.2 Minimally important change

The MIC of the POEM was analysed using two distribution-based methods and four anchor-based methods. Figure 7-3 summarizes the results of the MIC estimates for each method.

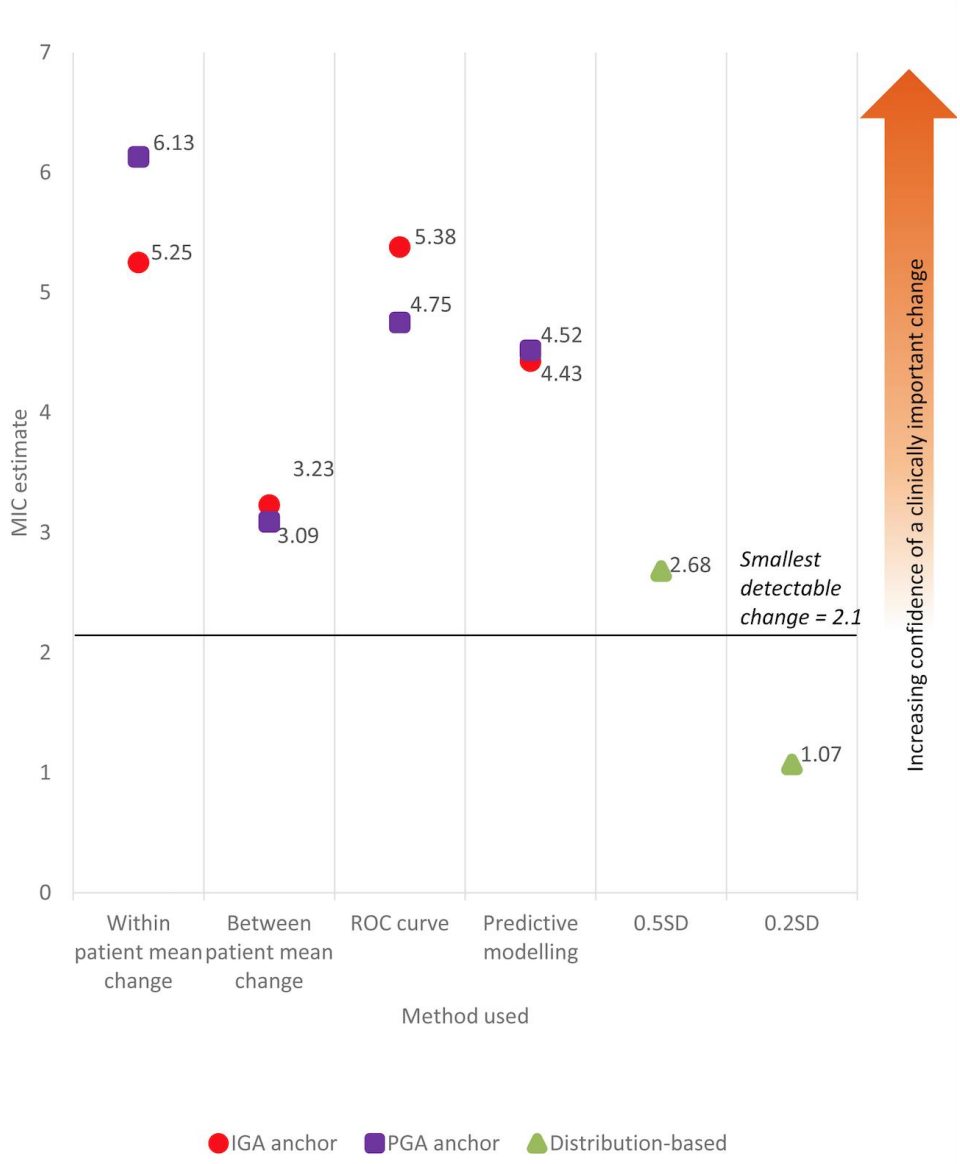


Figure 7-3 A visual representation of the study results

Note. POEM scores range from 0-28.

7.3.2.2.1 Distribution-based methods

As shown in Table 7-2, the SD of POEM scores at baseline was 5.36. Using 0.5 SD of baseline scores gave an MIC of 2.68 points (95% CI 2.5–2.89) and using 0.2 SD of baseline scores gave an MIC of 1.07 points (95% CI 1.00–1.16).

7.3.2.2.2 Anchor-based methods

Table 7-3 shows how the IGA anchor and PGA anchor were created.

7.3.2.2.2.1 Mean change methods – the 'within-patient' score change and the 'between-patient' score change

As presented in Table 7-3, when using the IGA as the anchor, the mean change score of the POEM was 2.02 for those with a change score of zero (no change on the IGA) and 5.25 for those with a score of one (defined as minimal improvement on the IGA). Therefore, for the 'within-patient' score change method the MIC is 5.25 (95% CI 4.04–6.46) and for the 'between-patient' score change method the MIC is 3.23 (95% CI 1.51–4.95).

As presented in Table 7-3, when using the PGA as the anchor, the mean change score of the POEM was 3.04 for those with a change score of zero (no change on the PGA) and 6.13 for those with a score of one (defined as minimal improvement on the PGA). Therefore, for the 'within-patient' score change method the MIC is 6.13 (95% CI 4.82–7.44) and for the 'between-patient' score change method the MIC is 3.09 (95% CI 1.25–4.94).

Table 7-3 Mean POEM change scores for participants classified according to change on anchors

Investigator global assessment (IGA)			Patient/parent global assessment (PGA)		
Change in score on IGA	N (%)	Mean POEM change score (SD)	Change in score on PGA	N (%)	Mean POEM change score (SD)
-5	0 (0)	n/a	-5	0 (0)	n/a
-4	0 (0)	n/a	-4	0 (0)	n/a
-3	0 (0)	n/a	-3	0 (0)	n/a
-2	1 (0.4)	-1 (n/a)	-2	9 (3.3)	-3 (9.64)
-1	24 (8.8)	1.54 (6.77)	-1	38 (13.9)	-0.79 (4.82)
0	95 (34.8)	2.02 (6.01)	0	81 (29.7)	3.04 (5.88)
1	111 (40.7)	5.25 (6.43)	1	92 (33.7)	6.13 (6.33)
2	33 (12.1)	10.67 (6.01)	2	38 (13.9)	9.05 (5.92)
3	9 (3.3)	15.89 (7.17)	3	10 (3.7)	11.8 (4.57)
4	0 (0)	n/a	4	5 (1.8)	18.2 (5.63)
5	0 (0)	n/a	5	0 (0)	n/a

7.3.2.2.2.2 ROC curve method

The MIC using the IGA anchor was 5.38 (95% CI 1.5–8.5) and the MIC using the PGA anchor was 4.75 (95% CI 3.5–6.5).

Predictive modelling method

Table 7-4 contains the results of the logistic regression analyses that were used to calculate the MIC and confidence intervals. Using IGA as the anchor, the MIC estimate using predictive modelling was 4.43 (95% CI 2.21–6.74). Using the PGA anchor, the MIC estimate was 4.52 (95% CI 2.81–6.29).

Table 7-4 Logistic regression results used for predictive modelling MIC estimates

	IGA	PGA
Pre-odds*	1.257	1.133
C (S.e)	-0.293 (0.159)	-0.598 (0.171)
B (S.e)	0.121 (0.022)	0.160 (0.025)
Correlation of C and B**	-0.570	-0.615

Note. S.e = standard error * odds of improvement based on anchor only ** this was calculated using IBM SPSS Statistics 22 and was required as part of the calculation of 95% confidence intervals

7.4 Study 2

7.4.1 Methods and Materials

Secondary analysis was conducted on datasets from five clinical trials including children from the UK, referred to as CLOTHES, SWET, BATHE, COMET and CREAM (Thomas et al., 2017, Thomas et al., 2011, Santer et al., 2015b, Ridd et al., 2015, Francis et al., 2016). As this study made secondary use of existing trial datasets, further ethics approval was not required, and this was confirmed by the University of Nottingham's Faculty of Medicine & Health Sciences Research Ethics Committee (Ref: 258-1712). The protocol was prospectively registered on the CEBD protocol registration portal (https://www.nottingham.ac.uk/research/groups/cebd/documents/researchdocs/protocol-mic-poem-patient-characteristics-v1.0.pdf).

POEM measures patient-reported frequency of itch, sleep disturbance, bleeding, weeping/oozing, cracking, and flaking and dryness/roughness over the past week. Each item is weighted equally and scored as 0 (no days), 1 (1–2 days), 2 (3–4 days), 3 (5–6 days) or 4 (every day). Age, gender and ethnicity variables in each trial dataset were transformed into the a priori defined categories outlined in Table 7-5.

Table 7-5 Summary of participant characteristics for each clinical trial dataset

	CLOTHES Clothes for the relief of Eczema trial	SWET Softened Water Eczema Trial	BATHE Bath additives for the Treatment of cHildhood Eczema trial	COMET Choice of Moisturiser for Eczema Treatment trial	CREAM ChildRen with Eczema Antibiotic Manageme nt trial
Trial Registratio n number	ISRCTN: 77261365	ISRCTN: 71423189	ISRCTN: 84102309	ISRCTN: 21828118	ISRCTN: 96705420
Population trial aimed to recruit from	Moderate – severe eczema	Moderate – severe eczema	All severities of eczema (except very mild)	All severities of eczema	Clinically infected eczema
Setting of recruitment	Secondary care and self-referral	Secondary care, primary care and self-referral	Primary care	Primary care	Primary care
Location of recruitment	UK	UK	UK	UK	UK
N	300	336	482	196	112
Age N (%)					
0-2	125 (42)	131 (39)	191 (40)	170 (87)	65 (58)
3-7	114 (38)	136 (40)	221 (46)	26 (13)	47 (42)
8-17	61 (20)	69 (21)	70 (14)	0	0
Ethnicity N (%)					
Not white	63 (21)	74 (22)	75 (16)	29 (15)	19 (17)
White	237 (79)	260 (77)	397 (82)	155 (79)	91 (81)
Not reported	0	2 (1)	10 (2)	12 (6)	2 (2)
Female N (%)	126 (42)	143 (43)	244 (51)	85 (43)	51 (46)

The baseline POEM scores and the patient characteristic variables from each trial were combined in one dataset in STATA Version 14. The POEM score that corresponds to the standardized effect size (0.5 multiplied by the SD at baseline) was calculated. To explore the impact of disease severity, the overall mean, standard deviation and POEM scores corresponding to the standardised effect size were calculated for each trial separately (as each trial recruited participants with different eczema severities). The individual participant data was also combined from the five

trials to calculate the mean, standard deviation and 0.5 SD of POEM scores for the overall sample and for each age, ethnicity and gender category.

Half of a SD for the trial of methotrexate versus azathioprine for severe atopic dermatitis (MAcAD) was also calculated (Dutch trial register: NTR1916) (Schram et al., 2011). However, the MAcAD trial dataset was not combined with the other datasets as the patients were not sufficiently similar (Dutch adults in MAcAD and UK children in all other studies).

7.4.2 Results

Data from 1,426 participants across five UK trials were combined. All five trials included children in the categories 0-2 years and 3-7 years. The CLOTHES, SWET and BATHE trials included some children in the 8-17 years category, however this was the least frequent age in all three trials. Gender was roughly equally distributed within all trials. The majority of participants in all trials were white (79%). A minority from four of the trials had participants who did not report ethnicity (Table 7-5).

Table 7-6 shows that multiplying 0.5 by the standard deviation of POEM scores at baseline from each of the five trials of children in the UK ranged from 2.68 to 2.95. When the individual participant data from these five trials were combined and categorised according to age, ethnicity and gender, multiplying 0.5 by the standard deviation of POEM scores at baseline ranged from 3.25 to 3.45.

The MAcAD trial included 43 adults with severe eczema, 20 (46.5%) female and 38 (88.4%) reported white ethnicity. Multiplying 0.5 by the standard deviation of POEM scores at baseline for this sample was 2.32.

Table 7-6 Summary of results for assessing POEM scores across datasets

	N	0.5 $SD_{(baseline)}$	Mean (SD)	Min. POEM score	Max. POEM score	Q_1, Q_3*
Total sample	1426	3.37	13.23 (6.74)	0	28	8, 18
Severity of eczema in the population						
Moderate – severe eczema (CLOTHES)	300	2.68	16.95 (5.36)	4	28	13, 20
Moderate – severe eczema (SWET)	336	2.95	16.87 (5.90)	0	28	12.5, 22
All severities of eczema (BATHE)	428	2.88	9.77 (5.76)	0	28	6, 13
All severities of eczema (COMET)	196	2.94	8.80 (5.87)	0	28	4, 12
Clinically infected eczema (CREAM)	112	2.69	14.99 (5.38)	6	28	11, 18.5
Age						
0-2	682	3.25	12.46 (6.50)	0	28	8, 17
3-7	544	3.40	13.41 (6.80)	0	28	8, 18
8-17	200	3.44	15.39 (6.87)	0	28	9.5, 20.5
Ethnicity						
White	1140	3.45	14.36 (6.89)	0	28	8, 18
Not white	260	3.34	13.05 (6.68)	0	28	9, 20
Gender						
Male	777	3.43	13.55 (6.85)	0	28	8, 19
female	649	3.30	12.85 (6.59)	0	28	8, 17

Note. Age, Ethnicity and Gender are the combined data from all five trials. *Q_1 = lower quartile, Q_3 = upper quartile

7.5 Discussion

The two studies presented in this chapter have contributed to assigning qualitative meaning to changes in POEM scores, hence improving the designing and interpreting of clinical trials using POEM as an outcome measure.

In study 1, the smallest detectable change on the POEM was suggested as 2.13 points. Therefore, only MIC estimates above 2.13 can be considered a change beyond measurement error. Using the CLOTHES trial dataset, the MIC estimates of study two ranged from 1.07 (using 0.2 SD of baseline POEM scores) to 6.13 (using an IGA as the anchor for the 'within-patient' score change method). The method used should be considered when interpreting published MIC values as it has clearly had an impact on estimates in this study. There is still debate over which method should be used, hence the pluralistic approach used in this study. There is a trade-off between convenient, hence widely used, distribution-based approaches and more theoretically appropriate anchor-based approaches.

Anchor-based approaches, unlike distribution-based approaches, include an explicit judgement of the importance of the change (de Vet et al., 2006a). Nevertheless, it has been suggested MIC estimates using 0.5 SD corresponded well with anchor-based methods (Turner et al., 2010). Within anchor-based approaches, the methods are evolving to become increasingly sophisticated, first with the development of the ROC curve method and more recently the development of the predictive modelling method.

The MIC estimates using the two anchors, the IGA and the PGA, were very close for some methods, but not others. For example, the MIC estimates using the two anchors were very close for the predictive modelling approach, but not for the ROC curve methods. It is not possible to provide firm conclusions as to why this is the case, but it could be because the ROC curve method is more sensitive to random sampling variation (Terluin et al., 2015).

