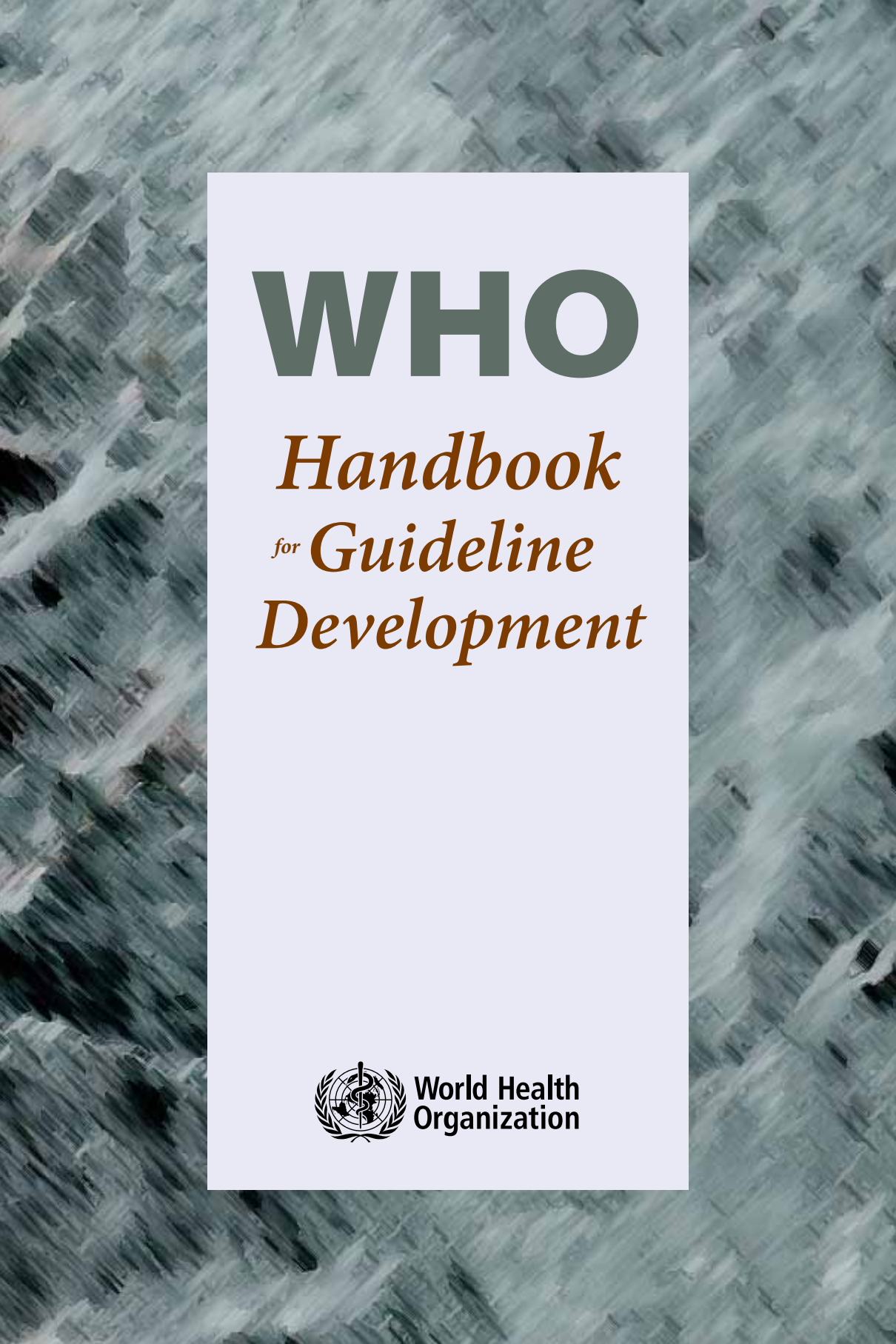


WHO

Handbook for Guideline Development

World Health Organization

WHO

Handbook for Guideline Development

World Health Organization

WHO Library Cataloguing-in-Publication Data

WHO handbook for guideline development.

1.Guidelines as a topic – standards. 2.Review literature. 3.Meta-analysis. 4.Peer review. 5.Evidence-based medicine. 6.World Health Organization. I.World Health Organization.

ISBN 978 92 4 154844 1 (NLM classification: WA 39)

Contents

1. Introduction

What is a WHO guideline?

A WHO guideline is any document containing recommendations about health interventions, whether these are clinical, public health or policy recommendations. A recommendation provides information about what policy-makers, health-care providers or patients should do. It implies a choice between different interventions that have an impact on health and that have implications for the use of resources. Guidelines are recommendations intended to assist providers and recipients of health care and other stakeholders to make informed decisions. WHO has adopted internationally recognized standards and methods for guideline development to ensure that guidelines are free from bias, meet a public health need and are consistent with the following principles.

- Recommendations are based on a comprehensive and objective assessment of the available evidence.
- The process used to develop the recommendations is clear. That is, the reader will be able to see how a recommendation has been developed, by whom, and on what basis.

What is the aim of this handbook?

This handbook provides stepwise advice on the technical aspects of developing a WHO guideline and the methods used. It aims to provide a clear path through the process and seeks to ensure that the resulting guidelines have credibility and meet WHO's criteria for content, methods and presentation, while remaining accessible and useful.

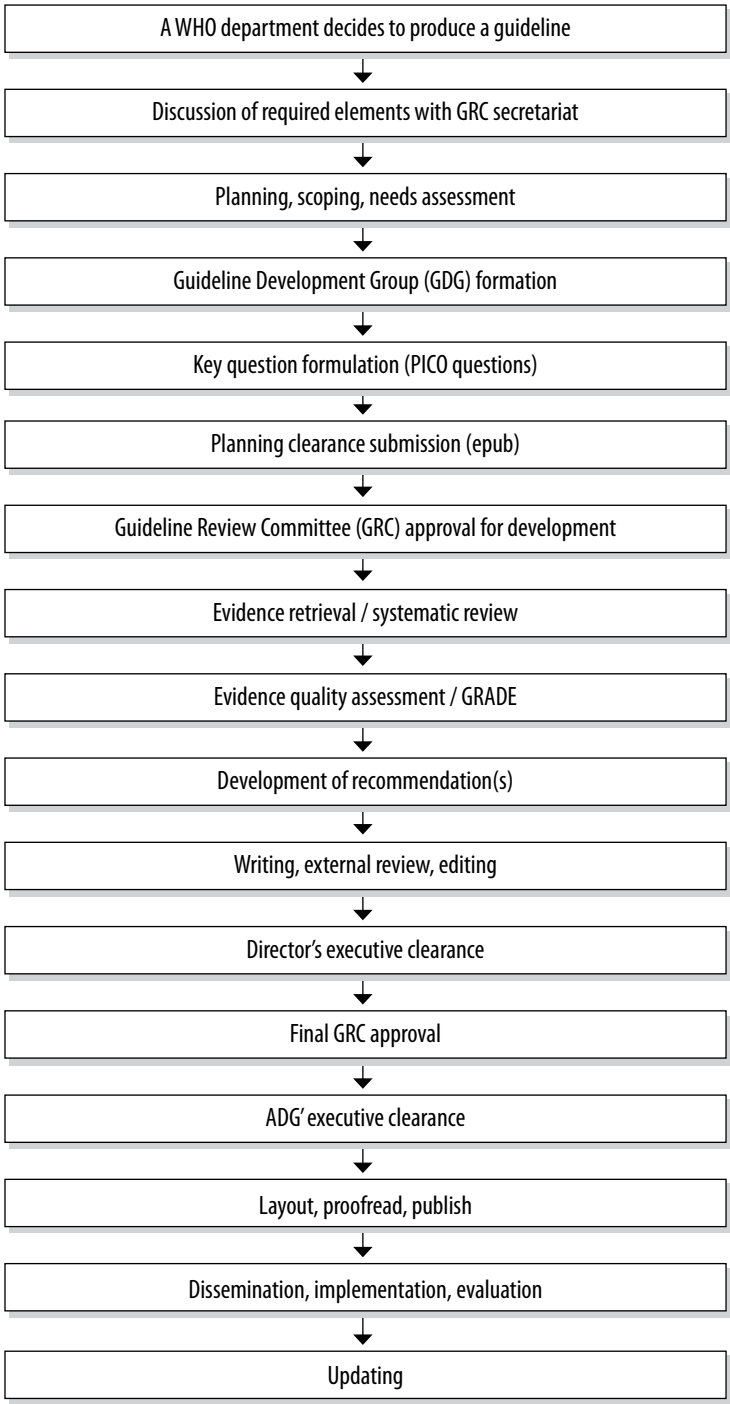
Who is the handbook for?

- Any WHO department who decides to produce a guideline.
- Members of the WHO steering group.
- Members of the guideline development groups (GDGs).
- Members of the external review group.
- Anyone interested in understanding how WHO develops guidelines.

How to use this handbook?

The structure follows the development of a WHO guideline from start through to publication. The guideline development process is summarized in Figure 1.1.

Figure 1.1 Overview of guidelines development process



Types of guidelines

If you are planning to produce a guideline, consider which of the following types of product best fits your purpose. The type of product will determine the methods and timeframe for development.

Rapid advice guidelines

A rapid advice guideline is produced in response to a public health emergency (such as pandemic influenza) in which WHO is required to provide rapid global leadership and guidance. This type of document needs to be produced within 1–3 months and will be evidence-informed, but it may not be supported by full reviews of the evidence. It will be prepared mainly by the responsible WHO staff members with external consultation and peer review. It must be published with a review-by date that indicates when the guidance will become invalid, or when it will be updated or converted to a standard guideline.

Standard guidelines

A standard guideline is produced in response to a request for guidance in relation to a change in practice or controversy in a single clinical or policy area – such as treatment of postpartum haemorrhage or minimum requirements for safe delivery of HIV care. A standard guideline is not expected to cover the full scope of the condition or public health problem. This guideline will usually take 9–12 months to complete and should be prepared after consultation on the scope of the guideline and the issue that it covers. It should be supported by systematic reviews of the evidence and one or two meetings of the guideline development group for consultation. A standard guideline may have a specified review-by date depending on the expected rate of change of evidence in the topic area. Most WHO guidelines fall into this category.

Full guidelines

A full guideline is one that provides complete coverage of a health topic or disease, such as dengue fever. It would be expected to include recommendations in relation to all aspects of the topic (e.g. surveillance, diagnosis, public health and clinical interventions) and to be fully based on systematic reviews of the evidence for each aspect. These are likely to take 2–3 years to complete, and will require several meetings of a guideline development group. Given the time and expense of producing full guidelines, the need for doing these in WHO needs to be carefully justified.

Compilations of guidelines

A compilation of guidelines contains current recommendations from WHO and other sources, but does not include any new recommendations. Compilations of guidelines are subject to Guidelines Review Committee (GRC) approval. All recommendations included must be current and should be referenced thoroughly and accurately. Producing a compilation of guidelines can be complex and updating may be difficult since individual recommendations may go out of date at different times.

In principle, all recommendations used in a compilation should be updated by WHO. However, recognizing that WHO resources are limited, this may not be realistic. Members of the guideline development group should discuss and agree on an acceptable level of quality and document their decisions carefully. The GRC recommends using the Appraisal of Guidelines for Research and Evaluation (AGREE) tool (available at http://www.agreetrust.org/) to do this.

It is also important that recommendations used in a compilation are of adequate quality. WHO recommendations are considered of adequate quality for use in a compilation if they were cleared by the GRC from 2009 onwards. If compiled recommendations have not been cleared by the GRC, an explicit and systematic process must be in place to ensure the quality of the compiled guidance. Production times for compilations of guidelines vary widely.

