TECHNISCHE UNIVERSITÄT DRESDEN

UniversitätsCentrum Evidenzbasierte Gesundheitsversorgung

Centre for Evidence-based healthcare, Dresden, Germany

The ***Harmonizing Outcomes Measures for Eczema*** initiative

Prof. Jochen Schmitt, MD MPH

COMET III meeting

Manchester, June 20, 2013

Atopic dermatitis, (atopic) eczema

The problem

Poor standardisation of outcome measurement:

In 94 RCTs published between 1994-2001

56 different outcome assessments for “objective severity” of AE were applied

(Charman & Williams, JID 2003)

Systematic review:

Named eczema outcome measures identified (2007)

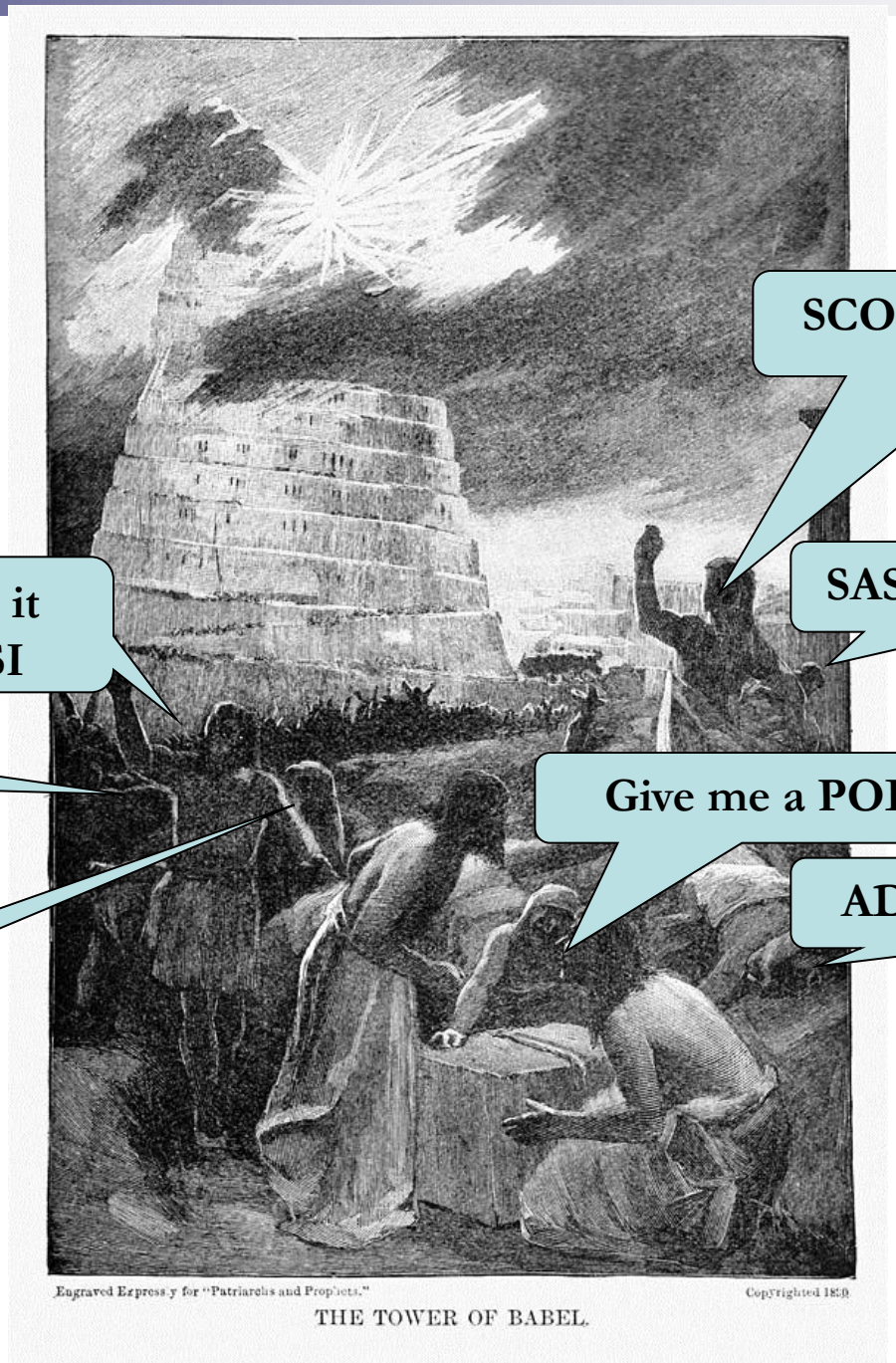
ADAM	Atopic Dermatitis Assessment Measure
ADASI	Atopic Dermatitis Area and Severity Index
ADSI	Atopic Dermatitis Severity Index
BCSS	Basic Clinical Scoring System
EASI	Eczema Area and Severity Index
FSSS	Four Step Severity Score
IGADA	Investigators´ Global Atopic Dermatitis Assessment
Leicester	Leicester index
NESS	Nottingham Eczema Severity Score
OSAAD	Objective Severity Assessment of Atopic Dermatitis
POEM	Patient-Oriented Eczema Measure
RL Score	Rajka and Langeland Score
SA-EASI	self-administered Eczema Area and Severity Index
SASSAD	Six Area, Six Sign Atopic Dermatitis severity score
SCORAD	Severity Scoring of Atopic Dermatitis index
SIS	Skin Intensity Score
SSS	Simple Scoring System
TBSA	6-area Total Body Severity Assessment
TISS	Three Item Severity Score
WAZ-S	(Polish acronym for atopic dermatitis severity score)