The results of the distribution-based methods suggest that an effect size of 0.5 corresponds better to the anchor-based methods than using an effect size of 0.2 (Norman et al., 2003, Norman et al., 2004). As the MIC result from the 0.2 SD method was below the smallest detectable change and much smaller than the other estimates, it is suggested this result is an outlier and it is not recommended that this

method be used in future studies and the 0.5 SD method was chosen to be used in study 2. However, both distribution-based methods produced lower MIC estimates than anchor-based methods. Either distribution-based methods are underestimating the MIC, or it is quite possible that the anchors used here were too broad to capture smaller yet important changes.

In study 2, multiplying 0.5 by the standard deviation of POEM scores at baseline was remarkably consistent across children of varying ages, gender, ethnicity and disease severity from five trials. It has been cautioned that the SD used in this calculation needs to reflect the SD for the population of interest for the results to be used appropriately (Cook et al., 2014, Cook et al., 2015). This study should give reassurance that these participant characteristics do not appear to influence the SD.

The results presented in this chapter are broadly in line with previous findings. Gaunt et al. (2016). found results that followed a similar pattern to the studies presented in this chapter when they used multiple methods to calculate the MIC and reported a variation of estimates ranging from 2.5 to 4.27 points in a sample of children from the CREAM study, who have milder eczema than the sample in the study 1. Schram et al. (2012) used an anchor-based method (within-patient mean change) and reported an MIC of 3.4 in adults with severe eczema. Synthesising the findings from these previously published studies and the findings in this chapter resulted in recommendations of how to interpret changes on the POEM (Table 7-7).

Table 7-7 A guide to enhance interpretations of change on the POEM

Change in the POEM	Suggested interpretation
2 points or less	Unlikely to be a change beyond measurement error
2.1 to 2.9 points	A small change detected that is likely to be beyond measurement error but unlikely to be clinically important
3 to 3.9 points	Probably a clinically important change
4 points +	Very likely to be a clinically important change

There is a balance to be struck when estimating sample sizes for powering trials (Cook et al., 2015). Powering to detect a non-clinically important change is unethical and wasteful, as it will result in overly large trials. However, an underpowered trial based on detecting a large change in the POEM that provides inconclusive results and wide confidence intervals is also unacceptable. The recommendations provided in this chapter should remain as guidelines. Rather than relying on fixed values to interpret the importance of a change on the POEM, researchers and clinicians should consider the context within which they are using the POEM. A small improvement in many individuals could result in a large reduction in burden at a societal level.

7.5.1 Strengths and limitations

Study 1 estimated the MIC of the POEM using a broad array of methods but only in one dataset, which may limit the generalizability of the results beyond children in the U.K. with moderate-to-severe eczema. However, Gaunt et al. (2016) and Schram et al. (2012) used similar methods to calculate the MIC, which has allowed us to compare the results across these different populations. This study also included methods that have not previously been used to determine the MIC of the POEM (i.e. the predictive modelling approach).

The anchors used in study 1 were determined by what was available in the CLOTHES dataset, which may not best conceptualize a MIC. Typically, anchors used in the MIC literature have asked participants to retrospectively assess the amount of change that has occurred. This is what has been used in previous studies assessing the MIC of the POEM, for example, Gaunt et al. (2016) asked “How is your child’s eczema compared with one month ago?” with responses much better, better, no difference, worse, or much worse. The anchors used in this study were somewhat unconventional as they calculated a change score based on the difference in a global assessment of severity at baseline and 6 months later. The anchors used in this study did not ask about the importance of the change to the patients/parents or the investigator, however this has also been a criticism of anchors generally used to

calculate MIC estimates (Terwee et al., 2010). One other problem with the anchors used in the literature that ask about change retrospectively is that they may be subject to complicated issues with recall bias, which is one possible advantage of the anchor method used in this study (Kamper et al., 2009, Streiner et al., 2015).

There may be concern that the anchor-response categories used in this study may be too broad to capture the smallest amount of change that is important to patients, as MIC is equated to a change on severity banding in a global assessment, and it could be argued that smaller changes are clinically meaningful. The anchors ask about today, whereas POEM asks about the last week. If the anchors used in this study are not considered to be an adequate indicator of minimally important change, it could mean the MIC has been over or underestimated within this study.

By combining individual participant data from five trials, study 2 was able to explore the role of age, gender and ethnicity more comprehensively than if this had been explored separately within the individual datasets, which can inform sample size calculations in different populations. There are limits on how confidently the results from this study can be applied to wider populations beyond children in the UK. For the MAcAD trial (where participants were Dutch adults), multiplying 0.5 by the standard deviation of POEM scores at baseline was 2.32, which was slightly lower than that of other trials. Since this trial included only patients with very severe disease, this is likely to explain why the variability in POEM scores amongst the participants was limited. It is acknowledged that using anchor-based approaches within study 2 would have been advantageous, but it was not possible as there was not enough consistency in terms of variables collected and time points assessed across the different trials.

7.5.2 Future directions and implications

The studies presented in this chapter should improve the interpretability of change scores for users of the POEM, be this for clinical trial sample size calculations and

interpretation of findings or for use in clinical practice. Whilst it would have been beneficial to assess interpretability of change in POEM scores in datasets beyond children in the UK, and we did aim to include the MAcAD trial which included Dutch adults, but difficulties obtaining other datasets limited this possibility and this remains an area for further research.

The ICC of the POEM has only been calculated in one study, and this was used to assess the smallest detectable change (Charman et al., 2004). The sample age from the original study completing the test–retest reliability ranged from 12 months to 62 years, compared with the CLOTHES trial age range of 1–15 years. Further investigation of the test–retest agreement would be useful to increase the reliability of the ICC estimate. If the ICC of the POEM used here is found to be inaccurate in subsequent studies, the smallest detectable change of the POEM should be replicated. In the meantime, users of the POEM should be cautious in claiming any clinical importance of a change in POEM scores of two points or less.

7.6 Summary of Chapter 7

In conclusion, 0.5 multiplied by the SD at baseline of POEM scores does not appear to be influenced by age, gender, ethnicity or disease severity of the population, which provides reassurance for those designing or interpreting the results of eczema clinical trials that have used POEM as their primary outcome, that changes in POEM scores can be consistently interpreted across a variety of eczema populations and settings. However, given the wide spread of MIC estimates generated using a variety of methods, the method used to calculate the MIC should be given careful consideration. Interpretation of POEM scores has been improved by understanding that a change in score that is two points or under is consistent with potential measurement error, and it is suggested that changes in scores should be three points or more before the change is deemed to be clinically important.

Section D

Overall reflections and discussion

This final section aims to synthesise the body of work presented in this thesis.

The selection of studies presented in Section B and Section C all contributed to a broader goal of improving the quality of patient-reported outcome measures used in eczema clinical trials and to inform the development of the core outcome set for eczema clinical trials.

Chapter 8 will firstly summarise the findings of section B and section C, outline how they have contributed to the progress of the HOME initiative, and consider future research directions. Secondly, it will consider the broader landscape of the field by turning to some key challenges and offering some final reflections.

Chapter 8

Looking to the future for eczema outcome measures in clinical trials

8.1 Summary of thesis findings

Summary of Section B (Informing the HOME initiative domain of long-term control of eczema):

This thesis proposes that eczema control is a complex construct that is best described as a multifaceted experience that is caused by changes in disease activity, the treatment and management of the condition, and psychological, social and physical functioning. Eczema control is an individual experience and the level of control that is acceptable can vary among people. This thesis developed a patient-reported outcome measure to capture 'eczema control' called Recap of atopic eczema (RECAP). Initial testing suggests that it is appropriate to use in eczema clinical trials. It was designed to be a candidate instrument for the HOME initiative core outcome set.

Summary of Section C (Informing the HOME initiative domain of patient-reported symptoms):

This thesis has also presented studies assessing the measurement properties of the POEM. This thesis suggests POEM is best considered to have a formative measurement model and that it is therefore not appropriate to assess the structural validity or internal consistency of this outcome measure. Guidance on interpreting changes in POEM scores in clinical trials was also developed based on the studies in Chapter 7. Interpretation of POEM scores has been improved by understanding that a change in score that is two points or under is consistent with potential measurement error, and that changes in scores should be three points or more before the change is deemed to be clinically important.

8.2 Thesis contributions

Contributions to the long-term control of eczema domain:

Chapters 2-4 have informed progress regarding the core outcome set domain long-term control of eczema. Chapters 3 and 4 served to gather stakeholder perspectives of 'eczema control', which were then presented at the HOME V meeting to inform consensus discussions. The HOME initiative was able to reach consensus that long-term control for the purposes of the core outcome set should be defined as disease severity (signs and symptoms), quality of life, and a patient-global assessment (Chalmers et al., 2018).

The output of the HOME V meeting suggested that there may need to be a multi-item global measure of 'eczema control', which initiated the project presented in Chapter 5. The findings from Chapter 3 and 4 were used to inform the conceptual framework for the development of this new global measure of 'eczema control'.

The development (and subsequent validation studies) of RECAP was presented at the HOME VII meeting in Tokyo, Japan April 2019 (Figure 8-1 is a photograph taken from the meeting). The measurement properties of all potential instruments to measure eczema control were presented using the COSMIN checklist. When voting on which instruments could be included in the core outcome set, less than 30% disagreed to including the RECAP and less than 30% disagreed to including another instrument called the Atopic Dermatitis Control Test (ADCT). Discussion about how both newly-developed instruments had similar content and similar measurement properties, but little known about how they would perform in trial settings led to the suggestion that it was too early to choose between them to agree on a single core instrument. Therefore, it was agreed that ADCT and RECAP would be included as the core instrument(s) to assess long-term control (75% agreed, 23% disagreed, 2% unsure).

Figure 8-1 Attendees at the HOME VII meeting, Tokyo

Contributions to the patient-reported symptoms domain:

Chapters 6-7 have informed progress regarding the core outcome set domain patient-reported symptoms by evaluating the POEM, which has been recommended as the core outcome instrument. Establishing that the POEM has a formative measurement model has two main benefits. Firstly, it helps people designing and interpreting trials have a greater understanding of what is being measured. Secondly, it guides decisions about which psychometric tests are appropriate to conduct when assessing the measurement properties of the POEM. Improving the interpretability of POEM by providing clear recommendations on the smallest detectable change and minimally important change can enhance the design and interpretation of studies using the POEM, which could in turn increase uptake of the core outcome set recommendations.

8.3 Future directions for research

This section will make some suggestions for future research on the HOME core outcome domains that have been addressed in this thesis: long-term control of eczema and patient-reported symptoms.

Future directions for the long-term control of eczema domain:

Regarding the instrument RECAP, a study by Bhanot et al. (manuscript in preparation) has already been conducted to assess the measurement properties of RECAP in a community-based UK sample. The study suggested RECAP had no floor or ceiling effects, good construct validity, good test-retest reliability and good responsiveness in this population. However, further assessment relating to the reliability, validity, responsiveness and interpretability of the instrument is required to improve our understanding of the instrument and its appropriateness for use in different settings (such as eczema clinical trials and routine healthcare), as well as different populations. There needs to be efforts to translate RECAP for use in different cultures and languages. It may also be a useful avenue of research to develop modes of administration that could improve the feasibility and acceptability of answering the questionnaire (e.g. development of online and smartphone application modes) and assess if it can be adapted to increase the ability of younger children to provide information about their eczema control.

There is further work required to assess what instrument should be in the core outcome set in the future, since there are currently two instruments (RECAP and ADCT) recommended for the long-term control of eczema domain. Including two instruments is not in line with the COSMIN/COMET guidance, which specify one instrument per domain should be chosen, as this allows for maximum ability to synthesis trial results (Prinsen et al., 2016). Although the HOME roadmap does not rule out having more than one core instrument, it does suggest that one best instrument is the ideal scenario (Schmitt et al., 2015).

The ADCT was designed to capture patient-perceived eczema control in clinical and non-clinical settings (Pariser et al., under review). It was developed via three main phases including 1) a targeted literature review and interviews with clinical experts to develop preliminary items, 2) concept elicitation and cognitive interviewing with adults with eczema to refine items, 3) validation of the pilot tool against established patient-reported outcome instruments (Pariser et al., under review). The resulting instrument is a six-item instrument where each item asks about one of the following concepts: eczema-related symptoms, intense episodes of itching, how much bother caused by the eczema, trouble falling or staying asleep, affecting daily activities, and affecting mood or emotions (Pariser et al., under review). methodological approach to development and in the conceptualisation of eczema control, the RECAP and ADCT are very similar. One difference is that RECAP is designed for use in all trials including people of all ages including children and adults, whilst the ADCT has initially been designed for adults use only. The similarities between the two instruments contributed to the difficulties HOME VII delegates had in selecting one instrument instead of the other. Further work to assess the measurement properties of these instruments, but also compare these two instruments, will be required.

Furthermore, throughout the HOME VII meeting it was discussed how overlap with content from other core domains (patient-reported symptoms and quality of life) could be problematic in terms of patient burden and trial design. Therefore, future research is needed to establish if there is a suitable single item that exists or can be created that may be suitable for inclusion in the core set.

Future directions for the patient-reported symptoms domain:

Regarding the POEM, there remain some gaps in the assessment of measurement properties that could be investigated further. Given the inclusion of POEM in the core outcome set of the HOME initiative, it is particularly relevant that future efforts are directed to ensuring that the POEM is adapted appropriately for use in different cultures and languages. However, with this in mind, consideration to the

comparability of translated versions of POEM should be given due attention. This will require a substantial body of work, which is likely to involve the efforts of multiple research groups across the globe. However, it would benefit from a co-ordinated effort that standardises methodology and collates information in an accessible way.

8.4 Reflections

8.4.1 Patient-centred outcomes

'Patient-reported outcome measures' have become commonplace among medical and health research, with increasing recognition that patient input is required to ensure appropriate content (Patrick et al., 2011a, Patrick et al., 2011b). Both RECAP and POEM have been developed to include content that is important to people with eczema/caregivers and reflects outcomes that stakeholders want to be measured in trials. However, a standardised questionnaire may never be fully patient-centred as it does not preserve the form of an individual's response (Long and Dixon, 1996). Although patient-centeredness is considered important in the current paradigm of healthcare delivery and health research, the degree of patient-centeredness is often balanced against what is practical within the constraints of the research methodology.

As clinical trials require questionnaires take a standardised format to aid assessment of differences in outcomes between groups, they are by nature limiting the ability for the measure to capture the essence of the individual's experience. This tension was highlighted by participants in cognitive interviews in Chapter 5 expressing a desire to have a qualitative way of sharing their experience of eczema control alongside the standardised questions, and this tension was eloquently described by one participant when they said: "You are having to fit yourself into somebody else's box. The box isn't the shape of me, it's somebody else's shape and it's got to fit a lot of different people".

8.4.2 Stakeholder / patient & public involvement

Stakeholder involvement, including PPI, took multiple forms throughout this body of work, and was considered to be very effective and influential on the research agenda, design and delivery, recruitment, data analysis, writing up and dissemination. Details of specific PPI methods and outcomes can be found in sections 3.3.3, 4.3.4, and 5.3.4. Contextual factors played a large role in providing access to stakeholders/patients and creating a climate for PPI. The Centre of Evidence Based Dermatology (CEBD), where this PhD took place, has a strong and long-standing ethos of patient and public involvement in research, with patients at the centre of the research strategy (Figure 8-2). As well as involving patients in specific research projects, in 2009, the CEBD Patient Panel was established to create a more effective research environment and to be able to provide more training and support to patients involved in CEBD research projects (Layfield and Roberts, 2012).

