Harmonising Outcome Measures for Eczema

HOME IV Meeting

23rd-24th April 2015

Malmö, Sweden

Minutes

List of Attendees

	Name	Stakeholder group	Country
1	Katrina Abuabara	Clinician - Dermatology / Methodologist	USA
2	Valeria Aoki	Clinician - Dermatology	Brazil
3	Christian Apfelbacher	Methodologist	Germany
4	Marius Ardeleanu	Pharmaceutical Industry Representative	USA
5	Maren Awici-Rasmussen	Patient/Carer/Patient Representative	Norway
6	Sebastien Barbarot	Clinician - Dermatology	France
7	Linda Beckman	Researcher	Sweden
8	Julie Block	Patient/Carer/Patient Representative	USA
9	Anthony Bragg	Pharmaceutical Industry Representative	
10	Tim Burton	Patient/Carer/ Patient Representative	UK
11	Joanne Chalmers	Methodologist	UK
12	Kim Katrine Clemmensen	Clinician - Dermatology	Denmark
13	Amanda Creswell-Melville	Patient/Carer/Patient Representative	Canada
14	Maj Dinesen	Pharmaceutical Industry Representative	Denmark
15	Aaron Drucker	Clinician - Dermatology	Canada
16	Laurent Eckert	Pharmaceutical Industry Representative	France
17	Andrew Finlay	Clinician - Dermatology	UK
18	Carsten Flohr	Clinician - Dermatology	UK
19	Madhur Garg	Pharmaceutical Industry Representative	
20	Louise Gerbens	Clinician - MD PhD student - Dermatology	Netherlands
21	Lykke Graff	Pharmaceutical Industry Representative	
22	Jon Hanifin	Clinician - Dermatology	USA
23	Daniel Heinl	Clinician - MD PhD student	Germany
24	Rosemary Humphreys	Patient/Carer/Patient Representative	UK
25	Henrique Ishii	Patient/Carer/Patient Representative	Brazil
26	Yoko Kataoka	Clinician - Dermatology	Japan
27	Willem Kouwenhoven	Patient/Patient Representative	Netherlands
28	Yael Leshem	Clinician - Dermatology	USA
29	Teresa Løvold Berents	Clinician - Dermatology	Norway
30	Bromwyn Lund	Pharmaceutical Industry Representative	
31	Burchard Marquort	Patient/Carer/Patient Representative	Sweden
32	Marie-Anne Massuel	Pharmaceutical Industry Representative	France
33	Stephanie Merhand	Patient/Carer/Patient Representative	France
34	Hitoshi Mizutani	Clinician - Dermatology	Japan
35	Hiroyuki Murota	Clinician - Dermatology	Japan
36	Dedee Murrell	Clinician - Dermatology	Australia
37	Takeshi Nakahara	Clinician - Dermatology	Japan
38	Ibrahim Nasr	Clinician - Dermatology	UK
39	Kristine Nograles	Pharmaceutical Industry Representative	
40	Yukihiro Ohya	Clinician - Pediatrician	Japan
41	Ian Osterloh	Pharmaceutical Industry Representative	UK
42	Jan Pander	Pharmaceutical Industry Representative	Netherlands
43	Cecilia (Sanna) Prinsen	Clinical Epidemiologist, Methodologist	Netherlands
44	Lynn Purkins	Pharmaceutical Industry Representative	
45	Matthew Ridd	Clinician (GP) & academic	UK
46	Tracey Sach	Methodologist	UK
47	Jochen Schmitt	Clinician - Dermatology	Germany
48	Marie-Louise Schuttelaar	Clinician - Dermatology	Netherlands
49	Shoko Shindo	Clinician - Dermatology	Japan
50	Eric Simpson	Clinician - Dermatology	USA
51	Jasvinder Singh	Methodologist	USA
52	Jevenija Smirnova	Clinician - Junior Doctor	Sweden
53	Phyllis Spuls	Clinician - Dermatology	Netherlands
54	Anne Sulzer	Pharmaceutical Industry Representative	France
55	Åke Svensson	Clinician - Dermatology / Molecular epidemiology	Sweden
56	Eli Synnøve Gjerde	Patient/Carer/Patient Representative	Norway
57	Roberto Takaoka	Clinician - Dermatology	Brazil

58	Helle vestby Talmo	Patient/Carer/Patient Representative	Norway
59	Marie Tauber	Clinician - Dermatology	France
60	Kim Thomas	Methodologist	UK
61	Florent Torchet	Patient/Carer/Patient Representative	France
62	Laura von Kobyletzki	Clinician - Other (General practitioner)	Sweden
63	Annika Volke	Clinician - Dermatology	Estonia
64	Carl-Fredrik Wahlgren	Clinician - Dermatology	Sweden
65	Stephan Weidinger	Clinician - Dermatology / Molecular epidemiology	Germany
66	Elke Weisshaar	Clinician - Dermatology	Germany
67	Hywel Williams	Clinician - Dermatology	UK
68	Andreas Wollenberg	Clinician - Dermatology	Germany
69	Kosuke Yamaga	Clinician - Dermatology	Japan
70	Cathy Zhao	Clinician - Dermatology	

Attendees of the HOME IV meeting

Glossary

AD	Atopic Dermatitis
ADQ	Atopic Dermatitis Quickscore
COS	Core outcome set
COSMIN	Consensus Based Standards for the Selection of Health Measurement Instruments
DIELH	Deutsches Instrument zur Erfassung der Lebensqualitat bei Hauterrankungen
DLQI	Dermatology Life Quality Index
EASI	Eczema Area and Severity Index
FLQA	Freiburg Life Quality Index
HOME	Harmonising Outcome Measures for Eczema
ISDL	Impact of chronic skin disease on daily life
ISS	Itch severity scale
KM	Kaplan Meier
LTC	Long term control
NESS	Nottingham Eczema Severity Score
OMERACT	Outcome Measures in Rheumatology
POEM	Patient–oriented Eczema Measure
PO-SCORAD	Patient Oriented SCORing Atopic Dermatitis
PRO	Patient reported outcome
QoL	Quality of Life
QoLIAD	Quality of Life Index for Atopic Dermatitis
RCT	Randomised Controlled Trial
SA-EASI	Self Administered–Eczema Area and Severity Index
SCORAD	SCORing Atopic Dermatitis
Skindex	Eczema-specific QoL measure
VAS	Visual Analogue Scale

Thursday 23rd April (09:00 – 17:00)

Session 1 – Introduction (Chair Jochen Schmitt)

Welcome

Åke Svensson (AS) and Laura von Kobyletzki (LvK) welcomed everyone to Malmo and outlined the programme for the next two days. Hywel Williams (HW) asked different stakeholder groups to stand up to give a sense of balance of attendees, and commented that there was a good balance with 12 patients / patient representatives, 7 methodologists, 38 clinicians (mainly dermatologists) and 13 representatives of the pharmaceutical industry covering North and South America, Europe, Japan and Australia.

