# Minutes of the HOME III Meeting
## 6-7 April 2013, San Diego, USA

## Attendees

| | | |
|---|---|---|
| Shehla Admani | Dermatologist | USA |
| Valeria Aoki | Dermatologist | Brazil |
| Christian Apfelbacher | Methodologist | Germany |
| Marius Ardeleanu | Pharmaceuticals Industry Representative | USA |
| Sebastien Barbarot | Dermatologist | France |
| Tim Berger | Dermatologist | USA |
| James Bergman | Dermatologist | Canada |
| JulieBlock | Patient Advocacy Organisation | USA |
| Nicola Borok | Nurse | USA |
| Tim Burton | Consumer | United Kingdom |
| Joanne Chalmers | Methodologist | United Kingdom |
| Sarah Chamlin | Dermatologist | USA |
| Therese Cosan | Nurse | USA |
| Stefanie Deckert | Methodologist | Germany |
| Cynthia DeKlotz | Dermatologist | USA |
| Magdalene Dohil | Dermatologist | USA |
| Lawrence F Eichenfield | Dermatologist | USA |
| Holli Fultz | Pharmaceutical Industry Representative | USA |
| Lykke Graff | Pharmaceutical Industry Representative | Denmark |
| Jon Hanifin | Dermatologist | USA |
| Adelaide Hebert | Dermatologist | USA |
| Rosemary Humphreys | Consumer | United Kingdom |
| Raegan Hunt | Dermatologist | USA |
| Christelle Jost | Consumer | USA |
| Yoko Kataoka | Dermatologist | Japan |
| Norito Katoh | Dermatologist | Japan |
| Renata Kisa | Pharmaceutical Industry Representative | USA |
| Sara Kuppuswami | Consumer | USA |
| Kathy Langevin | Dermatologist | USA |
| Saravanapriva Loganathan | Consumer | USA |
| David Margolis | Dermatologist | USA |
| Stephanie Merhand | Consumer | France |
| Rebecca Minnillo | Professional (Research) Society | USA |
| Hitoshi Mizutani | Dermatologist | Japan |
| Kaspar Mossman | Consumer | USA |
| Helen Nankervis | PhD Student | United Kingdom |
| Yukihiro Ohya | Paediatrician | Japan |
| Pamela Rodgers | Pharmaceutical Industry Representative | USA |
| Jochen Schmitt | Dermatologist/Methodologist | Germany |
| Mandy Schram | Dermatologist | Netherlands |
| Elaine Siegfried | Dermatologist | USA |
| Eric Simpson | Dermatologist | USA |
| Jas Singh | Rheumatologist | USA |
| Phyllis Spuls | Dermatologist | Netherlands |
| Jean-Francois Stalder | Dermatologist | France |
| Ake Svensson | Dermatologist | Sweden |
| Roberto Takaoka | Dermatologist | Brazil |
| Ariel Teper | Pharmaceutical Industry Representative | USA |
| Kim Thomas | Methodologist | United Kingdom |
| Wynnis Tom | Dermatologist | USA |
| Tom Volkman | Consumer | USA |
| Laura von Kobyletzki | Dermatologist | Sweden |
| Elke Weisshaar | Dermatologist | Germany |
| Hywel Williams | Dermatologist | United Kingdom |
| Gil Yosipovitch | Dermatologist | USA |
| Susan Zelt | Pharmaceutical Industry Representative | USA |

# Glossary

| | |
|---|---|
| AD | Atopic Dermatitis |
| AE | Atopic Eczema |
| AUC | Area under the Curve |
| BSA | Body Surface Area |
| CADIS | Childhood Atopic Dermatitis Impact Scale |
| CDLQI | Children's Dermatology Life Quality Index |
| COMET | Core Outcome Measures for Effectiveness Trials |
| COSMIN | Consensus Based Standards for the Selection of Health Measurement Instruments |
| DFI | Dermatitis Family Impact |
| DLQI | Dermatology Life Quality Index |
| EASI | Eczema Area and Severity Index |
| EQ-5D | EuroQoL – generic health-related quality of life measure |
| FDA | Food and Drug Administration |
| GREAT | Global Resource of Eczema Trials |
| HOME | Harmonising Outcome Measures for Eczema |
| HrQoL | Health Related Quality of Life |
| IDQoL | Infants' Dermatitis Quality of Life Index |
| IGA | Investigator Global Assessment |
| MAcAD | Trial comparing methotrexate with azathioprine in adult patients with severe Atopic Dermatitis |
| MCID | Minimal Clinically Important Difference |
| OMERACT | Outcome Measures in Rheumatology |
| PASI | Psoriasis Area and Severity Index |
| POEM | Patient–oriented Eczema Measure |
| PRO | Patient reported outcome |
| PROMIS | Patient-reported Outcomes Measurement Information System |
| PROVE | Trial comparing prednisolone with cyclosporine in adult patients with severe Atopic Dermatitis |
| 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 |
| SASSAD | Six Area, Six Sign Atopic Dermatitis |
| SCORAD | SCORing Atopic Dermatitis |
| Skindex | AD-specific QoL measure |
| SWET | Softened Water Eczema Trial |
| SSS | Simple Scoring System |
| TIS | Three Item Severity score |
| VAS | Visual Analogue Scale |
| VRS | Verbal Rating Scale |
| WCW | Well Controlled Weeks |