Figure 8-2 Visual representation of the CEBD's research strategy

The content of this PhD has been designed to meet the needs of the HOME initiative, which is ultimately driven by the needs of the involved stakeholders. Stakeholders including patients, parents/caregivers, people working at patient organisations, methodologists, pharmaceutical industry representatives, and healthcare professionals have all been in attendance in tasks and meetings prioritising outcome

domains and measurement instruments for the core outcome set. Stakeholders are from multiple countries and anyone with an interest is eczema outcome measures is welcome to join HOME and take part in tasks or meetings. Funding is secured to help patients and parents/caregivers attend the meeting free of charge and training events and pre-meeting activities are organised specifically for patients to support them during the consensus meetings. The stakeholders involved in the HOME initiative are prioritising outcomes for inclusion in clinical trials, hence influencing the research agenda in this area.

PPI can take place on multiple levels. It is theorised that a more collaborative approach may have the greater potential to impact research outcomes (Staniszewska et al., 2012). With regards to 'patient-reported outcome measures' development, guidance currently suggests involvement of patients with regards to incorporating their perspective or experiences, however it has been proposed that this often remains at a consultative level (Staniszewska et al., 2012). Researchers often use patient experiences to analyse and synthesise information, but ultimately the researchers determine the shape of the final measure (Staniszewska et al., 2012). The experience from this body of work used multiple approaches, some of which aimed for a fully collaborative approach (such as the expert panel members co-designing RECAP in Chapter 5). However, there were some situations where it was not necessary or feasible to have this approach (for example, the input from multiple stakeholders in Chapter 4 would not have been possible to conduct without some element of the researchers analysing and synthesising the final results).

With regards to assessment of measurement properties of instruments, there has been very little evidence of PPI to date (Staniszewska et al., 2012). This is a challenging area to get involvement due to it being largely focused on statistical analyses which can be quite complex. For example, Chapter 6 and Chapter 7 did not include PPI. However, on reflection this was perhaps due to a lack of knowledge about the best way to approach PPI with regards to these studies and lack of

examples within current literature, but there could be a vital role PPI can play, particularly regarding the interpretation and dissemination of results.

8.4.3 The HOME initiative and core outcome set development

Core outcome sets have presented a promising solution for improving our ability to systematically collate and compare findings from clinical trials, which is perhaps demonstrated by the widespread development of core outcome sets across many fields of healthcare research (Gargon et al., 2014). Through informing the HOME initiative, this thesis can improve standardisation of outcome measures in clinical trials, which in turn may improve the evidence base for eczema treatments. If the core outcome measurement instruments are measured in all clinical trials, people trying to make sense of the results (whether that be to systematically review the data and conduct meta-analyses or make clinical decisions on the basis of results) will no longer be comparing apples and pears. Furthermore, if the community expects to see these outcomes reported in all trial results, they can confidently use the findings from these outcome measures without concerns about selection bias (Williamson et al., 2005, Clarke, 2007).

However, challenges remain for the HOME initiative and other core outcome sets being developed across different areas of health, which include 1) incorporating all stakeholder perspectives effectively, 2) limited resources, 3) implementation of the core outcome set, 4) the demands placed on the design and reporting of trials, and 5) responding to new developments in outcome measures research.

Firstly, it is important that core outcome sets are developed in a way that reflects measurement of outcomes that are important to key stakeholders. Core outcome set development presents an opportunity to guide researchers designing clinical trials to measure the outcomes that are agreed as most important across stakeholders. However, there are challenges that remain in ensuring that the perspective of all stakeholders is clearly represented in the core outcome set (Young

and Bagley, 2016). Attention to the process of patient involvement and engagement in the process needs to be given careful attention.

Secondly, core outcome sets require considerable resources, expertise from different professionals and coordination, often with little funding. As multiple core outcome sets are being developed in similar conditions (e.g. across different dermatological conditions) there is increasing need to consider how to streamline the process of core outcome set development and look for efficient ways to develop core outcome sets to ensure duplicate efforts are not unnecessarily committed.

Thirdly, the benefit of a core outcome set in facilitating evidence synthesis and reducing outcome selection bias relies on community uptake and implementation of the core outcome set. Research can be done to assess the uptake of core outcome sets. Attempts to look at this in arthritis trials suggests uptake of core outcome sets is occurring, but that there is still potential for further uptake (Kirkham et al., 2013, Kirkham et al., 2017). Research is ongoing to assess uptake of the HOME initiatives' recommendations. Dissemination of core outcome set recommendations to relevant stakeholders is essential to ensure the potential of core outcome sets to reduce research waste will be left untapped. There needs to be ongoing engagement with stakeholders and promotion of the core outcome set recommendations.

Fourthly, core outcome set developers need to be mindful of the demands they place on trial design and reporting to ensure it is acceptable and feasible for those conducting or taking part in a trial. Core outcome sets aim to improve evidence synthesis for meta-analysis and are not designed to influence trial design. Many core outcome sets developed have only recommended the outcome domains, whilst groups including the HOME initiative have endeavoured to agree on core measurement instruments to further aid evidence synthesis. However, the level of measurement can be even more specific regarding what participant level analysis metric to use (e.g. change in POEM score from baseline), the method of aggregation to summarise information within each group of patients in the trial (e.g. mean

change score from baseline) and the time point of interest should be specified (e.g. POEM scores at baseline and 6 months). The HOME initiative recently developed guidelines on reporting of the currently agreed core measurement instruments of EASI and POEM to aid transparency of information available for meta-analysis (Grinich et al., 2018). However, the guidelines purposefully didn't address analysis methods or timing of assessment as this was felt to be an issue for trial design that was beyond the remit of the core outcome set.

Finally, a core outcome set may need to respond to new developments in outcome measures research. This has been reflected in the HOME initiative's outlook that a core outcome set is "forever preliminary" (Schmitt and Williams, 2010). However, there remains a challenge for core outcome sets to incorporate emerging evidence into existing recommendations, whilst continuing to fulfil the purpose of improving evidence synthesis.

8.5 Concluding Remarks

This thesis has resulted in the development and initial testing of RECAP, an instrument to measure the experience of eczema control designed to be appropriate for use in eczema clinical trials as well as routine care. The RECAP is now a recommended core measurement instrument for the HOME domain long-term control of eczema. This thesis has also evaluated the POEM, a HOME recommended core measurement instrument for the domain patient-reported symptoms.

The focus on the development and testing of patient-reported outcome measures has contributed to efforts to improve and standardise outcome measures used in eczema clinical trials. I would like to leave the reader with the four proposals for how research in outcome measures can best contribute to improving the evidence base provided by clinical trials. I hope these notions have been evident throughout this thesis, and I believe they should be the hallmarks of future research in this field.

Firstly, there needs to be continued efforts to address the detrimental impact of the lack of consistency of outcome measures across eczema clinical trials. The development of core outcome sets is being embraced across health research areas and are initiatives that hold promise for overcoming this problem via harmonisation of outcome measures across clinical trials in a given condition.

Secondly, another important requirement for outcome research to benefit clinical trial data is to ensure that the data collected are able to answer questions that are important and relevant to key stakeholders. There are multiple stakeholders that may be relevant for ensuring outcome measures used in clinical trials are important and relevant. However, if one carries the belief that research should ultimately benefit the patient it seems appropriate to highlight patients, and where relevant their caregivers, as a key stakeholder group that should be involved, engaged and participating in outcome research to ensure that outcomes used in eczema clinical trials are important and relevant to patients.

Thirdly, to maximise the benefit of clinical trials, researchers should be able to understand the data they are collecting and to be able to convey the meaning of their output to others. To do so requires transparent approaches to outcome measure development and rigorous assessment of measurement properties. As well as being able to present this information to a wide audience.

Finally, the above three objectives will be best achieved with a collaborative approach to outcome research. Where possible, this collaboration should take place across stakeholders and across nations, which was aimed to be achieved within this thesis and within the HOME initiative more widely. Whilst this can present various monetary, practical and political challenges, it is a notion that if embraced, can produce results that would not be possible without it.

References

SurveyMonky Inc. [Online]. San Mateo, California, USA. Available: www.surveymonkey.com [Accessed].

ABRAHAMSSON, T. R., JAKOBSSON, T., BÖTTCHER, M. F., FREDRIKSON, M., JENMALM, M. C., BJÖRKSTÉN, B. & OLDAEUS, G. 2007. Probiotics in prevention of IgE-associated eczema: a double-blind, randomized, placebo-controlled trial. *Journal of Allergy and Clinical Immunology,* 119**,** 1174-1180.

ABUABARA, K., YE, M., MCCULLOCH, C. E., SULLIVAN, A., MARGOLIS, D. J., STRACHAN, D. P., PATERNOSTER, L., YEW, Y. W., WILLIAMS, H. C. & LANGAN, S. M. 2019. Clinical onset of atopic eczema: Results from 2 nationally representative British birth cohorts followed through midlife. *Journal of Allergy and Clinical Immunology,* 144**,** 710-719.

ABUABARA, K., YU, A. M., OKHOVAT, J. P., ALLEN, E. & LANGAN, S. M. 2017. The prevalence of atopic dermatitis beyond childhood: A systematic review and meta-analysis of longitudinal studies. *Allergy*.

ALTMAN, D. G. 1990. *Practical statistics for medical research*, CRC press.

APFELBACHER, C. J. & NELSON, P. A. 2017. Patient-reported outcome measures and qualitative research in dermatology: the quest for authenticity. *British Journal of Dermatology,* 176**,** 285-287.

ARMITAGE, P., BERRY, G. & MATTHEWS, J. N. S. 2008. *Statistical methods in medical research*, John Wiley & Sons.

BALLARDINI, N., KULL, I., LIND, T., HALLNER, E., ALMQVIST, C., OSTBLOM, E., MELEN, E., PERSHAGEN, G., LILJA, G., BERGSTROM, A. & WICKMAN, M. 2012. Development and comorbidity of eczema, asthma and rhinitis to age 12: data from the BAMSE birth cohort. *Allergy,* 67**,** 537-44.

BARBAROT, S., ROGERS, N. K., ABUABARA, K., AUBERT, H., CHALMERS, J., FLOHR, C., HANIFIN, J., NALDI, L., MARGOLIS, D. J., PAUL, C., RIDD, M. J., ANNA SCHUTTELAAR, M. L., SIMPSON, E., TAUBER, M., VOLKE, A., WEIDINGER, S., WILKES, S. R., WOLLENBERG, A. & THOMAS, K. S. 2016. Strategies used for measuring long-term control in atopic dermatitis trials: A systematic review. *J Am Acad Dermatol,* 75**,** 1038-1044.

BEATON, D. E. 2003. Simple as possible? Or too simple?: possible limits to the universality of the one half standard deviation. *Medical care,* 41**,** 593-596.

BEATON, D. E., BOERS, M. & WELLS, G. A. 2002. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol,* 14**,** 109-14.

BLOOR, M., FRANKLAND, J., THOMAS, M. & ROBSON, K. 2001. Focus groups in social research.

BOWLING, A. 2005. Just one question: If one question works, why ask several? : BMJ Publishing Group Ltd.

BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology,* 3**,** 77-101.

BRENNINKMEIJER, E., SCHRAM, M., LEEFLANG, M., BOS, J. & SPULS, P. I. 2008. Diagnostic criteria for atopic dermatitis: a systematic review. *British Journal of Dermatology,* 158**,** 754-765.

BRENNINKMEIJER, E. E., LEGIERSE, C. M., SILLEVIS SMITT, J. H., LAST, B. F., GROOTENHUIS, M. A. & BOS, J. D. 2009. The course of life of patients with childhood atopic dermatitis. *Pediatric dermatology,* 26**,** 14-22.

BROD, M., TESLER, L. E. & CHRISTENSEN, T. L. 2009. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research,* 18**,** 1263.

CHALMERS, I. & GLASZIOU, P. 2009. Avoidable waste in the production and reporting of research evidence. *The Lancet,* 374**,** 86-89.

CHALMERS, J. R., SCHMITT, J., APFELBACHER, C., DOHIL, M., EICHENFIELD, L. F., SIMPSON, E. L., SINGH, J., SPULS, P., THOMAS, K. S., ADMANI, S., AOKI, V., ARDELEANU, M., BARBAROT, S., BERGER, T., BERGMAN, J. N., BLOCK, J., BOROK, N., BURTON, T., CHAMLIN, S. L., DECKERT, S., DEKLOTZ, C. C., GRAFF, L. B., HANIFIN, J. M., HEBERT, A. A., HUMPHREYS, R., KATOH, N., KISA, R. M., MARGOLIS, D. J., MERHAND, S., MINNILLO, R., MIZUTANI, H., NANKERVIS, H., OHYA, Y., RODGERS, P., SCHRAM, M. E., STALDER, J. F., SVENSSON, A., TAKAOKA, R., TEPER, A., TOM, W. L., VON KOBYLETZKI, L., WEISSHAAR, E., ZELT, S. & WILLIAMS, H. C. 2014. Report from the third international consensus meeting to harmonise core outcome measures for atopic eczema/dermatitis clinical trials (HOME). *Br J Dermatol,* 171**,** 1318-25.

CHALMERS, J. R., SIMPSON, E., APFELBACHER, C. J., THOMAS, K. S., VON KOBYLETZKI, L., SCHMITT, J., SINGH, J. A., SVENSSON, A., WILLIAMS, H. C., ABUABARA, K., AOKI, V., ARDELEANU, M., AWICI-RASMUSSEN, M., BARBAROT, S., BERENTS, T. L., BLOCK, J., BRAGG, A., BURTON, T., BJERRING CLEMMENSEN, K. K., CRESWELL-MELVILLE, A., DINESEN, M., DRUCKER, A., ECKERT, L., FLOHR, C., GARG, M., GERBENS, L. A., GRAFF, A. L., HANIFIN, J., HEINL, D., HUMPHREYS,

R., ISHII, H. A., KATAOKA, Y., LESHEM, Y. A., MARQUORT, B., MASSUEL, M. A., MERHAND, S., MIZUTANI, H., MUROTA, H., MURRELL, D. F., NAKAHARA, T., NASR, I., NOGRALES, K., OHYA, Y., OSTERLOH, I., PANDER, J., PRINSEN, C., PURKINS, L., RIDD, M., SACH, T., SCHUTTELAAR, M. A., SHINDO, S., SMIRNOVA, J., SULZER, A., SYNNOVE GJERDE, E., TAKAOKA, R., VESTBY TALMO, H., TAUBER, M., TORCHET, F., VOLKE, A., WAHLGREN, C. F., WEIDINGER, S., WEISSHAAR, E., WOLLENBERG, A., YAMAGA, K., ZHAO, C. Y. & SPULS, P. I. 2016. Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol*.