Presentation 1.1: Hywel Williams - Introduction and Background

HW set the scene for the meeting by describing the extent of the problem of too many outcome measures for atopic dermatitis/eczema, what core outcomes are and why they are needed and summarised the progress made so far by HOME. He reminded the group about the philosophy of HOME whereby everyone should put aside prejudices and allegiances with specific instruments in order to achieve the greater good for patient care. He mentioned the previously agreed consensus rules that would be applied to voting at this meeting which are:

- ***Consensus is reached where fewer than 30% of the voters disagree.***
- ***Undecided votes are counted in the agree group.***

HW then invited Jas Singh (JS) to give his reflections from OMERACT and explain his role in the meeting.

Presentation 1.2: Jas Singh – Reflections from OMERACT

JS stated he was attending as an observer from the OMERACT group, and would share experiences of OMERACT with the group. He emphasised that everyone present is an expert in a different way, but all have in common that they are passionate about improving the situation in eczema research.

JS asked the group to keep in mind the following points during the meeting:

1. The distinction between an outcome domain and an outcome measure instrument; as an example, a fever is a domain, and a thermometer used to measure temperature is an outcome measure instrument.
2. The HOME group have already agreed that the four core domains of signs, symptoms, quality of life and long-term control should be measured in clinical trials in eczema. It is the role of the group to now agree upon core outcome measurement instruments for the remaining domains of symptoms, quality of life and long–term control, having already agreed EASI as the core outcome measure instrument for clinical signs at HOME III in San Diego.
3. The voice of patients in agreeing upon outcome instruments for the patient reported outcomes (PRO's) of symptoms and quality of life is vital. He noted that there is good patient

representation at this meeting and emphasised that patients should feel free to speak up and steer the group.

4. He encouraged the group to listen to disagreement, particularly where there are low levels of agreement and reminded attendees to share their views *during* the meeting, rather than waiting until afterwards.
5. It is highly unlikely that there will be 100% agreement on anything and only up to 30% disagreement required to pass a vote. As the group progresses through the voting process, everyone should ask themselves - "Can I live with it?" rather than "is this my preferred solution?" and if the answer is yes then one should preferably vote in agreement.
6. Core outcome domains and instruments should be included in every eczema clinical trial but they don't have to be the primary outcome and researchers can include as many other outcomes as they wish. Inclusion of core outcomes doesn't mean changing the objectives of the research.
7. Because this meeting is regarding clinical trials, not routine practice record keeping, it was important to remember that recommended instruments don't need to be suitable for fitting into a short routine clinic appointments.

JS added that developing the core outcome set led to an increase in number of treatments available for rheumatoid arthritis, at least in part because it enabled the effective comparison of the new treatments with existing treatments. He also explained that some of the driving force behind OMERACT was the regulatory authorities as they wanted to know the best outcome measures for rheumatoid arthritis. Representatives from regulatory authorities attend approximately 50% of OMERACT meetings. HW agreed that HOME needs to keep regulators informed and engaged.

Before moving on to the symptoms domain, HW then took a few moments to ensure the group were happy with what was planned to be covered over the next two days.

Session 2 – Symptoms domain (Chairs Eric Simpson and Jas Singh)

Eric Simpson (ES) introduced the aims and summarised the proposed format of the session.

Presentation 2.1: Phyllis Spuls - Introduction

Phyllis Spuls (PS), lead for the symptoms working group, began by declaring no conflicts of interest then described how the HOME roadmap relates to the work that has been done by the symptoms working group and what data will be presented at this meeting. PS then stated that the goal is to agree a core outcome instrument to measure the symptoms of eczema in clinical trials.

PS then stated the standard definition of symptoms is:

- A departure from normal function or feeling which is noticed by a patient, indicating the presence of disease or abnormality

The definition of a symptom used at HOME III was:

- Any feature observed by the patient whereas a sign is observed by others. Some features such as itch can only be a symptom because it can only be measured by the patient

themselves (and not directly observed by other people), whereas others such as a blood cell count can only be measured by others so can only be a sign. Some features such as skin rash can be sign or a symptom depending on who observes it.

PS explained that to determine what symptoms were important to be included in the construct symptoms information had been obtained from several sources prior to the discussion at the HOME IV meeting. These were:

- Global Survey of which eczema signs and symptoms are important to patients? (See 2.2)
- Input from patients at the HOME IV patient pre-meeting (22nd April 2015)
- Systematic review of what symptoms are measured and reported in clinical trials (See 2.3)

Summary of whole group discussion

There followed a discussion about the definition of symptoms versus signs:

There was general agreement that, by the HOME definition, anything that patients report or complain of is a symptom.
Concerns were expressed that this definition does not sufficiently distinguish symptoms from quality of life.
It was recognised that some clinicians would normally refer to what is classed here as a symptom e.g. redness, as a patient reported sign.

Voting

Voting to establish stakeholder groups showed 51.5% were clinicians, 18.2% were patients, 12.1% were methodologists and 18.2% were pharmaceutical industry representatives.

The group then voted on whether they agreed with this HOME definition of symptoms: *anything that patients report or complain of is a symptom.*

	n	%
Agree	52	78.8
Disagree	10	15.2
Unsure	4	6.1
Total	66	

The group then voted on whether they agreed to use the proposed stricter definition of a symptom: *departure from normal function, appearance or feeling which is noticed by the patient, indicating the presence of disease or abnormality*

	n	%
Agree	61	95.3
Disagree	2	3.1
Unsure	1	1.6
Total	64	

Therefore it was agreed that this second definition would be used for this consensus meeting.

Presentation 2.2: Laura von Kobyletzki - Eczema Signs and Symptoms: what is important to patients?

LvK gave an overview of the methods used in the survey which had been distributed to patients by members of the HOME group. Respondents were given a list of symptoms and asked to give each a rating on a 5-point Likert scale, ranging from "very important" to "not relevant to me" in response to the question "How important are these features in deciding whether or not a treatment is working?"

A total of 1104 usable responses were received from 35 countries, with the majority of responses coming from Europe and North America. It was noted that in some countries the survey was only available in English language format. A wide range of severity and skin colours were represented, although the majority were light or slightly coloured skinned.

Results showed that itch and pain were the symptoms that were most frequently rated as very or quite important. There were no significant differences between different disease severities.

Summary of whole group discussion

There was a short discussion about the survey results:

The survey was not designed to develop an instrument so didn't ask about symptoms in an open-ended way; instead the survey asked patients and carers to state which symptoms from a predefined list they felt were important (although respondents could enter "other" symptoms) to get a global overview.
The frequency and prevalence of individual symptoms may be critical in terms of content validity for a measurement instrument, but the survey did not address this issue.
Patients from different countries have different understanding of symptoms, so may have interpreted these questions differently.

Presentation 2.3: Louise Gerbens - A systematic review of how symptoms are reported in RCTs of eczema treatments

Most eczema treatment trials reported symptoms (295 out of 378 trials, 78%). Itch and sleep loss were most commonly reported, but often these were only reported as part of a total composite instrument score so the treatment effect on symptom burden frequently remained unclear. For instance, where SCORAD was used, the total score was reported in 66% of trials, with the itch and sleep-loss VAS reported separately in only 23%. VAS scales and numerical scales were most often used but there was a lot of variety in both of these. There was discussion around whether sleep loss is really a symptom or is part of the effect on quality of life. Also, the question of how symptoms could be rated in very young children was discussed and it was pointed out that it is often the parents and carers that can do this until the patient is old enough.