Some guideline compilations do not require GRC review. These are:

- documents in which all the recommendations have previously been cleared by the GRC under its full (not transitional) requirements;
- documents that are clearly limited to operational guides for such guidelines.

Guideline compilations that require GRC review are documents in which any of the recommendations were initially published without GRC review.

Adaptations of guidelines

Guidelines originally intended for one setting may be adapted for use in another, such as routine obstetric care in emergency settings. Adaptations of guidelines must follow standard GRC procedures.

Guidelines prepared in collaboration with other organizations

Health-care guidelines are produced by many organizations, including national agencies, intergovernmental organizations and specialist medical societies. From time to time, it may be appropriate for WHO to collaborate with these groups to produce a joint guideline. However, national agency guidelines usually have a much narrower focus than those produced by WHO, and international society guidelines may have inherent problems owing to conflicts of interest in the funding of their development. The GRC will make case-by-case assessments of these types of proposals. However, joint guidelines must follow current WHO guideline development standards as outlined in this handbook. In addition to being aware of potential problems with regard to copyright, it is important to note that:

- adaptation or endorsement of another organization's guideline should be initiated by the WHO department concerned and not by the external group;
- adaptation or endorsement of another organization's guideline can be considered when no WHO guideline exists or an existing WHO guideline is outdated;
- minimum standards for WHO guidelines should be met (no funding from commercial sources, evidence systematically reviewed, conflicts of interest declared and reported, and methods of developing the guideline reported);

- the approach to reviewing and summarizing evidence should be consistent with that recommended for WHO guidelines;
- WHO should ensure global representation of experts in the development of the recommendations;
- the recommendations should be appropriate for a global audience.

Information products that are NOT considered guidelines

- Documents containing standards for manufacturing health technologies, such as pharmaceuticals and vaccines.
- 'How to' documents, or operational manuals (e.g. how to set up a research project or how to implement a service).
- Documents that describe standard operating procedures for organizations or systems.
- Documents that state established principles (e.g. ethics, human rights, WHO constitutional issues).
- Documents that provide information on different options for interventions without recommending any particular intervention.

If you are not sure whether your proposed document is a guideline, please submit it to the GRC for review.

The Guidelines Review Committee

Why was the GRC set up?

The Guidelines Review Committee (GRC) was established by the Director-General in 2007 to ensure that WHO guidelines are of high quality and are developed through a transparent, evidence-based decision-making process.

Since this date, all WHO publications containing recommendations must be approved by the GRC. Such publications are required to meet an unmet need, to be developed using internationally accepted best practices, including the appropriate use of evidence. This handbook provides guidance on the development of documents or publications containing WHO recommendations, and sets out the procedures to follow when such a document is submitted to the GRC for approval. To facilitate ease of reading, the term 'guideline' is used to refer to any document containing WHO recommendations.

The GRC reviews every WHO guideline twice during its development – once after the scope of the guideline has been defined at the initial planning stage, and again after the recommendations have been developed and the guideline document has been edited. The GRC meets on a monthly basis to review both initial proposals for guideline development and final versions of guidelines prior to their publication. The review of the initial proposals includes an assessment of whether the proposed guideline development process is consistent with the steps described in this handbook. The review of final submissions is done to ensure that the approved process has been followed and that the final guideline

document meets all reporting requirements and contains clear and actionable recommendations. The GRC also offers suggestions and advice on how to improve the quality of the guidelines at any stage of the process.

To allow adequate time for review, all relevant documents must be submitted to the GRC, through the publication clearance system no later than two weeks before the date of the next meeting.

The GRC can be contacted at grcinfo@who.int.

The GRC Secretariat

The principal aims of the Secretariat are to:

- coordinate and provide technical support on guidelines development to WHO departments, headquarters and regional offices;
- organize training on guideline production for WHO staff;
- provide administrative support for the work of the GRC;
- collaborate with other organizations and international networks that provide methodological expertise in relation to guideline development, adaptation and implementation;
- maintain the database of the GRC submissions.

2. **Planning guidelines**

Good planning will yield good guidelines

The first and most important step when planning is to ask a single question: *is this guideline really needed?*

WHO guidelines should meet a defined global need, have a public health perspective and not duplicate existing advice. Consult other departments at the beginning and decide, as early as possible, who should have primary responsibility for developing the guideline and who should be involved.

Who wants it? Is it a request from one or more WHO Member States? WHO guidelines generally should meet a global need, have a public health perspective and not duplicate existing resources. If an existing guideline meets the need, a new one is not required.

Why now? Is it required by WHO's governing bodies? Are there already guidelines on the same topic from other organizations or other WHO departments? Is the best advice on this topic available only from WHO?

Is it part of a departmental programme of work? Implementation of a guideline by WHO headquarters or by countries will be much easier if it fits with a programme or project. If no programme or project exists, is it really necessary to prepare the guideline?

Implementation? Who is likely to implement it? If you cannot identify a process for implementation, then you should not start.

What will it achieve? Will the guideline address poor practice or to try to change clinical programme approaches, or health policy? This should be the focus of most guidelines, and it is what differentiates guidelines from textbooks or reference works.

When is it needed? Is the guideline a response to a situation where need for advice is urgent? If so, consider producing rapid advice guidelines. These guidelines usually need to be produced and published as quickly as possible, ideally in 1–3 months and therefore the requirements and processes are different from those of other guidelines.

Agreement? Do you have agreement from your director? You will need to have formal approval from your director before your proposal can be considered by the GRC and your Assistant Director-General (ADG) will need to approve the proposal and final product.

Collaboration? Are there other departments that should be involved, or that might be producing similar products? The answer to this is nearly always yes. Avoid duplicating earlier or current work by consulting other relevant WHO

departments, the GRC secretariat and the WHO library. Make a preliminary search of published work relevant to your planned guideline. Once you have identified the relevant departments, decide which department should have primary responsibility for the guideline and who will be involved in developing it.

If you cannot answer all these questions, it is probably best not to start.

Practical planning

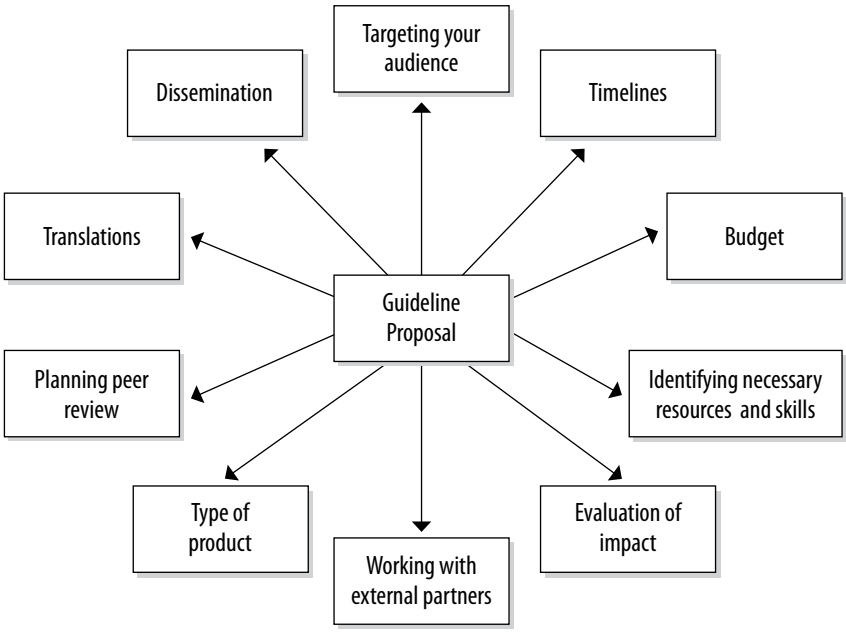
Having established that there is good reason to develop a guideline the next step is to answer some practical questions (Figure 2.1).

a. **Setting objectives.** Why are you doing this? What is the need you are responding to and why does WHO need to be producing this document? Set clear, achievable objectives that will govern development of your guideline.

b. **Targeting your audience.** Who is your target audience? Most WHO guidelines need to speak to multiple audiences, which makes them challenging to produce. If you can identify the key target audience, your task will be easier. Writing documents to meet the needs of policy-makers, health-care managers and clinicians simultaneously is not straightforward and should be avoided wherever possible.

c. **Timelines.** When does it need to be completed? Realistically, a good quality guideline will take at least 9–12 months to produce if all the evidence has already been synthesized and you have someone to write it. If the guideline is going to cover a large number of questions, it may take up to 2–3 years to produce.

Figure 2.1 Planning your guideline



d. **Funding.** Do you have adequate funds? For a standard WHO guideline, assuming that you will need to commission systematic reviews, an evidence synthesis and assessment, hold at least one consultation meeting, pay for writing, editing and layout of your document, please allow at least US$ 300 000. Note that WHO may not accept money from commercial bodies for guideline development and sources of funding for guidelines may need to be approved by the legal department.

e. **Existing guidance and resources.** Are there existing guideline documents that cover the same issue? If so, what is the added value and justification for the proposed document? If a WHO version is needed to build on an area covered by an existing guideline from a recognized national developer (e.g. the UK's National Institute for Health and Clinical Excellence), the existing guidelines can be used as a starting point. To be considered for adaptation, third-party guidelines should have been developed using standards equivalent to those described in this Handbook to ensure transparency and freedom from bias. Consider updating existing WHO recommendations if they are out of date or of low quality. When examining existing guidelines, do a quality assessment using a tool such as the Appraisal of Guidelines for Research and Evaluation (AGREE) tool available at http://www.agreetrust.org/

f. **The evidence base.** What existing scientific evidence can guide the recommendations? Do you know of existing systematic reviews? If not, it is worth doing a preliminary literature search at this stage to get a sense of what information is available. For standard and full WHO guidelines, a systematic search for evidence should be completed before developing the recommendations. If there is no evidence, what will be the basis of your guideline?

g. **Who should be involved?** It is worth spending some time at the beginning of the process to draw up a list of key external organizations, experts and stakeholders who will need to be consulted or involved in the process.

   - **First,** identify your WHO guideline steering group (WHO staff members responsible for guideline development).
   - **Second,** identify members of your guideline development group who will be actively involved in the development of the guidelines (usually 10–20 persons).
   - **Third,** you should establish an external review group made up of experts and stakeholders whom you may wish to consult on the scope of the document, the questions it covers and the choice of important outcomes for decision-making. Members of this group should also review the completed draft guideline. The external review group may include groups likely to oppose or criticize the output on the basis of scientific or philosophical differences. While it may not be possible to reach agreement with them, it is important to consider their input.

In addition, many of these groups and experts will play a key role in the implementation of the recommendations in the guideline; they are more likely to help implement the recommendations if they are involved from the beginning.

h. **Type of publication.** Consider what format will be most useful for your guideline users. Electronic versions may be more practical and cheaper, perhaps accompanied by short paper publications, wall charts, pamphlets, etc.

i. **Translations.** Are you planning translations? Which languages are spoken by those most in need of the advice in your guideline? Consider the implications for your budget and time frame and choose your guideline language carefully.

Having established all this, you are ready to move to the next stage, scoping your guideline.

Scoping the guideline

Scoping the guideline is the process of defining what the guideline will include and what it will not include. The scope should describe:

- the area of practice or policy to which the guideline applies
- those whom the recommendations are intended to affect
- the actions and interventions of interest
- the outcomes that may result – both positive and negative.

The scope should yield questions that will govern the data search, and help frame likely recommendations. It should ensure that the guideline is of manageable size and adequately focused.

Scoping is considered one of the most difficult but important aspects of guideline development. If you get the scope right, the guideline should be manageable.