CHALMERS, J. R., THOMAS, K. S., APFELBACHER, C., WILLIAMS, H. C., PRINSEN, C. A., SPULS, P. I., SIMPSON, E., GERBENS, L. A. A., BOERS, M., BARBAROT, S., STALDER, J. F., ABUABARA, K., AOKI, V., ARDELEANU, M., ARMSTRONG, J., BANG, B., BERENTS, T. L., BURTON, T., BUTLER, L., CHUBACHI, T., CRESSWELL-MELVILLE, A., DELOZIER, A., ECKERT, L., EICHENFIELD, L., FLOHR, C., FUTAMURA, M., GADKARI, A., GJERDE, E. S., VAN HALEWIJN, K. F., HAWKES, C., HOWELLS, L., HOWIE, L., HUMPHREYS, R., ISHII, H. A., KATAOKA, Y., KATAYAMA, I., KOUWENHOVEN, W., LANGAN, S. M., LESHEM, Y. A., MERHAND, S., MINA-OSORIO, P., MUROTA, H., NAKAHARA, T., NUNES, F. P., NYGAARD, U., NYGARDAS, M., OHYA, Y., ONO, E., REHBINDER, E., ROGERS, N. K., ROMEIJN, G. L. E., SCHUTTELAAR, M. L. A., SEARS, A. V., SIMPSON, M. A., SINGH, J. A., SROUR, J., STUART, B., SVENSSON, A., TALMO, G., TALMO, H., TEIXEIRA, H. D., THYSSEN, J. P., TODD, G., TORCHET, F., VOLKE, A., VON KOBYLETZKI, L., WEISSHAAR, E., WOLLENBERG, A. & ZANIBONI, M. 2018. Report from the fifth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol*.

CHAMLIN, S. L., FRIEDEN, I. J., WILLIAMS, M. L. & CHREN, M.-M. 2004. Effects of atopic dermatitis on young American children and their families. *Pediatrics,* 114**,** 607-611.

CHARMAN, C., CHAMBERS, C. & WILLIAMS, H. 2003. Measuring Atopic Dermatitis Severity in Randomized Controlled Clinical Trials: What Exactly Are We Measuring? *Journal of Investigative Dermatology,* 120**,** 932-941.

CHARMAN, C., VENN, A., RAVENSCROFT, J. & WILLIAMS, H. 2013. Translating Patient-Oriented Eczema Measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. *British Journal of Dermatology,* 169**,** 1326-1332.

CHARMAN, C. R., VENN, A. J. & WILLIAMS, H. C. 2004. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol,* 140**,** 1513-9.

CLARKE, M. 2007. Standardising outcomes for clinical trials and systematic reviews. *Trials,* 8**,** 39.

COE, R. 2002. It's the effect size, stupid: What effect size is and why it is important.

COLLINS, D. 2003. Pretesting survey instruments: an overview of cognitive methods. *Quality of life research,* 12**,** 229-238.

COLTMAN, T., DEVINNEY, T. M., MIDGLEY, D. F. & VENAIK, S. 2008. Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research,* 61**,** 1250-1262.

COOK, C. E. 2008. Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. *J Man Manip Ther,* 16**,** E82-3.

COOK, J. A., HISLOP, J., ADEWUYI, T. E., HARRILD, K., ALTMAN, D. G., RAMSAY, C. R., FRASER, C., BUCKLEY, B., FAYERS, P., HARVEY, I., BRIGGS, A., NORRIE, J. D., FERGUSSON, D., FORD, I. & VALE, L. D. 2014. Assessing methods to specify the targeted difference for a randomised controlled trial – DELTA (Difference ELicitation in TriAls) review. *Health Technol Assess,* 18.

COOK, J. A., HISLOP, J., ALTMAN, D. G., FAYERS, P., BRIGGS, A. H., RAMSAY, C. R., NORRIE, J. D., HARVEY, I. M., BUCKLEY, B., FERGUSSON, D., FORD, I. & VALE, L. D. 2015. Specifying the target difference in the primary outcome for a randomised controlled trial: guidance for researchers. *Trials,* 16**,** 12.

COPAY, A. G., SUBACH, B. R., GLASSMAN, S. D., POLLY, D. W., JR. & SCHULER, T. C. 2007. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J,* 7**,** 541-6.

COUTANCEAU, C. & STALDER, J.-F. 2014. Analysis of correlations between patient-oriented SCORAD (PO-SCORAD) and other assessment scores of atopic dermatitis severity and quality of life. *Dermatology,* 229**,** 248-255.

CROSBY, R. D., KOLOTKIN, R. L. & WILLIAMS, G. R. 2003. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol,* 56**,** 395-407.

DE BOER, A. G., VAN LANSCHOT, J. J., STALMEIER, P. F., VAN SANDICK, J. W., HULSCHER, J. B., DE HAES, J. C. & SPRANGERS, M. A. 2004. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research,* 13**,** 311-320.

DE VET, H. C., TERWEE, C. B., MOKKINK, L. B. & KNOL, D. L. 2011. *Measurement in medicine: a practical guide*, Cambridge University Press.

DE VET, H. C., TERWEE, C. B., OSTELO, R. W., BECKERMAN, H., KNOL, D. L. & BOUTER, L. M. 2006a. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and Quality of Life Outcomes,* 4**,** 1-5.

DE VET, H. C. W., OSTELO, R. W. J. G., TERWEE, C. B., VAN DER ROER, N., KNOL, D. L., BECKERMAN, H., BOERS, M. & BOUTER, L. M. 2006b. Minimally important change determined by a visual method integratingan anchor-based and a distribution-based approach. *Qual Life Res,* 16**,** 131-142.

DECKERS, I. A., MCLEAN, S., LINSSEN, S., MOMMERS, M., VAN SCHAYCK, C. & SHEIKH, A. 2012. Investigating international time trends in the incidence and prevalence of atopic eczema 1990–2010: a systematic review of epidemiological studies. *PloS one,* 7**,** e39803.

DIEFENBACH, M. A. & LEVENTHAL, H. 1996. The common-sense model of illness representation: Theoretical and practical considerations. *Journal of social distress and the homeless,* 5**,** 11-38.

DRUCKER, A. M., WANG, A. R., LI, W.-Q., SEVETSON, E., BLOCK, J. K. & QURESHI, A. A. 2017. The Burden of Atopic Dermatitis: Summary of a Report for the National Eczema Association. *Journal of Investigative Dermatology,* 137**,** 26-30.

EDWARDS, J. R. 2011. The fallacy of formative measurement. *Organizational Research Methods,* 14**,** 370-388.

EDWARDS, J. R. & BAGOZZI, R. P. 2000. On the nature and direction of relationships between constructs and measures. *Psychological methods,* 5**,** 155.

ELLIS, C. N., MANCINI, A. J., PALLER, A. S., SIMPSON, E. L. & EICHENFIELD, L. F. 2012. Understanding and managing atopic dermatitis in adult patients. *Semin Cutan Med Surg,* 31**,** S18-22.

ELMOSE, C. & THOMSEN, S. F. 2015. Twin Studies of Atopic Dermatitis: Interpretations and Applications in the Filaggrin Era. *J Allergy (Cairo),* 2015**,** 902359.

EMERSON, R., WILLIAMS, H. & ALLEN, B. 1998. Severity distribution of atopic dermatitis in the community and its relationship to secondary referral. *British Journal of Dermatology,* 139**,** 73-76.

EMERSON, R. M., CHARMAN, C. R. & WILLIAMS, H. C. 2000. The Nottingham Eczema Severity Score: preliminary refinement of the Rajka and Langeland grading. *Br J Dermatol,* 142**,** 288-97.

EMERSON, R. M., WILLIAMS, H. C. & ALLEN, B. R. 2001. What is the cost of atopic dermatitis in preschool children? *Br J Dermatol,* 144**,** 514-22.

EYERICH, K., EYERICH, S. & BIEDERMANN, T. 2015. The multi-modal immune pathogenesis of atopic eczema. *Trends in immunology,* 36**,** 788-801.

FAYERS, P. M. & MACHIN, D. 2013. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*, John Wiley & Sons.

FEINSTEIN, A. R. 1987. Clinimetric perspectives. *Journal of chronic diseases,* 40**,** 635-640.

FISHBEIN, A. B., VITATERNA, O., HAUGH, I. M., BAVISHI, A. A., ZEE, P. C., TUREK, F. W., SHELDON, S. H., SILVERBERG, J. I. & PALLER, A. S. 2015. Nocturnal eczema: Review of sleep and circadian rhythms in children with atopic dermatitis and future research directions. *J Allergy Clin Immunol,* 136**,** 1170-7.

FITZPATRICK, R., DAVEY, C., BUXTON, M. J. & JONES, D. R. 1998. Evaluating patient-based outcome measures for use in clinical trials.

FLOHR, C. 2011. Atopic dermatitis diagnostic criteria and outcome measures for clinical trials: still a mess. *J Invest Dermatol,* 131**,** 557-9.

FLOHR, C., JOHANSSON, S. G., WAHLGREN, C. F. & WILLIAMS, H. 2004. How atopic is atopic dermatitis? *J Allergy Clin Immunol,* 114**,** 150-8.

FLOHR, C., PASCOE, D. & WILLIAMS, H. C. 2005. Atopic dermatitis and the 'hygiene hypothesis': too clean to be true? *Br J Dermatol,* 152**,** 202-16.

FOOD AND DRUG ADMINISTRATION 2009. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims *In:* SERVICES, U. S. D. O. H. A. H. (ed.).

FOX, F. E., MORRIS, M. & RUMSEY, N. 2007. Doing synchronous online focus groups with young people: methodological reflections. *Qual Health Res,* 17**,** 539-47.

FRANCIS, N. A., RIDD, M. J., THOMAS-JONES, E., SHEPHERD, V., BUTLER, C. C., HOOD, K., HUANG, C., ADDISON, K., LONGO, M., MARWICK, C., WOOTTON, M., HOWE, R., ROBERTS, A., HAQ, M. I., MADHOK, V. & SULLIVAN, F. 2016. A randomised placebo-controlled trial of oral and topical antibiotics for children with clinically infected eczema in the community: the ChildRen with Eczema, Antibiotic Management (CREAM) study. *Health Technol Assess,* 20**,** 1-84.

FUSCH, P. I. & NESS, L. R. 2015. Are we there yet? Data saturation in qualitative research. *The Qualitative Report,* 20**,** 1408.

FUTAMURA, M., LESHEM, Y. A., THOMAS, K. S., NANKERVIS, H., WILLIAMS, H. C. & SIMPSON, E. L. 2016. A systematic review of Investigator Global Assessment

(IGA) in atopic dermatitis (AD) trials: Many options, no standards. *J Am Acad Dermatol,* 74**,** 288-94.

GAISER, T. J. 1997. Conducting On-Line Focus Groups A Methodological Discussion. *Social Science Computer Review,* 15**,** 135-144.

GARGON, E., GURUNG, B., MEDLEY, N., ALTMAN, D. G., BLAZEBY, J. M., CLARKE, M. & WILLIAMSON, P. R. 2014. Choosing important health outcomes for comparative effectiveness research: a systematic review. *PLoS One,* 9**,** e99111.

GAUNT, D. M., METCALFE, C. & RIDD, M. 2016. The Patient-Oriented Eczema Measure in young children: responsiveness and minimal clinically important difference. *Allergy,* 71**,** 1620-1625.

GERBENS, L. A., CHALMERS, J. R., ROGERS, N. K., NANKERVIS, H. & SPULS, P. I. 2016. Reporting of symptoms in randomised controlled trials of atopic eczema treatments: a systematic review. *Br J Dermatol*.

GERBENS, L. A., PRINSEN, C. A., CHALMERS, J. R., DRUCKER, A. M., VON KOBYLETZKI, L. B., LIMPENS, J., NANKERVIS, H., SVENSSON, A., TERWEE, C. B., ZHANG, J., APFELBACHER, C. J. & SPULS, P. I. 2017. Evaluation of the measurement properties of symptom measurement instruments for atopic eczema: a systematic review. *Allergy,* 72**,** 146-163.

GLASS, G. V. 1976. Primary, secondary, and meta-analysis of research. *Educational researcher,* 5**,** 3-8.

GORE, C., JOHNSON, R., CARESS, A., WOODCOCK, A. & CUSTOVIC, A. 2005. The information needs and preferred roles in treatment decision-making of parents caring for infants with atopic dermatitis: a qualitative study. *Allergy,* 60**,** 938-943.

GREEN, S. & HIGGINS, J. 2005. Cochrane handbook for systematic reviews of interventions. Version.

GRINICH, E., SCHMITT, J., KÜSTER, D., SPULS, P., WILLIAMS, H., CHALMERS, J., THOMAS, K., APFELBACHER, C., PRINSEN, C. & FURUE, M. 2018. Standardized Reporting of the Eczema Area and Severity Index (EASI) and the Patient-Oriented Eczema Measure (POEM): A Recommendation by the Harmonising Outcome Measures for Eczema (HOME) Initiative. *British Journal of Dermatology*.

HAGGER, M. S. & ORBELL, S. 2003. A meta-analytic review of the common-sense model of illness representations. *Psychology and health,* 18**,** 141-184.

HANIFIN, J. & RAJKA, G. 1980. Diagnostic features of atopic eczema. *Acta Dermatol Venereol (Stockh),* 92**,** 44-7.

HARGREAVES, D. S., SIZMUR, S., PITCHFORTH, J., TALLETT, A., TOOMEY, S. L., HOPWOOD, B., SCHUSTER, M. A. & VINER, R. M. 2018. Children and young people's versus parents' responses in an English national inpatient survey. *Archives of disease in childhood,* 103**,** 486-491.

HARRELL, F. E. 2015. *Regression modelling strategies. With applications to linear models, logistic regression, and survival analysis.,* New York, Heidelberg: Springer.

HEINL, D., CHALMERS, J., NANKERVIS, H. & APFELBACHER, C. J. 2016. Eczema Trials: Quality of Life Instruments Used and Their Relation to Patient-reported Outcomes. A Systematic Review. *Acta Derm Venereol,* 96**,** 596-601.

HEINZE, G., WALLISCH, C. & DUNKLER, D. 2018. Variable selection–A review and recommendations for the practicing statistician. *Biometrical Journal,* 60**,** 431-449.

HERD, R. M., TIDMAN, M. J., PRESCOTT, R. J. & HUNTER, J. A. 1996a. The cost of atopic eczema. *Br J Dermatol,* 135**,** 20-3.

HERD, R. M., TIDMAN, M. J., PRESCOTT, R. J. & HUNTER, J. A. 1996b. Prevalence of atopic eczema in the community: the Lothian Atopic Dermatitis study. *Br J Dermatol,* 135**,** 18-9.

HORNE, R., WEINMAN, J. & HANKINS, M. 1999. The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychology and health,* 14**,** 1-24.

HOWELLS, L., CHALMERS, J. R., COWDELL, F., RATIB, S., SANTER, M. & THOMAS, K. S. 2017. "When it goes back to my normal I suppose": a qualitative study using online focus groups to explore perceptions of 'control'amongst people with eczema and parents of children with eczema in the UK. *BMJ Open*.

HOWELLS, L., RATIB, S., CHALMERS, J. R., BRADSHAW, L. E., THOMAS, K. S. & TEAM, O. B. O. T. C. T. 2018a. How should minimally important change (MIC) scores for the Patient Oriented Eczema Measure (POEM) be interpreted? A validation using varied methods *British Journal of Dermatology,* Accepted.