Summary of whole group discussion

The whole group discussed the construct symptoms, summarised below:

Patients gave their input by commenting on what they felt were the most important symptoms from all those identified through the survey, the patient pre-meeting and the review of symptoms in clinical trials and discussing their reasons why.

The inclusion of pain as a symptom raised a number of issues: i) There were differing views in terms of the presence of pain relating to eczema. Some felt that there was pain related to eczema, usually due to skin cracks, but others reported a general level of pain associated with their eczema, unrelated to skin cracking. Others felt that there was no pain involved in eczema at all. ii) It was thought that some people would interpret the simple term "pain" differently to others. iii) Is it possible for parents to judge pain in their very young children? iv) Most instruments for eczema don't measure pain.
The group needs to accept there will be no scale that measures ***all*** essential symptoms well – this would mean one instrument per symptom which would not be feasible for all trials. The group should focus on identifying the minimum essential symptoms to be included in the instrument and accept the limitations of how each essential symptom will be measured in the core outcome instrument i.e. strike a balance between the feasibility and validity of an instrument.
Some symptoms that have been identified are probably related and therefore describe the same underlying symptom. This overlap needs to be addressed when deciding what items are essential.

Presentation 2.4: Cecilia.A.C. (Sanna) Prinsen – The COSMIN checklist

Cecilia Prinsen (CP) presented summary overview of the COSMIN checklist that was developed to evaluate the methodological quality of studies, including the COSMIN definitions for the measurement properties, the taxonomy, and its scoring system. In addition, quality criteria to evaluate the adequacy of the measurement instruments (i.e., measurement properties) were discussed. It was discussed why COSMIN is used in core outcome instrument selection and how the levels of evidence are used to give each instrument a quality rating to select the best instrument.

Presentation 2.5: Phyllis Spuls – Systematic review of the measurement properties of instruments designed to capture eczema symptoms

PS presented the preliminary results of a recently completed systematic review of the measurement properties of eczema symptoms instruments. Validation and development studies were included if at least 50% of the patients used to test the instrument had eczema or data on eczema patients were presented separately. Studies were only included if the full text was available and purely linguistic validation studies were excluded. A total of 26 studies were included and 3 other language studies were awaiting data extraction. From the results of the systematic review, each instrument was assigned a rating of A, B, C or D as shown in the table below. This rating was developed previously for the signs domain consensus decision, and the purpose is to help prioritise which scales assessed in the systematic review should be the focus of the discussions. It takes into account the quality of the measurement instrument and the methodological quality of the validation studies, and is intended as a guide for the group, but does not dictate the output of the consensus discussions.

Rating	Definition	Recommendation
A	Symptom measurement instrument meets all required quality items.	Could be recommended for use.

B	Symptom measurement instrument meets two or more required quality items, but performance in all other required quality items is unclear.	Has the potential to be recommended in the future depending on the results of further validation studies.
C	Symptom measurement instrument has low quality in at least one required quality criteria.	Not recommended for use.
D	Symptom measurement instrument has very little validation work so the performance in all or most relevant quality items is unclear.	Not recommended to be used until further validation has been performed. Future recommendation would depend on the results of further validation studies.

PS explained that no A-rated instrument were available. This is in line with expectations and comparable with outcomes in other areas of medicine. Three instruments were rated "B":

- Itch severity scale (ISS)
- Patient Oriented Eczema Measure (POEM)
- Self-administered Eczema Area and Severity Index (SA-EASI).

All other instruments were rated C or D. More detail was given on how these ratings were achieved on the presentation and in the handouts available (which was also emailed to participants before the meeting).

Summary of whole group discussion

There was some clarification around the methodology used; i) where there is more than one study, the one with the best methodological quality is used to obtain the rating ii) validation of an instrument is for the whole measure and it is not appropriate to use one part of the measure and assume the validation still applies
Some instruments e.g. Patient-oriented SCORAD (PO-SCORAD) and self-assessed (SA) EASI are essentially asking patients to perform the clinical signs ratings rather than being a true measurement of patient reported symptoms.

Voting

After the discussion, the group voted on whether to exclude scales rated "C" as these have been demonstrated to not pass validation tests. Firstly, voting to establish stakeholder groups showed 49.2% were clinicians, 20.0% were patients, 12.3% were methodologists and 18.5% were pharmaceutical industry representatives.

The group then voted on: *Do we agree to exclude the measurement instruments of category C?*

	n	%
Agree	47	78.3
Disagree	5	8.3
Unsure	8	13.3
Total	60	

Therefore, it was decided that only instruments rated B or D would be considered for the core outcome set (there are no A rated instruments).

Summary of whole group discussion

There was discussion then around the construct symptoms during which a long list of symptoms was identified taken from the international survey (2.2), the systematic review (2.3) and patients / carers attending the meeting.

Voting

Voting to establish stakeholder groups showed 49.2% were clinicians, 18.0% were patients, 13.1% were methodologists and 19.7% were pharmaceutical industry representatives. The group then voted on whether this long list included all essential symptoms of eczema.

Are there any essential symptoms missing in the symptom domain?

	n	%
Yes	2	3.1
No	50	76.9
Unsure	13	20.0
Total	65	

Therefore, the group had agreed that all essential symptoms were present in the long list.

Small group discussions

The meeting then split into smaller breakout groups. Groups had been pre-determined to ensure there was a mix of stakeholder groups and countries in each group. Each group was asked to consider the following in their discussions:

1. Which symptoms are considered "essential" to be included from the long list of all symptoms.
2. Which is the preferred measurement instrument(s) taking into account content (important symptoms) and the validation of instruments.

Each group then presented the results of their discussions in turn to the whole group, and the symptoms that were deemed to be essential by at least one group are detailed in the table below:

Task 1: Which symptoms are considered "essential" to be included:

	Group						
	1[a]	2[a]	3[c]	4	5	6[d]	Total
Itch	X	X		X	X	X	5
Redness / inflamed skin	X	X		X	X	X	5
Irritation	X			X	X	X	4
Dry skin		X		X	X	X	3
Sleep loss	X			X	X	X	3

Tight skin	X						1
Sensitivity to hot and cold	X						1
Pain	X						1
Pigmentation / lichenification	X						1
Involvement of visible sites				X			1

[a] **Group 1** - Pigmentation / lichenification may be appropriate only for long term trials.
[b] **Group 2 -** Determined their three most important symptoms. Felt sleep-loss was a function of the other symptoms
[c] **Group 3 –** Rather than discuss which were the most important symptoms, the group used the long list to inform discussions around their preferred instrument(s).
[d] **Group 6 –** Considered pain and bleeding to be a consequence of scratching so not essential if itch is included. Considered sleep loss to be a consequence of other symptoms so not essential.

Task 2: preferred measurement instrument(s)

Preferred Instrument (by group) and reasons

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
POEM	POEM	POEM	(POEM)	POEM	POEM
Contains 4 of the 8 symptoms deemed essential by this group.	Preferred instrument.	Includes 7 of the important symptoms. Single instrument for children and adults. Doesn't include intensity but this may be sufficiently related to frequency.	Best validated instrument	Rated between 2 and 5. Highly relevant and simple to use. Lacks a measure of intensity and inflamed skin not included.	Very good instrument but some validation gaps.