Domains included in 20 outcome measures

Scale	Intensity	Extent	Symptoms	Course	Epidermal function
ADAM	●	●	●		
ADASI	●	●	●		
ADSI	●		●		
BCSS		●			
EASI	●	●			
FSSS	●	●		●	
IGADA	●	●			
Leicester	●	●			
NESS		●	●	●	
OSAAD	●	●			●
POEM	●		●		
RL Score		●	●	●	
SA-EASI	●	●	●		
SASSAD	●	●			
SCORAD	●	●	●		
SIS	●		●		
SSS	●	●	●		
TBSA	●	●			
TISS	●				
WAZ-S	●	●	●		

Items used to measure intensity of lesions

Scale	erythema	edema / papulation	oozing/ crusting	excoriation	Licheni-fication	dryness	scaling	fissuring	vesicles	(de)pigmen-tation	flaking	bleeding	erosions
ADAM	•			•	•		•						
ADASI	•	•	•		•		•						
ADSI	•		•	•	•								
BCSS													
EASI	•	•		•	•								
FSSS	•	•	•	•	•	•							
IGADA	•	•	•	•	•		•						
Leicester	•			•	•	•		•					
NESS													
OSAAD	•		•	•	•								
POEM			•			•		•			•	•	
RL Score													
SA-EASI	•	•		•		•							
SASSAD	•		•	•	•	•		•					
SCORAD	•	•	•	•	•	•							
SIS	•					•							
SSS	•	•	•	•	•		•		•	•			
TBSA	•	•	•	•		•	•	•	•				
TISS	•	•		•									
WAZ-S	•	•	•				•		•	•			•


SCORAD scores again
SASSAD rules OK
Take it EASI
TIS but a scale
Give me a POEM
ADASI tonight?
Me too!
Engraved Expressly for "Patriarchs and Prophets."
Copyrighted 1859
THE TOWER OF BABEL.

Name of quality item	Definition of quality item	Measurement of quality item	Criteria for rating "adequate"	Criteria for rating "acceptable"
Construct validity:	Does the scale measure the hypothetical construct (objective severity of AE) it should?			
(a) convergent	(a) Are 2 outcome measurements that are presumed to measure the same latent construct correlated?	(a) and (b) Confirmatory factor analysis, Structural equations modeling (correlation of coefficients)	(a) Factor loading/ correlation coefficient >0.70	(a) Factor loading/ correlation coefficient 0.60-0.69
(b) divergent	(b) Are 2 outcome measurements that are presumed to measure different constructs not (highly) related?		(b) factor loading/ correlation coefficient <0.70	(b) Factor loading/ correlation coefficient 0.71-0.85
Content validity	Are the domains adequate to measure the construct in question? Are the items representative of the domain they are supposed to measure?	Rating by experts and consumers	Expert/consumer says yes for at least 90% of all items	Expert/consumer says yes for 70% to 89% of all items
Internal consistency	Are the different domains/items of the scale interrelated?	Cronbach α*	≥0.90 (individual patients) ≥0.70 (groups)	0.70-0.89 (individual patients) 0.60-0.69 (groups)
Interobserver reliability	Do 2 or more independent investigators achieve the same result?	(a) Correlation coefficient (b) κ† (c) Coefficient of variation (d) ANOVA (% variance explained by observer)	(a) >0.80 (b) >0.60 (c) <20% (d) <10%)	(a) 0.60-0.80 (b) 0.41-0.60 (c) 20% to 30% (d) 10% to 20%
Test-retest reliability	Do 2 assessments by one investigator in the same patient yield the same result?	(a) Correlation coefficient (b) Percentage variation (c) Coefficient of variation	(a) 0.90 (b) <5% (c) <10%	(a) 0.80-0.90 (b) 5% to 10% (c) 10% to 20%
Sensitivity to change	Can clinically relevant changes be detected by this measurement?	Correlation of changes in 2 or more outcome measurements of the same construct	>0.80	0.60-0.80
Acceptability	Is the measurement practical enough to be applied in: (a) everyday clinical practice (b) clinical trials	Time to administer	 (a) <3 min (b) <7 min	 (a) 3-5 min (b) 7-10 min