How to scope the guideline

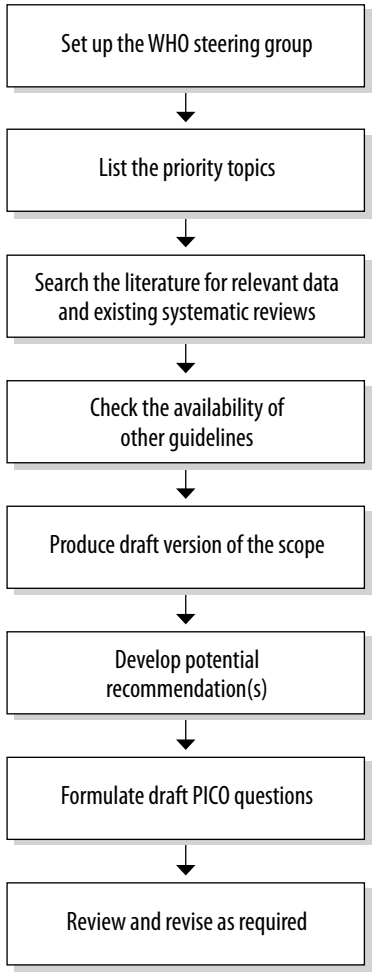
Set up a steering group. Convene a small group of WHO staff to define the scope of the guideline, including representatives of all relevant departments and ask this group to provide feedback on your assessment of priority topics, reference documents, questions and potential recommendations, as follows (Figure 2.2 Scoping procedure).

1. **List the priority topics.** What must be included? Identifying the key issues is crucial because this determines the breadth and depth of the work. Do not try to include everything; resist the temptation to write a textbook. Concentrate on the interventions or policies where change in practice is desired, and areas where there is controversy. Also consider the feasibility of implementing potential

recommendations. Although some background information may be useful, try to avoid repeating standard information (e.g. epidemiology, pathology, pharmacology) on the topic unless this is the area of controversy you wish to resolve in the guideline.

2. **Search the literature.** Do a preliminary search of the literature to identify relevant sources. This includes existing guidelines and systematic reviews, health technology assessment reports and economic evaluations relevant to the guideline topic. At this stage the search should not be exhaustive; once questions and draft recommendations have been formulated, rigorous systematic reviews will be done to retrieve the appropriate evidence.

Figure 2.2 Scoping procedure



3. **Draft potential recommendations.** Considering the potential final form of the guideline makes it easier to focus the development work.
4. **Sharpen the focus.** Take a step back and ask if you need to include all of these topics, questions and recommendations. The group should try to restrict the final list to the minimum at this stage, as it tends to expand during the development of the guideline.
5. **Formulate questions**. Use the topic list and possible recommendations to formulate the key questions to be answered in the guideline. These questions will guide the evidence synthesis and are best developed using the questions using the population, intervention, comparison, outcome (PICO) format (see section on formulating questions below).
6. **Review**. Once your group has finalized the scope, it should be circulated to the external review group for comments. (They should be reminded that WHO is producing a guideline, not a textbook, as the responses will almost always tend to expand the planned scope.)
7. **Reconsider**. Once you have the external feedback, check again. Is what you are trying to do feasible? Is your time frame reasonable? Do you have sufficient financial and human resources?

Further reading

Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development. 2. Priority setting. *Health Research Policy and Systems,* 2006, 4:14 (http://www.health-policy-systems.com/content/4/1/14, accessed 8 June 2012).

3. **Setting up guideline groups**

When developing WHO guidelines, three groups need to be set up:

- the WHO steering group
- the external review group
- the guideline development group.

Each group has different roles but the membership of each needs to be balanced so that members' technical interests, skills, expertise, values and knowledge of regional considerations complement one another and, ideally, negate potential biases.

The WHO steering group

The first group that needs to be set up is the WHO steering group. This should include members from any department or regional office that works directly on the topic of your guideline, though it is wise to keep it small (less than 10 members) to maximize efficiency. The steering group will assist with:

- scoping the guideline (described in Chapter 2);
- developing potential recommendations;
- drafting the PICO questions and overseeing evidence retrieval;
- selecting members of the guideline development group and external review groups;
- organizing guideline development meetings;
- overseeing the writing and finalization of the guidelines.

The external review group

This group is composed of people with an interest in the subject of the guideline. Members can be asked to review different stages of the development process. They may review the scope, the draft recommendations, and the PICO questions during the earlier stages. They will also be asked to review the guideline document when recommendations have been finalized. This group should be geographically and gender-balanced and include stakeholders and content experts.

The guideline development group

The guideline development group (GDG) is made up of external experts whose central task is to come up with evidence-based recommendations. The group can hold online or teleconference meetings but will usually need to have at least one face-to-face meeting. The group should be small enough for

effective group interaction but large enough to ensure adequate representation of relevant views. A group of 10–20 is usually feasible and affordable.

The role of this group is to:

- determine the PICO questions that the guideline addresses;
- choose and rank outcomes;
- provide advice, as required, on any modifications of the scope as established by the WHO steering group;
- appraise the evidence used to inform the guideline;
- advise on the interpretation of this evidence, with explicit consideration of the overall balance of benefits and harms;
- formulate the final recommendations, taking into account diverse values and preferences.

Composition of the guideline development group

The guideline development group should be multidisciplinary, gender and geographically- balanced - members should come from regions likely to use the guideline. There are different ways of finding, nominating and selecting members of guideline development groups. In addition to drawing members from established technical networks, collaborating centres and formally appointed expert advisory panels, you may wish to consider publishing an open call for nominees. Established guideline development groups within WHO have nomination procedures that you may wish to consider. Whichever mix of methods is used, the decision-making process should be documented. The aim is to have a balanced group (Figure 3.1) that includes:

- relevant technical expertise;
- implementers of the guideline such as programme managers and health professionals;
- representatives of groups most affected by the guideline, such as patients;
- methodologists (experts in assessing evidence and developing guidelines, health economists, statisticians, as required).

The group needs to identify a writer for the guideline – this person may need to be appointed in addition to the other members. A clearly written guideline and a well-documented process is critical to final clearance and subsequent use, so the writer needs to be involved throughout all the planning and development stages. Inclusion of end-users, either in the guideline group and/or in the external review group, increases the likelihood of producing a guideline that is appropriate to their needs and that will be implemented effectively.

The chair

The selection of the chair of the group is a key decision. Look for a chair who is expert in facilitating groups and interpreting evidence. While content knowledge is important, content experts with strong views about particular interventions should not chair the group. Where the best choice is a content expert, options for reducing risk of bias include ensuring that the chair does not have a veto within the group.

Technical experts

Technical content experts are selected for their expertise in the subject of the guideline. A balanced group includes a range of expertise and affiliations, with representatives from professional groups who will be implementing the guideline in each region.

End-users

These members represent groups affected by the likely recommendations (e.g. people with diabetes if the guidelines are about the management of diabetes, labour union representatives if the guidelines are about human resources for health) and/or groups likely to implement the guidelines (e.g. palliative care nurses for guidelines about pain management). Although it can be challenging to find such representatives for global guidelines, an increasing number of consumer groups are operating at international level. Many countries have nongovernmental organizations with members who may be able to participate, either as observers on behalf of their organizations, or in their individual capacity, as full members.

Involving consumers in guideline groups helps to ensure that:

- the questions addressed are relevant to consumers
- relevant aspects of the experience of illness are considered
- critical outcomes are identified and prioritized
- the final guideline can be understood by those it affects.

Barriers to consumer participation include:

- the lack of suitable consumer groups
- time constraints
- the complexity of scientific terminology used by committees.

Experience from organizations such as the United Kingdom's National Institute for Health and Clinical Excellence shows that consumers, provided with training and support, make critical contributions to guideline development.

Methodologist

Many WHO guideline groups include a methodologist, an expert in guideline development processes, to complement the technical expertise of the subject-matter experts. Methodologists should be consulted during the planning stage, before the guideline development group has been formed, as they can often provide valuable advice on group composition.

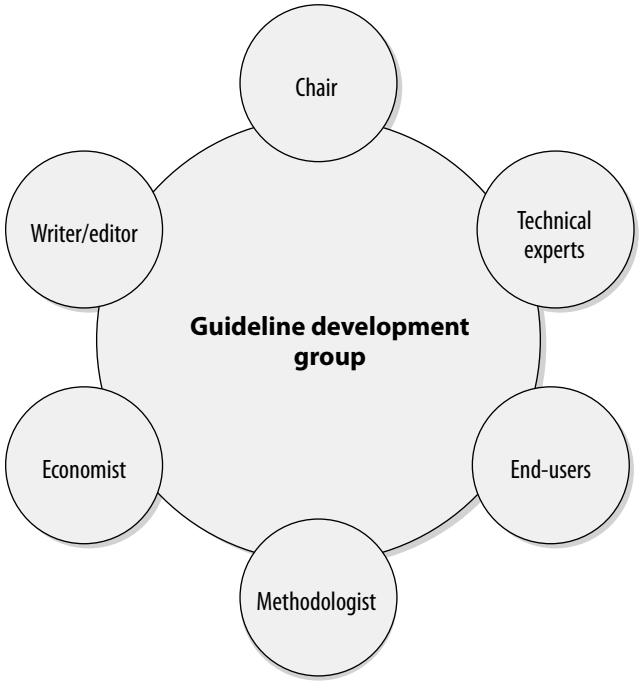
Economist

An economist should be able to advise on the potential economic benefits and drawbacks of the recommendations the group is considering making. The economist should also advise on how to best search the evidence on resource use and costs associated with likely recommendations and interpret that evidence for the group.

Designated writer/editor

It is strongly recommended that one person should be responsible for writing the guideline, while the rest of the group reviews and endorses the document. Later, when external review comments are received, this same person should finalize the document. This will help to ensure coherence, clarity and accuracy.

Figure 3.1 Guideline development group composition



Running an effective guideline development group

The chair

During meetings, the chair should ensure that each member is able to present their views, that assumptions can be debated and that the discussions are open and constructive. The chair should keep the group focused on the agenda and the timescale of the project.

Managing conflicts of interests

Chapter 4 explains how to manage and report conflicts of interest. The basic principles are that declarations are collected and reviewed before appointments are made and any changes need to be reported to the secretariat. At the meeting, each participant should verbally report potential conflicts of interests. Any changes to a member's declaration of interests should be recorded in the minutes of the meeting.

Planning for effective meetings

The guideline development group meeting needs to cover a lot of material in a short time. Ensure that everyone understands his or her role and the expected outputs by providing clear information about the how the meetings will run, including scope, roles, tasks and processes.

Scope of the meeting

- What is expected from meeting participants in terms of advance preparation?
- What needs to be achieved during the meeting?
- What can be done afterwards?
- What follow-up will take place with meeting participants?

Roles and process

- How the guideline will be developed.
- The roles of all guideline development group members, methodologists and observers.
- Declarations and management of conflicts of interest.
- How evidence will be retrieved and assessed.

Achieving meeting objectives

If the purpose of the meeting is to sign off on questions for guideline development.

- Prepare a draft set of questions as formulated by the steering group.
- Circulate the questions ahead of time to all meeting participants.

If the purpose of the meeting is to formulate recommendations.

- Distribute the evidence profiles at least a week prior to the meeting.
- At the meeting, present draft recommendations that have been prepared by the WHO steering group, so that participants can review and revise as necessary.

Further reading

Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health policy and systems research,* 2006, 4:15 (doi: 10.1186|1478-4505-4-15).

Hutchings A et al. A comparison of formal consensus methods used for developing clinical guidelines. *Journal of Health Services Research and Policy,* 2006, 11:218–224.

The guidelines manual 2007. Chapter 9. Making group decisions and reaching consensus. London, National Institute for Health and Clinical Excellence, 2007: 20–29; 56–59 (http://www.nice.org.uk/aboutnice/howwework/developingniceclinicalguidelines/clinicalguidelinedevelopmentmethods/theguidelinesmanual2007/the_guidelines_manual___all_chapters.jsp, accessed 8 June 2012).