Other instruments considered and reasons

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
ISS					
Ranked 3. Only 1 symptom. Sexual function question not relevant to all.		Only one symptom. Sexual function question may not always be appropriate.	Doesn't include essential symptoms so excluded.	Rated 7-8. Only covers itch.	Ranked 3. Sexual function question not always appropriate. More like a quality of life scale.
SA EASI					
Ranked 2		Essentially a patient assessed signs scale. Patients need training so not suitable for the core outcome set.	Needs more validation studies to be considered.	Rated 7-8 - complicated for patients and makes the patient an assessor of clinical signs.	Ranked 2. Different score for acute and chronic disease. Sleep loss not included. Body surface area is more appropriate for signs.
PO-SCORAD					
	Excluded due to the nature of the scale.	Liked VAS but validation studies are for the whole instrument so couldn't make a judgement or recommendation on VAS alone.			Too complicated for patients to use easily.
ADQ					
			Face validity good so could consider in future but further validation needed.		Mixes concepts.
NESS					
					Good for epidemiology studies but not felt to be suitable for clinical trials.

Summary of whole group discussion

The group then had further discussions on what should constitute the construct symptoms as it is essential to do this prior to deciding which instrument should be recommended.

Some of the symptoms listed including itch and sleep loss can only be assessed by the patient and ***cannot*** be measured by a clinician.
Should bear in mind the acute and chronic phases of eczema and instrument should cover all severities.
Opinion differed as to whether pain and irritation are function of other symptoms or symptoms in their own right. "Pain and soreness" was ranked second in the global symptoms survey (2.2) in importance, but the true prevalence of the symptom was not clear. Decided the group should vote on this issue. There are pain scales that could be considered for use.

Some patients felt unable to comment on the instruments that they had not used and some only felt able to discuss symptoms they had experienced personally.

Because of the high number of eczema symptoms and, with the exception of itch, there is huge variation between patients, it is difficult to determine a definitive list of essential symptoms. Also, there is overlap between them in terms of how they are interpreted by patients e.g. irritation covers tightness and soreness for some patients. Therefore, the group decided to discuss and then vote on the symptoms that were agreed by all the breakout groups to be the most important, and whether any further symptoms are essential.

Voting

Voting to establish stakeholder groups showed 49.2% were clinicians, 18.0% were patients, 13.1% were methodologists and 19.7% were pharmaceutical industry representatives. The group then voted on each symptom in turn.

*Is **itch** essential for symptoms domain?*

	n	%
Yes	64	98.5
No	0	0
Unsure	1	1.5
Total	65	

*Is **sleep loss** essential for symptoms domain?*

	n	%
Yes	48	73.8
No	12	18.5
Unsure	5	7.7
Total	65	

*Is **dryness** essential for symptoms domain?*

	n	%
Yes	52	78.8
No	8	12.1
Unsure	6	9.1
Total	66	

*Is **redness / inflamed skin** essential for symptoms domain?*

	n	%
Yes	43	65.2
No	11	16.7
Unsure	12	18.2
Total	66	

*Is **irritated skin** essential for symptoms domain?*

	n	%
Yes	38	59.4
No	14	21.9
Unsure	12	18.8
Total	64	

*Is **pain** essential for symptoms domain?*

	n	%
Yes	20	30.8
No	24	36.9
Unsure	21	32.3
Total	65	

*Is **more research needed** to understand what is meant by **pain** in eczema?*

	n	%
Yes	58	92.1
No	2	3.2
Unsure	3	4.8
Total	63	

In summary, itch, sleep loss, dryness, redness/inflamed skin and irritated skin are all essential for the construct symptoms. Pain was not considered essential, and it was agreed that further research is needed to fully understand what is meant by pain in eczema.

Summary of whole group discussion

The groups were then given the opportunity to voice concerns about the results of the voting and these are summarised below:

Although it is helpful to canvas the opinion of the group in this way, it should be noted that there have been published studies into which symptoms are essential to differentiate between disease states.
Only white skin is represented at this meeting – in the study by Charman et al, redness was not included in the instrument development because not able to detect this in dark skin. However, the vote here was for including red / inflamed skin.
Cracking of the skin has not been included in the vote and this may be an important symptom, but is covered by excoriation in the signs domain.
Sleep loss may be a consequence of other symptoms.
There was discussion about whether C rated instruments should remain excluded from consideration or whether this issue should be re-opened for discussion given the results of the voting on the construct and none of the B or D rated instruments contained all of the agreed symptoms.

Voting

Should we re-open the discussion on C rated instruments?

	n	%
Yes	16	25.4
No	42	66.7
Unsure	5	7.9
Total	63	

Therefore, it was agreed by voting that instruments rated "C" would remain excluded from consideration for the core set. The group will consider POEM, PO-SCORAD and SA-EASI.

The chair then drew the meeting to a close for the day.

Friday 24th April (08:30 – 16:00)

Session 2 – Symptoms domain (continued)

The group continued with the symptoms session, moving onto considering which instruments(s) could be recommended for the core set. To recap the information used the previous day on how to base the voting, a summary table of the three B rated instruments (POEM, PO-SCORAD, SA-EASI) was shown. The table detailed which items each instrument includes relative to the agreed essential symptoms list and measurement property ratings. Subjective SCORAD was not considered as it is only part of the full SCORAD and no validation data is available on the subjective elements only.

Members of the group who were conflicted with the scales being considered stood up so that it was clear to all meeting participants who were conflicted.

Voting

Voting to establish stakeholder groups showed 55.4% were clinicians, 15.4% were patients, 12.3% were methodologists and 16.9% were pharmaceutical industry representatives. The group then voted on each instrument:

*Is **POEM** is an adequate instrument to measure the domain of symptoms?*

	n	%
Yes	56	87.5
No	2	3.1
Unsure	6	9.4
Total	64	

*Is **PO-SCORAD** is an adequate instrument to measure the domain of symptoms?*

	n	%
Yes	10	15.4
No	40	61.5
Unsure	15	23.1
Total	65	

*Is **SA-EASI** is an adequate instrument to measure the domain of symptoms?*

	n	%
Yes	3	4.7
No	46	71.9
Unsure	15	23.4
Total		

Therefore, POEM was the only instrument voted as being an adequate to measure the domain of symptoms so no further discussion of other instruments or voting was required.

POEM is recommended for inclusion in the core outcome set for clinical trials to measure patient reported symptoms

Post meeting note

For details of how to use the POEM and downloadable POEM forms please visit http://www.nottingham.ac.uk/homeforeczema/resources.aspx

Summary of whole group discussion

The group were then given the opportunity to voice concerns and reflect on the decision that had been made, and these are summarised below:

There are no instruments that are rated as "A" so should there be more work to improve POEM? It is highly unlikely that an instrument will achieve an A rating because the COSMIN criteria on which the rating is based are very stringent. The COSMIN group have not seen any instrument that would meet the requirements to achieve an A rating from their work in many other areas of medicine.
Although structural validity of POEM has not been shown, the instrument generally meets the OMERACT filter of truth, discrimination and feasibility. However, it was agreed that the validation gaps for POEM should be addressed as per the HOME roadmap. If the results are negative and result in re-categorisation to C-rated, the group can either improve the instrument or, because the core set may evolve over time, POEM may be replaced with a better instrument if one is available. However, there should not be a delay in recommending the instrument until more studies are done, especially given that there are not many gaps in validation.
JS reminded the group that although POEM may not be perfect, it takes 8 to 10 years to develop a new instrument and it is better to have something in place even if it is an imperfect measure.