Predefined criteria for recommendation

QUALITY ITEM

‘‘adequately met’’→ full credit (100%)

‘‘acceptably met’’ → half credit (50%)

‘‘not acceptably met’’ / not assessed → no credit (0%)

→ Calculation of total relative score ranging from 0 - 100%

Score	Recommendation*	Reason
> 90%	highly recommended	measurement is valid & reliable
70-90%	recommended	measurement meets most validity criteria
50-69%	acceptable but not recommended	validity criteria only partly met
30-49%	not recommended	significant validity criteria are not met or have not been evaluated
< 30%	not acceptable	measurement is invalid or has not been validated

* based on: content validity, construct validity, internal consistency, interobserver reliability, test-retest reliability, sensitivity to change

Psychometric properties & recommendations

Outcome	Content validity		Construct validity		Internal consistency	Inter-observer reliability	Test-retest reliability	Sensitivity to change
	Expert	Consumer	convergent	divergent				
ADAM	●	●	●			○		
ADASI	●	●	○			○		
ADSI	●	●						
BCSS	●	●	○			●		
EASI	●	●	●	●	○	○	○	●
FSSS	●	●						
IGADA	●	●	●					●
Leicester	●	●						○
NESS	●	●	○	●		●		
OSAAD			○			●		○
POEM		●	●	○	●	n.a.	●	○
RL Score	●	●				○		
SA-EASI	●	●	○	●		n.a.		○
SASSAD	●	●	○			●	○	○
SCORAD	●	●	●	●		●	○	●
SIS	●	●						
SSS	●	●	○			○		○
TBSA	●	●						
TISS	●	●	●	●		○		
WAZ-S		●						

JACI 2007; 120:1389-98

What are the best outcome measurements for atopic eczema? A systematic review

Jochen Schmitt, MD, MPH,[a] Sinead Langan, MD,[b] and Hywel C. Williams, PhD, FRCP,[b] on behalf of the European Dermato-Epidemiology Network *Dresden, Germany, and Nottingham, United Kingdom*

Conclusions from systematic review

- From 20 outcome measures aiming to assess the severity of eczema, only SCORAD, EASI, and POEM have adequate psychometric properties to be recommended
- Most outcomes have not been tested properly or perform inadequately when tested
- **The continuing use of non-validated outcome measures for eczema hampers scientific communication and is a significant threat to evidence-based dermatology**

→ International Delphi exercise to define core sets of outcome domains for clinical trials and for routine care

Three stage web-based international Delphi consensus exercise

conducted between June 2008 and March 2010.

Delphi consensus panel

- **Multi-professional collaboration involving the views of different stakeholder groups**
  - Consumers: Speakers of eczema self-help groups (n=6)
  - Clinical experts: Major interest in eczema; scientific advisory board ISAD Kyoto 2008; scientific committee IDEA Nottingham 2008
  - Representatives of regulatory agencies: EMEA, FDA
  - Journal editors: JACI, JID, Arch Dermatol, JAAD, Brit J Dermatol, Acta Derm Venereol, JEADV, JDDG
- **Exclusion criteria**
  - Involvement in development of named outcome measure for eczema
  - Affiliation with pharmaceutical industry

Delphi questionnaire

- Background information provided, problem addressed
- Indication of the importance of outcome domains for eczema on a 9-point Likert scale (rounds 1 and 2)
  - Scores 1-3: domain is not important
  - Scores 4-6: equivocal
  - Scores 7-9: domain is important
- 2 different contexts / settings
  - Clinical trials
  - Recordkeeping in daily practice

Delphi questionnaire (cont.)