# Saturday 6th April (1pm – 5pm)

## Session 1

### Welcome

Professor Larry Eichenfield (LE) opened the meeting by welcoming the group to San Diego. Professor Hywel Williams (HW) added his welcome and then asked everyone to introduce themselves. The attendees of the meeting were a mix of people who had attended previous HOME meetings and new members. Most had been involved with designing or recruiting into clinical trials (86%). About half of attendees had experience using SCORAD and EASI but few had used TIS or SASSAD. Attendees were asked to indicate which stakeholder group they felt best described them.

| Which of these subgroups best describes you? | % |
|---|---|
| Clinician | 55 |
| Patient / patient representative / carer | 11 |
| Methodologist | 19 |
| Regulatory agency representative | 0 |
| Pharmaceutical Industry representative | 15 |

The previously agreed consensus rules were applied to voting at this meeting:

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

### Presentation 1.1: Introduction and Background

#### Hywel Williams

HW presented the background to HOME and why core outcomes are such an essential part of good clinical research. He then summarised the results of the Delphi exercise and previous HOME meetings. HW explained the HOME roadmap highlighting the steps needed to progress each HOME workstream. HW mentioned that we should look to OMERACT to guide our process as they are 20 years ahead of HOME. He thanked Dr Jas Singh from the OMERACT Group for coming to HOME III as an external advisor to help keep the meeting focussed and fair.

HW acknowledged that the majority of HOME is being conducted on a voluntary basis, then summarised the aims of the HOME III meeting which were:

- To discuss and interpret new research since HOME II from the four working groups.
- To make decisions about which tools should be used to measure the essential four domains.
- To prioritise topics for further research.

HW stressed that it was crucial that all present needed to put aside preferences and allegiances and work together as a global community for the benefit of patients.

## Presentation 1.2: The Gift of the OMERACT filter

### Jasvinder Singh

Dr Jasvinder Singh (JS) explained why OMERACT was needed, highlighting the similarities between the situation in rheumatoid arthritis research 20 years ago and the situation with atopic eczema now. The remit of OMERACT has now been extended to include many other inflammatory arthritis conditions.

Consensus should be viewed as "something you can live with". It may not be perfect but think "is it preferable to the status quo?" In the interests of achieving consensus the group must consider the question "can we live with it?" when voting. He stressed that it is important to listen to the voices of dissent in the consensus process.

JS made the point that once an instrument has been included in the core set it is relatively easy to collect a lot of data on that instrument as it becomes more widely used. The core outcomes can be the primary or secondary outcome but be aware that it may be underpowered if it is not the primary outcome. The core set can be updated to reflect new findings about disease or if a new stakeholder group is added. The measures included in the core set can also be different for children and adults.

JS then moved on to present details of an update on the existing OMERACT filter. Version 2.0 has a broader conceptual basis than the previous version. The concept of filter 2.0 was shared in confidence with the HOME III meeting as it is not yet published. The OMERACT group has not yet applied this new version of the filter to an outcome set.

## Presentation 1.3: Feasibility in all settings

### Kim Thomas

Dr Kim Thomas (KT) highlighted that core outcomes need to be feasible in all settings and trial designs ranging from early phase trials with intensive follow up to large multi-national pragmatic trials with minimal follow-up.

The main aspects of feasibility to be considered are:

- The outcome set needs at least one objective measure because not all interventions / trial settings can be blinded.
- The core set should be relatively quick and easy to perform with minimal training required to use the instruments so they can be included even when they are not the most important thing for that particular trial.
- The instruments should perform well, be consistent and reliable and the data entry not too onerous.
- Instruments need to measure things that are important to patients and clinicians.

The group discussed in pairs whether signs should be investigator assessed and how a representative site should be assessed. It was clear that there was considerable difference of opinion in both of these areas.

KT then went on to talk about some aspects of feasibility that are specific to the signs and symptoms domains:

- The Three Item Severity Score (TIS) stands alone from other signs scales because it is so quick and simple to complete compared with other scales and has been shown to be associated with worsening disease.
- We should consider whether it is necessary to include all eczema signs in the core outcome instrument; acute signs can be more responsive to change, but other signs may also be important (e.g. lichenification).
- Itch cannot be judged by an Investigator – it must be patient reported. It is even difficult for a parent / carer to judge this for their child.

## Presentation 1.4: The Use of Global Assessments in Atopic Dermatitis Research - A Systematic Review of Randomized Controlled Trials:

Masutaka Futamura, Kim Thomas, Helen Nankervis, Joanne Chalmers , Hywel Williams, Eric Simpson

Dr Eric Simpson (ES) presented data on the use of Investigator Global Assessment (IGA) and called for the IGA to be standardised. The FDA require IGA to be included in trials but ES presented data from the GREAT database of eczema trials (www.greatdatabase.org.uk) which showed there is huge variation on how IGA is conducted:

- One quarter of trials included instructions on which signs to consider in the IGA.
- When instructions are given, only signs are described, but where no instructions are given, further domains may be taken into account such as itch, quality of life and BSA.
- Many different scales are used to rate the IGA.
- The definition of success varies across trials.

There is a concern that by providing detailed instructions on which signs to include and how to rate them, the IGA effectively become a distillation of signs scores rather than a true global assessment. To demonstrate this, ES showed examples of patients which should (according to the IGA instructions) be classed as moderate but were actually very different patients "globally" with respect to the BSA involvement. ES posed the question of whether an algorithmic approach, a new numerical composite or a gestalt approach is needed to standardise IGA.

This paper was then opened up for discussion and the main points summarised below:

- Although IGA is not included as a core domain for HOME (because a numerical scale is required) the HOME initiative should be involved with helping to standardise and better validate the IGA and engage with the FDA to offer a solution to this problem.
- The IGA should be conducted by thinking like a patient, not a clinician or psychometrician.
- The IGA should not become a composite of the core domains.