There was some concern over the speed of progress in the proceedings from considering instruments as potentially suitable to being accepted for the core set.
POEM has unclear cross-cultural validity and it is crucial that good quality translations are produced for use around the world. Experience from OMERACT suggests that once an instrument is widely used then translations will follow.

The symptoms session was then brought to a close.

Summary of symptoms session

Symptoms are usually measured in trials but often reported separately from other domains
There are a large number of symptoms, not all experienced by all patients, and sometimes only experienced at certain times by an individual patient. Itch, however, is the universal symptom of eczema.
A systematic review identified a number of instruments to measure symptoms that had the potential to be recommended for inclusion in the core set.
Itch, sleep loss, dryness, redness/inflamed skin, and irritated skin are all essential for measuring eczema.
Further work is required to establish how pain is part of the symptoms of eczema.
The POEM scale was voted as adequate for measuring the symptoms of eczema in clinical trials and therefore is to be recommended for inclusion in the core outcome set.
Some further validation work is needed on the POEM scale.

Session 3 – Quality of Life domain (Chair: Hywel Williams & Jas Singh)

Presentation 3.1: Christian Apfelbacher - Introduction

Christian Apfelbacher (CA), lead for the Quality of Life (QoL) working group opened the session by introducing some conceptual considerations.

- QoL can mean different things to different people. An individual's QoL level is related to their expectations and so it is not possible to assign an absolute value.
- There is a response shift in QoL because people with a chronic disease tend to adapt their expectations over time.
- QoL is multi-dimensional but the relevance of some subdomains of QoL may not be universal; for instance, relevance of spirituality may depend on culture.
- There are three types of QoL instruments; generic, skin specific and disease specific.
- Wilson and Cleary is a widely used conceptual framework.
- Most QoL instruments measure functional limitations, but there are some in which a needs based approach is alternative to this i.e. life gains quality when needs are fulfilled.
- Proposed a definition of QoL.

Summary of whole group discussion

There followed a discussion which is summarised below:

The development of QoL instruments has become more sophisticated over time; instruments are now usually developed using conceptual models whereas this was not the case for older instruments.
Important to let patients guide this topic as the construct of QoL is centred wholly around patients.
Acknowledged that deciding on a QoL instrument for the core outcome set is more challenging than for the symptoms domain and the group may have to decide whether it is preferable to adopt a scale for use immediately or spend several years developing an improved scale using modern theory and techniques.
Computer adaptive testing versus classical static paper questionnaire is not ready yet but should be considered in the future.

Presentation 3.2: Daniel Heinl - Systematic review of how quality of life is measured in eczema clinical trials

This scoping review was carried out in studies on adults and children. A total of 22 different instruments were identified. Most of these were skin specific, with a few generic and eczema specific instruments. Approximately 1 in 5 clinical trials in eczema report on quality of life.

Summary of whole group discussion

After the presentation, there was opportunity for questions and discussion, which mainly focussed on the differences and similarities between adults and children, summarised below:

The QoL working group has completed the systematic review on validation of QoL scales for adults, but the equivalent for children is yet to be done. It was made clear to the group that this means that data was only available on the measurement properties of QoL instruments in adults for the meeting discussions. HW confirmed with the group that they were happy to proceed in this knowledge.
Some felt uncomfortable discussing scales for children with no data available, but it was agreed it would be helpful to at least get patients input on face validity for scales in children during this meeting. There were mixed views from patients as to whether QoL issues were sufficiently similar in children and adults or whether there was little overlap.
There needs to be discussion around whether an eczema specific instrument is needed or whether a dermatology specific instrument is sufficient.
Important to ensure progress is made rapidly on the children scales after this meeting.

The group then moved on to discuss what subdomains of QoL are essential. HW summarised that the subdomains so far as stated in the introduction presentation (3.1) were:

- Psychological
- Functional
- Social

The patient representatives were asked for their views on what about their eczema affects their QoL. It was pointed out that the group should remember that it can be difficult for patients to talk about their personal feelings and relationships in front of a large group like this.

The issues that patients raised as affecting their quality of life were:

Can't go out to play
Can't do what you want to do
Have to think about your skin condition 24 / 7 and it can dominate your life.
Awareness that you have to do things all the time
Time lost looking after skin
Allergies can limit where you can go and can stops you socialising with friends e.g. can't visit a friend because they have animals in the house.
Prejudice from other people e.g. people won't shake hands when say goodbye.
When seasons change have to anticipate the effect on eczema and do things you wouldn't normally have to do.
Feeling guilty because of effect on other people e.g. not being able to do an activity.
Interferes with choice of activities.
Causes arguments with partner because of limitations on what you can do.
When have eczema on the face it is very visible and affects how you feel about yourself.
Females can be more affected by eczema on face and they can't wear cosmetics.
Perfect skin is "normalised" nowadays.
Don't want to be a burden to people but feel they are sometime
Affects self esteem
Depression
Loss of intimacy with other people
Feel stigmatised
Affects job choice and prospects
There are financial considerations
Affects what clothing they can wear
Affects people differently at different life stages

Presentation 3.3: Daniel Heinl - systematic review of measurement properties of QoL instruments in adults

Daniel Heinl explained that the protocol was published and registered with PROSPERO. He summarised the methods used to conduct this review of the measurement properties of QoL in adults and how each instrument is given a rating of A, B, C or D (more details of how these ratings are derived can be found in the symptoms section of this report). A total of 15 articles were included. He then explained the rating that each instrument had been given and the reasons for that rating. He made the point that although DLQI is widely used, it only has a C rating because content validity and structural validity both obtained a negative rating.

Degree of recommendation	Instrument(s)
A	None
B	English QoLIAD (UK) English QoLIAD (US) French QoLIAD German QoLIAD Spanish QoLIAD
C	English DLQI (UK) ISDL Dutch QoLIAD
D	DIELH Danish DLQI German DLQI Spanish DLQI FLQA-c FLQA-d Italian QoLIAD German Skindex-29

Summary of whole group discussion

The group then discussed these results and several issues were raised, summarised below:

Different language versions were treated separately because the measurement properties relate to the data, not the instrument itself, so not methodologically sound to collapse the results.
Only validation or development studies were included in the review, not indirect evidence such as responsiveness collected in trials.
There are published studies on interpretability for DLQI not included in the review but these were excluded because the population was not at least 50% eczema patients or data on the eczema not presented separately
Validation studies need to be done in the different languages.
The number of questions in the instrument is important for feasibility.

