WHO EVIDENCE-INFORMED GUIDELINE DEVELOPMENT PROCESS

Ludovic Reveiz, MD, MSc, PhD
(on behalf of Dr Juan Pablo Peña-Rosas)
Advisor Knowledge Translation and Evidence
Knowledge Management, Bioethics and research
Pan American Health Organization

March 16–17, 2015
### Disclosures for: Ludovic Reveiz

| Affiliation / Financial interests | Pan American Health Organization  
Member of the WHO Guideline Review Committee  
No financial interests to declare |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intellectual:</td>
<td>Author SRs Cochrane groups, Internal referee of Cochrane Pregnancy and Childbirth Group</td>
</tr>
</tbody>
</table>
WHO Evidence-Informed Guideline Development Process

- WHO continually provide health recommendations
- Many concerns were raised related with the variety of methods, format of presentation and type of recommendations
WHO Evidence-Informed Guideline Development Process

Use of evidence in WHO recommendations

Andrew D Oxman, John N Lewis, Afti Fretheim

Summary

Background: WHO regulations, dating back to 1951, emphasise the role of expert opinion in the development of recommendations. However, the organisation's guidelines, approved in 2003, emphasise the use of systematic reviews for evidence of effects, processes that allow for the explicit incorporation of other types of information (including values), and evidence-informed dissemination and implementation strategies. We examined the use of evidence, particularly evidence of effects, in recommendations developed by WHO departments.

Methods: We interviewed department directors (or their delegates) at WHO headquarters in Geneva, Switzerland, and reviewed a sample of the recommendation-containing reports that were discussed in the interviews (as well as related background documentation). Two individuals independently analysed the interviews and reviewed key features of the reports and background documentation.

Findings: Systematic reviews and concise summaries of findings are rarely used for developing recommendations. Instead, processes usually rely heavily on experts in a particular speciality, rather than representatives of those who will have to live with the recommendations or on experts in particular methodological areas.

Interpretation: Progress in the development, adaptation, dissemination, and implementation of recommendations for member states will need leadership, the resources necessary for WHO to undertake these processes in a transparent and defensible way, and close attention to the current and emerging research literature related to these processes.

Introduction

Every year, WHO develops a large number of recommendations aimed at many different target audiences, including the general public, healthcare professionals, managers working in health facilities (e.g., hospitals) or regions (e.g., districts), and public policymakers in member states. These recommendations address a wide range of clinical, public health, and health policy topics related to achieving health goals. WHO's regulations emphasise the role of expert opinion in the development of such recommendations. However, relatively few recommendations integrate fully the collection, analysis, and interpretation of the results. However, systematic reviews are only as good as the evidence that they summarise. There might be no evidence. When there is evidence, judgments are still needed about the quality and, especially for public health and health policy topics, its applicability in different contexts.

Evidence of effects needs to be complemented by information about needs, factors that could affect whether effectiveness will be realised in the field, such as the following:
In response to concerns about the quality of WHO guidelines, and following up on recommendations by The Advisory Committee on Health Research (ACHR) and resolution EB120.R15 of the 120th Session of the Executive Board, this note announces the establishment of a WHO Guidelines Review Committee (GRC). The GRC will develop and implement standards and procedures for guideline development that ensure that WHO guidelines are consistent with internationally accepted best practice, including appropriate use of evidence.
GRC: Terms of Reference

1. Defining appropriate and standardized processes related to guideline development
2. Ensuring that all guidelines prepared by WHO comply with the WHO Handbook for Guideline Development
3. Developing and implementing a plan to build capacity of WHO staff to develop guidelines
4. Develop collaboration and cooperation with other organizations and international networks that have methodological expertise
Quality: 123 WHO Guidelines

WHO Guidelines Production Process

1. A WHO department decides to produce a guideline
2. Initial review by GRC
3. Initial approval for development
4. Final approval by GRC
5. Relevant approvals are obtained (ADG or DGO)

ADVICE AND SUPPORT
- from GRC Secretariat
- from GRC members
- from WHO Collaborating Centres
- from GRC through WHO lists of technical experts
- from external experts on guideline production
### Characteristics of the types of guidelines produced by WHO