- How many domains should be included into core sets for clinical trials and for daily recordkeeping?
- What are the top three most important outcome domains for clinical trials and for daily recordkeeping?
- Final round: Explicit question on whether or not to include outcome domain into the core set for clinical trials and for daily recordkeeping
- Feedback: previous rating, group response (median, IQR)
- Three rounds conducted by electronic mail

Outcome domains to be considered

Domains identified by SR:

- Clinical signs (physician/patient)
- Symptoms
- Disease extent
- Course of disease
- Global disease severity (physician/patient)

Additional domains (panel)

- Involvement of visible areas
- Treatment utilisation

Additional domains

- General quality of life
- Dermatology-specific quality of life
- Control of disease flares (short term/long term)
- Time to/ duration of remission
- Health utilities
- Work/school limitations
- Consequences of pruritus,
- Cost-effectiveness
- Direct / indirect cost
- Work productivity loss
- Compliance

Definition of consensus

- *A priori* defined in study protocol

- INCLUSION OF DOMAIN INTO CORE SET

  ≥ 60% of all members of at least three stakeholder groups *including consumers* recommended including a domain in the core set of outcomes.

Delphi Panel

	No of participants invited	No participated in round 1; response rate (%)	No participated in round 2; response rate (%)	No participated in round 3; response rate (%)
Stakeholders				
Consumers	6	6 (100%)	6 (100%)	6 (100%)
Clinical experts	41	32 (78%)	32 (100%)	29 (91%)
Regulatory agency representatives	2	1 (50%)	1 (100%)	1 (100%)
Journal editors	8	7 (88%)	7 (100%)	7 (100%)
Sex				
Female	18	14 (78%)	14 (100%)	14 (100%)
Male	39	32 (82%)	32 (100%)	29 (91%)
Total	**57**	**46 (81%)**	**46 (100%)**	**43 (93%)**

Results

- Main effect of feedback process was reduction of variability in scores assigned to each domain
- Little change in the median score of each domain
- Great variety of domains was considered important by the panel
- Median number of different domains to be included in the core set: 3

Results round 3:
Core set of outcome domains: **Clinical trials**

Outcome domain	Proportion recommending including outcome domain into the **CORE SET** of outcomes for eczema that should be routinely assessed **in every CLINICAL TRIAL** on eczema?				Consensus to include domain into core set		
	Consumers (n=6)	Experts (n=29)	Agency (n=1)	Editors (n=7)	YES	Un-clear	NO
Clinical signs (physican)	100%	100%	100%	100%	•		
Clinical signs (patient)	17%	21%	0%	0%			•
Investigator global assessment	33%	59%	0%	57%			•
Patient global assessment of	17%	34%	0%	29%			•
Symptoms	83%	76%	0%	57%	•		
Quality of life (specific)	33%	72%	100%	86%		•	
Quality of life (general)	17%	3%	0%	0%			•
Short term control of flares	33%	7%	0%	0%			•
Long term control of flares	67%	62%	100%	43%	•		
Cost	17%	3%	0%	0%			•
Overall extent of disease	17%	21%	0%	14%			•
Involvement of high expr. areas	17%	7%	0%	14%			•
Treatment utilization	17%	31%	0%	14%			•

Results round 3:
Core set of outcome domains: **Recordkeeping**

Outcome domain	Proportion recommending including outcome domain into the **CORE SET** of outcomes for eczema that should be routinely assessed in **DAILY PRACTICE, i.e. to be used AT EVERY PHYSICIAN VISIT**				Consensus to include domain into core set		
	Consumers (n=6)	Experts (n=29)	Reg. agency (n=1)	Editors (n=7)	YES	Un-clear	NO
Clinical signs (physician)	83%	34%	0%	43%		•	
Clinical signs (patient)	33%	14%	0%	0%			•
Investigator global assessment	17%	66%	100%	71%		•	
Patient global assessment	50%	28%	0%	43%		•	
Symptoms	100%	83%	0%	86%	•		
Consequences of itching	67%	17%	0%	0%		•	
Quality of life (specific)	17%	10%	0%	0%			•
Quality of life (general)	0%	7%	0%	0%			•
Short term control of flares	33%	14%	100%	0%			•
Long term control of flares	67%	41%	100%	29%		•	
Compliance	33%	31%	0%	0%			•
Work/school limitations	17%	14%	0%	0%			•
Overall extent of disease	17%	21%	0%	29%			•
Involvement of high expr. areas	17%	17%	0%	14%			•
Treatment utilization	0%	34%	100%	14%			•

ORIGINAL ARTICLE

Core Outcome Domains for Controlled Trials and Clinical Recordkeeping in Eczema: International Multiperspective Delphi Consensus Process

Jochen Schmitt[1], Sinéad Langan[2], Tanja Stamm[3] and Hywel C. Williams[4], on behalf of the Harmonizing Outcome Measurements in Eczema (HOME) Delphi panel[5]

Preliminary core set of outcome domains

Clinical trials

- Measurement of eczema symptoms
- Physician-assessed clinical signs using a score
- Measurement for long term control of flares

Recordkeeping in daily practice

- Measurement of eczema symptoms

Journal of Investigative Dermatology advance online publication, 14 October 2010; doi:10.1038/jid.2010.303