SIGN 50: a guideline developer's handbook. Section 5. The guideline development group. Revised edition 2008. Edinburgh, Scottish Intercollegiate Guidelines Network, 2008 (http://www.sign.ac.uk/pdf/sign50.pdf, accessed 8 June 2012).

4. Declaration and management of interests

A conflict of interest occurs when a set of conditions in which professional judgement concerning a primary interest (such as a patient's welfare or the validity of research) tends to be unduly influenced by a secondary interest (such as financial gain). The declaration of a secondary interest does not automatically mean the presence of a conflict of interest that precludes participation in a guideline development group or expert review group.

In WHO, a 'conflict of interest' can be defined as any interest held by an expert that may affect or reasonably be perceived to affect the expert's objectivity and independence in providing advice to WHO. The conflict of interest rules are designed to avoid potentially compromising situations that could undermine or otherwise affect the work done by WHO. Consequently, the scope of the inquiry is any interest that could reasonably be perceived to affect the function that the expert is performing.

This chapter describes the main principles of how interests should be managed during a guideline development process. More detailed information can be obtained from the GRC secretariat and the office of the legal counsel.

Who should declare interests?

According to the rules in the WHO guidelines for declaration of interests, all experts participating in WHO meetings must declare any interest relevant to the meeting before their participation. In the case of guideline development this means that all members of the guideline development group and the expert review panel, as well as any other experts or advisers invited to guideline development meetings, should fill in a declaration of interests (DOI) form.

In addition, anyone invited to participate in a substantive way in the development of a guideline must also complete a DOI form, and must agree to the publication of the declaration in the guideline. Preparation of systematic reviews specifically for the guideline panel and evidence profiles, or contributing to the formulation of recommendations and writing the guideline are considered substantial contributions.

How to manage interests?

The WHO process of managing interests is as follows:

1. Experts and advisers complete the DOI forms before the guideline meeting.

2. The WHO steering group assesses the declared interests prior to the person participating in the meeting to determine whether a conflict exists that may preclude or limit the participation of the person in the guideline group.
3. At each meeting of the guideline group, the declaration of interest forms are summarized and presented to the entire group, so that the group can be aware of any interests that exist among the members. Each member is offered the opportunity to update and/or amend their declaration. The management strategy for each member is also presented to the entire panel.
4. All declared interests are reported in the final guideline document.

What needs to be declared?

WHO collects and reports interests in three categories: financial, academic and public positions.

A financial conflict of interest arises when the expert receives income or support that is related to, or could be affected by, the outcome of the WHO meeting or activity in which they are involved. This includes both personal interests and interests of immediate family members of the expert. Financial interests include:

- personal financial gain (paid work, consulting income or honoraria) or research, proprietary interests and patents;
- grants or fellowships from a commercial entity that has an interest in the topic or the outcomes of the guideline group's work;
- shares or bonds in a related commercial entity;
- employment or consultancies.

Academic conflicts and public positions may be more difficult to recognize but the principle is to include any interest that could be reasonably perceived to affect an individual's objectivity and independence while working with WHO.

Assessing declarations of interest

Declarations of interests are required for potential members of both the external review group and the guideline development group before these groups are finalized and invitations issued. The WHO technical officer needs to collect and review these declarations, in collaboration with the WHO guideline steering group. Further advice can be sought from legal counsel as required.

The aim is to have a chair and a majority of guideline group members with no conflicts of interest.

The first question is whether any declared interests constitute a conflict of interest. What constitutes a potentially significant conflict of interest is a matter of judgement. Some examples of interests that are clearly a conflict and that should preclude participation in developing recommendations include:

- owning shares in a company that manufactures a product or technology that may be recommended for use in the guideline (note that there is a financial threshold specified in the reporting form);
- holding a patent on a product or technology that may be recommended for use in the guideline;
- having a family member who works for a company that manufactures a product or technology that may be recommended for use in the guideline;
- current or past involvement in a major academic programme of work that concerns a product or technology likely to be considered in a recommendation, including conducting trials or systematic reviews that recommend a particular product or technology;
- receiving funding from, being or have recently been employed by, consulting for, or acting as an adviser, paid speaker, or opinion leader for a company or organization with an interest in a specific product related to the guideline – this involves receiving any support for travel, professional training or similar.

If members declare interests that are relevant to the meeting, the WHO technical officer and steering group, assisted by legal counsel, decides whether and to what extent they can participate in the guideline development. These decisions are made on a case-by-case basis, but in general, participants should not participate at all if they declare significant personal financial interests in a single company with a commercial interest in the outcome of the guideline.

Participants can participate in the discussion, but are recused during the development of recommendations if:

- they have links with multiple companies that have commercial interests in the outcome of the guideline;
- they have received research funding from companies that have commercial interests in the outcome of the guideline.

A person with a conflict of interest should not chair a guideline group meeting. Guideline group members who are involved in either primary research or conducting systematic reviews relating to the recommendations in question, should declare these activities as academic interests. All decisions on how to manage declared interests need to be documented prior to the meeting and included in the final guideline. Legal counsel will provide advice on how to handle individual cases.

Management decisions may be:

- the conflict of interest requires no action beyond declaration at the meeting and reporting in the final guideline;
- the conflict of interest is significant but related to only some areas of the guideline development group's work in which case the participant cannot participate when the group considers these areas, and will not have access to the relevant documents;
- the conflict of interest is such as to preclude participation;
- the conflict of interest is such that participation in the discussion is

appropriate, but the member will be recused for development and ratification of recommendations.

Open declaration at the meeting

All declarations of interest made by guideline development group members should be provided to all participants at the meeting as one of the first items on the agenda. If there are any changes to previously declared interests, WHO staff will then need to make a judgement as to whether the revised declarations of interests are of potential importance with respect to likely recommendations and if so, how to manage the declared conflicts. All decisions made should be clearly documented and shared with the entire guideline group (i.e. everyone should know how the conflicts of interest will be managed for the individual members).

Reporting DOIs in the guideline

A summary of how conflicts of interest declarations were collected, any declared conflicts and a brief description of how they were managed must be included in the actual guideline document. If no conflict was declared, this information needs to be provided as well. The GRC will not clear a guideline document that does not contain this information.

Declared conflicts of interest should be reported in the guideline according to the following examples. The wording of this part of the guideline document must be approved by legal counsel before final review by the GRC.

> *Dr N.C. reported being an investigator on trials for GlaxoSmithKline, Quintiles, Uriach and Biomarin but not for any products or products related to those being considered at the meeting, and also holding shares in Biota. She therefore was excluded from discussion of the late item on antivirals.*

> *Dr M.R. reported having been a consultant for Roche on drug research and development. He is currently a member of a data safety and monitoring board for Roche; receives royalties through the US National Institutes of Health from the use of gossypol for cancer; and is a consultant to several start-up companies, none of which have products on the market. As there were no products related to any of these items on the agenda, no action was required.*

> *Dr A.F. reported having a family member who is an employee of Merck, Sharpe and Dohme, Brazil. He therefore excluded himself from review or discussion of the product applications from Merck on this agenda.*

What to do when there are too many conflicts?

WHO has traditionally relied on experts to develop its recommendations, on the assumption that experts' advice is objective and free from bias. Unfortunately, research has shown that there is an association between the

financial interest of guideline group members and decisions that support those interests. If you need to consider input from experts who have conflicts, the GRC secretariat can advise on how to use an evidence jury. To do this, you need at least two thirds of your group to have no conflicts of interest. They become your jury. The experts present their evidence and views to the jury. The jury then develops the recommendations in the absence of the experts. This requires dividing the meeting into two parts: a first session with all members present for the presentation of the evidence and a second closed session, with only the jury, to develop and ratify recommendations.

Further reading

Bekelman JE, Li Y, Gross C. Scope and impact of financial conflicts of interest in biomedical research: A systematic review. *JAMA,* 2003, 289:454-65 (http://jama.ama-assn.org/cgi/reprint/289/4/454, accessed 8 June 2012).

Boyd EA, Bero LA. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Research Policy and Systems, 2006, 4:16* (http://www.health-policy-systems.com/content/4/1/16, accessed 8 June 2012).

Boyd EA, Lipton S, Bero LA. Implementation of financial disclosure policies to manage conflicts of interest. *Health Affairs, 2004, 23:206–214* (http://content.healthaffairs.org/content/23/2/206.full.html, accessed 8 June 2012).

Choudhry NK, Stelfox HT, Detsky AS. Relationships between authors of clinical practice guidelines and the pharmaceutical industry. *JAMA,* 2002, 287:612–617 (http://jama.ama-assn.org/cgi/reprint/287/5/612, accessed 8 June 2012).

Lexchin J et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ,* 2003, 326:1167–1170 (http://www.bmj.com/cgi/reprint/326/7400/1167, accessed 8 June 2012).

Lo B, Field MJ. *Conflict of interest in medical research, education and practice. National Academy of Sciences, 2009.*

Norris SL et al. Conflict of interest in clinical practice guideline development: a systematic review. *PLoS ONE, 2011, 6(10):e25153* (http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0025153, accessed 8 June 2012).

Sniderman AD, Furberg CD. Why guideline-making requires reform. *JAMA,* 2009, 301:429–431 (http://jama.ama-assn.org/cgi/reprint/301/4/429, accessed 8 June 2012).

Steinbrook R. Controlling conflict of interest: proposals from the Institute of Medicine. *NEJM,* 2009, 360:2160–2163 (http://content.nejm.org/cgi/content/full/360/21/2160, accessed 8 June 2012).

Williams MJ, Kevat DAS, Loff B. Conflict of interest guidelines for clinical guidelines. MJA, 2011, 195:442–445 (https://www.mja.com.au/journal/2011/195/8/conflict-interest-guidelines-clinical-guidelines, accessed 8 June 2012).

5. **Formulating questions and choosing outcomes**

The choice of questions that need to be addressed by the guideline strongly influences the final recommendations – so getting this stage right is crucial. The questions should be used to systematically search the evidence for answers to controversial areas the guideline is trying to address. When the scope of the document has been developed and potential recommendations identified, the questions that need to be asked should become clear. Because these questions drive the evidence search and form the basis of your recommendations, they should be clear and well defined.

When developing these questions it helps to look at the type of information needed. Usually the information leads to two types of questions: background and foreground questions.

Background questions. These relate to the subject of the guideline and provide important background information on the issues under consideration. However, they do not provide direct evidence informing recommendations. They include questions on definitions, the prevalence of the problem or disease and mechanisms underlying possible interventions (i.e. how the intervention might work).

> *e.g. How is human papilloma virus infection associated with cervical cancer?*

Questions about mechanisms may be answered with a wide range of information ranging from basic scientific data to theoretical frameworks such as behavioural change theories underpinning public health interventions. A full review is not usually required for background information referred to in the guideline, such as a section on the epidemiology or pathology of a disease.

Foreground questions. These address the effectiveness of an intervention that the guideline development group is considering recommending. They usually include questions about the efficacy of the intervention but can also provide information on negative consequences, social acceptability, or the cost-effectiveness of an intervention under consideration, helping provide an evidence base for values, preferences and economic implications that should be considered when making a recommendation.

> *e.g. What impact does human papilloma virus vaccination have on the incidence of cervical cancer?*

The foreground questions are the most important ones for a guideline. They are used to inform the recommendations and they will require a systematic

review and quality assessment of the evidence using the GRADE (grading of recommendations, assessment, development and evaluation) approach. Because the answers to foreground questions will form the evidence base upon which the recommendations will be made, these questions should be framed in a way that enables a systematic search of the literature. The PICO format is an effective way to do this.

Formulating PICO questions

PICO refers to four elements that should be in a question governing a systematic search of the evidence: Population, Intervention, Comparator and Outcomes (see Box 5.1 PICO question components).

Population

Who is targeted by the action being recommended?

