Why harmonise outcome measures?

Hywel Williams
Nottingham
The problem
Outcome measures for eczema/atopic dermatitis – a mess

- Too many – over 20 named scales
- Many not tested at all
- Some are only partly tested (validity, repeatability, sensitivity change, consistency, interpretability)
- Some that are tested do not pass the tests
SCORAD scores again

SASSAD rules OK

Give me a POEM

ADASI tonight?

IGADA bad headache

What’s all the FSSS about?

Take it EASI

TIS a right mess

Me too!

Meet my SIS

My name is ADAM
What we need are *core* outcomes that are used in all trials
**What are core outcomes?**

- **Minimum** set for all clinical trials
- Need to be relevant to patients
- Relevant to those making decisions about health care
- Maybe different for clinical trials and routine care
- Need to measure what they’re supposed to measure, be repeatable, sensitive to change, and be easy to use
Why?
– so that we can compare

Standard Cochrane Skin Group abstract:

“We found 27 studies, but outcomes were heterogeneous precluding any formal meta-analysis”

And it goes on.....
If only we could do more of this...

<table>
<thead>
<tr>
<th>Study</th>
<th>Pimecrolimus 1% BID vs Vehicle BID</th>
<th>Relative Risk (Random)</th>
<th>Weight n/N</th>
<th>Relative Risk (Random) n/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 1 week: CASN041C2322 2005 25/168</td>
<td>12/168</td>
<td>-</td>
<td>100.0</td>
<td>2.00 [1.06, 3.76]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>168</td>
<td>168</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>Total events: 26 (Pimecrolimus 1% BID), 18 (Vehicle BID)</td>
<td>Text: heterogeneity: not applicable</td>
<td>Test for overall effect τ=1.15, p=0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02 2 weeks: CASN041C2322 2005 38/168</td>
<td>24/168</td>
<td>-</td>
<td>100.0</td>
<td>1.58 [1.00, 2.52]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>168</td>
<td>168</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>Total events: 26 (Pimecrolimus 1% BID), 24 (Vehicle BID)</td>
<td>Text: heterogeneity: not applicable</td>
<td>Test for overall effect τ=1.04, p=0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>03 2 weeks: Kanna 2002</td>
<td>38/71</td>
<td>8/35</td>
<td>-</td>
<td>28.5</td>
</tr>
<tr>
<td>Eichenfield (a) 2002</td>
<td>35/120</td>
<td>2/68</td>
<td>7.5</td>
<td>3.15 [2.07, 4.87]</td>
</tr>
<tr>
<td>Eichenfield (b) 2002</td>
<td>37/127</td>
<td>8/68</td>
<td>25.5</td>
<td>2.30 [1.13, 4.65]</td>
</tr>
<tr>
<td>Ho 2003</td>
<td>54/223</td>
<td>11/52</td>
<td>35.5</td>
<td>2.51 [1.42, 4.46]</td>
</tr>
<tr>
<td>Lugner 2001</td>
<td>5/25</td>
<td>0/13</td>
<td>1.0</td>
<td>10.52 [0.60, 191.72]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>377</td>
<td>377</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>Total events: 169 (Pimecrolimus 1% BID), 25 (Vehicle BID)</td>
<td>Text: heterogeneity ch-square=4.58 df=4 p=0.33 P=0.12,74</td>
<td>Test for overall effect τ=3.01, p=0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>04 4 weeks: CASN041C2322 2005 54/160</td>
<td>36/160</td>
<td>-</td>
<td>100.0</td>
<td>1.42 [1.00, 2.02]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>160</td>
<td>160</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>Total events: 56 (Pimecrolimus 1% BID), 38 (Vehicle BID)</td>
<td>Text: heterogeneity: not applicable</td>
<td>Test for overall effect τ=1.94, p=0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05 6 weeks: Eichenfield (a) 2002</td>
<td>49/130</td>
<td>11/58</td>
<td>-</td>
<td>28.5</td>
</tr>
<tr>
<td>Eichenfield (b) 2002</td>
<td>44/127</td>
<td>14/52</td>
<td>32.6</td>
<td>2.56 [1.19, 5.52]</td>
</tr>
<tr>
<td>Ho 2003</td>
<td>57/123</td>
<td>17/53</td>
<td>46.5</td>
<td>2.39 [1.45, 3.96]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>239</td>
<td>199</td>
<td>-</td>
<td>100.0</td>
</tr>
<tr>
<td>Total events: 100 (Pimecrolimus 1% BID), 40 (Vehicle BID)</td>
<td>Text: heterogeneity ch-square=1.42 df=2 p=0.49 P=0.90</td>
<td>Test for overall effect τ=1.51, p=0.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