HOME I
Munich 2010

- *Is there enough interest, enthusiasm and commitment to sort our core outcomes for atopic eczema/atopic dermatitis?* - YES
- Are you willing to set aside your preferences/prejudices/allegiances to work as a group? - YES

HOME II
Amsterdam, 2011

- 43 people came from around the world
- Included 4 consumers
- Presentations, discussions and key pad voting
- Impartial guidance from Maarten Boers

- Consensus on Consensus rule – *"if less than 30% disagree"*
- Consensus to focus on clinical trials (first)
- Main objective: to clarify the role of HRQL

HOME II (2011)

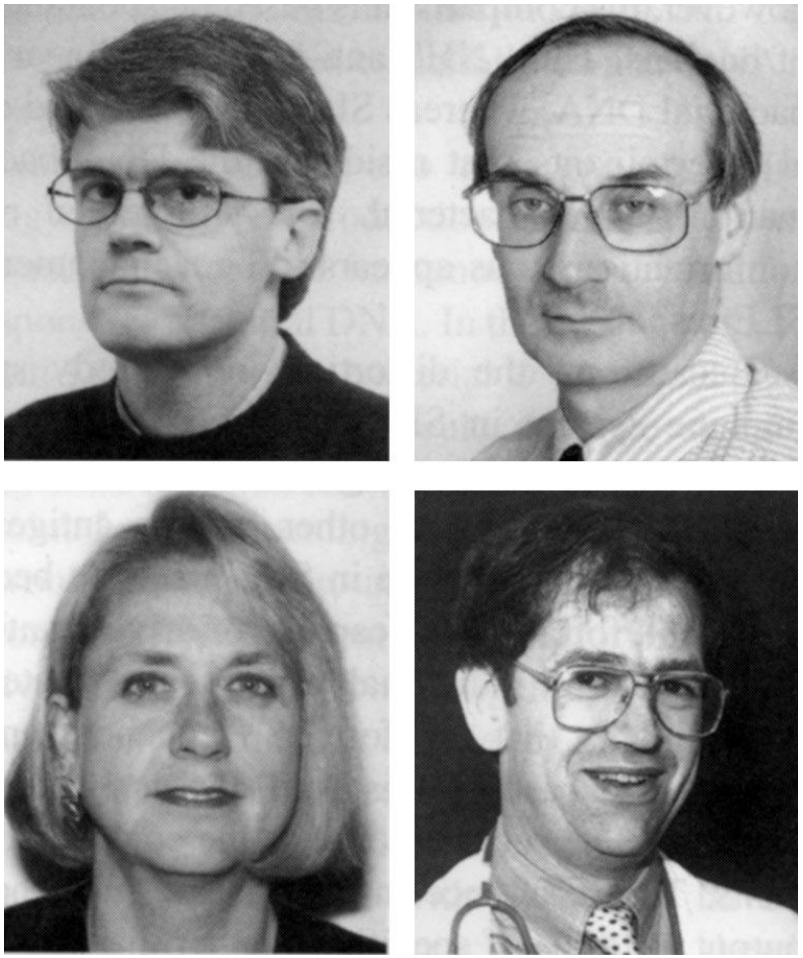
- Symptoms
- Clinical signs using a score
- Long term control of flares
- Quality of life

HOME II (2011)

→ Set up four working groups on identifying best instrument for:

1. Symptoms (Phyllis Spuls)
2. Signs (Jochen Schmitt)
3. HRQL (Magdalene Dohil)
4. Long-term control (Kim Thomas)

Adoption of the OMERACT filter



Truth, Discrimination and Feasibility

Structure of HOME

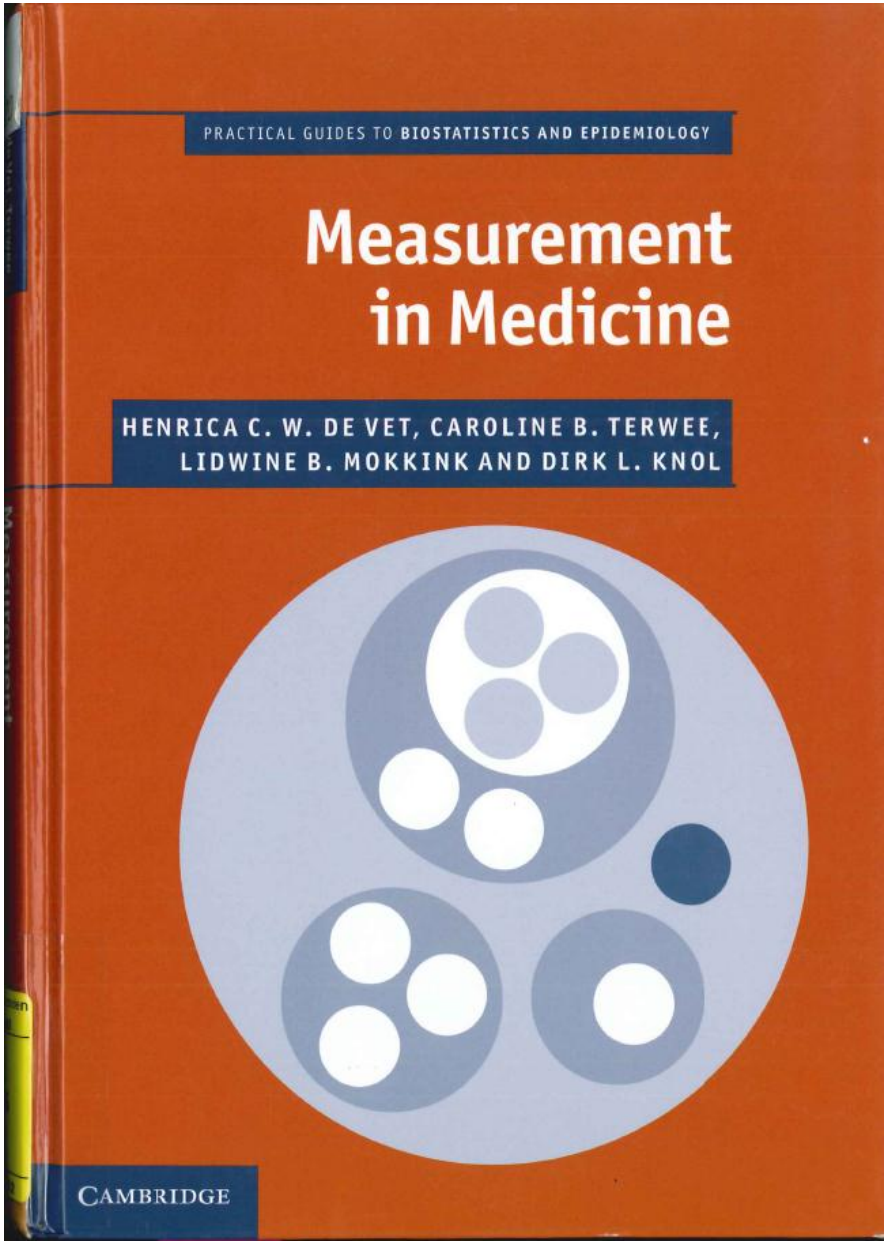
HOME Executive Board		Group lead
Hywel Williams	UK	
Jochen Schmitt	Germany	Signs
Masutaka Furue	Japan	
Magdalene Dohil	USA	
Christian Apfelbacher	Germany	Quality of Life
Eric Simpson	USA	
Phyllis Spuls	Netherlands	Symptoms
Kim Thomas	UK	Long term

HOME Scientific Advisory Board

Jon Hanifin (Chair)	USA
Maarten Boers	Netherlands
Uwe Gieler	Germany
Jean-Francois Stalder	France
Carsten Flohr	UK
Christian Apfelbacher	Germany
Amy Paller	USA
Stephan Weidinger	Germany
Sue Lewis-Jones	UK
Mira Pavlovic	France
Gil Yosipovitch	USA
Carolyn Charman	UK
Mary-Margaret Chren	USA
Roberto Takaoka	Brazil
Yukihiro Ohya	Japan
Elizabeth Hoff	USA
Hidehisa Saeki	Japan
Kefei Kang	China
Kam-Ium Ellis Hon	Hong Kong
John Masenga	Africa
Dedee Murrell	Australia