- How can they be best described? What are the relevant demographic factors? Please consider age groups, sex, ethnicity, social identities, behavioural characteristics, etc.
- What is the setting? For example, hospitals, communities, schools.
- Are there any subgroups that might need to be considered?
- Are there groups or subgroups that should be excluded?

Intervention

What action is being considered?

- Which treatment, procedure, diagnostic test, prognostic factor, risk factor, lifestyle change, social activity, screening test, preventive measure, or approach is being evaluated?
- Are there variations you might want to consider? (dosage, frequency, delivery or administration, personnel and delivery channels, timing and duration, etc.).
- Where interventions are complex, consider which components are of most interest to your guideline group and how they might best be described.

Comparator

What are the alternative choices of action?

- This may be what is currently being done – including no specific treatment – or another measure the guideline panel may be considering in comparison.
- Comparisons may be made to placebo, no intervention, standard care, current standard diagnostic, variations of the intervention or a different one.

Outcomes

What is the purpose of the recommendation?

- What will it achieve?
- What harms could it lead to?

- Possible outcomes both positive and negative need to be selected carefully with input from experts, implementers and those most affected by the recommendations.

Box 5.1 PICO question components	
Population	(Who is targeted by the action being recommended?)
In girls aged 9–13 years	
Intervention	(What action is being considered?)
does HPV vaccine	
Comparator	(What are the alternative choices of action?)
compared with no vaccination	
Outcome	(What is the purpose of the recommendation?)
reduce the incidence of cervical carcinoma?	

Examples of PICO questions

In a rural population in a low-income country (Population), does paying higher salaries to health workers (Intervention), compared with paying standard salaries (Comparator), increase the number of health workers in rural areas (Outcome)?

This format can also be used for questions on diagnosis, prevention, aetiology and resource use. For example:

- In babies born to HIV-positive women (P), does screening with a new rapid diagnostic test (I), compared with standard diagnostic methods (C) accurately detect disease (O)?
- In an urban population (P), is exposure to an environmental chemical (I), compared to no exposure (C) associated with an increased risk of cancer (O)?
- Is intervention A (I) as cost-effective in preventing mortality (O) compared to intervention B (C)?
- In a national population (P), how does one intervention (I), compared to another (C), perform in terms of costs per quality-adjusted life year gained (O)?

PICO questions may be broad or narrow in scope. While a broad question will lead to a comprehensive summary of a larger body of evidence and more generalizable findings, it may also require more resources to answer. A broad question may also yield greater heterogeneous evidence, making interpretation difficult. A narrow question may be easier to manage, but the evidence might

be sparse and findings less generalizable. Depending on the scope of the guideline and the availability of information the steering group may decide to split a broad question into a number of narrow questions.

Example of a broad PICO question:

> *Do financial incentives (I) compared to no financial incentives (C) improve the retention (O) of health workers (P) in rural areas?*

Example of a narrow PICO question:

> *Does a housing allowance (I) compared to no housing allowance (C) improve the retention (O) of health workers (P) in rural areas?*

Choosing and rating outcomes

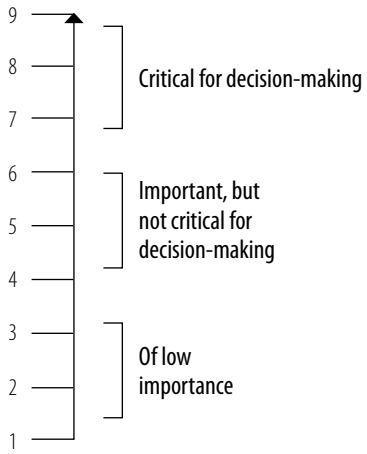
The purpose of any recommendation is to achieve a desirable effect or outcome. Choosing the most important outcome is therefore critical to producing a useful guideline but different groups value outcomes differently. For this reason it is essential that the external review group (which should contain end-users, implementers and policymakers, as well as technical experts) be asked to identify the key outcomes that need to be considered when the recommendations are made.

Rating outcomes

The WHO steering group should make an initial list of relevant outcomes, including desirable and undesirable effects and ask both the guideline development group and the external review group to identify any other outcomes that have not been listed.

Once a workable list of outcomes has been collected, an effective means of prioritizing these is to ask group members (this can be both the guideline development group and the external review group) to rate them. Group members are asked to give outcomes a score from 1–9, where 7–9 rates the outcome as critical for a decision, 4–6 indicates that it is important and 1–3 indicates that it is not important (Figure 5.1). The average score for each outcome can then be used to determine the relative importance of each outcome, although it is helpful to provide the range of results as well.

Figure 5.1 Scale for rating outcomes



If necessary, the final rating of outcomes can be reviewed and confirmed at a later stage when the guideline group meets.

Finalizing the questions

Questions should be finalized by the WHO steering group after input from all the relevant experts, including end-users (e.g. programme managers, partner agencies, and consumer and patient groups), who should, ideally, be members of either the guidelines development group or the external review group. Because the number of questions that need systematic reviews will be a major determinant of the time and resources needed to complete the guideline, the steering group should aim to reduce the number of questions to those dealing with the most controversial and least understood areas.

Step 1: Generate initial list of questions

The WHO steering group develops an initial list of questions based on the scope of the guideline. It helps to divide these into background and foreground questions.

Step 2: Draft PICO questions

The WHO steering group, with input from the guideline development group, applies the PICO framework to the foreground questions.

Step 3: List relevant outcomes

The WHO steering group should list relevant outcomes, including both desirable and undesirable effects. The guideline development group reviews this and may add additional important outcomes.

Step 4: Comment and revise

The list of questions and outcomes of interest should be sent to the external review group for review and revision and inclusion of any omissions.

Step 5: Rate outcomes

Selecting relevant outcomes is critical to producing an effective guideline. Outcomes should be rated in order of importance by a wide group comprising the guideline development group, the external review group and relevant stakeholders. To make this workable, a formal rating process, such as that described above, should be used.

Step 6: Prioritize questions

Prioritize questions and determine which questions need systematic reviews. This is done by the WHO steering group using input from the guideline development group and the external review group.

Further reading

Question formulation for clinical practice guidelines. In: *Handbook for the preparation of explicit evidence-based clinical practice guidelines.* Wellington, New Zealand Guidelines Group, 2001:15–21 (http://www.nzgg.org.nz/download/files/nzgg_guideline_handbook.pdf, accessed 8 June 2012).

Kaifeng AK. Principles of evidence based medicine. *Archives of Disease in Childhood,* 2005, 90:837–840.

O'Connor D, Green S, Higgins JPT, eds. Defining the review question and developing criteria for including studies. In: Higgins JPT, Green S, eds. *Cochrane handbook of systematic reviews of intervention. Version 5.0.1* (updated September 2008). The Cochrane Collaboration, 2008 (http://www.cochrane-handbook.org, accessed 8 June 2012).

Murphy MK et al. Consensus development methods and their use in clinical guideline development. *Health Technology Assessment,* 1998, 2(3) (http://www.ncchta.org/fullmono/mon203.pdf, accessed 8 June 2012).

6. **Evidence retrieval and synthesis**

Systematic reviews

WHO recommendations need to be based on the best evidence available. Ensuring that all the relevant evidence has been sought and presented is not always easy. An effective approach is to perform a systematic review using specific questions about the intervention(s) likely to be recommended in the guideline. Systematic reviews, if conducted properly, reduce the risk of selective citation and improve the reliability and accuracy of decisions.

The key characteristics of a systematic review are:

- a specific and clearly focused question (in PICO format)
- an explicit, reproducible method including pre-defined eligibility criteria
- a comprehensive, exhaustive and systematic search for primary studies
- a selection of studies using clear and reproducible eligibility criteria
- critical appraisal of included studies for quality
- systematic presentation and synthesis of the characteristics and findings of the included studies.

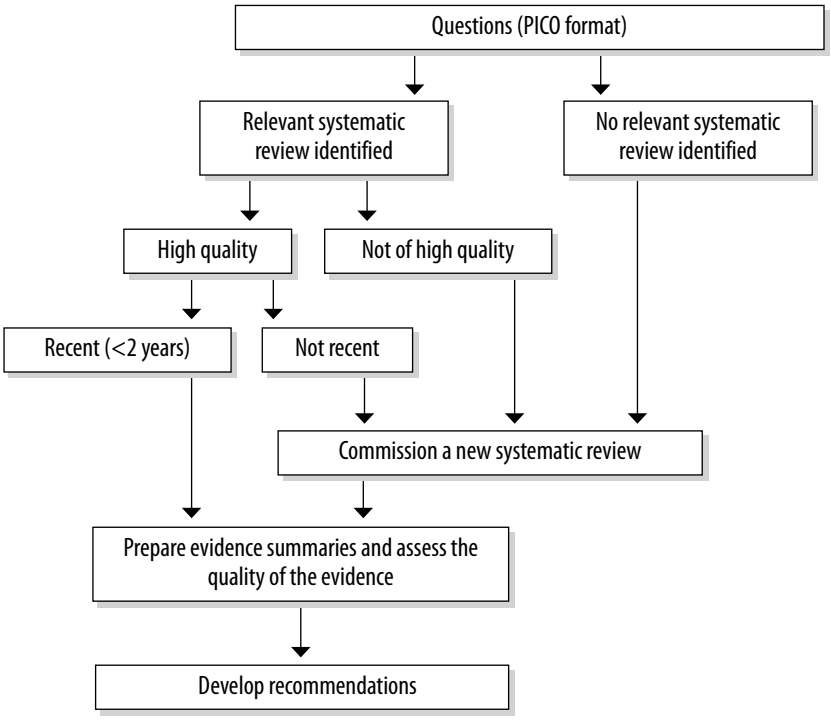
Is a new review needed?

While systematic reviews should be used to assemble all the evidence, it is not always necessary to commission new ones (see Figure 6.1 Evidence retrieval decision diagram). If current, relevant and high quality systematic reviews exist, these should be used. Updates, if needed, are less expensive and time-consuming than new reviews. The search for existing systematic reviews can be done by the WHO steering group, or can be sub-contracted to a group preparing the evidence summaries.

The Cochrane Collaboration – a large global network that produces systematic reviews – is a nongovernmental organization in official relations with WHO. The Cochrane Collaboration may be able to identify existing or forthcoming systematic reviews on the guideline topic. The GRC secretariat can refer you to WHO's Cochrane Collaboration focal point who will liaise with relevant Cochrane groups. For reviews on complex interventions such as behavioural change, these groups include:

- The Cochrane Qualitative Research Methods Group
  http://cqrmg.cochrane.org/
- The Cochrane Consumers and Communication Review Group
  http://www.latrobe.edu.au/chcp/cochrane/
- The Cochrane Effective Practice and Organisation of Care Group
  http://epoc.cochrane.org/

Figure 6.1 Evidence retrieval decision diagram



The Campbell Collaboration http://www.campbellcollaboration.org/ has a database of reviews of effectiveness of social and educational policies and practices.

The search for existing systematic reviews should be done in a systematic way and documented in a protocol. The protocol should describe the databases used and the search strategy applied to each. Please have your protocol reviewed by a WHO librarian or other expert in information retrieval to ensure that you have included all the necessary databases and search terms.

Your protocol may start by reviewing the reference lists of existing guidelines on the topic, before moving on to the major biomedical databases, such as PubMed. PubMed's clinical queries or special queries options permit specific searches to identify systematic reviews of different types of studies. You can also now use the publication limits in PubMed to limit your search to all reviews or only to meta-analyses. Systematic reviews of policy interventions may be difficult to find, and other search strategies will be needed. You will need to include a brief description of how and when this search was actually done, as well as the results, in the final guideline.