HOWELLS, L., THOMAS, K. S., SEARS, A. V., NASR, I., WOLLENBERG, A., SCHUTTELAAR, M. L., ROMEIJN, G. L., PALLER, A. S., MUELLER, K. & DOYTCHEVA, K. 2018b. Defining and measuring "eczema control": An international qualitative study

to explore the views of those living with and treating atopic eczema. *Journal of the European Academy of Dermatology and Venereology*.

IRWIN, D. E., VARNI, J. W., YEATTS, K. & DEWALT, D. A. 2009. Cognitive interviewing methodology in the development of a pediatric item bank: a patient reported outcomes measurement information system (PROMIS) study. *Health and Quality of Life Outcomes,* 7**,** 3.

JOBE, J. B. & HERRMANN, D. J. 1996. Implications of models of survey cognition for memory theory. *Basic and applied memory research,* 2**,** 193-205.

JOHANSSON, S. G., BIEBER, T., DAHL, R., FRIEDMANN, P. S., LANIER, B. Q., LOCKEY, R. F., MOTALA, C., ORTEGA MARTELL, J. A., PLATTS-MILLS, T. A., RING, J., THIEN, F., VAN CAUWENBERGE, P. & WILLIAMS, H. C. 2004. Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *J Allergy Clin Immunol,* 113**,** 832-6.

KAMPER, S. J., MAHER, C. G. & MACKAY, G. 2009. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy,* 17**,** 163-170.

KANTOR, R., THYSSEN, J. P., PALLER, A. S. & SILVERBERG, J. I. 2016. Atopic dermatitis, atopic eczema, or eczema? A systematic review, meta-analysis, and recommendation for uniform use of 'atopic dermatitis'. *Allergy*.

KEELEY, T., WILLIAMSON, P., CALLERY, P., JONES, L. L., MATHERS, J., JONES, J., YOUNG, B. & CALVERT, M. 2016. The use of qualitative methods to inform Delphi surveys in core outcome set development. *Trials,* 17**,** 230.

KIRKHAM, J. J., BOERS, M., TUGWELL, P., CLARKE, M. & WILLIAMSON, P. R. 2013. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials,* 14**,** 324.

KIRKHAM, J. J., CLARKE, M. & WILLIAMSON, P. R. 2017. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials. gov: findings from a review of randomised controlled trials of rheumatoid arthritis. *bmj,* 357**,** j2262.

KIRWAN, J. R., MINNOCK, P., ADEBAJO, A., BRESNIHAN, B., CHOY, E., DE WIT, M., HAZES, M., RICHARDS, P., SAAG, K. & SUAREZ-ALMAZOR, M. 2007. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *The Journal of Rheumatology,* 34**,** 1174-1177.

KROSNICK, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology,* 5**,** 213-236.

KROSNICK, J. A. & PRESSER, S. 2010. Questionnaire Design. *In:* MARSDEN, P. V. & WRIGHT, J. D. (eds.) *Handbook of Survey Research.* West Yorkshire, UK: Emerald Group.

KUHN, T. 1964. A function for thought experiments.

LANGAN, S. M., BOURKE, J. F., SILCOCKS, P. & WILLIAMS, H. C. 2006a. An exploratory prospective observational study of environmental factors exacerbating atopic eczema in children. *Br J Dermatol,* 154**,** 979-80.

LANGAN, S. M., SCHMITT, J., WILLIAMS, H. C., SMITH, S. & THOMAS, K. S. 2014. How are eczema 'flares' defined? A systematic review and recommendation for future studies. *Br J Dermatol,* 170**,** 548-56.

LANGAN, S. M., SILCOCKS, P. & WILLIAMS, H. C. 2009. What causes flares of eczema in children? *Br J Dermatol,* 161**,** 640-6.

LANGAN, S. M., STUART, B., BRADSHAW, L., SCHMITT, J., WILLIAMS, H. C. & THOMAS, K. S. 2017. Measuring long-term disease control in atopic dermatitis: a validation study of well controlled weeks. *Journal of Allergy and Clinical Immunology*.

LANGAN, S. M., THOMAS, K. S. & WILLIAMS, H. C. 2006b. What is meant by a "flare" in atopic dermatitis? A systematic review and proposal. *Arch Dermatol,* 142**,** 1190-6.

LANGAN, S. M. & WILLIAMS, H. C. 2006. What causes worsening of eczema? A systematic review. *Br J Dermatol,* 155**,** 504-14.

LAWSON, V., LEWIS-JONES, M. S., FINLAY, A. Y., REID, P. & OWENS, R. G. 1998. The family impact of childhood atopic dermatitis: the Dermatitis Family Impact Questionnaire. *Br J Dermatol,* 138**,** 107-13.

LAYFIELD, C. & ROBERTS, A. 2012. The Centre of Evidence Based Dermatology (CEBD) Patient Panel: Wider benefits for patient health? *INVOLVE Conference 2012.* East Midlands Conference Centre, Nottingham.

LEWIS-JONES, S. 2006. Quality of life and childhood atopic dermatitis: the misery of living with childhood eczema. *Int J Clin Pract,* 60**,** 984-92.

LEWIS-JONES, M. 2006. A comparative study of impairment of quality of life in children with skin disease and children with other chronic childhood diseases. *British Journal of Dermatology,* 155**,** 145-151.

LONG, A. F. & DIXON, P. 1996. Monitoring outcomes in routine practice: defining appropriate measurement criteria. *Journal of Evaluation in Clinical Practice,* 2**,** 71-78.

MANZONI, A. P., PEREIRA, R. L., TOWNSEND, R. Z., WEBER, M. B., NAGATOMI, A. R. & CESTARI, T. F. 2012. Assessment of the quality of life of pediatric patients with the major chronic childhood skin diseases. *An Bras Dermatol,* 87**,** 361-8.

MARGOLIS, J. S., ABUABARA, K., BILKER, W., HOFFSTAD, O. & MARGOLIS, D. J. 2014. Persistence of mild to moderate atopic dermatitis. *JAMA dermatology,* 150**,** 593-600.

MCDOWELL, I. 2006. *Measuring health: a guide to rating scales and questionnaires*, Oxford University Press, USA.

MEYER, D., LEVENTHAL, H. & GUTMANN, M. 1985. Common-sense models of illness: the example of hypertension. *Health psychology,* 4**,** 115.

MOKKINK, L., DE VET, H., PRINSEN, C., PATRICK, D., ALONSO, J., BOUTER, L. & TERWEE, C. 2017. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research***,** 1-9.

MOKKINK, L. B., TERWEE, C. B., KNOL, D. L., STRATFORD, P. W., ALONSO, J., PATRICK, D. L., BOUTER, L. M. & DE VET, H. C. 2010a. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC medical research methodology,* 10**,** 22.

MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L., ALONSO, J., STRATFORD, P. W., KNOL, D. L., BOUTER, L. M. & DE VET, H. C. 2009. COSMIN checklist manual. *Amsterdam: VU University Medical Centre*.

MOKKINK, L. B., TERWEE, C. B., PATRICK, D. L., ALONSO, J., STRATFORD, P. W., KNOL, D. L., BOUTER, L. M. & DE VET, H. C. 2010b. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology,* 63**,** 737-745.

MOORE, K., DAVID, T., MURRAY, C., CHILD, F. & ARKWRIGHT, P. 2006. Effect of childhood eczema and asthma on parental sleep and well-being: a prospective comparative study. *British Journal of Dermatology,* 154**,** 514-518.

MOORE, T., MCKEE, K. & PAULINE, M. 2015. Online focus groups and qualitative research in the social sciences: their merits and limitations in a study of housing and youth. *People Place and Policy Online*.

NANKERVIS, H., BAIBERGENOVA, A., WILLIAMS, H. C. & THOMAS, K. S. 2012. Prospective registration and outcome-reporting bias in randomized controlled trials of eczema treatments: a systematic review. *Journal of Investigative Dermatology,* 132**,** 2727-2734.

NANKERVIS, H., THOMAS, K. S., DELAMERE, F. M., BARBAROT, S., ROGERS, N. K. & WILLIAMS, H. C. 2016. Programme Grants for Applied Research. *Scoping systematic review of treatments for eczema.* Southampton (UK): NIHR Journals Library.

NATHAN, R. A., SORKNESS, C. A., KOSINSKI, M., SCHATZ, M., LI, J. T., MARCUS, P., MURRAY, J. J. & PENDERGRAFT, T. B. 2004. Development of the asthma control test: a survey for assessing asthma control. *Journal of Allergy and Clinical Immunology,* 113**,** 59-65.

NEERGAARD, M. A., OLESEN, F., ANDERSEN, R. S. & SONDERGAARD, J. 2009. Qualitative description–the poor cousin of health research? *BMC medical research methodology,* 9**,** 52.

NICE 2007. National Institute for Health and Care Excellence. Atopic eczema in under 12s: diagnosis and management. National Institute for Health and Care Excellence

NOERRESLET, M., JEMEC, G. B. E. & TRAULSEN, J. M. 2009. Involuntary autonomy: patients' perceptions of physicians, conventional medicines and risks in the management of atopic dermatitis. *Soc Sci Med,* 69.

NORMAN, G. R., SLOAN, J. A. & WYRWICH, K. W. 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical care,* 41**,** 582-592.

NORMAN, G. R., SLOAN, J. A. & WYRWICH, K. W. 2004. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res.,* 4**,** 581-585.

NORQUIST, J. M., GIRMAN, C., FEHNEL, S., DEMURO-MERCON, C. & SANTANELLO, N. 2012. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Quality of Life Research,* 21**,** 1013-1020.

NUTTEN, S. 2015. Atopic Dermatitis: Global Epidemiology and Risk Factors. *Annals of Nutrition and Metabolism,* 66(suppl 1)**,** 8-16.

PARISER, D. M., SIMPSON, E., GADKARI, A., BIEBER, T., MARGOLIS, D. J., BROWN, M., NELSON, L., MAHAJAN, P., REANEY, M., GUILLEMIN, I., MALLYA, U. G. & ECKERT, L. under review. Design, Validation and Scoring of the Atopic Dermatitis Control Tool (ADCT).

PATRICK, D. L., BURKE, L. B., GWALTNEY, C. J., LEIDY, N. K., MARTIN, M. L., MOLSEN, E. & RING, L. 2011a. Content Validity—Establishing and Reporting the Evidence in Newly Developed Patient-Reported Outcomes (PRO) Instruments

for Medical Product Evaluation: ISPOR PRO Good Research Practices Task Force Report: Part 1—Eliciting Concepts for a New PRO Instrument. *Value in Health,* 14**,** 967-977.

PATRICK, D. L., BURKE, L. B., GWALTNEY, C. J., LEIDY, N. K., MARTIN, M. L., MOLSEN, E. & RING, L. 2011b. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. *Value in Health,* 14**,** 978-988.

PATTON, M. Q. 1990. *Qualitative evaluation and research methods*, SAGE Publications, inc.

PETERSON, C. H., GISCHLAR, K. L. & PETERSON, N. A. 2017. Item construction using reflective, formative, or Rasch measurement models: Implications for group work. *The Journal for Specialists in Group Work,* 42**,** 17-32.

POLS, D. H., WARTNA, J. B., MOED, H., VAN ALPHEN, E. I., BOHNEN, A. M. & BINDELS, P. J. 2016. Atopic dermatitis, asthma and allergic rhinitis in general practice and the open population: a systematic review. *Scand J Prim Health Care,* 34**,** 143-50.

PRINSEN, C. A., MOKKINK, L. B., BOUTER, L. M., ALONSO, J., PATRICK, D. L., DE VET, H. C. & TERWEE, C. B. 2018. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research,* 27**,** 1147-1157.

PRINSEN, C. A., VOHRA, S., ROSE, M. R., BOERS, M., TUGWELL, P., CLARKE, M., WILLIAMSON, P. R. & TERWEE, C. B. 2016. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials,* 17**,** 449.

REVICKI, D., HAYS, R. D., CELLA, D. & SLOAN, J. 2008. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol,* 61**,** 102-9.

RIDD, M. J., REDMOND, N. M., HOLLINGHURST, S., BALL, N., SHAW, L., GUY, R., WILSON, V., METCALFE, C. & PURDY, S. 2015. Choice of Moisturiser for Eczema Treatment (COMET): study protocol for a randomized controlled trial. *Trials,* 16**,** 304.

RITCHIE, J., LEWIS, J., NICHOLLS, C. M. & ORMSTON, R. 2013. *Qualitative research practice: A guide for social science students and researchers*, sage.

RITCHIE, J. & SPENCER, L. 2002. Qualitative data analysis for applied policy research. *The qualitative researcher's companion,* 573**,** 305-329.

RODRIGUEZ, E., BAURECHT, H., HERBERICH, E., WAGENPFEIL, S., BROWN, S. J., CORDELL, H. J., IRVINE, A. D. & WEIDINGER, S. 2009. Meta-analysis of filaggrin polymorphisms in eczema and asthma: robust risk factors in atopic disease. *J Allergy Clin Immunol,* 123**,** 1361-70.e7.

ROYSTON, P. & SAUERBREI, W. 2008. *Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables*, John Wiley & Sons.

SACH, T. H., MCMANUS, E., MCMONAGLE, C. & LEVELL, N. 2016. Economic evidence for the prevention and treatment of atopic eczema: a protocol for a systematic review. *Syst Rev,* 5**,** 90.

SANDELOWSKI, M. 2000. Whatever happened to qualitative description? *Research in nursing & health,* 23**,** 334-340.

SANDELOWSKI, M. 2010. What's in a name? Qualitative description revisited. *Research in nursing & health,* 33**,** 77-84.

SANTER, M., BURGESS, H., YARDLEY, L., ERSSER, S., LEWIS-JONES, S., MULLER, I., HUGH, C. & LITTLE, P. 2012. Experiences of carers managing childhood eczema and their views on its treatment: a qualitative study. *Br J Gen Pract,* 62**,** e261-7.

SANTER, M., BURGESS, H., YARDLEY, L., ERSSER, S. J., LEWIS-JONES, S., MULLER, I., HUGH, C. & LITTLE, P. 2013. Managing childhood eczema: qualitative study exploring carers' experiences of barriers and facilitators to treatment adherence. *J Adv Nurs,* 69**,** 2493-501.

SANTER, M., MULLER, I., YARDLEY, L., BURGESS, H., ERSSER, S. J., LEWIS-JONES, S. & LITTLE, P. 2015a. 'You don't know which bits to believe': qualitative study exploring carers' experiences of seeking information on the internet about childhood eczema. *BMJ open,* 5**,** e006339.

SANTER, M., RUMSBY, K., RIDD, M. J., FRANCIS, N. A., STUART, B., CHOROZOGLOU, M., WOOD, W., ROBERTS, A., THOMAS, K. S., WILLIAMS, H. C. & LITTLE, P. 2015b. Bath additives for the treatment of childhood eczema (BATHE): protocol for multicentre parallel group randomised trial. *BMJ Open,* 5**,** e009575.