<table>
<thead>
<tr>
<th>Primary types of guidelines</th>
<th>Purpose</th>
<th>Scope</th>
<th>Developer</th>
<th>New or existing recommendations</th>
<th>Development period</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard</strong></td>
<td>To provide recommendations on a specific topic or condition</td>
<td>Focused or comprehensive</td>
<td>WHO technical staff</td>
<td>Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate</td>
<td>6 months to 2 years</td>
</tr>
<tr>
<td><strong>Consolidated</strong></td>
<td>To aggregate all the existing guidance on a disease or condition</td>
<td>Comprehensive</td>
<td>WHO technical staff</td>
<td>Existing recommendations that have been evaluated and found to be up to date; may contain some new recommendations</td>
<td>1 to 2 years</td>
</tr>
<tr>
<td><strong>Interim</strong></td>
<td>To provide guidance when new interventions, exposures or diseases arise or when new evidence becomes available or data are likely to be incomplete</td>
<td>Focused</td>
<td>WHO technical staff</td>
<td>New</td>
<td>6 to 9 months</td>
</tr>
<tr>
<td><strong>Guidelines produced in response to an emerging or urgent need</strong></td>
<td>To meet an emergent or urgent public health need when the short timeline mandates a modified process</td>
<td>Focused</td>
<td>WHO technical staff</td>
<td>Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate</td>
<td>1 to 3 months</td>
</tr>
</tbody>
</table>

#### Other types of guidelines

| Developed in collaboration with (an) external organization(s) | To provide recommendations on a specific topic or condition when organizations have a shared interest or remit | Focused or comprehensive | WHO technical staff and staff from the external organization(s) | Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate | 1 to 2 years |
| Developed by (an) external organization(s) | To provide recommendations on a specific topic or condition when a guideline produced by an external organization already exists | Focused or comprehensive | External organization(s) | Existing recommendations may be updated | 1 to 3 months |
| Adaptation of existing WHO guidelines | To develop recommendations specific to the local context where they will be implemented | Focused or comprehensive | Policy-makers and programme managers in WHO Member States | Reflect the content of the original guideline | 1 to 3 months |
WHO Guidelines...

- Must meet the highest quality standards for evidence-informed guidelines
- Must be based on high-quality systematic reviews of all relevant evidence
- Use GRADE, which provides an explicit approach to:
  - Assessing the quality of the evidence across studies and outcomes
  - Translating evidence to recommendations
- Incorporate multiple processes to minimize bias and optimize usability
- Must incorporate transparency in all judgments and decision making processes
WHO Evidence-Informed Guideline Development Process

WHO handbook for guideline development. 2 edition (2014)
WHO Evidence-Informed Guideline Development Process

**WHO Guideline Steering Committee**
- WHO Departments
- Directors or alternate appointee

**Guideline Development Group**
- Nutrition Actions
  - Geographic representation
  - Multi-disciplinary
  - Gender-balanced
  - Un-conflicted as possible
  - 20-25 members overall

**External Peer Review**
- 2 external reviewers
- + open documented process
- • WHO Nutrition Mailing List
- • SCN Mailing List
- • WHO Nutrition Website
- • Social Media

Slide  Content provided by Dr Juan Pablo Peña-Rosas.
Slide content provided by Dr Alonso Carrasco – Grade Working group
Evidence Retrieval, Assessment and Synthesis [Systematic Review(s)]

WHO has followed four approaches to retrieve, assess and synthesise the evidence:

1. Use of existing systematic reviews, and contact authors if not current
2. Build on the systematic reviews developed by other groups
3. Commission "tailored" systematic reviews
4. Lead review teams to undertake the systematic reviews

Slide Content provided by Dr Juan Pablo Peña-Rosas.
The Grading of Recommendations Assessment, Development and Evaluation Approach

Clear separation of two issues:

1) Quality of the evidence (high, moderate, low, very low)
   - methodological quality of evidence
   - likelihood of bias
   - by outcome

• Ideally, people who grade evidence should have available to them systematic reviews of the evidence regarding the benefits and risks of the alternative management strategies they are considering

• Better research gives better confidence in the evidence (and the following decisions)
# The Quality of the Evidence

The extent to which one can be confident that an estimate of effect or association is correct.

<table>
<thead>
<tr>
<th>Quality</th>
<th>Further Research Description</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Further research <strong>is very unlikely to change</strong> our confidence in the estimate of effect</td>
<td>✦✦✦✦</td>
</tr>
<tr>
<td>Moderate</td>
<td>Further research <strong>is likely</strong> to have an important impact on our confidence in the estimate of effect and may change the estimate.</td>
<td>✦✦✦</td>
</tr>
<tr>
<td>Low</td>
<td>Further research <strong>is very likely</strong> to have an important impact on our confidence in the estimate of effect and <strong>is likely</strong> to change the estimate.</td>
<td>✦✦✦</td>
</tr>
<tr>
<td>Very low</td>
<td>Any estimate of effect <strong>is very uncertain</strong></td>
<td>✦✦✦✦</td>
</tr>
</tbody>
</table>
The Quality of the Evidence

Author(s): Ignacio Neumann, Luz M Letelier, Gabriel Rada, Juan Carlos Claro, Janet Martin, Colin W Howden, Yuhong Yuan, Grigoris I Leontiadis

Date: 2013-08-11

Question: Should high dose regimen vs non-high dose regimen be used for acute peptic ulcer bleeding?