Evaluating the quality of systematic reviews

Once the systematic reviews are retrieved, the WHO steering group will need to assess the relevance, timeliness and quality of these reviews before making a decision as to whether new reviews will need to be commissioned. To assess relevance, compare the PICO question of the systematic review to the PICO questions for the guideline that were developed during the scoping. If a relevant systematic review is less than two years old, and is judged to be of sufficient quality, it can be used. If it is older than two years it needs to be updated to include more recent evidence. If it is a Cochrane Review, the relevant review group could be contacted to determine if an update is planned. If there are several systematic reviews, use the most recent one of high quality.

The following checklists may be used in assessing quality:

- Systematic review critical appraisal sheet
  http://www.cebm.net/index.aspx?o=1157
- Assessment of multiple systematic reviews (AMSTAR)

Please note that a checklist merely provides a list of items that should be appraised and that the final decision on whether a systematic review is of high quality is a judgement based on a combination of all items. The following five aspects are important quality indicators of a systematic review.

1. PICO question and eligibility criteria

The review question should specify the types of population, interventions, comparisons and outcomes of interest and the types of study that will be included in the review. Together, these form the basis of the pre-specified eligibility criteria.

Is the question being addressed clearly and explicitly stated with reference to participants, interventions, comparisons and outcomes? Does the review identify which study designs are included? Are inclusion and exclusion criteria clearly defined?

2. Information sources and search for original articles

A comprehensive search includes major bibliographic databases (e.g. Medline, EMBASE), but also a search of reference lists from relevant studies, conference abstracts and other grey literature and contact with experts to identify additional studies. Searches limited to the English language only are likely to miss relevant papers. Both text words and subject headings (e.g. Medical Subject Headings – MeSH terms) should be used.

Is the search strategy, including search terms used, clearly described? Is it comprehensive? Was a thorough search of appropriate databases done and are other sources of information used? Are there language limitations?

3. Study selection and data extraction

Assessment of eligibility of studies, and extraction of data from study reports, should have been done by at least two people, independently.

Were the title and abstracts scanned for eligibility by two independent researchers? Is the total number of titles and abstracts reviewed indicated? Is the number of excluded articles reported and are reasons for exclusion given? Were data extracted in duplicate?

4. Study quality and risk of bias

The risk of bias in included studies depend on the type of studies included. Examples of criteria for assessing risk of bias include randomization, blinding, completeness of follow-up.

Does the review describe how the risk of bias for each study was assessed? Were predetermined criteria used?

5. Synthesis and reporting of results

The review should describe how data were handled and the results of studies combined. A test for heterogeneity should be done if meta-analyses are presented. An assessment of risk of bias across studies – due to publication or reporting bias – should be reported.

If a meta-analysis was done, was a test for heterogeneity reported? If there was significant heterogeneity between studies, were possible reasons explored? Are the conclusions supported by the data?

If the identified systematic reviews are all of low quality, you will need to commission a new one.

How to commission a systematic review

A new systematic review is needed if relevant existing systematic reviews could not be identified, or when existing systematic reviews are of low quality. If you find a high-quality review that is more than two years old, you may be able to commission an update to include more recent evidence.

Systematic reviews take time, expertise and resources to do well, and are best commissioned from external suppliers by the WHO steering group. Members of the Cochrane Collaboration may be able to do or update the systematic reviews required. Regardless of supplier, you should estimate a minimum of US$ 20 000 per review, although this amount will vary depending on the complexity of the review needed.

In commissioning the reviews, the WHO Steering Group will need to:

- disseminate a request for proposals to established suppliers of systematic reviews;
- provide clear terms of reference to the suppliers selected;
- review and approve the suppliers' protocol before the evidence search is started; and
- request regular updates from the suppliers on the progress of the review.

Search strategies

Regardless of the supplier chosen, systematic reviews used to inform WHO recommendations must be developed according to the standards outlined by the Cochrane Collaboration in the Cochrane handbook. The handbook has a specific chapter dealing with reviews in public health and health promotion.

For a WHO guideline, it is important to search for studies from low- and middle-income countries in all regions as well as from more standard literature sources. Some journals are not well represented in PubMed and commercial databases such as EMBASE and CAB Abstracts. Regional databases grouped under the general heading of the Global Health Index contain unique citations and full-text articles. WHO's regional offices have supported the development of these indices to highlight the health research of developing countries. Most journals indexed by regional databases are not indexed in PubMed. Please ensure that the supplier of your systematic review consults a WHO librarian to ensure that the search strategy includes these databases. Your supplier is also likely to need a way of searching for and assessing evidence in WHO's six official languages and should specify how the relevant grey literature will be identified.

Including qualitative research

Systematic reviews on the effect of an intervention can be complemented by qualitative evidence. Qualitative evidence can help explain, interpret and apply the quantitative results of a systematic review. A synthesis of qualitative research can be done as part of the scoping of the guideline and can also help to define and refine the questions. The relevant chapter in the Cochrane handbook explains in detail how to include qualitative research and provides additional reading and relevant web sites.

A mixed methods synthesis brings together a meta-analysis of quantitative data and a qualitative analysis. In this case, systematic reviewers will search for qualitative research imbedded in original studies evaluating health interventions. This is most appropriate if the intent is to identify why the intervention might or might not work.

A qualitative evidence synthesis uses specific searches for – and synthesis of – evidence from qualitative studies. This type of synthesis can address questions on effectiveness of the intervention, such as contextual barriers and facilitators, or values and preferences of those receiving the intervention.

Evidence synthesis

The results of the systematic reviews will be presented to the guideline development group in a meeting in which the total body of evidence is assessed and recommendations are developed. The two most common ways of presenting the evidence are briefly described here, but for more details, please consult the Cochrane handbook.

Meta-analysis

If the data extracted from the systematic review meet certain requirements (the most important one being a high level of homogeneity of effect measures across studies), then the data can be combined using meta-analyses. A meta-analysis is the use of statistical methods to summarize the results of independent studies. By combining information from all relevant studies, meta-analyses can provide more precise estimates of the effects of health care than those derived from the individual studies included within a review. The results of a meta-analysis are usually displayed in a figure called a forest plot.

Narrative synthesis

If a meta-analysis is not feasible – due to heterogeneity – or not sensible because different types of interventions are covered, the evidence can be presented in a narrative synthesis. The method used to produce this synthesis needs to be specified before starting and followed rigorously to avoid introducing bias. The results of each individual study can be presented in a table. Irrespective of the way in which the results are presented, it is important that the same elements of information are included in the same order. Grouping the studies can help if a large number have been included in the review.

Further reading

Grimshaw J et al. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Quality and Safety in Health Care,* 2003, 12:298–303.

Harden A et al. Applying systematic review methods to studies of people's views: an example from public health research. *Journal of Epidemiology and Community Health,* 2004, 58:794–800.

Jackson N, Waters E for the Guidelines for Systematic Reviews in Health Promotion and Public Health Taskforce. Criteria for the systematic review of health promotion and public health interventions. *Health Promotion International,* 2005, 20:367–374.

Moher D et al. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine,* 2009, 6(7): e1000097

Liberati A et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine,* 2009, 6(7): e1000100

Ogilvie D et al. Systematic reviews of health effects of social interventions: 1. Finding the evidence: how far should you go? *Journal of Epidemiology and Community Health,* 2005, 59:804–808.

Ogilvie D et al. Systematic reviews of health effects of social interventions: 2. Best available evidence: how low should you go? *Journal of Epidemiology and Community Health,* 2005, 59:886–892.

Shepperd S et al. Can we systematically review studies that evaluate complex interventions? *PLoS Medicine,* 2009, 6(8): e1000086

Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011 (http://www.cochrane.org/training/cochrane-handbook, accessed 8 June 2012).

7. **Evidence assessment**

GRADE

The evidence that has been retrieved and synthesized in a systematic review needs to be assessed for quality. Quality of evidence is defined as the "extent to which one can be confident that an estimate of the effect or association is correct."

WHO uses the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) approach to assess the quality of a body of evidence, develop and report recommendations. GRADE methods are used by WHO because these represent internationally agreed standards for making transparent recommendations. Detailed information on GRADE is available through the GRC secretariat and on the following sites:

- GRADE working group: http://www.gradeworkingroup.org
- GRADE online training modules: http://cebgrade.mcmaster.ca/
- GRADE profile software: http://ims.cochrane.org/revman/gradepro

The GRC secretariat offers GRADE workshops throughout the year. Please contact the secretariat, or check the GRC intranet web page for training dates.

Assessing the evidence and developing evidence summaries is a specialized task that is best done by a methodological expert. WHO guideline development groups usually include, or work with, a methodologist. The GRC secretariat can recommend GRADE experts and methodologists from the Cochrane Collaboration. The methodologist is contracted to do the evidence assessment, prepare the evidence summaries, and present them at the guideline development group meeting. The guideline development group uses these summaries as the basis for their discussions and recommendations.

Guideline development group members that have no previous experience of working with GRADE should be briefed on the process by WHO prior to the guideline meeting. This can be done with a combination of the online training modules, publications and presentations listed above. Additionally, many guideline groups find it useful to start their meetings with an introduction to GRADE presented by the methodologist or the GRC secretariat.

Evidence profiles

The methodologist will present GRADE evidence profiles for each PICO question for which a systematic review was done (Table 7.1). Evidence profiles are tables that contain the quality assessment and the summary of findings.

Table 7.1 Components of a GRADE evidence profile

Question (PICO format)										
Quality assessment							**Summary of findings**			
							Study event rates (%)		Effects	
Participants (studies)	Risk of bias	Inconsistency	Indirectness	Imprecision	Reporting bias	Overall quality of evidence	Intervention	Control	Relative effect (95% CI)	Absolute effect
Outcome A (evidence from randomized trials)										
Outcome A (evidence from observational studies)										
Outcome B (evidence from randomized trials)										

The quality assessment

GRADE categorizes the quality of evidence as high, moderate, low or very low (Table 7.2). These quality ratings apply to the body of evidence assessed for the PICO question, not to individual studies.

Table 7.2 Significance of the four levels of evidence

Quality	Definition	Implications
High	The guideline development group is very confident that the true effect lies close to that of the estimate of the effect	Further research is very unlikely to change confidence in the estimate of effect
Moderate	The guideline development group is moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate
Low	Confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the true effect	Further research is very likely to have an important impact on confidence in the estimate of effect and is unlikely to change the estimate
Very low	The group has very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of the effect	Any estimate of effect is very uncertain

The starting point for rating the quality of evidence is always the study design, broadly classified into two types:

- randomized controlled trials (RCTs);
- observational studies, including interrupted time-series (or quasi-experimental design), cohort studies and case-control studies, and other types of design such as case series and case reports.

The design is the baseline for rating the quality of evidence. If you have studies of more than one design reporting the outcome, you should have a separate row in your table for each type. Although randomized controlled trials are the preferred source of evidence for measures of effectiveness, in many instances guideline developers rely on information from observational studies. This situation arises when RCTs are not ethical, appropriate or feasible, when few RCTs are available and observational studies are needed to estimate the effect size, or when information on the feasibility of the intervention in different settings is needed.

Evidence based on randomized controlled trials is given a high-quality rating and evidence from observational studies is given a low-quality rating. These initial ratings can be adjusted by the following factors.

Five factors that can lower the quality of evidence

1. Study limitations

For randomized controlled trials, the main criteria for assessing limitations are:

- whether concealment of allocation to treatment group is adequate;
- whether participants and investigators were blinded, especially if the outcomes are measured subjectively and subject to bias;
- whether an intention-to-treat analysis is reported;
- whether all withdrawals and patients lost to follow-up are accounted for;
- whether the trial was stopped early for benefit.

For studies of diagnostic accuracy, additional limitations include:

- whether patients were consecutively recruited and not classified by disease state;
- whether both the new test and the reference standard were done in all patients;
- whether evaluators were blinded to the results of the alternative test and reference standard.

