Presenting research evidence in clinical guidelines – can we do a better job?

Shaun Treweek
Quality, Safety & Informatics Research Group
CPSE
streweek@mac.com
Decisions should be influenced not only by the best estimates of the expected effect but also by the confidence in these estimates.
An example - hormone replacement therapy

The American College of Physicians recommended hormone replacement therapy* for postmenopausal women for over a decade based on very low quality evidence

An example - hormone replacement therapy

Later evidence from trials* showed that doctors following these recommendations were not providing better clinical care and may have inadvertently been doing harm.

The recommendations should have considered the quality of evidence, not just the estimate of effect.

The DECIDE EC FP7 Project

Developing and Evaluating Communication strategies to support Informed Decisions and practice based on Evidence
DECIDE: the basics

• Coordinated by the University of Dundee
• 10 partners in 7 countries + WHO
• €3 million budget over 5 years
• 8 Workpackages (5 science, 1 tools, 1 dissemination, 1 management)
• Lots of external collaborators, especially the GRADE Working Group
DECIDE: the ten partners

- University of Dundee
- Norwegian Knowledge Centre for the Health Services
- Iberoamerican Cochrane Centre
- Italian Cochrane Centre
- University of Amsterdam
- World Health Organisation (WHO)
- German Cochrane Centre
- National Institute for Health and Clinical Excellence (NICE)
- Scottish Intercollegiate Guidelines Network (SIGN)
- Finnish Medical Society Duodecim

Kick-off meeting was two weeks ago in Geneva.
Cut to the chase - what will we have in five years’ time?

- A range of dissemination strategies for evidence-based recommendations that are tailored to particular audiences (eg. health professionals, the public)
- Evidence that these strategies are effective
- Have tried out strategies in real guidelines eg. those of WHO, NICE, SIGN
- Tools to support effective dissemination strategies
Science WPs: a picture

Phase 1: strategy development and user testing

DECIDE strategies A, B and C

atable:

| Brainstorming workshops | Stakeholder feedback | Survey on current dissemination strategies | User testing |

RCT A vs. B

RCT B vs. C

RCT C vs. A

RCT X vs. Convent.

Phase 2: Evaluation of strategies

Phase 2: Testing strategies with real guidelines

Before / after study
Real guideline testing
Strategy A

Before / after study
Real guideline testing
Strategy C

Friday, 25 February 2011
Welcome

The Grading of Recommendations Assessment, Development and Evaluation (short GRADE) Working Group began in the year 2000 as an informal collaboration of people with an interest in addressing the shortcomings of present grading systems in health care. The working group has developed a common, sensible and transparent approach to grading quality of evidence and strength of recommendations. Many international organizations have provided input into the development of the approach and have started using it. » learn more
A GRADE Evidence Profile

<table>
<thead>
<tr>
<th>No of studies (Design)</th>
<th>Limitations</th>
<th>Inconsistency</th>
<th>Indirectness</th>
<th>Imprecision</th>
<th>Publication bias</th>
<th>Placebo</th>
<th>Antibiotics</th>
<th>Absolute risk (95% CI)</th>
<th>Control risk (95% CI)</th>
<th>Risk difference (95% CI)</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain at 24h 5 (RCT)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>241/605</td>
<td>223/624</td>
<td>RR 0.9 (0.78–1.04)</td>
<td>367/1,000</td>
<td>Not Significant</td>
<td>★★★★ High</td>
</tr>
<tr>
<td>Pain at 2–7 d 10 (RCT)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>No serious indirectness</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>303/1,366</td>
<td>228/1,425</td>
<td>RR 0.72 (0.62–0.83)</td>
<td>257/1,000</td>
<td>72 fewer per 1,000 (44–98)</td>
<td>★★★★ High</td>
</tr>
<tr>
<td>Hearing, inferred from the surrogate outcome abnormal tympanometry—1 mo 4 (RCT)</td>
<td>No serious limitations</td>
<td>No serious inconsistency</td>
<td>Serious indirectness (because of indirectness of outcome)</td>
<td>No serious imprecision</td>
<td>Undetected</td>
<td>168/460</td>
<td>153/467</td>
<td>RR 0.89 (0.75–1.07)</td>
<td>350/1,000</td>
<td>Not Significant</td>
<td>★★★★ Moderate</td>
</tr>
</tbody>
</table>
**Summary of Findings for policymakers**

### Routine childhood immunizations

**Patients or population:** Children up to the age of 7  
**Settings:** Diverse, some low income, in USA (11 studies) and Australia (1 study)  
**Intervention:** Reminder and recall interventions to promote immunization uptake  
**Comparison:** Usual care, except one study which used a printed schedule of routine immunizations