The HOME Roadmap

	Stage 1 →	Stage 2 →	Stage 3 →			Stage 4 →	Stage 5
Task	**Identify all instruments** previously used to measure the domain.	Establish the **extent and quality of testing** of the identified instruments.	Determine **which instruments** are good enough quality **meet the requirements of the OMERACT filter** and be shortlisted for further consideration.			Carry out **validation studies** on shortlisted scales.	**Finalise core outcome(s)** for domain.
Methodology	Systematic review of outcome instruments used.	Systematic review of validation studies of the long-list of identified instruments. Highlight any gaps in validation.	**Apply OMERACT filter;** Truth, discrimination and feasibility**:**			Consensus discussion and voting to determine what validation studies will be conducted on short-listed instruments. Gaps in testing were highlighted in stage 2 (systematic review). Appropriate methods used to fill the gaps in validation.	Re-apply the OMERACT filter with the results of the completed validation studies. Consensus discussion and voting on core outcome to be recommended.
			Truth *"Is the measure truthful, does it measure what it intends to measure? Is the result unbiased and relevant?"*	**Discrimination** *"Does the measure discriminate between situations that are of interest?"*	**Feasibility** *"Can the measure be applied easily in it's intended setting, given constraints of time, money, and interpretability?"*		
			Consensus discussion and voting on truth: 1. Face validity 2. Content validity 3. Construct validity 4. Criterion validity	Consensus discussion and voting on discrimination: 1. Reliability 2. Sensitivity to change	Consensus discussion and voting on feasibility: 1. Time taken 2. Cost 3. Interpretability		
Output	**Long-list of all instruments previously used to measure the domain.**	**Summary of which instruments have been tested and the quality, extent and results of any testing.**	**Short-list of potential instruments that meet the requirements of the OMERACT filter.**			**Short-list of fully tested instruments.**	**Recommended core outcome(s) for the domain.**



- COSMIN
- Checklist
- Conceptual framework
- Reliability
- Search strings
- …

Systematic review on eczema signs measures

1. To systematically assess measurement properties of outcome measurements for atopic dermatitis **signs**

2. To identify outcome measures for atopic dermatitis signs
   - that meet the predefined criteria (*OMERACT Filter*) to be recommended for the measurement of signs in future atopic dermatitis trials
   - that have the potential to be recommended in the future depending on the results of further validation studies
   - that do **not** meet the predefined criteria to be recommended and therefore should **not** be used any more.

3. To provide the evidence base
   - for an international consensus process to further standardize the assessment of atopic dermatitis signs in clinical trials.
   - for an international consensus process to prioritize further research concerning atopic dermatitis signs outcome assessment.

Data extraction and quality assessment

- Data extraction and assessment for each “substudy”
- independent quality assessment
  - **methodological quality of included studies** based on COSMIN checklist → rating: a 4-point scale → “worse score counts”
  - **rating of scale quality**

Four categories of recommendation

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.

Summary of psychometric properties of measures for clinical signs of eczema

Quality item (name)	SCORAD	TIS	EASI	SASSAD	POEM	BCSS	ADAM	ADASI	ADQ	OSAAD	SSS	W-AZS	Unnamed sale 1	Unnamed sale 2
Content validity	+	+	+	+/-	-	-	+/-	+/-	-	-	+/-	+/-	+/-	-
Construct validity	+	+/-	+	+	n.r.	-	n.r.	n.r.	+/-	+/-	?	n.r.	n.r.	n.r.
Internal consistency	+/-	+/-	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Intra-observer reliability	n.r.	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Inter-observer reliability	+	+/-	+/-	+/-	n.a.	+	+/-	n.r.	n.r.	+	?	n.r.	+	+
Sensitivity to change	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	?	n.r.	n.r.	n.r.	n.r.
Floor or ceiling effects	+	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	+	+	n.r.	n.r.	n.r.
Interpretability	+	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Acceptability	+/-	+	n.r.	+	+	n.r.	n.r.	+	+	-	n.r.	n.r.	+	+
RECOMMENDATION	**B**	**B**	**B**	**B**	**C**	**C**	**D**	**D**	**C**	**C**	**D**	**D**	**D**	**C**

(+) positive rating indicating „adequate“ scale quality; (+/-) intermediate rating indicating “intermediate” scale quality; (-) negative rating indicating “inadequate” scale quality
n.r.: not reported

Categories of recommendation:

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 (as a measurement of eczema signs) 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.

Summary of psychometric properties of measures for clinical signs of eczema

Quality item (name)	SCORAD	TIS	EASI	SASSAD	POEM	BCSS	ADAM	ADASI	ADQ	OSAAD	SSS	W-AZS	Unnamed sale 1	Unnamed sale 2
Content validity	+	+	+	+/-	-	-	+/-	+/-	-	-	+/-	+/-	+/-	-
Construct validity	+	+/-	+	+	n.r.	-	n.r.	n.r.	+/-	+/-	?	n.r.	n.r.	n.r.
Internal consistency	+/-	+/-	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Intra-observer reliability	n.r.	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Inter-observer reliability	+	+/-	+/-	+/-	n.a.	+	+/-	n.r.	n.r.	+	?	n.r.	+	+
Sensitivity to change	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	?	n.r.	n.r.	n.r.	n.r.
Floor or ceiling effects	+	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	+	+	n.r.	n.r.	n.r.
Interpretability	+	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Acceptability	+/-	+	n.r.	+	+	n.r.	n.r.	+	+	-	n.r.	n.r.	+	+
RECOMMENDATION	**B**	**B**	**B**	**B**	**C**	**C**	**D**	**D**	**C**	**C**	**D**	**D**	**D**	**C**

(+) positive rating indicating „adequate“ scale quality; (+/-) intermediate rating indicating “intermediate” scale quality; (-) negative rating indicating “inadequate” scale quality
n.r.: not reported

Categories of recommendation:

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 (as a measurement of eczema signs) 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.