SAUERBREI, W. 1999. The use of resampling methods to simplify regression models in medical statistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 48**,** 313-329.

SCHMITT, J., APFELBACHER, C., SPULS, P. I., THOMAS, K. S., SIMPSON, E. L., FURUE, M., CHALMERS, J. & WILLIAMS, H. C. 2015. The Harmonizing Outcome

Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. *J Invest Dermatol,* 135**,** 24-30.

SCHMITT, J., LANGAN, S., STAMM, T. & WILLIAMS, H. C. 2011. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. *J Invest Dermatol,* 131**,** 623-30.

SCHMITT, J., LANGAN, S. & WILLIAMS, H. C. 2007. What are the best outcome measurements for atopic eczema? A systematic review. *J Allergy Clin Immunol,* 120**,** 1389-98.

SCHMITT, J., SPULS, P., BOERS, M., THOMAS, K., CHALMERS, J., ROEKEVISCH, E., SCHRAM, M., ALLSOPP, R., AOKI, V., APFELBACHER, C., BRUIJNZEEL-KOOMEN, C., BRUIN-WELLER, M., CHARMAN, C., COHEN, A., DOHIL, M., FLOHR, C., FURUE, M., GIELER, U., HOOFT, L., HUMPHREYS, R., ISHII, H. A., KATAYAMA, I., KOUWENHOVEN, W., LANGAN, S., LEWIS-JONES, S., MERHAND, S., MUROTA, H., MURRELL, D. F., NANKERVIS, H., OHYA, Y., ORANJE, A., OTSUKA, H., PAUL, C., ROSENBLUTH, Y., SAEKI, H., SCHUTTELAAR, M. L., STALDER, J. F., SVENSSON, A., TAKAOKA, R., WAHLGREN, C. F., WEIDINGER, S., WOLLENBERG, A. & WILLIAMS, H. 2012. Towards global consensus on outcome measures for atopic eczema research: results of the HOME II meeting. *Allergy,* 67**,** 1111-7.

SCHMITT, J., SPULS, P. I., THOMAS, K. S., SIMPSON, E., FURUE, M., DECKERT, S., DOHIL, M., APFELBACHER, C., SINGH, J. A. & CHALMERS, J. 2014. The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. *Journal of Allergy and Clinical Immunology,* 134**,** 800-807.

SCHMITT, J. & WILLIAMS, H. 2010. Harmonising Outcome Measures for Eczema (HOME). Report from the First International Consensus Meeting (HOME 1), 24 July 2010, Munich, Germany. *Br J Dermatol,* 163**,** 1166-8.

SCHOFIELD, J., GRINDLAY, D. & WILLIAMS, H. 2009. Skin conditions in the UK: a health care needs assessment.

SCHRAM, M. E., ROEKEVISCH, E., LEEFLANG, M. M. G., BOS, J. D., SCHMITT, J. & SPULS, P. I. 2011. A randomized trial of methotrexate versus azathioprine for severe atopic eczema. *Journal of Allergy and Clinical Immunology,* 128**,** 353-359.

SCHRAM, M. E., SPULS, P. I., LEEFLANG, M. M., LINDEBOOM, R., BOS, J. D. & SCHMITT, J. 2012. EASI, (objective) SCORAD and POEM for atopic eczema:

responsiveness and minimal clinically important difference. *Allergy,* 67**,** 99-106.

SCHRAM, M. E., TEDJA, A. M., SPIJKER, R., BOS, J. D., WILLIAMS, H. C. & SPULS, P. I. 2010. Is there a rural/urban gradient in the prevalence of eczema? A systematic review. *British Journal of Dermatology,* 162**,** 964-973.

SCHUTTELAAR, M. L., VERMEULEN, K. M. & COENRAADS, P. J. 2011. Costs and cost-effectiveness analysis of treatment in children with eczema by nurse practitioner vs. dermatologist: results of a randomized, controlled trial and a review of international costs. *Br J Dermatol,* 165**,** 600-11.

SILVERBERG, J., THYSSEN, J., PALLER, A., DRUCKER, A., WOLLENBERG, A., LEE, K., KABASHIMA, K., TODD, G., SCHMID-GRENDELMEIER, P. & BIEBER, T. 2017. What's in a name? Atopic dermatitis or atopic eczema, but not eczema alone. *Allergy,* 72**,** 2026-2030.

SILVERBERG, J. I., GARG, N. K., PALLER, A. S., FISHBEIN, A. B. & ZEE, P. C. 2015. Sleep Disturbances in Adults with Eczema Are Associated with Impaired Overall Health: A US Population-Based Study. *Journal of Investigative Dermatology,* 135**,** 56-66.

SIMPSON, E. L., CHALMERS, J. R., HANIFIN, J. M., THOMAS, K. S., CORK, M. J., MCLEAN, W. H. I., BROWN, S. J., CHEN, Z., CHEN, Y. & WILLIAMS, H. C. 2014. Emollient enhancement of the skin barrier from birth offers effective atopic dermatitis prevention. *J Allergy Clin Immunol,* 134**,** 818-23.

SMITH, F. J., IRVINE, A. D., TERRON-KWIATKOWSKI, A., SANDILANDS, A., CAMPBELL, L. E., ZHAO, Y., LIAO, H., EVANS, A. T., GOUDIE, D. R., LEWIS-JONES, S., ARSECULERATNE, G., MUNRO, C. S., SERGEANT, A., O'REGAN, G., BALE, S. J., COMPTON, J. G., DIGIOVANNA, J. J., PRESLAND, R. B., FLECKMAN, P. & MCLEAN, W. H. 2006. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nat Genet,* 38**,** 337-42.

SPULS, P. I., GERBENS, L. A. A., SIMPSON, E., APFELBACHER, C. J., CHALMERS, J. R., THOMAS, K. S., PRINSEN, C. A. C., VON KOBYLETZKI, L. B., SINGH, J. A., WILLIAMS, H. C. & SCHMITT, J. 2017. Patient-Oriented Eczema Measure (POEM), a core instrument to measure symptoms in clinical trials: a Harmonising Outcome Measures for Eczema (HOME) statement. *Br J Dermatol,* 176**,** 979-984.

STANISZEWSKA, S., HAYWOOD, K. L., BRETT, J. & TUTTON, L. 2012. Patient and public involvement in patient-reported outcome measures. *The Patient-Patient-Centered Outcomes Research,* 5**,** 79-87.

STEWART, K. & WILLIAMS, M. 2005. Researching online populations: the use of online focus groups for social research. *Qualitative Research,* 5**,** 395-416.

STRACHAN, D. P. 1989. Hay fever, hygiene, and household size. *Bmj,* 299**,** 1259-60.

STREINER, D. L. 2003. Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of personality assessment,* 80**,** 217-222.

STREINER, D. L., NORMAN, G. R. & CAIRNEY, J. 2015. *Health measurement scales: a practical guide to their development and use*, Oxford University Press, USA.

TERLUIN, B., EEKHOUT, I. & TERWEE, C. B. 2017. The anchor-based minimal important change (MIC), based on ROC analysis or predictive modeling, may need to be adjusted for the proportion of improved patients. *J Clin Epidemiol*.

TERLUIN, B., EEKHOUT, I., TERWEE, C. B. & DE VET, H. C. W. 2015. Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. *J Clin Epidemiol,* 68**,** 1388-1396.

TERWEE, C. B., BOT, S. D., DE BOER, M. R., VAN DER WINDT, D. A., KNOL, D. L., DEKKER, J., BOUTER, L. M. & DE VET, H. C. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol,* 60**,** 34-42.

TERWEE, C. B., ROORDA, L. D., DEKKER, J., BIERMA-ZEINSTRA, S. M., PEAT, G., JORDAN, K. P., CROFT, P. & DE VET, H. C. 2010. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol,* 63**,** 524-34.

THOMAS, K. S., ARMSTRONG, S., AVERY, A., PO, A. L. W., O'NEILL, C., YOUNG, S. & WILLIAMS, H. C. 2002. Randomised controlled trial of short bursts of a potent topical corticosteroid versus prolonged use of a mild preparation for children with mild or moderate atopic eczema. *BMJ,* 324**,** 768.

THOMAS, K. S., BRADSHAW, L. E., SACH, T. H., BATCHELOR, J. M., LAWTON, S., HARRISON, E. F., HAINES, R. H., AHMED, A., WILLIAMS, H. C., DEAN, T., BURROWS, N. P., POLLOCK, I., LLEWELLYN, J., CRANG, C., GRUNDY, J. D., GUINESS, J., GRIBBIN, A., MITCHELL, E. J., COWDELL, F., BROWN, S. J., MONTGOMERY, A. A. & TEAM, U. K. D. C. T. N. S. C. T. 2017. Silk garments plus standard care compared with standard care for treating eczema in children: A randomised, controlled, observer-blind, pragmatic trial (CLOTHES Trial). *PLOS Med,* 14**,** e1002280.

THOMAS, K. S., CHARMAN, C., NANKERVIS, H., RAVENSCROFT, J. C. & WILLIAMS, H. C. 2014. Chapter 24: Atopic Eczema. *In:* WILLIAMS, H. C. (ed.) *Evidence-Based Dermatology.* West Sussex: BMJ Publishing Group Limited.

THOMAS, K. S., DEAN, T., O'LEARY, C., SACH, T. H., KOLLER, K., FROST, A. & WILLIAMS, H. C. 2011. A randomised controlled trial of ion-exchange water softeners for the treatment of eczema in children. *PLoS Med,* 8**,** e1000395.

THOMAS, K. S., STUART, B., O'LEARY, C. J., SCHMITT, J., PAUL, C., WILLIAMS, H. C. & LANGAN, S. 2015. Validation of treatment escalation as a definition of atopic eczema flares. *PloS one,* 10**,** e0124770.

THORNELOE, R. J., GRIFFITHS, C. E., ASHCROFT, D. M. & CORDINGLEY, L. 2017. The challenges of assessing patients' medication beliefs: a qualitative study. *BMC health services research,* 17**,** 119.

THORNLEY, B. & ADAMS, C. 1998. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *Bmj,* 317**,** 1181-1184.

THORNLEY, B., RATHBONE, J., ADAMS, C. E. & AWAD, G. 2003. Chlorpromazine versus placebo for schizophrenia. *Cochrane Database of Systematic Reviews*.

TOURANGEAU, R., RIPS, L. J. & RASINSKI, K. 2000. *The psychology of survey response*, Cambridge University Press.

TRUJOLS, J., PORTELLA, M. J., IRAURGI, I., CAMPINS, M. J., SIÑOL, N. & COBOS, J. P. D. L. 2013. Patient-reported outcome measures: are they patient-generated, patient-centred or patient-valued? *Journal of mental health,* 22**,** 555-562.

TURNER, D., SCHUNEMANN, H. J., GRIFFITH, L. E., BEATON, D. E., GRIFFITHS, A. M., CRITCH, J. N. & GUYATT, G. H. 2010. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol,* 63**,** 28-36.

VON KOBYLETZKI, L. B., THOMAS, K. S., SCHMITT, J., CHALMERS, J. R., DECKERT, S., AOKI, V., WEISSHAAR, E., OJO, J. A. & SVENSSON, Å. 2016. What factors are important to patients when assessing treatment response: an international cross-sectional survey. *Acta Derm Venereol***,** 00015555-2480.

WALKER, J. 2001. *Control and the psychology of health*, Open University Press Buckingham.

WARD, D. J., FURBER, C., TIERNEY, S. & SWALLOW, V. 2013. Using framework analysis in nursing research: a worked example. *Journal of Advanced Nursing,* 69**,** 2423-2431.

WEIDINGER, S., ILLIG, T., BAURECHT, H., IRVINE, A. D., RODRIGUEZ, E., DIAZ-LACAVA, A., KLOPP, N., WAGENPFEIL, S., ZHAO, Y., LIAO, H., LEE, S. P., PALMER, C. N., JENNECK, C., MAINTZ, L., HAGEMANN, T., BEHRENDT, H., RING, J., NOTHEN, M. M., MCLEAN, W. H. & NOVAK, N. 2006. Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin Immunol,* 118**,** 214-9.

WEIDINGER, S. & NOVAK, N. 2016. Atopic dermatitis. *Lancet,* 387**,** 1109-1122.

WELLER, K., GROFFIK, A., CHURCH, M. K., HAWRO, T., KRAUSE, K., METZ, M., MARTUS, P., CASALE, T. B., STAUBACH, P. & MAURER, M. 2014. Development and validation of the Urticaria Control Test: A patient-reported outcome instrument for assessing urticaria control. *Journal of Allergy and Clinical Immunology,* 133**,** 1365-1372.e6.

WILLIAMS, H., JBURNEY, P., HAY, R., ARCHER, C., SHIPLEY, M., AHUNTER, J., BINGHAM, E., FINLAY, A., PEMBROKE, A. & CGRAHAM-BROWN, R. 1994a. The UK Working Party's Diagnostic Criteria for Atopic Dermatitis. *British journal of dermatology,* 131**,** 383-396.

WILLIAMS, H., ROBERTSON, C., STEWART, A., AÏT-KHALED, N., ANABWANI, G., ANDERSON, R., ASHER, I., BEASLEY, R., BJÖRKSTÉN, B. & BURR, M. 1999. Worldwide variations in the prevalence of symptoms of atopic eczema in the International Study of Asthma and Allergies in Childhood. *Journal of allergy and clinical immunology,* 103**,** 125-138.

WILLIAMS, H. C. & BURRELL-MORRUS, C. 2000. 13 - Atopic dermatitis in migrant populations. *In:* WILLIAMS, H. C. (ed.) *Atopig Dermatitis The Epidemiology, Causes and Prevention of Atopic Eczema.* Cambridge: Cambridge University Press.

WILLIAMS, H. C., STRACHAN, D. P. & HAY, R. J. 1994b. Childhood eczema: disease of the advantaged? *Bmj,* 308**,** 1132-1135.

WILLIAMSON, P. R., ALTMAN, D. G., BLAZEBY, J. M., CLARKE, M., DEVANE, D., GARGON, E. & TUGWELL, P. 2012. Developing core outcome sets for clinical trials: issues to consider. *Trials,* 13**,** 132.

WILLIAMSON, P. R., ALTMAN, D. G., BLAZEBY, J. M., CLARKE, M. & GARGON, E. 2011. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative. *Trials,* 12**,** A70.

WILLIAMSON, P. R., GAMBLE, C., ALTMAN, D. G. & HUTTON, J. L. 2005. Outcome selection bias in meta-analysis. *Statistical Methods in Medical Research,* 14**,** 515-524.

WILLIS, G. B. Cognitive interviewing: A “how to” guide. Research Triangle Institute. 1999 Meeting of the American Statistical Association. Research Triangle Park, NC: Research Triangle Institute, 1999.

WILLIS, G. B. 2015. *Analysis of the cognitive interview in questionnaire design*, Oxford University Press.

WILLIS, G. B. & ARTINO JR, A. R. 2013. What do our respondents think we're asking? using cognitive interviewing to Improve Medical Education surveys. *Journal of graduate medical education,* 5**,** 353-356.