Small group discussions

The meeting then split into the same smaller breakout groups as for the symptoms session. Each group was asked to consider the following in their discussions:

1. What are the essential subdomains for the domain QoL.
2. Which is the preferred measurement instrument(s).

Each group then presented the results of their discussions in turn to the whole group, collated in the tables below:

Task 1: Which aspects of Quality of Life are "essential" to be included:

	Group						
	1 [a]	2	3 [b]	4	5 [c]	6	Total
Emotions	X		X	X	X	X	5
Treatment burden	X		X		X	X	4
Personal relationships	X		X	X	X		4
Impact of symptoms	X			X		X	3
Work / study	X		X		X		3
Activity daily living	X		X				2
Leisure	X		X				2
Coping	X				X		2
Physical functioning		X				X	2
Social functioning		X				X	2
Appearance / visibility	X						1
Psychological functioning		X					1
Stigma				X			1

[a] **Group 1** – Severity not included as symptoms are a separate domain.

[b] **Group 3** – Considered treatment burden to include time and cost. Considered that other subdomains including appearance should be covered by those domains if worded appropriately.

[c] **Group 5** – Regarding the effect on work, it is often other peoples prejudice rather than a true inability to do the work.

Task 2: preferred measurement instrument(s)

Preferred Instrument (by group) and reasons

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
DLQI	**QoLIAD**	**DLQI**	**-**	**-**	**(DLQI)**
Preferred instrument but would like more validation data. Question on sexual difficulties not always appropriate.	Preferred instrument. Concerned about responsiveness because of dichotomous answers and all weighted equally. But questions have good relevance to eczema patients e.g. question about stopping doing things.	More support for DLQI but felt unable to recommend because "C" rated. However, need to fully understand why this is. Content and face validity good. Widely used already.	Didn't feel able to recommend an instrument.	Group did not state a preferred instrument.	Could not recommend based on data presented but very widely used in trials so should use existing datasets to improve validation.

Other instruments considered and reasons

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
DLQI					
	Some of the group favoured DLQI. Less relevant to patients. Sexual difficulties questions may be difficult for some patients.		Further validation data needed to be able to recommend.	Good face validity and patients felt questions were relevant and sensible. Liked that it measures intensity.	
QoLIAD					
Aggressive tone and depressing to complete.		Too negative and didn't like this instrument. Focussed on emotions and could be upsetting to complete. Dichotomy is not good. Patients in group didn't think this instrument is acceptable therefore rejected.	Could accept this as the recommended instrument if it is shown to be better than the other scales or if it can be improved. But doesn't cover all the subdomains.	Validation suggests QoLIAD better than DLQI and may perform better in trials, but lack responsiveness is crucial so wouldn't recommend. No measure of intensity and content more limited – focussed on emotions (e.g. no measure of impact on work).	Reflects the preferred subdomains well but dichotomous responses will miss changes. So wouldn't recommend for the core set.
Skindex					
	Not enough time to also discuss Skindex	Hard to look at because didn't have copy of the scale. Skindex 29 too long but 16 ok – more validation would be needed. Couldn't reject without seeing the scales. needs more validation.	Not enough information / data	Couldn't evaluate not much validation data and not available to see	
DIELH					
			Much more validation work needed before consideration for the core outcome set.		

Summary of whole group discussion

There was discussion across the whole group regarding the feedback, particularly around the reasons why DLQI was not rated more highly than a C:

- ***Content validity*** - because DLQI is a skin specific rather than an eczema specific instrument, it was developed with patients with many different skin diseases. One study found that there were not enough items for patients with mild eczema and that there was a high proportion of "not relevant" answer to DLQI items 6,7 and 9 (>30%, almost 20%, >20%, respectively) in eczema patients.
- ***Structural validity*** - items work differently in different subgroups which is why it was rated negative (5 of 10 items had differential item functioning so the score may not mean the same for different genders and different age groups). The DLQI fits a Rasch model for eczema patients, although item residual statistics were indicative of model misfit for eczema patients.

Andrew Finlay, originator of the DLQI, stated that he had decided not to vote on the instrument recommendation due to his significant conflict of interest.

Voting

Voting to establish stakeholder groups showed 56.7% were clinicians, 13.3% were patients, 10.0% were methodologists and 20.0% were pharmaceutical industry representatives. The group then voted on what subdomains are essential for the construct QoL.

*Is **psychological functioning** an essential subdomain for the domain QoL?*

	n	%
Yes	57	95
No	1	1.7
Unsure	2	3.3
Total	60	

*Is **social functioning** an essential subdomain for the domain QoL?*

	n	%
Yes	60	100
No	0	0
Unsure	0	0
Total	60	

*Is **physical functioning** an essential subdomain for the domain QoL?*

	n	%
Yes	57	95
No	0	0
Unsure	3	5
Total	60	

*Are there **any other essential subdomains missing** for assessing QoL in atopic dermatitis?*

	n	%
Yes	9	15
No	39	65
Unsure	12	20
Total	60	

*Is **DLQI** an adequate instrument to measure the domain of QoL in AD clinical trials?*

	n	%
Yes	27	45
No	21	35
Unsure	12	20
Total	60	

*Is **QoLIAD** an adequate instrument to measure the domain of QoL in AD clinical trials?*

	n	%
Yes	15	25.9
No	34	58.6
Unsure	9	15.5
Total	58	

*Is there sufficient information to vote on **Skindex** today?*

	n	%
Yes	1	1.7
No	56	94.9
Unsure	2	3.4
Total	59	

Therefore, although the DLQI was the preferred potential instrument for this group, fewer than 70% were in agreement and therefore no QoL instrument for adults can be recommended for the core outcome set at this stage.

Summary of whole group discussion

The group discussed two options for how to proceed:

Option 1: Further work on testing QoLIAD
Option 2: Conditional recommendation for DLQI if the issues are fixable

Regarding option 2, CA made some comments about what might be appropriate:

- ***Content validity*** – if the content of the instrument is changed, the measurement properties may change and so the validation studies need to be repeated.
- ***Structural validity*** – only one study which may not be enough (although if it is of excellent methodological quality according to COSMIN, then one may be enough). However, given that structural validity problems had also occurred in studies on other skin diseases, it seems unlikely that there will be no problems in further studies in eczema.
- ***Responsiveness*** – DLQI is known to be very responsive, but this is in ALL skin conditions. The evidence for eczema specifically is not available (other than for the Spanish version).

Given the importance of responsiveness, the group questioned why the validation studies of the Spanish DLQI can't be extrapolated to other languages. CA explained that it cannot be assumed that the measurement properties remain exactly the same after translation. Validation data should be generated for different language versions. The underlying reason is that in fact measurements, not the instruments themselves, are valid, reliable and responsive (or not).

The point was made that being able to compare the effect on QoL in eczema with other skin conditions and being able to compare new studies with older studies are both important issues that the group need to take into account. Additionally, the length of time it takes to develop a new instrument, and the need to compare it to the currently widely used DLQI, should be factored into any decision when voting on an instrument.

There are some methodological issues around using effect size for measuring responsiveness.