Summary of psychometric properties of measures for clinical signs of eczema

Quality item (name)	SCORAD	TIS	EASI	SASSAD	POEM	BCSS	ADAM	ADASI	ADQ	OSAAD	SSS	W-AZS	Unnamed sale 1	Unnamed sale 2
Content validity	+	+	+	+/-	-	-	+/-	+/-	-	-	+/-	+/-	+/-	-
Construct validity	+	+/-	+	+	n.r.	-	n.r.	n.r.	+/-	+/-	?	n.r.	n.r.	n.r.
Internal consistency	+/-	+/-	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Intra-observer reliability	n.r.	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Inter-observer reliability	+	+/-	+/-	+/-	n.a.	+	+/-	n.r.	n.r.	+	?	n.r.	+	+
Sensitivity to change	+	n.r.	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	?	n.r.	n.r.	n.r.	n.r.
Floor or ceiling effects	+	+	n.r.	+	n.r.	n.r.	n.r.	n.r.	n.r.	+	+	n.r.	n.r.	n.r.
Interpretability	+	+	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.
Acceptability	+/-	+	n.r.	+	+	n.r.	n.r.	+	+	-	n.r.	n.r.	+	+
RECOMMENDATION	**B**	**B**	**B**	**B**	**C**	**C**	**D**	**D**	**C**	**C**	**D**	**D**	**D**	**C**

(+) positive rating indicating „adequate“ scale quality; (+/-) intermediate rating indicating “intermediate” scale quality; (-) negative rating indicating “inadequate” scale quality
n.r.: not reported

Categories of recommendation:

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 (as a measurement of eczema signs) 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.

HOME III
San Diego, 2013

- 56 participants (clinicians, patients, methodologists, industry representatives)
- Support / moderation by Jas Singh

→ Sessions on each core outcome domain
→ Presentations, (small group) discussion, voting

HOME III
San Diego, 2013

Aims

- 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

HOME III
San Diego, 2013

Clinical signs session

- Discussion and consensus
  - on the specification of the construct
  - on the specification of relevant items for the domain
- Presentation and discussion of SR

- Consensus which SINGLE instrument to use as core outcome measure for clinical signs of eczema

Philosophy of HOME

- Working hard together
- Respecting all stakeholder viewpoints
- Putting prejudices and allegiances aside in order to achieve the greater good for patient care
- Evidence-based and evidence-generating
- Pragmatic
- To have fun
- With very little money

Summary: HOME research since 2005

1. Systematic review on validity, reliability, sensitivity to change, and ease of use of named outcomes for AD →describe the problem, give evidence-based recommendations

2. International delphi consensus on core coutcome domains for AD → identify *preliminary* core set of outcome domains

3. HOME I: Munich 2010 → the scientific community expressed interest to form outcomes research group – HOME

4. HOME II: Amsterdam 2011 → (1) start working as a group
   (2) consolidate consensus on core outcome domains
   (3) base recommendations of outcome measures on OMERACT filter
   (4) form project groups for future HOME work

5. 2012: HOME roadmap; HOME projects

6. HOME III: San Diego April, 2013 →consensus on core outcome instrument for clinical signs

Structure of HOME

HOME Executive Board		Group lead
Hywel Williams	UK	
Jochen Schmitt	Germany	Signs
Masutaka Furue	Japan	
Magdalene Dohil	USA	
Christian Apfelbacher	Germany	Quality of Life
Eric Simpson	USA	
Phyllis Spuls	Netherlands	Symptoms
Kim Thomas	UK	Long term

HOME Scientific Advisory Board

Jon Hanifin (Chair)	USA
Maarten Boers	Netherlands
Uwe Gieler	Germany
Jean-Francois Stalder	France
Carsten Flohr	UK
Christian Apfelbacher	Germany
Amy Paller	USA
Stephan Weidinger	Germany
Sue Lewis-Jones	UK
Mira Pavlovic	France
Gil Yosipovitch	USA
Carolyn Charman	UK
Mary-Margaret Chren	USA
Roberto Takaoka	Brazil
Yukihiro Ohya	Japan
Elizabeth Hoff	USA
Hidehisa Saeki	Japan
Kefei Kang	China
Kam-Ium Ellis Hon	Hong Kong
John Masenga	Africa
Dedee Murrell	Australia