# Session 2: Signs Domain

## Introduction

Professor Jochen Schmitt (JS), lead for the signs working group, opened the signs session by reiterating that it had previously been agreed that signs assessed by a score should be included as a core outcome. He then introduced the aims of this session which were:

- To clearly define the construct "clinical signs"
- To work towards consensus on criteria for content validity of AD-signs measures
- To determine a short list of instruments to measure AD signs
- To work towards consensus to recommend one scale to measure AD signs
- To prioritize outcomes research concerning AD-signs measurements

JS explained with reference to the book "Measurement in Medicine" (de Vet *et al* 2011) that when the conceptual framework was considered, signs were a formative model because the items of instrument form the construct (domain) to be measured. Therefore, content validity is extremely important (items included in the instrument need to be relevant and all relevant items need to be considered) but internal consistency is less important.

## Presentation 2.1: Measuring responsiveness (sensitivity to change) and minimally clinically important difference (MCID) of the Three Item Severity scale (TIS).

Mandy Schram, Phyllis Spuls, Kim Thomas, Jochen Schmitt

Dr Mandy Schram (MS) explained that the Three Item Severity (TIS) score is quick and easy to administer because it includes only erythema, oedema and excoriation, each measured at one representative site. MS presented TIS data extracted from three existing trial datasets; MAcAD, (Schram *et al* 2011), PROVE (Schmitt *et al* 2010) and SWET (Thomas *et al* 2011). Results showed that responsiveness measured by Area Under the Curve (AUC) was 0.71 (95% CI 0.64-0.76) when compared to IGA (0-5) in MAcAD and PROVE trials and 0.57 (85% CI 0.51-0.63) when compared to the bother score (0-10) in SWET. MCID measured as longitudinal data on absolute changes within individuals showed that MAcAD/ PROVE: 1.02 (SD 1.21) and SWET: 0.62 (SD 1.36).

## Presentation 2.2: Measurement properties of outcome measurements for atopic dermatitis.

Jochen Schmitt, Stefanie Deckert, Sinead Langan, Ake Svensson, Laura von Kobyletzki, Kim Thomas, Phyllis Spuls

This systematic review had been carried out by members of the signs working group led by JS, to assess the measurement properties of atopic eczema sign scales. These data would provide the evidence base for consensus on the signs domain.

Stephanie Deckert (SD) presented the methods used in the review. The methodological quality of the included validation studies was rated using the COSMIN checklist (Terwee *et al* 2012); each scale was given a percentage score for each measure. The performance of each scale was then categorised as follows:

A. Outcome measure meets all requirements to be recommended for use.
B. Outcome measure meets two or more quality items, but performance in all other required quality items is unclear, so that the outcome measure has the potential to be recommended in the future depending on the results of further validation studies.
C. Outcome measure has low quality in at least one required quality criteria (≥1 rating of "minus") and therefore is not recommended to be used any more
D. Outcome measure has (almost) not been validated. Its performance in all or most relevant quality items is unclear, so that it is not recommended to be used until further validation studies clarify its quality.

JS then presented the results. A total of sixteen instruments were evaluated in the review, including both uni-dimensional scales in which only signs are measured and multi-dimensional scales which also measure other items such as symptoms. Most of the validation studies had been carried out on SCORAD, TIS and EASI.

JS also summarised two other previously published studies on content validity:

- Charman *et al* (2005) systematically showed that three clinical signs are independent predictors of patient-rated disease severity: excoriations, erythema, and oedema/papulation. These three signs should be measured in a scale as a minimum for it to have adequate content validity.
- Schmitt *et al (2007)* previously reported that both intensity and extent of disease should be measured.

JS concluded that:

- (Objective) SCORAD and EASI have:
    - Adequate content validity (i.e. include the essential signs of excoriations, erythema, and oedema/papulation).
    - Have been sufficiently tested and have good psychometric properties.
- TIS and SASSAD would require further consensus on content validity as TIS lacks BSA involvement and SASSAD does not capture oedema/papulation.
- The Simple Scoring System (SSS) has intermediate content validity but very little validation work has been carried out.

It was therefore agreed by consensus that objective SCORAD and EASI should be taken forward for consideration for inclusion in the core set.

JS then proposed a definition for the construct "clinical signs" which was discussed by the group:

*Measurement of the intensity and extent of all relevant clinical signs of atopic dermatitis by a physician by either a uni-dimensional instrument reported as a score or a multidimensional construct reported as a profile.*

The group then discussed which signs should be included in the core outcome measure. HW explained this discussion would continue the next day and drew the meeting to a close for the day.

# Sunday 7th April (8.45am – 5pm)

The second day of the meeting opened with a continuation of the whole group discussion on which signs are essential to reflect the construct of the disease and therefore should be included in the core outcome. The main points arising were:

- The three signs which have previously been shown to be independent predictors of patient-rated disease severity (excoriations, erythema, and oedema/papulation) are important acute signs.
- Lichenification should also be considered for inclusion because it can be assessed better than some of the other signs in darker skin types. Also, lichenification is a chronic sign and may, therefore, reflect the chronic relapsing nature of the disease. Concerns were raised that lichenification may not change much during a short term trial, but it was agreed that lichenification can be reduced in the medium term.
- Many other important signs of eczema were suggested and discussed including crusting/oozing, xerosis /dryness, blanching, flaking.
- It is important that the core outcome" clinical signs" is an investigator-assessed measure. The other three domains (symptoms, quality of life and long term control) are primarily patient reported, so it is important to include an investigator assessed objective measure to reduce information bias particularly for trials in which patients are not blinded to treatment allocation.
- Some felt that patient reported signs should also be included in addition to the investigator reported signs previously agreed. Patient versions of scales are often not the same as investigator versions (e.g. EASI and SA-EASI) so should not be used interchangeably. When considering patients reporting signs, the intrusiveness into patients' lives needs to be considered.
- When deciding a core outcome measure, the focus should be on the design and setting of the *majority* of trials. It is not possible to have outcomes that are ideally suited to *all* designs of eczema trials.
- Where multiple terms are used in some scales (e.g. oedema /papulation/infiltration) these may not be exactly the same thing and could lead to inter-investigator variability. However, validation studies have shown that these descriptors are acceptable and so meet the OMERACT filter.

It was felt that sufficient progress had been made to permit a vote on which signs should be included in the core outcome measure. This would be followed by identifying a suitable scale to measure these, rather than taking the approach of starting from the scales that are available. It was noted that if a new scale should need to be developed as a result of this voting, then this would delay the process of recommending a core outcome by 3-5 years, so the question "can I live with this?" should be kept in mind when voting.

Results of signs voting:

| | Yes (%) | No (%) | Unsure (%) |
|---|---|---|---|
| Should this domain measure intensity of clinical signs? | 93 | 2 | 4 |
| Should this domain measure extent of clinical signs? | 89 | 4 | 7 |
| Should this domain measure both intensity and extent of clinical signs? | 96 | 4 | 0 |
| Should signs be assessed by an investigator? | 91 | 2 | 7 |
| Is it essential that erythema is included in the domain "clinical signs"? | 95 | 2 | 2 |
| Is it essential that excoriation is included in the domain "clinical signs"? | 81 | 7 | 12 |
| Is it essential that oedema / papulation is included in the domain "clinical signs"? | 84 | 5 | 11 |
| Is it essential that lichenification is included in the domain "clinical signs"? | 79 | 7 | 14 |
| Are there any other signs that you think are essential to be included? | 17 | 83 | - |

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

## Consensus:

The core outcome measure for the signs domain should measure:

- Both intensity and extent of clinical signs
- Erythema, excoriation, oedema / papulation & lichenification

*(fewer than 30% of the voters disagree)*

Because the consensus was that there are no other *essential* signs to be included, further individual signs were not voted on.

JS revisited the results of the systematic review in light of the voting to determine if there is a scale that could be recommended for the core outcome that measures the agreed four signs? Objective SCORAD and EASI include the four signs voted as essential for inclusion (erythema, excoriations, oedema/papulation and lichenification) and have been shown to perform well. These two scales were described in detail to the attendees by JS and ES and the main differences highlighted:

| Objective SCORAD | EASI |
|---|---|
| A representative site is selected for each of 6 signs | Assesses different areas of the body for each sign (4 signs and 4 body sites) |
| Gives more weight to intensity than extent so a large change in extent has little effect on the overall score | Each sign and extent are equally weighted |

Further details of the validation studies of these two scales were presented by JS. It was noted that:

- Some of the validation studies were on the full SCORAD and others on objective SCORAD
- Some aspects of validation were missing for EASI (interpretability and floor/ceiling effects, feasibility).
- Feasibility is better studied for SCORAD than EASI.
- Further discussion would be needed on selecting a representative site if SCORAD is recommended.

The group then split into 5 smaller groups, each with at least one patient representative, to discuss the advantages and disadvantages of these two scales and hold a preliminary vote (show of hands) within each group on which scale they would prefer to see included.

The results of the preliminary small group voting were:

| Within group voting | EASI | obj SCORAD | unsure | abstained |
|---|---|---|---|---|
| Group 1 | 11 | 0 | 0 | 0 |
| Group 2 | 5 | 2 | 0 | 1 |
| Group 3 | 5 | 1 | 1 | 0 |
| Group 4 | 8 | 1 | 0 | 0 |
| Group 5 | *Exact numbers not stated but voted in favour of EASI* | | | |

Each group reported on their discussions and the main points summarised in the table below:

| Summary of small group discussions |
|---|
| • EASI only includes the four essential signs so unnecessary data not collected. |
| • Investigator does not need to select a representative site for each sign in EASI which can be problematic when using SCORAD. |
| • EASI distinguishes between different body areas which is important for different age groups and possibly for future treatments which may target particular body sites. |
| • Extent has sufficient weighting in the EASI scale, whereas not given enough importance in SCORAD. |
| • Easier to present sub-analysis of EASI than SCORAD. |
| • EASI is very similar to the PASI used in psoriasis so many dermatologists will be familiar with the format. |
| • If additional validation studies are required for EASI then these can be done. |
| • Despite minor differences, both scales are essentially measuring the same thing (construct validity is good for both scales) and both are freely available at no cost. |

The whole group then voted on which scale should be recommended for inclusion in the core set.

Results of voting:

| | EASI (%) | Objective SCORAD (%) | Unsure (%) |
|---|---|---|---|
| Which of these instruments should be included in the core set? | 90 | 7 | 2 |

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

## Consensus:

EASI will be recommended as the core outcome for measuring signs.