WOLLENBERG, A., ORANJE, A., DELEURAN, M., SIMON, D., SZALAI, Z., KUNZ, B., SVENSSON, A., BARBAROT, S., VON KOBYLETZKI, L., TAIEB, A., DE BRUIN-WELLER, M., WERFEL, T., TRZECIAK, M., VESTERGARD, C., RING, J., DARSOW, U. & THE EUROPEAN TASK FORCE ON ATOPIC DERMATITIS, E. E. T. F. 2016. ETFAD/EADV Eczema task force 2015 position paper on diagnosis and treatment of atopic dermatitis in adult and paediatric patients. *Journal of the European Academy of Dermatology and Venereology,* 30**,** 729-747.

WORLD HEALTH ORGANISATION. *Clinical trials* [Online]. Available: https://www.who.int/topics/clinical_trials/en/ [Accessed 1 March 2019].

WRIGHT, A., HANNON, J., HEGEDUS, E. J. & KAVCHAK, A. E. 2012. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *J Man Manip Ther,* 20**,** 160-6.

WUTHRICH, B. 1999. Clinical aspects, epidemiology, and prognosis of atopic dermatitis. *Ann Allergy Asthma Immunol,* 83**,** 464-70.

YOUNG, B. & BAGLEY, H. 2016. Including patients in core outcome set development: issues to consider based on three workshops with around 100 international delegates. *Research Involvement and Engagement,* 2**,** 25.

ZUBERBIER, T., ORLOW, S. J., PALLER, A. S., TAIEB, A., ALLEN, R., HERNANZ-HERMOSA, J. M., OCAMPO-CANDIANI, J., COX, M., LANGERAAR, J. & SIMON, J. C. 2006. Patient perspectives on the management of atopic dermatitis. *J Allergy Clin Immunol,* 118**,** 226-32.

Appendices

Appendix A

Topic guide for UK online focus groups

Final topic guide

Note. This version was used in group of adults with eczema, but this was adopted to refer to "your child's eczema" in groups with parents

Stage 1: Set up (5 mins)

Informal greetings and indicate when will start

- Introduce self quickly (e.g. Hi I'm Laura and I will be moderating this discussion today.)
- Hi I'm Kim and I will be helping moderate.

You'll probably get adverts popping up to start with – these should stop after a while so please don't worry, just ignore them.

I would like to remind you all to respect the other people in the group. The questions we ask will hopefully lead to us all having an open discussion together. There are no right or wrong answers – we are interested in your experience and opinions!

Information you share will not be shared with your doctor or anybody else outside this discussion. The only exception is if we use quotes from you – if we do they will not have you're your name on them. Please remember that everyone in the discussion is able to see what you are writing, so try and not use confidential information such as your address or date of birth.

The chatroom does not like messages that are too long so if you are going to write more than roughly 100 words please send through the message in stages (I.e. two or three shorter messages). You will receive a pop up telling you if your message was too long. If this does happen just copy and paste parts of your answer into separate, shorter messages to ensure we all receive them.

Is everybody happy to continue?

Wait for responses

If you don't understand any of the questions as we go along, please do let us know so we can explain a bit more.

Stage 2: Introductions/openings (5-15 mins)

So let's all introduce ourselves. I'll make a start.

Hello, my name is Laura and I am a researcher at the Centre of Evidence Based Dermatology at the University of Nottingham. I am interested in your experiences of long-term control of your eczema. In previous groups we have discussed the different words used to describe their long-term control (e.g. flares, bad days, under control) and a whole range of different experiences people have. In the discussion group today we would like to build on what we have learnt so far by using your expertise as patients to focus on two topics in a bit more depth. We would like discuss how you decide if a treatment is working and how you think we should measure this.

Hello, my name is Kim and I am also a researcher at the Centre of Evidence Based Dermatology at the University of Nottingham.

Now it would be great for everyone else to introduce themselves to the group. You can introduce yourself in as much depth as you feel comfortable with. To start, it would be useful if you could tell us the name you would like to be addressed by and about how your eczema is at the moment.

Now we have all introduced ourselves, let's start our discussion on how you decide if a treatment is working.

Stage 3: Patient experiences (15-60 mins)

How would you decide if a treatment has been working well or not?

(Spend roughly 20-25 minutes on this question)

Possible prompts:

What would trigger you to make changes to your treatment regimen?

If you were having a period where your child's eczema was not well managed, what words would you use to describe this to others?

How would you describe it to your friends, family and colleagues?

What would be the best way to measure if a treatment is working well or not?

(Spend roughly 20-25 minutes on this question)

Possible prompts:

Do you think that a doctor or nurse should measure this or should you measure this?

How often do you think your eczema would need to be measured to fully capture how quickly your eczema changes? Would you say we would need to understand how eczema changes every day, every week or every month?

Why do you think that we need to measure it that often?

Would you be happy to fill out a measure (insert depending on what discussed – e.g. every week, every day)?

Stage 4: Ending (65-70 mins)

We are reaching the end of our discussion now.

Is there anything else you would like to add or any discussions you felt you didn't manage to fully state your opinion on before we end?

Wait for responses

Thank you all so much for taking the time to participate in our discussion. Your opinions are very important in guiding future eczema research. As a mark of our appreciation for you giving up your time today we will send you all a £20 Amazon gift voucher that you will receive via your email in the next few weeks. We will be contacting you via email to send you this gift voucher and provide you with some information in case the discussion today has raised any concerns for you.

(General goodbye)

If you could please close down your browser window now; this will allow you to exit the group.

Appendix B

Online focus group methods by country

Country	Number of focus groups	Dates focus groups took place	Website hosting the focus group	Interviewer(s) conducting focus groups	Interviewer characteristics and relationship with the participants	Recruitment methods
UK	6	August-October 2016	www.chatstep.com	LMH FC KST JRC SG	LMH, FC, KST and JRC have had training and experience using qualitative methods. KST had conduced online focus groups previously. The interviewers knew some participants personally due to their involvement in other research activities, but these relationships were not explicitly discussed during the focus groups. SG has had no previous relationship with the participants or experience of conducting qualitative work.	Participants were a community sample that were approached via social media, email and newsletters.
The Netherlands	2	February 2017	www.chatstep.com	MS GR	MS and GR have had training and experience using qualitative methods. MS and GR conducted the focus groups. MS and GR knew some of the patients/caregivers involved from the outpatient clinic.	Participants were patients/caregivers from the outpatient clinic. They were approached by telephone and directly at the outpatient clinic.
France	2	May 2017	www.chatstep.com	SB JD	SB has had training and experience using qualitative methods. SB and JD conducted the focus group. SB knew some of the patients/ families involved but these relationships were not explicitly discussed during the focus groups.	Participants were adults and parents of children with predominantly moderate to severe atopic eczema seen at an academic centre. Participants were approached by telephone and emails.

Sweden	2	May 2017	www.sunet.se	LK LB	LK and LB have had training at master courses and experience using qualitative methods in previous studies.	Participants were patients and parents to patients connected to the asthma and allergy foundation, Sweden.
USA	2	May-June, 2017	www.chatstep.com	AP KM KD	AP, KM, and KD have had training and experience using qualitative methods. KM and KD conducted the focus group and had no relationship with the patients and their families; AP knew the families involved and helped with recruiting but did not participate in data collection.	Participants were parents of children with predominantly moderate to severe atopic eczema seen at an academic centre.
Japan	2	April-May 2017	Goggle, Suite.	YK	YK knew some participants as her patients.	Participants were gathered through clinicians who were members of HOME or Japanese AD treatment research group.

Appendix C

Summary of the online survey of the HOME membership

- What does long-term control of eczema mean to you?

- What do you think are the advantages of collecting patient-reported outcomes (including quality of life) regularly during the trial, as a way of measuring long-term control?
- What do you think are the disadvantages?

- What do you think are the advantages of collecting clinician-reported outcomes regularly during the trial, as a way of measuring long-term control?
- What do you think are the disadvantages?

- What do you think are the advantages of capturing changes in eczema medication (such as amount of medication used, use of rescue medication, stepping up of medication) during the trial, as a way of measuring long-term control?
- What do you think are the disadvantages?

- What do you think are the advantages of collecting data on flares during the trial, as a way of measuring long-term control?
- What do you think are the disadvantages?

- Many patients and clinicians refer to the concept of a flare, but they can be difficult to capture in trials. Ways of defining flares include an arbitrary cut-off such as a change in score from a baseline measurement, a behavioural measure such as use of rescue medication, or a composite measure (e.g. IGA >4 AND the need for rescue medication). What do you think would be the best way to define a flare in clinical trials?
- How would you define the start of a flare?
- How would you define the end of a flare?

- Well controlled weeks is a concept used in asthma trials. What does the term "well controlled weeks" mean to you with regards to eczema?
- What do you think are the advantages of collecting well-controlled weeks regularly during the trial, as a way of measuring long-term control?
- What do you think are the disadvantages?

- Is there anything else you would like to add about measuring the long-term control of eczema in trials?

Appendix D

Interview guide for cognitive interviews

Interview guide

What are you thinking when you answer questions about your / your child's eczema control?

Introductions

- ☐ Greetings with participant/s
- ☐ If they have not already done so, ask participants to read information sheet and provide informed consent (for child as well if they are present)
- ☐ Give participants the opportunity to ask questions
- ☐ Explain the purpose of today's interview is to help us make sure the questions we ask in eczema research and clinical practice are meaningful to people. We want to know how easy the questions are to understand, what the questions mean to you and what they make you think of.
- ☐ Check they are happy for the audio recording to be turned on and turn it on.
- ☐ Give participants an opportunity to share some information about themselves (listen out for age, sex, ethnicity, disease duration, and severity related information).
- ☐ Give them an option to have a practice at "thinking out loud" if they would like. If they would like to ask them to "Try to visualize the place where you live, and think about how many windows there are in that place. As you count up the windows, tell me what you are seeing and thinking about." (can probe if needed)
- ☐ Explain that you are going to ask them to fill out a questionnaire measuring eczema control (N.B. the exact questions will be developed in stage 1 but will ask about how various aspects of their eczema have been over a recent time period). If also going to be asking about the Patient Oriented Eczema Measure (POEM), explain that you are going to ask about this questionnaire on eczema symptoms too. When they answer the questions, they will need to "think out loud" as they are filling them in. The interviewer or the participant can read the question out loud too if this is preferable. Explain that you may also ask them some questions about what they are thinking throughout the interview. Explain that there are no right or wrong answers.

Prompts that can be used are presented below. These have been informed by cognitive theory of self-report questionnaires and some of them have been adapted from different sources (Thorneloe et al., 2017, Willis and Artino Jr, 2013, Willis, 1999, Irwin et al., 2009, Collins, 2003, Jobe and Herrmann, 1996). They are categorised according to the purpose of the prompt. Other probes in response to what participants say may also be used.

To determine understanding/comprehension:

- ☐ What does the term "INSERT TERM" mean to you?
- ☐ What did you understand by the term "INSERT TERM"?
- ☐ Does "INSERT TERM" used in this question sound OK to you, or would you say something different?
- ☐ What were you thinking about when answering that question?

To determine how the respondent interprets the task demands of the question:

- ☐ What is this question asking?
- ☐ Did you think questions (insert number) were asking about the same thing or something different?

To determine relevance for this person:

- ☐ Is this something you would consider as being relevant to your child's eczema?
- ☐ Do you feel this question applies to your child?

To determine ability to retrieve information:

- ☐ What time period where you thinking about when you answered that question?
- ☐ What were you thinking about when answering that question?
- ☐ How did you remember that?
- ☐ Did you have a particular time period in mind when you were answering that question?

To determine how the respondent reconstructs information to form an answer:

- ☐ How did you manage to calculate that answer?
- ☐ How did you pull all of the information you had together to arrive at that answer?
- ☐ How did you arrive at that thought?

- ☐ How did you come up with your answer?

To determine how confident respondent is in their judgement:

- ☐ How easy or hard was it to come up with your answer?
- ☐ How sure are you about your answer?
- ☐ How well do you remember this?

To understand why they chose a response option:

- ☐ How did you decide on this answer?
- ☐ How did you decide to circle/tick that response?
- ☐ How well does the response that you circled/ticked apply to you?
- ☐ What do you think of the response choices here?
- ☐ Can you tell me why you chose the category "INSERT" rather than the category "INSERT"?
- ☐ Can you tell me why you chose to circle the category "INSERT" for this item, but you circled "INSERT" for this item?
- ☐ Did you change your mind at all during deciding what answer to pick?
- ☐ Did you have any difficulty deciding which response option to pick?
- ☐ Is there anything else you would have liked to have responded with that wasn't here as an option?

To encourage elaboration of "think aloud" process:

- ☐ What made you say that?
- ☐ You were hesitating then – what were you thinking about?
- ☐ Is there anything else that you were thinking about?

Overall questionnaire:

- ☐ Are there any questions that you think don't belong in this group?
- ☐ Are there things that we forgot to ask about that you think are important?
- ☐ What are your overall thoughts of these questions?
- ☐ Is there anything about the look of the questions that you don't like?
- ☐ Is there anything else you would want to change about these questions?

- Give the opportunity to add anything else they would like to say.

- Debrief – Reiterate the purpose of this study. If any concerns have been raised by the participant, provide information of where they could receive further support (GP, National Eczema Society – www.eczema.org)

Appendix E

Information in online survey about eczema control questions

Page 1: Participant Information Sheet and consent to taking part

Thank you for your interest in this online study.

Researchers at the University of Nottingham are inviting you to take part in a survey about eczema. The survey is to help make sure that questions asked about eczema control (during medical appointments and clinical trials) are meaningful.

Before deciding to take part in this study, please take time to read the following information carefully.

To take part:
- You or your child must have been diagnosed with eczema by a doctor
- Live in the UK and be able to speak English
- Be 16 years of age or older or be completing this with your parent

What's involved?
If you choose to take part, you will be asked to fill in a survey about your eczema. This should take between 10-15 minutes to complete.

Do I have to take part?
No. It is up to you whether or not you take part. Even if you do agree to take part, you may withdraw from the study at any time. Please be aware that if you do withdraw, the answers you have already provided on previous pages of the survey will have already been recorded and cannot be removed.

Will the information I provide be kept confidential?
Yes. Study data may be looked at by the team of researchers working on this project, however any personal details will be stored separately from the results. All data will be reported anonymously.

What will happen to the results of the study?
If you want to find out about the results of the study, you can share your email address with us at the end of the survey and we will share the findings with you when we reach the end of the project. The findings will hopefully result in a new questionnaire that can be used in eczema research and at medical appointments. The research team will write up the research and publish the results in scientific journals. The findings will also be presented for the public in a format that is clear and accessible.

Who is funding the study?

This project is funded by the British Skin Foundation.

What if I have more questions or concerns?
If you have any questions about this project, you may contact the research team
By email: eczema@nottingham.ac.uk
By phone: 0115 84 68634

If they are unable to resolve your concern or you wish to make a complaint regarding the study, please contact: FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, School of Medicine Education Centre, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH and/or
email: louise.sabir@nottingham.ac.uk

This study has been reviewed and given a favourable opinion by the University of Nottingham, Faculty of Medicine &
Health Sciences Research Ethics Committee [ref:18-1805].