Andrew Finlay, the originator of the DLQI, then left the room to allow an open discussion to include any shortcomings in the DLQI. Issues raised were:

Whilst the content validity of DLQI has been established as acceptable, there are some issues around the structural validity, particularly the redundancy of items or items that may be not applicable.
It is a principle of questionnaire development that each question should only ask one specific thing, whereas the DLQI asks about multiple things in one questions e.g. Question 3 asks about shopping, home and garden.
Concerns about the cultural validity as the scale was developed in the UK.
Content validity was thought to be good until studies showed that at least 3 different items are not relevant. For some this meant that DLQI could not be recommended.
Should the QoL scale be eczema specific or skin specific? A skin specific instrument would allow comparison with other diseases, and could be adequate providing it is adequately tested and performs well. However it should be remembered that comparison with other diseases is not the main purpose of the core outcome set and the performance of the instrument in measuring eczema is key.
Both DLQI and QoLIAD have children's versions so the need for use in all ages should not affect the decision between the two instruments.
The sensitivity of the instrument is crucial because if the instrument is not sensitive to change then patients could be denied a new treatment. Responsiveness it probably more important than structural validity.

Because the discussion showed there was clearly a lot of support for DLQI, the group were asked to decide whether the DLQI should be opened up to the vote again, perhaps allowing a vote on whether a conditional recommendation should be made.

Voting

Do we want to vote again on DLQI ?

	n	%
Yes	21	37.5
No	27	48.2
Unsure	8	14.3
Total	56	

Therefore, there was not enough support to reopen the vote on DLQI. Consensus on an instrument for QoL in adults was not reached.

The session on QoL was then brought to a close.

Summary of QoL session

QoL is difficult to define and often instruments lack underlying conceptual models.
A systematic review identified a number of generic, dermatology specific and eczema specific QoL instruments that have previously been used in eczema trials
A systematic review found that no existing instrument can be recommended without further validation work
Essential subdomains of a suitable QoL instrument for eczema include physical, psychological and social functioning.

Consensus was not achieved for any proposed instrument
The DLQI was preferred above others but there were significant problems with validity (both content and structural)

Session 4 – Long Term Control domain – Chair: Jochen Schmitt

Presentation 4.1: Kim Thomas - Introduction

Kim Thomas (KT) opened the session by explaining that the first step is to clarify what is meant by this domain: is it "long-term control (LTC)" or "long-term control of flares"? Is LTC really a separate domain, or repeated measurement of other core outcomes? It was previously agreed at HOME III that the LTC domain applies only to trials of more than 3 month duration.

There was only a short amount of time available for this session so rather than opening up for a long discussion, KT asked all participants to complete a questionnaire asking for their opinion which were collected at the end of the meeting.

Presentation 4.2: Sébastien Barbarot - How has long term control been captured in randomised controlled trials of eczema treatments?

Sébastien Barbarot presented a systematic review that had recently been carried out by members of the LTC working group. A recent paper showed that measuring flares is potentially a good way to capture LTC (Langan et al. *Br J Dermatol* 2014). However, there are some drawbacks; i) the threshold of what is considered a flare can greatly affect the number of flares, ii) need for frequent data collection to reflect the rapid changes seen in eczema and iii) patients can have moderate to severe eczema but without any flares so simply measuring flares would suggest their eczema is controlled. Other ways to measure LTC could include the use of standard medication or repeated measurement of other outcomes. Results of the review showed that most long term studies use other repeated outcome measures of disease severity (usually monthly clinician reported outcomes assessing disease severity), whereas less than a third used either flare data or standard medication use to assess LTC. When analysing long term outcomes, approximately 40% of studies did not take into account all time points. Sébastien concluded by reiterating that many different ways of measuring LTC have been used in trials.

Presentation 4.3: Andreas Wollenberg - Thoughts on long term control - Is LTC a separate domain or a function of the other 3 domains?

Andreas Wollenberg opened by suggesting the only measure of LTC that can be considered a truly separate domain is flares. He presented "clinically meaningful worsening of signs and symptoms of AD leading to therapeutic intervention" as a proposal for a flare definition and discussed some considerations for the use of flares as a measure of LTC; i) collecting time to first flare is relatively straightforward but makes sense only if active and control group are analysed, ii) collecting the number of flares needs to be much longer, and iii) the disease severity of the study population and the duration of the trial will affect the number of patients who experience a flare. Andreas proposed presenting flare data in a Kaplan Meier (KM) plot in which patients are removed as they experience a flare. He then proposed his own measure called "Fixed time profit" which was a measure of the

number of patients who have benefitted from the treatment after a fixed time using time to first flare. AW stated that this parameter is clinically meaningful, extractable from most published KM-plots and better controlled than the percentage of patients who experienced a flare.

Summary of whole group discussion

The group were reminded that the need to be able to measure the effects of potential proactive treatments (i.e. those aimed at preventing flares) was a contributing factor in the decision to include LTC in the core outcome set.
There are some potential limitations in using time to first flare that would need to be investigated before considering it as a core outcome measure including; i) is time to first flare predictive of the severity of eczema ii) how representative is a short study (i.e. time to first flare) for long term control and iii) would it be an appropriate measure for very mild disease?
Should the outcomes be different for during treatment and post treatment follow up?
There was a reminder that methods other than the number of flares have been used to capture LTC and all should be considered, but some felt strongly that the only truly distinct domain is a count of flares. It should also be remembered that researchers are free to include repeated measures of other domains such as signs and symptoms in trials if flares was the core outcome measure.
There was some support for the proposal of expressing time to first flare in KM plots
Behavioural changes, such as seeking medical help, can be a good outcome measure but they may not accurately reflect disease severity due to the relationship with adherence and practical issues.
Self- management and use of rescue medicine can be a good measure, although there can be a delay for patients of several days in getting the medication.
Future work could be a study of content validity to establish what patients think is important.

Voting

Voting to establish stakeholder groups showed 64.3% were clinicians, 9.5% were patients, 9.5% were methodologists and 16.7% were pharmaceutical industry representatives. The group then moved on to voting:

Do you think that long term control should be measured as a separate unique construct (e.g. flares) or can it be captured using repeated measurement of one of more of the other 3 domains?

	n	%
Flares	14	32.6
Repeated	7	16.3
Both	15	34.9
Other	2	4.7
Unsure	5	11.6
Total	43	

Therefore there no consensus was reached and this issue needs further discussion at HOME V.

The results of the questionnaires (n-=23) completed by the group at the end of the session were assessed after the meeting and this resulted in identification of the following next steps for the LTC domain:

1. Qualitative work to establish what long-term control means to patients (with reference to the existing qualitative literature).
2. An e-Delphi consensus study to agree what the domain of LTC should be, prior to HOME V.

Other projects of interest were validation studies of LTC outcomes and exploration of frequency of data collection.

The LTC session was then brought to a close.

Meeting close

Participants were made aware of the work to disseminate the outcome of the previous HOME meeting (HOME III) with publications, an EASI manual, EASI video, and EASi app available on the HOME website.

Hywel Williams then thanked everyone for coming and for their valued contributions to the meeting. Much had been achieved in terms of accepting a new core instrument for symptoms, but much remained to be done in terms of developing or further testing of instruments for QoL, and more conceptual work needed to be done on long term control. He then drew the meeting to a close.