*(fewer than 30% of the voters disagree)*

Because the two scales are similar and both perform well, a further discussion was held on whether both scales should be recommended. It was noted that there are no studies published that have used both EASI and objective SCORAD which would allow us to determine which is the most sensitive to change (published data on sensitivity to change for both scales have been based on IGA which is potentially the wrong anchor). It was suggested that if we recommended that both scales are used (where possible) for the next few years we would be able to answer this question. The potential disadvantages pointed out were that this may dilute the clear message of "one core outcome for signs" and research funders may reject inclusion of two "core" outcomes to measure the same domain because of the increased cost and risk to the trial. A vote was held:

| | Yes (%) | No (%) | Unsure (%) |
|---|---|---|---|
| Do we add an asterisk to our recommendation that objective SCORAD be measured in addition to EASI in future clinical trials where possible? | 57 | 36 | 7 |

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

This proposal was therefore rejected as more than 30% disagreed and EASI alone will be recommended as the core outcome for measuring signs. Hywel Williams thanked Jean-Francois Stalder (J-FS) for setting his personal interests in SCORAD aside by voting for EASI in the interests of good science and international harmony.

## Summary of signs domain session decisions made:

- The core outcome measure for the signs domain should:
    - Include intensity *and* extent of the clinical signs
    - Be assessed by an investigator
    - Include erythema, excoriation, oedema / papulation and lichenification
- No other signs were considered essential to reflect the construct of the disease.
- Objective SCORAD and EASI include the four essential signs and perform adequately in validation studies.
- Agreed by consensus that EASI will be recommended as the core outcome for measuring signs.
- The proposal to recommend that objective SCORAD be measured in addition to EASI in future clinical trials where possible was rejected.

## Next steps for signs domain:

- Publish the systematic review
- Publish the consensus statement on EASI as the core outcome.
- Publicise the use of EASI as the core outcome.

## Session 3: Quality of Life domain

This session began with two presentations, followed by an open whole group discussion.

### Presentation 3.1: Quality of Life (QoL) Measures in Atopic Dermatitis Clinical Care and Research

Shehla Admani and Magdalene Dohil

Dr Shehla Admani (SA) presented the results of a review showing which QoL tools are the most commonly used in AD.

| Type of Instrument | Name of Instrument |
|---|---|
| Generic | • EuroQol Quality of Life Scale (EQ-5D)<br>• Medical Outcomes Short Form-36 Health Survey (SF-36) |
| Dermatology Specific | • Dermatology Life Quality Index (DLQI)<br>• Children's Dermatology Life Quality Index (CDLQI)<br>• Skindex – 29 |
| Disease Specific | • Infant's Dermatitis Quality of Life (IDQoL) Index<br>• Childhood Atopic Dermatitis Impact Scale (CADIS)<br>• Quality of Life Index for Atopic Dermatitis (QoLIAD)<br>• Dermatitis Family Impact (DFI) [Family] |

This was followed by presenting details of the properties of each commonly used instrument and pros and cons of each of these instruments.

### Presentation 3.2: Measurement properties of eczema-specific measures of health-related quality of life: systematic review

Christian Apfelbacher

Dr Christian Apfelbacher (CA) presented details of a proposed protocol which will systematically assess measurement properties of eczema-specific measures of HrQoL and identify outcome measures for eczema-specific HrQoL. CA proposed using similar methods to that used for the signs systematic review; assess the methodological quality of studies using the COSMIN checklist then rating the quality of each instrument using the same four category definitions. A summary of the discussion points following the two presentations is presented in the table below:

- The AAD have done a systematic review which has identified other QoL scales not present in the review presented here by SA. Not all have been used in clinical trials yet. This search strategy could be used for this review.
- Generic scales can perform as well as specific scales with regards to sensitivity to change is but not popular with patients as can be lengthy to complete.
- Rather than including a dermatology specific QoL scale in the core outcome set, should we consider PROMIS (Patient Reported Outcomes Measurement Information System) – validated instruments that individually measure patient reported concepts such as pain, fatigue, physical function, depression, anxiety and social function (www.nihpromis.org). This is a US based initiative and would need internationalisation of the terminology used to be included in HOME.
- We should be aware of these different approaches but we should adhere to the previous agreement at HOME II that a dermatology specific measure of QoL will be included in the core outcome set and it should be validated in patients with AD.
- We need to be aware that different scales may perform better for specific age groups.
- It is particularly important that patients have a lot of input into the QoL group.

## Summary of QoL domain session:

- Generic scales can perform better than perhaps previously thought.
- Should look to new methodologies being used in this field such as PROMIS.
- Lots of work to be done in this domain.
- Extremely important to get patient input.

## Next steps for QoL domain:

- Get the systematic review team in place.
- Agree the protocol for the systematic review.
- Conduct the systematic review.

# Session 4: Symptoms domain

## Introduction

Dr Phyllis Spuls (PS), lead for the symptoms group, opened the session by distinguishing between signs and symptoms:

It is not necessarily the nature of the sign or symptom which defines it, but who observes it. A symptom is any feature which is observed by the patient whereas a sign is observed by other people.

- Some features of eczema can only be symptoms, because they cannot be directly observed by other people (such as pain and itch).
- Other features can only be signs (such as a blood cell count measured in a laboratory).
- However, some features can be a sign or a symptom depending on who observes it (such as skin rash which may be noticed by either a healthcare professional as a sign, or by the patient as a symptom).

## Presentation 4.1: The Use of Symptom Outcome Measures in Atopic Dermatitis Research - a Systematic Review of Randomized Controlled Trials

Joanne Chalmers, Helen Nankervis, Kim Thomas and Phyllis Spuls

PS showed what symptoms have been reported in eczema trials published between 2000 and 2012. Data for the review were obtained from the Global Resource of Eczema Trials (GREAT) database www.greatdatabase.org.uk. Symptoms were reported in three-quarters of trials, with itch and sleep loss being the most commonly reported. Assessment of itch and sleep loss was often performed as the subjective elements of SCORAD but not reported separately.