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Comparative risks*</th>
<th></th>
<th>Relative effect (95% CI)</th>
<th>Number of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without reminder/recall</td>
<td>With reminder/recall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immunized</td>
<td>30 per 100 (35 to 42)</td>
<td>38 per 100 (1.28 to 1.66)</td>
<td>OR 1.45</td>
<td>13 818 (12 studies)</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

CI: Confidence interval; OR: Odds ratio  
GRADE: GRADE Working Group grades of evidence (see above and last page)  
*Illustrative comparative proportions of children with up-to-date immunizations for an assumed proportion of 30 per 100 without reminders or recall, based on the overall relative effect (OR = 1.45).
Conclusions

• Guidelines are often inconsistent in how they deal with uncertainty

• Guidelines are generally one-size-fits-all

• DECIDE will tailor how research evidence is presented and evaluate the effectiveness of these new strategies

• Promising strategies will be used in real guidelines through partners such as NICE and SIGN
Thank you!

streweek@mac.com
Strength of recommendations

**Quality of evidence**
High quality
 Moderate quality
 Low quality
 Very low quality

**Strength of recommendation**
Strong recommendation for using an intervention
Weak recommendation for using an intervention
Weak recommendation against using an intervention
Strong recommendation against using an intervention
Methods for the science WPs

• Brainstorming
• Stakeholder (advisory) groups
• Surveys
• User testing
• Randomised controlled trials
• Testing in real guidelines
Decisions should be influenced not only by the best estimates of the expected effect but also by the confidence in these estimates.

Guidelines are typically developed as a one-size-fits-all package with no attempt at tailoring the guideline for particular audiences.
An example - hormone replacement therapy

Later evidence from trials* showed that doctors following these recommendations were not providing better clinical care and may have inadvertently been doing harm.

An example - hormone replacement therapy

The recommendations should have considered the *quality of evidence*, not just the estimate of effect.
The scope of the science WPs

- Aspects of how guidelines are developed that can affect their dissemination
- How evidence and other information is presented
- How recommendations and other judgements are formulated
- How information is disseminated
Our objective

To *improve the dissemination* of evidence-based recommendations by building on the work of the GRADE Working Group to *develop and evaluate* methods that address the targeted dissemination of guidelines.

*Kick-off meeting was two weeks ago in Geneva.*
July 2008–SUPPORT Summary of a systematic review

Does integration of primary healthcare services improve health care delivery and outcomes?

Primary healthcare in many low and middle-income countries is organised through a series of vertical programmes for specific health problems such as tuberculosis control or immunisation of children. Vertical programmes can help deliver particular technologies, but may lead to service duplication, inefficiency and service fragmentation. The World Health Organization and other organizations promote integration, where inputs, delivery, management and organization of particular service functions are brought together, as a solution to such problems.

Who is this summary for?
People deciding whether to integrate primary healthcare services

This summary is based on the following systematic review:
Briggs CJ, Garner P. Strategies for integrating primary health services in middle- and low-income countries at the point of delivery. Cochrane Database of Systematic Reviews 2006, Issue 2.

What is a systematic review?
A summary of studies addressing a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise the relevant research, and to collect and analyse data from the included studies.
WP7 Dissemination and coordination

Strategies for targeted dissemination to key stakeholders who determine what happens in clinical practice

WP1 Strategies for communicating evidence-based recommendations to clinicians

WP2 Strategies for communicating evidence-based recommendations to policymakers and managers

WP3 Strategies for communicating evidence-based recommendations to patients and the general public

WP4 Strategies for communicating evidence-based recommendations about diagnostic tests

WP5 Strategies for communicating evidence-based recommendations about public health and health systems policies

WP6 Strategies for collaboration among guideline developers and HTA agencies in Europe

Strategies for targeted dissemination of different types of recommendations
WP7 Dissemination and coordination

Strategies for targeted dissemination to **key stakeholders** who determine what happens in clinical practice

WP1 Strategies for communicating evidence-based recommendations to **clinicians**

WP2 Strategies for communicating evidence-based recommendations to **policymakers and managers**

WP3 Strategies for communicating evidence-based recommendations to **patients and the general public**

WP4 Strategies for communicating evidence-based recommendations about **diagnostic tests**

WP5 Strategies for communicating evidence-based recommendations about **public health and health systems policies**

WP6 Strategies for **collaboration** among guideline developers and HTA agencies in Europe

Strategies for targeted dissemination of **different types of recommendations**