For observational studies, the main criteria depend on the design (i.e. case-control or cohort studies). For both designs, the methods used to select the population in the study and the comparability of the two groups are important. For case-control studies the method of determining exposure to the factor of interest also needs to be evaluated. For cohort studies the method of measuring outcomes should be evaluated.

The evidence summaries should categorize the limitations as follows.

a. ***"No limitations"*** generally means that the majority of studies meet all the minimum quality criteria for the design. The implication of this is that the rating of quality of evidence remains the same as the initial assessment.

b. ***"Minor limitations"*** applies when minor flaws are found when analyzing how the available studies were designed and performed. If you decide there are minor limitations, these should be noted in a footnote in the evidence profile but they would not usually downgrade the quality.

c. ***"Serious limitations"*** means that one of the minimum criteria for quality is not met by the majority of studies in the review. This results in a -1 score for the overall quality rating (e.g. "high" becomes "moderate").

d. ***"Very serious limitations"*** means that at least two of the criteria proposed as potential study limitations are present in the majority of studies in the review. This results in a -2 score for quality.

2. Consistency

Consistency relates to whether the results are similar across studies. Differences in the direction, the size, and the significance of the differences in effect, guide the decision about whether important inconsistency exists. If all the results of the studies for one outcome are in the same direction with overlapping confidence intervals, there is unlikely to be significant consistency. To evaluate the degree of consistency of the results of the available studies, the direction and size of the effect for each outcome should be evaluated. If a formal meta-analysis was conducted, the result of the test for heterogeneity can be used to help assess consistency. Variability or inconsistency in results may arise from differences in the populations in the studies, in the interventions or in outcomes.

If there is inconsistency in the results, such as the largest trial showing results that contradict smaller trials, then a -1 score should be applied. If the results are very heterogeneous, "very serious" should be chosen, which will downgrade the evidence for this outcome by two levels. If only one study is present, consistency is not applicable as a criterion.

3. Directness

Directness, generalizability, external validity of study results and applicability are all synonyms. There are two types of indirectness.

- Indirect comparison occurs when a comparison of intervention A versus B is not available, but A was compared with C and B was also compared with C. Such trials allow indirect comparisons of the magnitude of effect of A versus B. Such evidence is of lower quality than direct comparisons of A and B would provide.
- Indirect population, intervention, comparator or outcome arise when the question being addressed by the guideline development group or by the authors of a systematic review is different from the available evidence regarding the population, intervention, comparator or outcome.

4. Imprecision

Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. In this case the quality of the evidence is lower than it otherwise would be because of uncertainty in the results.

When event rates are very low, 95 per cent confidence intervals around relative effects can be very wide, but 95 per cent confidence intervals around absolute effects may be narrow. In the latter case, the quality should not be downgraded for imprecision.

5. Reporting bias

Reporting bias (also called publication bias) is a systematic underestimate or overestimate of the underlying beneficial or harmful effect due to the selective publication of studies or selective reporting of outcomes. Reporting bias arises when investigators fail to report studies they have done (typically those that show no effect) or neglect to report outcomes that they have measured (typically those for which they observed no effect).

Despite methods to detect the possibility of publication bias, the authors of systematic reviews and guideline panels must often make assumptions about the extent of this bias. Reporting bias should be considered when published evidence is limited to a small number of trials, all of which were funded by a for-profit organization. In such a situation, consider the extent to which evidence about the magnitude of the effect is uncertain due to selective publication of studies or reporting of outcomes. If this is likely, downgrade the quality rating by one or even two levels.

The criteria that are used for downgrading the quality of evidence and the reason for the assessment should be explained in a footnote to the table.

Three factors that can increase the quality of evidence

1. Dose-response gradient

The presence of a dose-response gradient may increase confidence in the findings of observational studies and thereby increase the quality of evidence. However, this applies only to studies that are not downgraded for any reason. To rate the presence of a dose-response gradient:

- if there is no evidence of a dose-response gradient, there is no change;
- if there is evidence of a dose-response gradient, upgrade the evidence for this outcome by 1 level.

2. Direction of plausible bias

On occasion, all plausible biases from observational studies may tend to underestimate the true treatment effect. For instance, if only sicker patients receive an experimental intervention or exposure, yet they still improve, it is likely that the actual intervention or exposure effect is larger than the data suggest. Only studies with no threats to validity (not downgraded for any reason) can be upgraded. To rate the effect of all plausible residual confounding:

- if there is no evidence that the influence of all plausible residual confounding would reduce the observed effect, there is no change;
- if there is evidence that the influence of all plausible residual confounding would reduce the observed effect, upgrade the evidence for this outcome by 1 level.

3. Magnitude of the effect

When methodologically strong observational studies yield large or very large and consistent estimates of the magnitude of a treatment or exposure effect, we may have confidence in the results. In such situations, the weak study design is

unlikely to explain all the apparent benefit or harm, even though observational studies are likely to provide an overestimate of the true effect. The larger the magnitude of effect, the stronger the evidence becomes. Only studies with no threats to validity (not downgraded for any reason) can be upgraded.

The final category for the quality of evidence is determined by adding the additional ratings to the original baseline category (Table 7.3).

Table 7.3 How to upgrade or downgrade the quality of evidence

Downgrade in presence of	Upgrade in presence of
Study limitations -1 Serious limitations -2 Very serious limitations	Dose-response gradient +1 Evidence of a dose-response gradient
Consistency -1 Important inconsistency	Direction of plausible bias +1 All plausible confounders would have reduced the effect
Directness -1 Some uncertainty -2 Major uncertainty	Magnitude of the effect +1 Strong, no plausible confounders, consistent and direct evidence
Precision -1 Imprecise data	+2 Very strong, no major threats to validity and direct evidence
Reporting bias -1 High probability of reporting bias	

The summary of findings

In the summary of findings, the following information is presented.

- The number of patients (or, in the case of a policy intervention, units) studied: for the intervention and the control group the total number of patients and the number of patients who experienced the outcome are reported.
- The effect size: both absolute and relative effects are reported – these are obtained from the systematic reviews (e.g. from the meta-analysis).
- The quality of the evidence: this is the result of the quality assessment and is reported as one of the four categories of evidence (high, moderate, low or very low).
- The importance of the outcome: this is the result of the rating of the importance of the outcome that was done during the scoping process.

Statistical methods for combining results of observational studies are more complex than the methods used for combining randomized controlled trials. If it is possible to pool data, they can be reported in a standard GRADE evidence profile. If the results cannot be combined, the results can be presented in a narrative synthesis as described in Chapter 6.

Further reading

Barbui C et al. Challenges in developing evidence-based recommendations using the GRADE approach: the case of mental, neurological, and substance use disorders. *PLoS Medicine,* 2010, 7(8): e1000322

Bruce N et al. Enhancement to GRADE (termed GRADE+) for environmental health interventions. A proposal with special consideration of application to the development of WHO indoor air quality guidelines: household fuel combustion [manuscript in preparation].

Guyatt GH et al. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology,* 2010, 64:380–382.

Guyatt GH et al. Rating quality of evidence and strength of recommendations GRADE: an emerging consensus on rating quality of evidence and strength of recommendations**.** *BMJ,* 2008, 336:924–926.

Schünemann HJ et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ,* 2008, 336:1106–1110.

8. Developing recommendations

Factors that condition recommendations

Once the evidence is retrieved, synthesized and assessed, this evidence must now be used to develop recommendations. In the GRADE approach the quality of the evidence and the balance between benefits and harms determine whether the recommendation is for or against the recommendation. The strength of the recommendation is determined by a consideration of values and preferences, and resource implications.

The quality of the evidence – the degree of confidence in estimates of effects – is the first factor considered. The higher the quality of evidence, the more likely a strong recommendation is warranted. If the quality of the evidence is lower, it will create greater uncertainty about the size of the relative effects. This uncertainty can concern both beneficial and harmful effects and therefore makes a conditional recommendation more likely.

When considering the balance between benefits and harm, one should look at the magnitude of the effect as well as the importance of the outcomes. If the benefits clearly outweigh the harms, a strong recommendation is more likely. A conditional recommendation is more likely if there is uncertainty about the balance of benefits versus harms or when there are only marginal net benefits, that is, when the anticipated net benefits are small.

Values and preferences are based either on collected qualitative evidence or by the experience and opinion of various stakeholders present in the guideline development group. The greater the variability in values and preferences, or uncertainty in values and preferences the more likely a conditional recommendation is warranted. There might be uncertainty about the relative importance of the benefits and harms to those affected, or differences in how these benefits and harms are perceived.

The guideline group's consideration of resource implications can be informed either by a full formal economic evaluation or by estimates collected during the evidence retrieval. The more resources the intervention consumes, the less likely a strong recommendation is warranted. Uncertainty about resource use – whether the net benefits are worth the costs, lack of information about the cost, or questions about whether the resource expenditure is justified by the anticipated benefit – make a conditional recommendation more likely.

Decision tables

Decision tables can be used to record the guideline's group judgements about these factors and how they contributed to the development of the recommendation, as shown in Table 8.1.

Table 8.1 Decision table to support the development of recommendations

Recommendation:		
Population:		
Intervention:		
Factor	Decision	Explanation
Quality of the evidence (The higher the quality of the evidence, the more likely a strong recommendation is warranted.)	High Moderate Low Very low	
Balance of benefits versus harms and burdens (The larger the difference between the benefits and harms, the more likely a strong recommendation is warranted. The smaller the net benefit and the lower the certainty for that benefit, the more likely a conditional recommendation is warranted.)	Benefits clearly outweigh harms Benefits and harms are balanced Potential harms clearly outweigh potential benefits	
Values and preferences (The greater the variability or uncertainty in values and preferences, the more likely a conditional recommendation is warranted.)	No major variability Major variability	
Resource use (The higher the costs of an intervention, that is, the more resources consumed, the more likely a conditional recommendation is warranted.)	Less resource-intensive More resource-intensive	
Overall strength of the recommendation: (strong or conditional)		
Research gaps:		

The strength of the recommendation

The strength of a recommendation communicates the importance of adherence to the recommendation.

Strong recommendations

With strong recommendations, the guideline communicates the message that the desirable effects of adherence to the recommendation outweigh the undesirable effects. This means that in most situations the recommendation can be adopted as policy.

Conditional recommendations

These are made when there is greater uncertainty about the four factors above or if local adaptation has to account for a greater variety in values and preferences, or when resource use makes the intervention suitable for some, but not for other locations. This means that there is a need for substantial debate and involvement of stakeholders before this recommendation can be adopted as policy.

When not to make recommendations

When there is lack of evidence on the effectiveness of an intervention, it may be appropriate not to make a recommendation. The lack of evidence can be highlighted by stating: "No recommendation can be made because of insufficient evidence". Instead of providing a recommendation, the findings of the systematic review or an overview of interventions may be published. By doing so, a range of optional interventions can be presented without indicating a preference for one over the other. In other situations guidance from WHO might be needed, despite there being little or no evidence. In these instances the absence of evidence should be highlighted and the basis of the options presented such as case reports, national experience or opinion, should be clearly indicated.

Research recommendations

When there is a lack of evidence, or the available evidence is insufficient, research recommendations should be specified, and prioritized if appropriate. In formulating research needs, guideline groups should be as specific as possible about what is needed and why. Research recommendations can be structured as shown in Table 8.2.

Table 8.2 Suggested format for research recommendations

Core elements		
E	Evidence	What is the current state of the evidence?
P	Population	What is the population of interest?
I	Intervention	What are the interventions of interest?
C	Comparison	What are the comparisons of interest?
O	Outcome	What are the outcomes of interest?
T	Time stamp	Date of literature search or recommendation

continues

Optional elements		
d	Disease burden	Disease burden or relevance
t	Time	Time aspect of core elements of EPICOT
s	Study design	Appropriate study type according to local need

Reaching agreement on recommendations

The WHO steering group usually prepares the draft recommendations, including a justification and a reference to the relevant evidence profile for each recommendation.