The main points from the discussion that followed were:

- To progress towards consensus on a core outcome, the symptoms group should follow the HOME roadmap i.e.:
  i. Determine the important constructs in the symptoms domain from patient's perspective
  ii. Systematic review to identify the instruments used to measure these symptoms
  iii. Systematically review the validation studies on these instruments and assess the quality of the scales to determine whether they pass the OMERACT filter.
- It may be possible to use data from the many published trials that have reported symptoms to determine construct validity but will depend on how many have reported symptoms as a uni-dimensional score or only as part of a composite scale.
- The effects of sleep loss can be extremely important to patients as it affect the ability to drive and to concentrate at school/work). Straightforward sleep loss questions may not detect the real impact. Fatigue should also be considered as a separate symptom.
- The constructs quality of life and symptoms domains overlap so the two groups will need to work together.

## Presentation 4.2: Eczema Signs and Symptoms: what is important to patients?

Laura von Kobyletzki, Åke Svensson, Jochen Schmitt, Kim Thomas

Members of the HOME initiative distributed this web-based survey to patients and parents / carers in thirty-one countries asking them to rate the importance of a list of physical signs and symptoms in relation to a "worsening of the eczema". The randomly ordered list was drawn from a previous systematic review and published outcome measures including the POEM score and the importance of each symptom was rated on a 5 point Likert scale.

A total of eight-hundred and thirty-one responses were received, mainly from USA, UK and Sweden. At least 80% of patients rated the following signs and symptoms as important or very important: itch, pain, hot / inflamed skin, amount of skin affected by eczema, eczema on visible sites (e.g. hands and face), sensitive sites, bleeding, weeping, cracks, sleep difficulties, with itch and pain the highest rated. Patients and parents / carers were also able to add other signs or symptoms that were not on the list but they considered to be important.

The main points from the discussion that followed were:

- The purpose of this work was to determine the important constructs in the symptoms domain from a patient's perspective i.e. identify universal eczema symptoms rather than a list of all symptoms regardless of frequency. Given the number of responses, it is likely that saturation has been reached for the population that has been targeted.
- The majority of survey responses came from western countries and white skinned patients. Therefore the survey should be distributed to other target populations, especially dark skinned patients, by other members of HOME over the next few months. Translations may be required. The survey should then be closed and analysed to allow the symptoms group to move on to the next stage. Focus groups could be conducted to obtain more in depth information.
- The question of whether psychological symptoms should be included in the both the symptoms and QoL domains or only in the QoL domain needs to be addressed. Itch and psychological status are very closely linked; most itch studies also include an assessment of quality of life to show how itch is perceived and understood. Patients do not like to be asked to distinguish between itch and the impact of itch on the QoL.
- Although it would be useful to survey patients about the psychological effects of eczema, the survey should not be changed at this stage due to the significant number of responses already received.
- Parental input indicated that the age of the child can affect what symptoms are important. For instance sleep loss can be important during early childhood, but then bullying because of visible eczema can become an issue in later childhood.
- Differences in climate may affect the responses.
- The study by Charman *et al* which led to the development of the POEM score showed that eight variables explained most of the variability in eczema severity, suggesting that it is not sufficient to measure only itch and sleep loss.

## Presentation 4.3: Comparison of VAS and verbal rating scale in Japanese patients using VAS with 10 point end of "worst imaginable itch".

Norito Katoh, Makiko Nakahara, Hidehisa Saeki, Akihito Hagihara, Hitoshi Mizutani and Masutaka Furue

This project assessed the relationship between Visual analogue scale (VAS) and Verbal rating scale (VRS) in 949 Japanese patients with itchy skin diseases (approximately 40% with eczema) in four university hospitals. VAS significantly correlated with VRS. Each category of VRS (no pruritus, mild, moderate, severe, very severe) differed significantly from the others regarding VAS scoring. There were some differences in itch intensity according to skin diseases and between regions of Japan.

The main points from the discussion that followed were:

- It is not possible to calculate the MCID from these data
- It would be helpful to see some cross-validation with other scales.
- The work presented here is not on the HOME roadmap but is related to the output of HOME – other research groups should be encouraged to include related projects to HOME in a similar manner.

Following the presentations and discussions, a vote was held on whether itch and sleep loss were sufficient to reflect the construct of the disease or whether other symptoms should also be considered for inclusion.

Symptoms voting results:

| | Yes (%) | No (%) | Unsure (%) |
|---|---|---|---|
| Should symptoms include items in addition to itch and sleep loss? | 78 | 18 | 5 |

Therefore, the symptoms group will consider other symptoms in the work towards a core outcome measure.

### Summary of symptoms domain session:

- The measurement of symptoms is messy – many different outcomes currently used.
- Need to ensure that the overlap with QoL is considered at all times.
- Agreed that the symptoms domain needs to cover more than just itch and sleep loss.

### Next steps for symptoms domain:

- Complete the project to identify which symptoms are important to patients.
- Follow the HOME roadmap starting with a systematic review of all measures of symptoms and validation.
- Validation of the POEM scales has been considered in the signs review and this could be used to inform the symptoms group.