Settings: Hospital


<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>No of patients</th>
<th>Effect</th>
<th>Quality</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High dose regimen</td>
<td>Non-high dose regimen</td>
<td>Relative (95% CI)</td>
<td>Absolute</td>
</tr>
<tr>
<td>Mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Rebleeding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Further EHT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Length of hospital stay (Better indicated by lower values)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
</tr>
<tr>
<td>Blood transfusions (Better indicated by lower values)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 randomised trials</td>
<td>serious&lt;sup&gt;1&lt;/sup&gt;</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
</tr>
</tbody>
</table>

<sup>1</sup> Of the 22 trials included in this review, 17 had high risk of bias, 5 unclear risk of bias and none low risk of bias. The main limitation was the lack of blinding in 16 trials.

<sup>2</sup> It is not possible to exclude a clinically relevant benefit or harm.
### Self management for patients with chronic obstructive pulmonary disease

**Patient or population:** patients with chronic obstructive pulmonary disease

**Settings:** primary care, community, outpatient

**Intervention:** self management

**Comparison:** usual care

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks*</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| St George’s Respiratory Questionnaire. | The mean quality of life ranged across control groups from 38 to 60 points | The mean quality of life in the intervention groups was 2.58 lower (5.14 to 0.02 lower) | 698 (7) | $$
| Dyspnoea Borg Scale. Scale from: 0 to 10. | The mean dyspnoea ranged across control groups from 1.2 to 4.1 points | The mean dyspnoea in the intervention groups was 0.53 lower (0.96 to 0.1 lower) | 144 (2) | $$
| Number and severity of exacerbations* | See comment | See comment | Not estimable* (3) | See comment | Effect is uncertain |
| Respiratory-related hospital admissions (follow-up: 3 to 12 months) | Low risk population* | | | $$
| | 10 per 100 | 7 per 100 | (5 to 9) | OR 0.64 (0.47 to 0.89) | (8) | $$
| | High risk population* | | | $$
| | 50 per 100 | 39 per 100 | (32 to 47) | | | |
| Emergency department visits for lung diseases (follow-up: 6 to 12 months) | The mean emergency department visits for lung diseases ranged across control groups from 0.2 to 0.7 visits per 0.1 higher person per year | The mean emergency department visits for lung diseases in the intervention groups was 0.2 lower to 0.3 higher | 328 (4) | $$
| Doctor and nurse visits (follow-up: 6 to 12 months) | The mean doctor and nurse visits ranged across control groups from 1 to 5 visits per person per year | The mean doctor and nurse visits in the intervention groups was 0.02 higher (1 lower to 1 higher) | 629 (8) | $$

*The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

CI: Confidence interval; OR: Odds ratio;
The Grading of Recommendations Assessment, Development and Evaluation Approach

2) Two grades of recommendation: strong or conditional (*for or against*)
   - Quality of evidence only one factor
   - Considers values and preferences, trade-off of harms and benefits and feasibility of implementation
   - Evidence alone is never sufficient to make a clinical or public health decision
   - One qualifier per recommendation over critical outcomes

<table>
<thead>
<tr>
<th>Disease burden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evidence</td>
</tr>
<tr>
<td>Balance benefits/harms</td>
</tr>
<tr>
<td>Acceptability</td>
</tr>
<tr>
<td>Resource use</td>
</tr>
<tr>
<td>Feasibility</td>
</tr>
<tr>
<td>Equity</td>
</tr>
</tbody>
</table>

Slide Content provided by Dr Juan Pablo Peña-Rosas.
Recommendation 1

The expert panel recommends cryotherapy over no treatment for women who have histologically confirmed CIN2+ disease

(strong recommendation, ⭐⭐⭐⭐ evidence)

Remarks: This recommendation is strong, although the available evidence was very low quality. The expected benefit of cervical cancer prevention is very high and outweighs harms and any use of resources, but there is uncertainty related to preterm delivery in future pregnancies. However, the panel felt that women would prefer to be treated despite the uncertainty of these risks. This recommendation applies to women regardless of HIV status.