The guideline development group reviews and discusses the evidence profiles presented by the methodologist. The guideline development group considers values and preferences and resource implications of the intervention. If evidence on these was collected, this is reviewed and discussed.

The guideline development group agrees on the direction and the strength of the initial recommendation. Ideally the group should take decisions about recommendations by consensus, but guideline development groups need to decide how they will reach a decision if consensus cannot be reached. Voting rules should be agreed on before the meeting.

Writing recommendations

Recommendations needs to be clear and actionable, reflect the PICO format, and indicate the strength of the recommendation and the quality of the evidence. Wherever possible, the language should be consistent across all recommendations made in a guideline.

Further reading

Brown P et al. How to formulate research recommendations. *BMJ,* 2006, 333:804–806.

Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 5. Group processes. *Health Research Policy and Systems,* 2006, 4:17.

Greenhalgh T. How to formulate research recommendations: the pie or the slice? *BMJ,* 2006, 333:917.

Jaeschke R et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ,* 2008, 337:327–337.

Murphy NK et al. Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment,* 1998, 2(3).

9. **Producing and publishing your guideline**

Peer review

WHO guidelines should undergo peer review during development and before the draft is finalized for publication. There are several stages when peer review and external comment are sought.

- Drafts of the questions formulated for the guideline should be circulated for comments to experts and end-users at WHO headquarters, regional offices and externally.
- If systematic reviews are commissioned, the systematic review protocol (outlining search strategy and eligibility criteria) and included studies may be circulated to experts for comments on the methods and evidence identified.
- Draft evidence profiles can be circulated to experts for identification of any missing evidence and are reviewed at the guideline development group meeting.
- A final draft guideline with recommendations may be circulated for review before clearance.

The process of reviewing comments and responding to them should be recorded. It is not necessary to respond to every single comment individually. However, it is important to document how comments were handled, either as a version of the document with the changes, or as a separate summary.

If the guideline is circulated for comments after recommendations are finalized, be clear about what changes can be made. It is suggested that changes after finalization should be restricted to major errors of fact.

Different types of guidelines have a slightly different peer review requirements.

- Rapid advice guidelines: peer review can be limited to review of the complete draft only, immediately before final clearance, perhaps by 3–6 experts.
- Standard guidelines: a more complete peer review process is expected, including:
  - review of questions;
  - review of evidence tables and completed draft recommendations (after the guideline meeting);
  - a record of the response to the comments and any changes that are made.
- Full guidelines: peer review would be expected to be as above, with an optional additional review after a second draft.

Guideline format

All guidelines should have an executive summary, a main body and appendices. A general recommendation for the length of these sections is the 1–3–25 rule – i.e. an executive summary of 1 page, the main guideline of 3 pages, and appendices of 25 pages.

The executive summary should contain the key recommendations of the guideline. As executive summaries are often read as stand-alone documents, the quality of evidence for each recommendation should be specified in the executive summary as well in the main body of the guideline.

Summary of finding tables and descriptive evidence tables should be made publicly available but do not have to be included in the main guideline document. They can be published electronically as background documents, as long as they are cited in the guideline itself.

Prior to submission for clearance, the WHO steering group may wish to use the AGREE 2 appraisal instrument (http://www.agreetrust.org) to check whether the guideline meets international reporting criteria.

The production process

Production of WHO guidelines should follow the same process as other WHO publications. Detailed information on each step is available on the Intranet, a brief overview is provided here.

Writing

Identify a writer early in the process. This can be a WHO staff member or an external writer contracted on a freelance basis. If the writing will be done by a staff member, it is important to accurately estimate the demands that will be made on the person's time. Once you have an idea of the approximate length of your document, you can make a rough calculation of the time needed and can begin negotiations with an external writer if necessary. WHO does not have a standard writing pay scale but WHO Press usually advises a minimum of US$ 0.50 per word for writers, or a negotiated daily rate from current daily pay rates for consultants (available from the Human Resources Department). When negotiating fees and schedules, calculate a minimum of one week of full time work to produce 5000 words.

It is strongly recommended to avoid the 'committee' approach to writing a guideline. Asking experts to draft chapters for free may seem to be a cheap and efficient way of getting the job done, but unless you can guarantee quality, consistency and timely delivery, it will inevitably create more work than it eliminates.

Legal advice on proprietary products

Proprietary products should not be named. Devices and diagnostics used in interventions should be described generically avoiding identification

of specific products and trademarks. If in doubt, please contact WHO's legal counsel.

Editing and proofreading

You will also need an editor and a proof reader. WHO press maintains lists of approved freelance editors and proof readers, and provides sample terms of reference and standard rates of pay for these tasks. The best editors and proof readers are often booked up many months in advance, so plan production schedules as early as possible, and reserve their time accordingly.

Executive clearance and GRC approval

GRC review of final guideline documents occurs as part of the final executive clearance. Documents should be in a final edited form ready for layout and printing when they are submitted for final clearance.

Layout

Once you have an edited, proofed and cleared text, you will need to send it for layout. Again, WHO Press can advise on external typesetters and the specifications that you should include when contracting for this work. The WHO graphics team also provides an internal layout service. As many design decisions have major implications for the cost of production, printing, dissemination and subsequent translations, it is worth discussing the possibility of using an existing publication template with WHO Press before engaging an external designer. You will need a cover design, an ISBN (international standard book number) and a barcode, the latter two are issued by WHO Press.

Printing

Internal print will provide printing quotes and arrange for your files to be sent to the printer. You must have the printers' proofs checked again by your proof reader, so be sure to include this step in the initial proofreading contract. Once the print copies are delivered, you can focus on distribution and implementation.

Disseminating guidelines

Dissemination involves making guidelines accessible, advertising their availability and distributing them widely. Guideline developers should consult with WHO Press on priced and mandatory free distribution. When thinking about the dissemination of your guideline, consider the following options:

Online publication

There are different formats in which your guideline can appear on the Internet. At a minimum, you should contract your designer or typesetter to produce a web-ready PDF (portable document format) – a smaller file size than the PDFs produced for print – that is easier to download and navigate. Depending on the length of the guideline and its intended audience, you may

also wish to consider providing full-text HTML (hypertext mark-up language) and additional materials, both electronic and printed. The WHO web team is a good source of advice.

Translations

Because WHO guidelines target a global audience it might be appropriate to provide translations of the guideline in one or several languages, particularly the six official languages Arabic, Chinese, English, French, Russian and Spanish. To ensure accurate translation of technical content, experts should be involved in checking the translations. Translations must be planned in advance and the timing of the translations discussed with the translation suppliers or regional office involved. Do not forget to budget for translation costs. To reduce translation costs, translations may be limited to the executive summaries. Special care should be taken in the translation of the recommendations themselves: the meaning of the recommendation, or its strength should not change in translation.

Journals

The systematic reviews commissioned for the guideline may be submitted for publication in the Bulletin of the World Health Organization or other journals. Cochrane reviews are published in the Cochrane database. In order to increase awareness of the guideline, the process and/or recommendations may also be published in peer-reviewed journals.

Other forms of dissemination

Mobile phone applications for guideline dissemination and decision support are being developed. You may also wish to consider planning an official launch, a press release or press conference, an announcement on the WHO web site, distribution through regional offices or at meetings, or endorsement by stakeholders and interest groups.

Updating guidelines

Review-by date

WHO guidelines should be issued with a 'review-by' date to indicate how long the recommendations are expected to remain valid. There is no absolute rule about the length of validity. In deciding on the date by which a guideline should be reviewed, take account of the pace of change of research on the topic, areas where no evidence has been found, and the potential need for new advice. For standard and full guidelines a minimum of two years and a maximum of five years are suggested. The department that will be responsible for initiating the review should be named in the document.

Updating recommendations

All WHO recommendations that are not based on rigorous evidence review (particularly those published before the GRC was in established in 2007), should be updated as described in this handbook. Updating guidelines can be challenging if evidence has to be retrieved to support a large number of existing recommendations. In this situation it is important to give priority to controversial areas, or those in which new evidence has emerged.

Interim updates

Occasionally guideline developers may want to update guidelines before the 'review-by' date. For example, if new evidence supporting or contradicting the current recommendations is published. This new evidence should always be seen in the context of the total body of evidence supporting the recommendations. Therefore, it should be part of a new or updated systematic review. Any interim update that involves changing recommendations, needs to be reviewed by the guidelines review committee. Updates that add new evidence without changing the recommendations do not require review, although under certain circumstances, if the topic or new evidence is highly controversial, GRC review may be advisable.

10. **Implementation and evaluation**

Implementation

Implementation of a guideline should be taken into account right from the beginning of its development. A guideline project should ideally be in a departmental or other programme of work on the particular topic since that is more likely to lead to an effective plan for implementation.

Implementation will generally be the responsibility of regions and national or subnational groups, which is why they need to be involved in the development of the guideline. WHO headquarters and regional offices can support implementation activities by providing tools, support and coordination of efforts.

Implementation strategies need to be tailored to specific local circumstances. The basic steps for implementing a guideline are:

- analyse local needs and priorities (look for additional data on actual practice);
- identify all potential barriers and facilitating factors;
- determine available resources;
- design an implementation strategy (consider how to encourage the adoption of the recommendations and how to make the overall context favourable to the proposed changes).

There is a range of derivative documents or tools that can be developed to facilitate implementation. These can be distributed with the guideline, or they can be developed by local guideline implementers. Such documents or tools may include a slide set reflecting the guideline content; a 'how to' manual or handbook; a flow chart, decision aid or algorithm; fact sheets; quality indicators; checklists; application tools; templates, etc.

Implementation or operational research can help inform field testing and rollout strategies to promote the uptake of recommendations.

Evaluation and monitoring

An evaluation should be done to measure the impact of the guideline. The guideline should include outcome or performance measures that can be monitored for the main recommendations. Performance measures might be related to:

- guideline dissemination
- change in practice performance
- change in health outcomes
- change in end-user knowledge and understanding
- economic consequences.

Ideally, there should be baseline measures against which to assess performance in relation to the change induced by the guideline. Impact assessment of guidelines will need to be done in the places that they are implemented. WHO should work with Member States to evaluate the impact of the guidelines by coordinating efforts and providing advice and practical support.

Further reading

Grol R, Wensing M, Eccles M. *Improving patient care*. Edinburgh, Elsevier, 2005.

Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet,* 2003, 362:1225–1230.

Hearnshaw HM et al. Are audits wasting resources by measuring the wrong things? A survey of methods used to select audit review criteria. *Quality and Safety in Health Care,* 2003, 12:24–28 (http://qshc.bmj.com/cgi/reprint/12/1/24, accessed 8 June 2012).

How to put the evidence into practice: implementation and dissemination strategies. Handbook series on preparing clinical practice guidelines. Canberra, National Health and Medical Research Council, 2000 (http://www.nhmrc.gov.au/publications/synopses/_files/cp71.pdf, accessed 8 June 2012).

Wensing M, Wollersheim H, Grol R. Organizational interventions to implement improvements in patient care: a structured review of reviews. *Implementation Science,* 2006, 1:2 (http://www.implementationscience.com/content/1/1/2, accessed 8 June 2012).

Implementation of CPGs to change practice and outcomes. In: Davis D, Goldman J, Palda VA. *Handbook on clinical practice guidelines*. Ottawa, Canadian Medical Association, 2007 (http://www.cma.ca//multimedia/CMA/Content_Images/ClinicalResources/PDF/English/CPGHandbook.pdf, accessed 8 June 2012).

SIGN 50: a guideline developer's handbook. Section 10. Implementation. Revised edition 2008. Edinburgh, Scottish Intercollegiate Guidelines Network, 2008 (http://www.sign.ac.uk/pdf/sign50.pdf, accessed 8 June 2012).