A Simple Guideline Appraisal Instrument Based On IOM Standards
Matthew D. Mitchell, Brian Leas, Julia G. Lavenberg, David R. Goldmann, Craig A. Umscheid
Center for Evidence-based Practice (CEP)
University of Pennsylvania Health System
Philadelphia PA  USA

Why Do We Need this Instrument?
- We know some guidelines are more reliable than others. Good guidelines are free of bias, based on a thorough review of current clinical evidence, and written in a way that makes it easy for clinicians to act on the recommendations.
- CEP evidence reports frequently include summaries of published guidelines, and one of our tasks is to communicate to readers which guidelines are more reliable and why.
- A set of standards has already been published by the AGREE group (www.agreetrust.org), but it is geared more towards the needs of guideline developers than clinicians and other end-users. The Guidelines International Network (G-I-N) also has published a set of standards.
- We do not assume a guideline is unreliable just because its methods are not reported to the highest level of detail.

The IOM Trustworthy Guideline Standard
- The Institute of Medicine published “Clinical Practice Guidelines We Can Trust” in 2011: a report setting out principles for developing high-quality guidelines.
- The IOM standards were adopted by the Agency for Healthcare Research and Quality for use by the National Guideline Clearinghouse and other federal programs.

Developing Our Own Appraisal System
- Because the current instruments are time-consuming to complete and require assessment of methodologic details often best understood by guideline specialists, we developed a simplified appraisal instrument.
- With an “NR” (Not Reported) grade available for some domains, our grading scheme distinguishes between weaknesses in documentation of process and weaknesses that affect guideline reliability.
- We used the IOM standards as a starting point because of their broad acceptance.
- Objective grading criteria are provided for each standard. Validation studies are in progress.

Presenting Results of a CEP Trustworthy Guideline Appraisal
- Avoid counting up the number of “A” ratings or trying to create a point score from the results. That requires a subjective assumption that all domains are of equal importance and it requires a subjective decision about whether to count an “NR” as negative or neutral.
- Our grid of colored boxes is based on the study quality grids used in Cochrane reviews.
- You can see from our grid format which guidelines have more threats to their validity and which have fewer.
- You can also see which domains are common strengths or weaknesses of all the guidelines.

Comparing the CEP, AGREE II, and G-I-N Standards
- AGREE II has 28 individual criteria while the IOM standards and our assessment instrument are organized into 8 domains.
- The AGREE II criteria that are not included in our system are mostly concerned with process rather than the risk of bias or the risk of basing recommendations on incomplete evidence.
- Guidelines International Network (G-I-N) has its own guideline quality standards: They have 11 domains, with greater emphasis on methods and potential conflicts of interest. G-I-N also has a unique element on the scope of the guideline.
- Most of the important areas are covered by all three sets of standards, though IOM is the only standard to include an element of guideline 1:

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sample guideline 1</th>
<th>Sample guideline 2</th>
<th>Sample guideline 3</th>
<th>Sample guideline 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Transparency</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2. Conflict of interest</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>NR</td>
</tr>
<tr>
<td>3. Development group</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>4. Systematic review</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>5. Supporting evidence</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>6. Recommendations</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>7. External review</td>
<td>B</td>
<td>NR</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>8. Currency and updates</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

Sample Guideline Appraisal Domains

4. Systematic review

- A Guideline includes a systematic review of the evidence or links to a current review.
- B Guideline is based on a review which does not meet systematic review criteria, or cannot be readily obtained.
- C Guideline is not based on a review of the evidence.

7. External review

- A Guideline was made available to external groups for review.
- B Guideline was reviewed by members of the sponsoring body only.
- C Guideline was not externally reviewed.
- NR No external review process is described.

To obtain a copy of the full scale, drop a business card or e-mail mdmitchell@uphs.upenn.edu