Evidence-to-recommendation table

<table>
<thead>
<tr>
<th>Decision domain</th>
<th>Judgement</th>
<th>Summary of reason for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of evidence</td>
<td></td>
<td>There is low- to very-low-quality evidence from non-randomized studies with no independent control. There was also imprecision as a result of few events or participants in the studies, inconsistency, and/or risk of bias as a result of selective reporting of complications.</td>
</tr>
<tr>
<td>Is there high or moderate quality evidence?</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Balance of benefits versus harms and burdens</td>
<td></td>
<td>Residual/recurrence rates of CIN2+ are probably lower with cryotherapy resulting in lower risk of cervical cancer and related mortality compared to no treatment. These benefits outweigh the low risk of major bleeding and infection with cryotherapy, and the unclear risk of premature delivery or spontaneous abortion.</td>
</tr>
<tr>
<td>Are you confident that the benefits outweigh the harms and burdens for the recommended strategy?</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Values and preferences</td>
<td></td>
<td>A high value was placed on the risk of cervical cancer and mortality with no treatment. The panel felt that women would prefer to be treated despite the uncertainty of the risks related to reproductive outcomes.</td>
</tr>
<tr>
<td>Are you confident about the assumed or identified relative values and are they similar across the target population?</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Resource implications</td>
<td></td>
<td>The resources for cryotherapy when no other treatments are available are worth the net benefits.</td>
</tr>
<tr>
<td>Is the cost small relative to the net benefits for the recommended strategy?</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>
Challenges

- Paucity of evidence: especially policy and health system interventions in LMIC
- Low quality evidence, weak study designs
- Diverse user needs: applicability and implementation
- Identifying and building capacity for methodological expertise
- Remaining culture of expert opinion
- Guideline dissemination
- Creating global guidelines based on (nonlocal) evidence
- Guideline adoption and adaptation (contextualization)
- Implementation into policy into actions
- Process evaluation 2013:
  - Are WHO evidence-informed nutrition guidelines credible, useful and feasible to implement?
  - Need more translation knowledge on how to implement and how to integrate into broader programs
  - Is the process followed worthy?
Impact of WHO Guidelines

✓ Uptake of recommendations included in HIV and TB guidelines issued by WHO from 2009 to 2013 across guidelines from 20 low- and middle-income countries in Africa and Southeast Asia

✓ 82% of strong recommendations and 61% of conditional recommendations were adopted in national guidelines (overall 76% of WHO recommendations)

✓ An association with the strength of recommendation (strong) and with the higher the level of quality of evidence was found

Impact of WHO Guidelines

TRANSLATING RESEARCH INTO ACTION: WHO EVIDENCE-INFORMED GUIDELINES FOR SAFE AND EFFECTIVE MICRONUTRIENT INTERVENTIONS

Juan Pablo Peña-Rosas,1,2 Luz Maria De-Regil,1 Lisa M. Rogers,1 Ameya Bopardikar,1 and Ulysse Pasinset2

1Micronutrients Unit, Department of Nutrition for Health and Development, and 2Evidence and Networks for Health - EVIPNet, World Health Organization, Geneva, Switzerland

Abstract

In 2009 WHO adopted a new process by which recommendations for safe and effective micronutrient interventions are developed, ensuring the use of best practices and available evidence. This process includes nine steps ranging from establishing steering and guideline groups and prioritizing needs to planning the implementation and updating the guidelines. Systematic reviews of evidence are used to address critical outcomes for decision making, considering the balance among risks and benefits, values, preferences, and costs. Grading of Recommendations, Assessment, Development and Evaluation (GRADE) methodology is used to assess the overall evidence quality and establish the strength of the recommendations. Guideline development is underway for interventions covering iron and vitamin A supplementation, home fortification with multiple micronutrient powders, and fortification of staple foods. Global guidelines are disseminated through the WHO electronic Library of Evidence for Nutrition Actions, a resource of the evidence and tools for scaling-up micronutrient interventions. The WHO Department of Nutrition for Health and Development and the Evidence-Informed Policy Network will support countries to scale-up the delivery of micronutrient interventions by adopting these evidence-informed guidelines and policies to make them context specific. This will be accomplished by providing summaries of the best available evidence on micronutrient interventions, evidence on health systems, and effective delivery systems along with capturing the tacit knowledge of the countries’ realities. With a systematic approach that uses the WHO strategies on research for health as the connecting thread, the challenges to successfully implement safe and effective micronutrient programs can be addressed.

Introduction

High-quality research and evidence are critical for improving global health and equity and, ultimately, for all people to attain the highest possible level of health. For the WHO, improved response to them. This definition covers the full spectrum of research, which spans five generic areas of activity: measuring the problem; understanding its cause(s) and determinants; developing solutions; translating the solut-
Thank you