# Session 5: Long term control domain

## Introduction

Dr Kim Thomas (KT), lead for the Long term group, introduced the session and presented some issues that are particular to this domain:

- What exactly do we mean by "long term control?"
- Long term control has been voted as a core outcome, but is long term control truly a separate domain, or is it simply a repeated measurement of other core outcomes?
- How should we measure long term control? Suggestions are flares, escalation of treatment, well-controlled weeks, accessing of health resources.
- Can we learn from how other chronic disease areas have dealt with this? Researchers in asthma have recently published on using composite measures / exacerbations (Fuhlbrigge *et al* 2012). This may not be appropriate for capturing worsening of eczema because treatment escalation is not standardised like in asthma.

Whole group discussions took place and several definitions of long term control were proposed:

- The need to intervene with a treatment
- Escalation of treatment
- Serial measurement of signs, symptoms and QoL
- Number of bad days / clusters of bad days

Other points to note from the discussion were that the measure needs to suit trials of all durations and it needs to reflect that eczema is a chronic disease.

## Presentation 5.1: Systematic review of flare definitions used in prospective studies. "How should atopic dermatitis "flares" be defined? Implications for the design and conduct of trials.

Sinead Langan, Kim Thomas, Jochen Schmitt, Sherie Smith, Hywel Williams

KT presented an update of the systematic review on definition of flares first published in 2005. There is no consistency in how flares are defined; of the 26 studies that included a measure of flares 21 different definitions were used. The review team specified flare definitions *a priori* as; i) assessed by patients, ii) feasible to collect in all settings and iii) flares assessed at the time symptoms experienced. None of the definitions identified met all 3 of the criteria.

KT pointed out some of the pitfalls with different methods for collecting data on flares:

- Unscheduled visits can be biased because they are dependent on the patient being motivated and available to attend clinic.
- Daily diaries are resource intensive for data management and can be burdensome for patients.
- Scheduled trial visits may not coincide with the flare so are likely to be less accurate.

KT concluded that flares (as currently defined) are *not* a good contender for a core outcome measure for the HOME long-term control domain (particularly for trials with long follow up or minimal patient contact) although might be useful for short-term studies or studies looking at prevention of flares.

## Presentation 5.2: A validation study of a flare definition based on the need to escalate treatment as a result of worsening disease control.

Kim Thomas, Beth Stuart, Jochen Schmitt, Sinead Langan, Carle Paul, Hywel Williams

This study applied the OMERACT filter to data from two previously published trials; RCT of water softeners for eczema (SWET) (Thomas *et al* 2011*)*) and a cohort study of environmental triggers for flares (Langan *et al* 2009) to test the definition of flare proposed in the previously published systematic review (Langan *et al* 2006); Escalation of therapy due to worsening of disease, where escalation therapy is defined at baseline on individual patient basis. Results showed:

- Feasibility: This flare definition appears to have face validity. Although it requires completion of daily diaries, it is acceptable to patients. However, data management of daily diaries is resource intensive.
- Truth: Large floor effect seen - even in a population with moderate to severe eczema which may not be an important factor in this context.
- Validity: correlates moderately well with eczema severity scales POEM, TIS and SASSAD.

KT concluded that this flare definition could be useful in some circumstances, but probably *not* a good contender for a core outcome measure for the HOME long-term control domain, because it may be not suitable for all settings.

## Presentation 5.3: Validation Study of well-controlled weeks as a way of measuring long-term control.

Kim Thomas, Beth Stuart, Jochen Schmitt, Sinead Langan, Carle Paul, Hywel Williams

This study applied the OMERACT filter to data from the same two datasets as the previous presentation. The concept of well-controlled weeks is "borrowed" from asthma research and is defined as; **Treatment "escalated" for ≤ 2 days** plus **≤ 2 days with bother score>4.** Results showed:

- Feasibility: the concept of well controlled weeks (i.e. the number of weeks when eczema controlled) is intuitively understood. However, data collection and management is complex and resource intensive (combination of symptoms & escalation of treatment) so may not be suitable in all trial settings.
- Truth: a floor effect is observed.
- Validity: significant relationship with eczema severity scales POEM, TIS and SASSAD.

KT concluded that well controlled weeks as a measure of long term control could be useful for some trials, but probably *not* a good contender for a core outcome measure for the HOME long-term control domain because may not be suitable for all trial settings.

KT summarised the situation in long term group:

- Capturing disease control in "real time" is challenging.
- Intensive data collection may be suitable for some trials (particularly if short-term).
- "Well-controlled weeks" and "flares" seem to be intuitively useful concepts, but how to measure them is still unclear.

There was discussion about how the long term group should progress summarised below:

- All attendees submitted on paper what they considered to be long term control and there were a range of responses:
    - Stable disease over a period of months
    - Duration of the trial (long-term)
    - Lack of need for escalation of treatment / rescue therapy
    - Number and duration of flares
    - Repeated measurement of other HOME domains
- Therefore, before progress can be made, the definition of "long term control" needs to be determined.
- To follow the HOME roadmap, the first stage will be to conduct a systematic review to establish how long-term control has been captured in other trials (and possibly other long-term prospective studies) followed by a systematic review of validation studies if there are any.
- The group then can work on establishing how periodicity should be captured; should long term control be a separate domain / outcome measure or should it be serial measurements of the other domains? If it is a separate outcome, how should it be measured?