I have read and understood the above information, and I understand by clicking the NEXT button at the bottom of this page, I indicate my willingness to voluntarily take part in the study.
(You can click the OK button or scroll down to find the NEXT button.)

Page 2: Required question to assess eligibility

Is our survey right for you?

Do all of the following statements apply to you:

Either myself or my child has been diagnosed with eczema by a doctor.

I live in the UK and I speak English.

I am 16 years of age or older or I am completing this survey with my parent or guardian.

Yes No

(if answer No participants were redirected out of survey)

Page 3: Required assessment of which pages participant should be given to fill out

Deciding which questions it is best to ask you.

Please select which statement best describes who is completing this survey so that we can direct you to the most appropriate questions.

Adult - I am over 16 years old and have eczema. I am completing this survey about my eczema. (*Directed to pages 4-8)*

Child with parent - I am under 16 years old and have eczema. My parent is helping me complete this survey, but I would like to answer the questions about my eczema myself. *(Directed to pages 4-8)*

Parent on behalf of child - My child has eczema. I am completing this survey about my child's eczema. *(Directed to pages 9-13)*

Page 4: Participant characteristics

About you
Section 1 out of 5

Age:
Drop down menu

Sex:
Male
Female
Rather not say
Other (please specify)

Ethnicity
White
Black (other)
Black African
Black Caribbean
Indian
Pakistani
Bangladeshi
Chinese
Other Asian (non-Chinese)
Mixed Race
Rather not say
Other (please specify)

Page 5: Bother scale and one draft RECAP item

A couple of questions about this past week.

Section 2 out of 5
Please read the following two questions and click the response that you think best fits with your experience. Only click one response for each question. Try to respond to every question, but if you are unable to respond then leave it blank.
How much bother has your eczema been over the past week?
0 (no bother at all)
1
2
3
4
5
6
7
8
9
10 (as much bother as you can imagine)
Over the last week, how has your eczema been?
Very good Good OK Bad Very bad

Page 6: 14 draft RECAP items and assessment of relevance of items

RECAP of eczema control

Section 3 out of 5

This is the main section of the survey and it is the longest part of the survey to complete.

You will be asked 14 questions with three parts to each question (A, B and C).

Part A questions provide a snapshot of how your eczema has been over the last week from your point of view.

Part B questions include a statement about your eczema and asks if you have experienced that at any time in the past year.

Part C questions include the same statement as part B, and ask you how important it is when thinking about your eczema.

Please read the following questions and select the response that you think best fits with your experience. Only select one response for each question. Try to respond to every question, but if you are unable to respond then leave it blank.

Question 1 out of 14
1A
Thinking about all the eczema treatments you have used in the last week, on how many days has your treatment been enough to manage your eczema?
Every day 5-6 days 3-4 days 1-2 days No days

Question 1 out of 14
1B
The eczema treatments you have used being enough to manage your eczema. Have you experienced this at any time over the past year?
No Yes

Question 1 out of 14
1C
The eczema treatments you have used being enough to manage your eczema. How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 2 out of 14
2A
Over the last week, how acceptable has your eczema been to you?
Completely acceptable Mostly acceptable Quite acceptable Not very acceptable Not at all acceptable

Question 2 out of 14
2B
Eczema that is acceptable to you.
Have you experienced this at any time over the past year?
No Yes

Question 2 out of 14
2C
Eczema that is acceptable to you.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 3 out of 14
3A
Over the last week, on how many days has your skin been itchy because of your eczema?

No days 1-2 days 3-4 days 5-6 days Every day

Question 3 out of 14
3B
Your skin being itchy because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 3 out of 14
3C
Your skin being itchy because of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 4 out of 14
4A
Over the last week, how much has your sleep been disturbed because of your eczema?
Not at all A little bit Quite a lot A huge amount Completely

Question 4 out of 14
4B
Sleep disturbance because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 4 out of 14
4C
Sleep disturbance because of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 5 out of 14
5A
Over the last week, on how many days were you unable to stop scratching because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 5 out of 14
5B
Unable to stop scratching because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 5 out of 14
5C

Unable to stop scratching because of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 6 out of 14
6A
Over the last week, how much has your eczema been getting in the way of day to day activities?
Not at all A little bit Quite a lot A huge amount Completely

Question 6 out of 14
6B
Your eczema getting in the way of day to day activities.
Have you experienced this at any time over the past year?
No Yes

Question 6 out of 14
6C
Your eczema getting in the way of day to day activities.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 7 out of 14
7A
Over the last week, on how many days has your eczema affected how you have been feeling?
No days 1-2 days 3-4 days 5-6 days Every day

Question 7 out of 14
7B
Your eczema affecting how you have been feeling.
Have you experienced this at any time over the past year?
No Yes

Question 7 out of 14
7C
Your eczema affecting how you have been feeling.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 8 out of 14
8A
Over the last week, on how many days has your eczema stopped you doing something you wanted or needed to do?

No days 1-2 days 3-4 days 5-6 days Every day

Question 8 out of 14
8B
Your eczema stopping you doing something you wanted or needed to do.
Have you experienced this at any time over the past year?
No Yes

Question 8 out of 14
8C
Your eczema stopping you doing something you wanted or needed to do.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 9 out of 14
9A
Over the last week, on how many days have you felt self-conscious or embarrassed because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 9 out of 14
9B
Feeling self-conscious or embarrassed because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 9 out of 14
9C
Feeling self-conscious or embarrassed because of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 10 out of 14
10A
Over the last week, on how many days have you felt isolated because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 10 out of 14
10B
Feeling isolated because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 10 out of 14
10C

Feeling isolated because of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 11 out of 14
11A
Over the last week, on how many days have you experienced an eczema flare?
No days 1-2 days 3-4 days 5-6 days Every day

Question 11 out of 14
11B
Experiencing an eczema flare.
Have you experienced this at any time over the past year?
No Yes

Question 11 out of 14
11C
Experiencing an eczema flare.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 12 out of 14
12A
Over the last week, on how many days have you had any symptoms of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 12 out of 14
12B
Having any symptoms of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 12 out of 14
12C
Having any symptoms of your eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 13 out of 14
13A
Over the last week, on how many days has your skin felt painful or sore because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 13 out of 14
13B
Your skin feeling painful or sore because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 13 out of 14
13C
Your skin feeling painful or sore because of the eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 14 out of 14
14A
Over the last week, on how many days has your skin been intensely itchy because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 14 out of 14
14B
Your skin being intensely itchy because of your eczema.
Have you experienced this at any time over the past year?
No Yes

Question 14 out of 14
14C
Your skin being intensely itchy because of the eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Page 7: Global Severity Question

Section 4 out of 5

This section includes a single question. Please select one of the responses below. Leave it blank if you feel unable to answer it.

How has your eczema been over the past week?
Clear
Almost clear
Mild
Moderate
Severe

Page 8: POEM to assess eczema symptoms over same time period as RECAP items

Patient-Oriented Eczema Measure (POEM)

Section 5 out of 5

This is the final section.
Please select one response for each of the seven questions below about your eczema. Please leave blank any questions you feel unable to answer.

Question 1 out of 7
Over the last week, on how many days has your skin been itchy because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 2 out of 7
Over the last week, on how many nights has your sleep been disturbed because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 3 out of 7
Over the last week, on how many days has your skin been bleeding because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 4 out of 7
Over the last week, on how many days has your skin been weeping or oozing clear fluid because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 5 out of 7
Over the last week, on how many days has your skin been cracked because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 6 out of 7
Over the last week, on how many days has your skin been flaking off because of your eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 7 out of 7
Over the last week, on how many days has your skin felt dry or rough because of your eczema?

No days 1-2 days 3-4 days 5-6 days Every day

Page 9: Characteristics of participant's child

About your child (Section 1 out of 4)

Child's age
Drop down options
Child's sex:
Male
Female
Rather not say
Other (please specify)

Child's ethnicity
White
Black (other)
Black African
Black Caribbean
Indian
Pakistani
Bangladeshi
Chinese
Other Asian (non-Chinese)
Mixed Race
Rather not say
Other (please specify)

Page 10: Bother scale about child and one draft RECAP item

A couple of questions about this past week.
Section 2 out of 5

Please read the following two questions and click the response that you think best fits with your experience. Only click one response for each question. Try to respond to every question, but if you are unable to respond then leave it blank.

How much bother has your child's eczema been over the past week?
0 (no bother at all)
1
2
3
4

5
6
7
8
9
10 (as much bother as you can imagine)

Over the last week, how has your child's eczema been?
Very good Good OK Bad Very bad

Page 11: 14 draft RECAP items and assessment of relevance of items

RECAP of eczema control

Section 3 out of 5

This is the main section of the survey and it is the longest part of the survey to complete.

You will be asked 14 questions with three parts to each question (A, B and C).

Part A questions provide a snapshot of how your child's eczema has been over the last week from your point of view.

Part B questions include a statement about your child's eczema and asks if you think your child has experienced that at any time in the past year.

Part C questions include the same statement as part B, and ask you how important it is when thinking about your child's eczema.
Please read the following questions and select the response that you think best fits with your child's experience. Only select one response for each question. Try to respond to every question, but if you are unable to
respond then leave it blank.

Question 1 out of 14
1A
Thinking about all the eczema treatments your child has used in the last week, on how many days has their treatment been enough to manage their eczema?
Every day 5-6 days 3-4 days 1-2 days No days

Question 1 out of 14
1B
The eczema treatments your child has used being enough to manage their eczema.

Do you think your child has experienced this at any time over the past year?
No Yes

Question 1 out of 14
1C
The eczema treatments your child has used being enough to manage their eczema.
How important is this when thinking about how your eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 2 out of 14
2A
Over the last week, how acceptable has your child's eczema has been to you?
Completely acceptable Mostly acceptable Quite acceptable Not very acceptable Not at all acceptable

Question 2 out of 14
2B
Eczema that is acceptable to you.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 2 out of 14
2C
Eczema that is acceptable to you.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 3 out of 14
3A
Over the last week, on how many days do you think your child's skin has been itchy because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 3 out of 14
3B
Your child's skin being itchy because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 3 out of 14
3C
Your child's skin being itchy because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)
Question 4 out of 14
4A

Over the last week, how much do you think your child's sleep has been disturbed because of their eczema?
Not at all A little bit Quite a lot A huge amount Completely

Question 4 out of 14
4B
Your child having disturbed sleep because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes
Question 4 out of 14
4C
Your child having disturbed sleep because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 5 out of 14
5A
Over the last week, on how many days was your child unable to stop scratching because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 5 out of 14
5B
Your child being unable to stop scratching because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 5 out of 14
5C
Your child being unable to stop scratching because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 6 out of 14
6A
Over the last week, how much has your child's eczema been getting in the way of day to day activities?
Not at all A little bit Quite a lot A huge amount Completely

Question 6 out of 14
6B
Your child's eczema getting in the way of day to day activities.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 6 out of 14
6C
Your child's eczema getting in the way of day to day activities.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 7 out of 14
7A
Over the last week, on how many days do you think your child's eczema has affected how they have been feeling?
No days 1-2 days 3-4 days 5-6 days Every day

Question 7 out of 14
7B
Your child's eczema affecting how they have been feeling.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 7 out of 14
7C
Your child's eczema affecting how they have been feeling.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 8 out of 14
8A
Over the last week, on how many days do you think your child's eczema has stopped them doing something they wanted or needed to do?
No days 1-2 days 3-4 days 5-6 days Every day

Question 8 out of 14
8B
Your child's eczema stopping them doing something they wanted or needed to do.
Do you think your child has experienced this at any time over the past year?
No Yes
Question 8 out of 14

8C
Your child's eczema stopping them doing something they wanted or needed to do.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 9 out of 14
9A

Over the last week, on how many days do you think your child has felt self-conscious or embarrassed because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 9 out of 14
9B
Your child feeling self-conscious or embarrassed because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 9 out of 14
9C
Your child feeling self-conscious or embarrassed because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 10 out of 14
10A
Over the last week, on how many days do you think your child has felt isolated because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 10 out of 14
10B
Your child feeling isolated because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 10 out of 14
10C
Your child feeling isolated because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 11 out of 14
11A
Over the last week, on how many days has your child has experienced an eczema flare?
No days 1-2 days 3-4 days 5-6 days Every day

Question 11 out of 14
11B
Your child experiencing an eczema flare.
Do you think your child has experienced this at any time over the past year?

No Yes

Question 11 out of 14
11C
Your child experiencing an eczema flare.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 12 out of 14
12A
Over the last week, on how many days has your child had any symptoms of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 12 out of 14
12B
Your child having any symptoms of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 12 out of 14
12C
Your child having any symptoms of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 13 out of 14
13A
Over the last week, on how many days do you think your child's skin has felt painful or sore because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 13 out of 14
13B
Your child's skin feeling painful or sore because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 13 out of 14

13C
Your child's skin feeling painful or sore because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Question 14 out of 14

14A
Over the last week, on how many days do you think your child's skin has been intensely itchy because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 14 out of 14
14B
Your child's skin being intensely itchy because of their eczema.
Do you think your child has experienced this at any time over the past year?
No Yes

Question 14 out of 14
14C
Your child's skin being intensely itchy because of their eczema.
How important is this when thinking about how your child's eczema has been?
1 (not important) 2 3 4 5 (extremely important)

Page 12: Global Severity Question

Section 4 out of 5

One question in this section. Please select one of the responses below. Leave it blank if you feel unable to answer it.

How has your child's eczema been over the past week?
Clear
Almost clear
Mild
Moderate
Severe

Page 13: POEM to assess eczema symptoms over same time period as RECAP items

Patient-Oriented Eczema Measure (POEM)

Section 5 out of 5

This is the final section.

Please select one response for each of the seven questions below about your child's eczema. Please leave blank any questions you feel unable to answer.

Question 1 out of 7
Over the last week, on how many days has your child's skin been itchy because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 2 out of 7
Over the last week, on how many nights has your child's sleep been disturbed because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 3 out of 7
Over the last week, on how many days has your child's skin been bleeding because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 4 out of 7
Over the last week, on how many days has your child's skin been weeping or oozing clear fluid because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 5 out of 7
Over the last week, on how many days has your child's skin been cracked because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 6 out of 7
Over the last week, on how many days has your child's skin been flaking off because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Question 7 out of 7
Over the last week, on how many days has your child's skin felt dry or rough because of their eczema?
No days 1-2 days 3-4 days 5-6 days Every day

Page 14: Final page all participants directed to for debrief and opportunity to provide email address to receive study results.

This is the end of the study.

Thank you for taking part in this survey. The aim of this study was to identify what questions are most appropriate to ask people to help us gain an understanding of how well controlled their or their child's eczema is. We hope the findings of these studies will improve how we measure long-term control of eczema in research.

If taking part in this study has raised any concerns or questions for you or if you require support, please do not hesitate to visit your general practitioner to discuss these issues.

You can also receive support in a variety of ways via the National Eczema Society:

http://www.eczema.org/get-support

The results of this study will be posted on our website

https://www.nottingham.ac.uk/research/groups/cebd/projects/index.aspx once they are available, but if you would like a copy of the results to be sent to you directly then please enter your email address here: [open text response]