Funding

There was no charge to attend the meeting, but all participants were required to cover their own travel and accommodation costs. Expenses for patients and patient group representatives were met by either an eczema patient association, a HOME member or by a donation from a pharmaceutical company. Travel expenses for pharmaceutical company employees were met by their own company. Hywel Williams supported the travel and accommodation costs for Jas Singh to attend, representing OMERACT.

The local organisers (Åke Svennson and Laura von Kobyletzki) used an unrestricted educational grant from the LEO Foundation plus contributions from the Swedish Asthma and Allergy Foundation and the County of Skåne to support the local meeting arrangements.

Appendices

Appendix 1: HOME IV break-out groups

GROUP 1		
Name	**Stakeholder Group**	**Country**
Hywel Williams (facilitator for quality of life discussion)	Clinician	UK
Dedee Murrell (facilitator for symptoms discussion due to Hywel Williams conflict of interest)	Clinician - Dermatology	Australia
Valeria Aoki	Clinician - Dermatology	Brazil
Julie Block	Patient/Carer/Patient Representative	USA
Lykke Graff	Pharmaceutical Industry Representative	
Burchard Marquort	Patient/Carer/Patient Representative	Sweden
Kristine Nograles	Pharmaceutical Industry Representative	
Yukihiro Ohya	Clinician - Paediatrician	Japan
Jasvinder Singh	Methodologist	USA
Anne Sulzer	Pharmaceutical Industry Representative	France
Helle vestby Talmo	Patient/Carer/Patient Representative	Norway
Elke Weisshaar	Clinician - Dermatology	Germany

Group 2		
Name	**Stakeholder Group**	**Country**
Christian Apfelbacher (facilitator)	Methodologist	Germany
Katrina Abuabara	Clinician - Dermatology / Methodologist	USA
Marius Ardeleanu	Pharmaceutical Industry Representative	USA
Tim Burton	Patient/Carer/ Patient Representative	UK
Amanda Creswell-Melville	Patient/Carer/Patient Representative	Canada
Laurent Eckart	Pharmaceutical Industry Representative	France
Takeshi Nakahara	Clinician - Dermatology	Japan
Ibrahim Nasr	Clinician - Dermatology	
Marie-Louise Schuttelaar	Clinician - Dermatology	Netherlands
Tracey Sach	Methodologist (Health Economist)	UK
Annika Volke	Clinician - Dermatology	Estonia
Carl-Fredrik Wahlgren	Clinician - Dermatology	Sweden
Stephan Weidinger	Clinician - Dermatology / Molecular epidemiology	Germany

Group 3		
Name	**Stakeholder Group**	**Country**
Kim Thomas (facilitator)	Methodologist	UK
Maren Awici-Rasmussen	Patient/Carer/Patient Representative	Norway
Sebastien Barbarot	Clinician - Dermatology	France
Linda Beckman	Other - Researcher	Sweden
Anthony Bragg	Pharmaceutical Industry Representative	
Rosemary Humphreys	Patient/Carer/Patient Representative	UK
Yoko Kataoka	Clinician - Dermatology	Japan
Yael Lesham	Clinician - Dermatology	USA
Bronwyn Lund	Pharmaceutical Industry Representative	
Hiroyuki Murota	Clinician - Dermatology	Japan
Florent Torchet	Patient/Carer/ Patient Representative	France
Laura von Kobyletzki	Clinician - Other (General practitioner)	Sweden
Andreas Wollenberg	Clinician - Dermatology	Germany

Group 4		
Name	**Stakeholder Group**	**Country**
Phyllis Spuls (facilitator)	Dermatologist	Netherlands
Maj Dinesen	Pharmaceutical industry Representative	
Aaron Drucker	Clinician - Dermatology	Canada
Andrew Finlay	Clinician	UK
Louise Gerbens	Clinical - other MD PhD student - dermatology	Netherlands
Daniel Heinl	Student of Medicine	Germany
Marie-Anne Massuel	Pharmaceutical Industry Representative	France
Stephanie Merhand	Patient/Carer/Patient Representative	France
Jevgenija Smirnova	Clinician - Junior Doctor	Sweden
Åke Svensson	Clinician - Dermatology / Molecular epidemiology	Sweden

Group 5		
Name	**Stakeholder Group**	**Country**
Eric Simpson (facilitator)	Clinician	USA
Carsten Flohr	Clinician - Dermatology	UK
Henrique Ishii	Patient/Carer/Patient Representative	Brazil
Teresa Løvold Berents	Clinician - Dermatology	Norway
Ian Osterloh	Pharmaceutical Industry Representative	UK
Cecilia (Sanna) Prinsen	Clinical Epidemiologist, Methodologist	Netherlands
Lynn Purkins	Pharmaceutical Industry Representative	
Shoko Shindo	Clinician - Dermatology	Japan
Eli Synnove Gjerde	Patient/Carer/Patient Representative	Norway
Roberto Takaoka	Clinician - Dermatology	Brazil
Cathy Zhao	Clinician - Dermatology	
Jan Pander	Pharmaceutical Industry Representative	Netherlands

Group 6		
Name	**Stakeholder Group**	**Country**
Jochen Schmitt (facilitator)	Clinician - Dermatology	Germany
Madhur Garg	Pharmaceutical Industry Representative	
Jon Hanifin	Clinician - Dermatology	USA
Hitoshi Mizutani	Clinician - Dermatology	Japan
Matthew Ridd	Clinical – (GP) / academic	UK
Marie Tauber	Clinician - Dermatology	France
Willem Kouwenhoven	Patient/Patient Representative	Netherlands
Kosuke Yamaga	Clinician - Dermatology	Japan
Kim Katrine Clemmensen	Clinician - Dermatology	Denmark

Appendix 2: Conflicts of interest declared

Sebastien Barbarot	PO-SCORAD
Aaron Drucker	Burden of Disease in AE (BODE)
Andrew Finlay	DLQI, CDLQI, DFI, FDLQI, FROM-16, EDI, IDQoL
Jon Hanifin	EASI
Yoko Kataoka	ADQoL-J (Japan)
Yael Leshem	Investigator Global Signs Assessment (IGSA)
Hitoshi Mizutani	Scratch meter
Yukihiro Ohya	Japanese version of POEM, DLQI, CDLQI, FDI, IDQOL, QPCAD, PQCAD short form
Lynn Purkins	Ziarco Itch Diary
Marie-Louise Schuttelaar	QoLHEC
Elke Weisshaar	VAS and questionnaires
Hywel Williams	POEM
Andreas Wollenberg	PO-SCORAD

References

The Harmonising Outcomes for Eczema (HOME) roadmap: A methodological framework to develop core sets of outcome measurements in dermatology. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, Chalmers J, Williams HC. J Invest Dermatol 2015 Jan;135(1):24-30. doi: 10.1038/jid.2014.320. Epub 2014 Sep 4

The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. Schmitt J[1], Spuls PI[2], Thomas KS[3], Simpson E[4], Furue M[5], Deckert S[6], Dohil M[7], Apfelbacher C[8], Singh JA[9], Chalmers J[3], Williams HC[3]; HOME initiative collaborators. J Allergy Clin Immunol. 2014 Oct;134(4):800-7. doi: 10.1016/j.jaci.2014.07.043.

Website: www.homeforeczema.org