The two broad approaches to measuring long term control of eczema were discussed:

| Measure of flares / Well controlled weeks (WCW) | Repeated serial measurements of the other 3 domains (signs, symptoms and QoL) |
|---|---|
| Measurement of flares should be a patient reported outcome (PRO) because a flare is a significant event for patients so they are well placed to determine when a flare has occurred and many trials do not have enough clinic visits to enable flares to be measured by the Investigator. | There are many ways these data can be analysed to measure long term control (e.g. mixed models, fixed effect models) |
| There is variability between patients in how they define a flare so individual flare definitions are needed in a trial. Patients can have poor control without a flare. | The EASI scale only captures the eczema at that moment, so difficult to anticipate how signs could be captured as long term control. |
| A definition of a flare (including the end of a flare) needs to be agreed on if it to be used. | Does not include a measure of change in treatment. |
| Parents can get confused about whether well controlled weeks (WCW) refer to the eczema or the child's behaviour. | Scales would need to be completed frequently enough to capture the fluctuations in the eczema which may impact on the feasibility of this approach. |
| Need to determine whether or not the floor effect is important because trials are measuring a benefit. | |
| Existing trial data where a mixture of daily, weekly and monthly data has been collected should be used to establish whether or not detailed daily / weekly information is necessary. | |

Voting then took place on some of the issues that had been discussed:

| | Yes (%) | No (%) | Unsure (%) |
|---|---|---|---|
| Is a systematic review of long term control outcomes required as the next step? | 68 | 18 | 15 |
| Is long term control a unique concept that requires its own outcome tool ? | 68 | 18 | 15 |
| Should "long term control" be completed exclusively by patients? | 33 | 56 | 11 |

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

| | Daily (%) | Weekly (%) | Monthly (%) | Depends on trial design (%) | Unsure (%) |
|---|---|---|---|---|---|
| How often should long term control ideally be measured by patients? | 16 | 42 | 11 | 16 | 16 |

| | 1 month (%) | 2 months (%) | 3 months (%) | 6 months (%) | 12 months (%) |
|---|---|---|---|---|---|
| In trials that purport to measure long term control, what is the minimum recommended duration? | 3 | 0 | 41 | 38 | 18 |

## Summary of long term domain session:

- This domain presents an especially big challenge; neither repeated serial measurements of the other 3 domains or a measure of flares / WCW as ways of measuring long term control have much data to support their use.
- Flares or well controlled weeks (as currently defined) both have disadvantages which may make them unsuitable for the core outcome measure and it is not clear how they should be measured.
- Agreed that long term control is a unique concept that should be measured by a specific tool and should apply to trials of more than 3 months duration.

## Next steps for the long term group:

- Before progress can be made in this domain, what is meant by long term control needs to be clarified through more in depth discussion.
- A systematic review to establish how long term control has been captured and any validation studies should be performed.

## Summing up

Hywel Williams summed up each domain session and the meeting overall:

- A good mix of stakeholders from different countries were able to attend.
- Everyone has contributed to the meeting and we had good input from patients and parents.
- The group worked well together.
- The input throughout the meeting of Jas Singh from OMERACT has been invaluable.
- We should always remember that the purpose of the HOME initiative is to help patients.

## Next Steps

- The outputs from this meeting will be submitted for publication.
- Requested that members of HOME need to join the working groups and help progress the remaining domains.
- HOME would benefit from a methodology group.
- We need to persevere with engaging the FDA.

# References


Charman CR, Venn AJ, Williams H – *Measuring atopic eczema severity visually: which variables are most important to patients?* Arch Dermatol. 2005 Sep;141(9):1146-51.

De Vet HCW, Terwee CB, Mokkink LB, Knol KL. Measurement *in Medicine: A Practical Guide (Practical Guides to Biostatistics and Epidemiology)*. Publication Date: 11 Aug 2011 | ISBN-10: 0521133858

Fuhlbrigge A, Peden D, Apter AJ, Boushey HA, Camargo CA Jr, Gern J, Heymann PW, Martinez FD, Mauger D, Teague WG, Blaisdell C. *Asthma outcomes: Exacerbations* J Allergy Clin Immunol. 2012; 129:S34-48

Langan SM, Thomas KS, Williams HC. *What is meant by a "flare" in atopic dermatitis? A systematic review and proposal.* Arch Dermatol. 2006 Sep;142(9):1190-6.

Langan SM, Silcocks P, Williams HC. *What causes flares of eczema in children?* Br J Dermatol. 2009 Sep;161(3):640-6. doi: 10.1111/j.1365-2133.2009.09320.x. Epub 2009 Jun 5.

Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. *The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study*. Qual Life Res (2010) 19:539-549. DOI 10.1007/s11136-010-9606-8.

Schmitt J, Langan S, Williams HC; European Dermato-Epidemiology Network. *What are the best outcome measurements for atopic eczema? A systematic review*. J Allergy Clin Immunol. 2007 Dec;120(6):1389-98.

Schmitt J, Schäkel K, Fölster-Holst R, Bauer A, Oertel R, Augustin M, Aberer W, Luger T, Meurer M. *Prednisolone vs. ciclosporin for severe adult eczema. An investigator-initiated double-blind placebo-controlled multicentre trial.* Br J Dermatol. 2010 Mar;162(3):661-8. doi: 10.1111/j.1365-2133.2009.09561.x. Epub 2009 Oct 26.

Schram ME, Roekevisch E, Leeflang MM, Bos JD, Schmitt J, Spuls PI. *A randomized trial of methotrexate versus azathioprine for severe atopic eczema.* J Allergy Clin Immunol. 2011 Aug;128(2):353-9. doi: 10.1016/j.jaci.2011.03.024. Epub 2011 Apr 22.

Thomas KS, Dean T, O'Leary C, Sach TH, Koller K, Frost A, Williams HC; SWET Trial Team. *A randomised controlled trial of ion-exchange water softeners for the treatment of eczema in children*. PLosMed. 2011 Feb 15;8(2):e1000395. doi: 10.1371/journal.pmed.1000395