..and compare something that is meaningful rather than just measurable

- Good outcomes that cover the key domains of a particular skin condition

- Instruments that pass the test of truth, discrimination and feasibility
What is happening elsewhere?

- **OMERACT** [http://www.omeract.org/](http://www.omeract.org/)

- Pain – IMMPACT: [www.immpact.com](http://www.immpact.com)

- **COMET initiative**: Core Outcome Measures in Effectiveness Trials

  Tugwell P BM et al. OMERACT: An initiative to improve outcome measurement in rheumatology. Trials. 2007;8(38).

The world of medicine is moving on – what about skin diseases?
Some signs of hope in skin
..starting with HOME

- Agreed that core outcomes needed
- Started with clinical trials
- Domains identified
- Instrument for clinical signs determined
- Now about to tackle symptoms and QoL in Malmo
- Not easy....
**AIM of HOME:** To agree a set of core outcome measures for eczema for use in all clinical trials. Ultimately, the aim is to have just *one instrument* per domain for:

1. Signs
2. Symptoms
3. Quality of Life
4. Measure of long term control of flares

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
<th>Stage 4</th>
<th>Stage 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task</strong></td>
<td>Identify all instruments previously used to measure the domain.</td>
<td>Establish the extent and quality of testing of the identified instruments.</td>
<td>Determine which instruments are good enough quality meet the requirements of the OMERACT filter and be shortlisted for further consideration.</td>
<td>Carry out validation studies on shortlisted scales.</td>
</tr>
<tr>
<td><strong>Methodology</strong></td>
<td>Systematic review of outcome instruments used.</td>
<td>Systematic review of validation studies of the long-list of identified instruments. Highlight any gaps in validation.</td>
<td>Apply OMERACT filter; Truth, discrimination and feasibility: <strong>Truth</strong> <em>&quot;Is the measure truthful, does it measure what it intends to measure? Is the result unbiased and relevant?&quot;</em> <strong>Discrimination</strong> <em>&quot;Does the measure discriminate between situations that are of interest?&quot;</em> <strong>Feasibility</strong> <em>&quot;Can the measure be applied easily in it’s intended setting, given constraints of time, money, and interpretability?&quot;</em></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Output</strong></td>
<td>Long-list of all instruments previously used to measure the domain.</td>
<td>Summary of which instruments have been tested and the quality, extent and results of any testing.</td>
<td>Short-list of potential instruments that meet the requirements of the OMERACT filter.</td>
<td>Short-list of fully tested instruments.</td>
</tr>
</tbody>
</table>

**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

**Recommended core outcome(s) for the domain.**
And other skin diseases listed in COMET website…

- Vitiligo
- Acne
- Incontinence-dermatitis
- Vulval skin conditions
- Leishmaniasis
- HS
- Psoriasis (IDEOM – not sure)
Progress will vary...
Adoption of the OMERACT filter

Truth, Discrimination and Feasibility
Use COSMIN to rate studies that rate outcomes

CO\textit{nsensus-based S\textit{tandards for the selection of health M\textit{easurement I\textit{nstruments}}}

http://www.cosmin.nl/
Philosophy of CSG-COUSIN

- Working hard together to help patients
- 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 inform clinical decisions
- To have fun
- With very little money
What’s all the FSSS about?

Take it EASI

TIS a right mess

Me too!

Meet my SIS

My name is ADAM

SCORAD scores again

SASSAD rules OK

Give me a POEM

ADASI tonight?

IGADA bad headache
Cochrane reviews of the future..

- We found 27 acne studies for our intervention, 25 of which used our primary outcome of inflammatory lesion count.
- We were able to directly and indirectly compare them as a result.
- Treatment A is clearly very helpful, treatment B and C are only moderately so, and treatments D and E are pretty useless.
Why do it?