WHO

Handbook for Guideline Development

2nd edition

World Health Organization

WHO Library Cataloguing-in-Publication Data
WHO handbook for guideline development – 2nd ed.

1.Guidelines as Topic – standards. 2.Review. 3.Meta-Analysis. 4.Peer Review. 5.Evidence-Based Medicine. 6.World Health Organization. I.World Health Organization.

ISBN 978 92 4 154896 0 (NLM classification: WA 39)

Acknowledgements

This handbook was prepared by Susan L. Norris. Laragh Gollogly and Charles Penn provided advice and support throughout the development process. Contributions in the areas of equity, human rights, gender and social determinants were made in this edition of the handbook (including Chapter 5) by the WHO Gender, Equity and Human Rights mainstreaming team (acknowledged in alphabetical order): Erik Blas, Theadora Koller, Veronica Magar, Rebekah Thomas, and Joanna Vogel. The other main contributors were Holger J. Schünemann (McMaster University, Hamilton, Canada) to Chapters 9 and 10, Chantelle Garritty (Ottawa Hospital Research Institute, Ottawa, Canada) to Chapters 8 and 11, and Ahmed Abou-Setta (University of Manitoba, Winnipeg, Canada) to Chapter 11. Maria Luisa Clark edited this handbook, Myriam Felber provided technical support and Amanda Milligan proofread.

Valuable inputs in the form of peer review and suggestions were received from members of the WHO Guideline Review Committee (2014). (An updated list of Committee members is available to WHO staff at http://intranet.who.int/homes/ker/grc). The following individuals provided peer review comments (acknowledged in alphabetical order): Ahmed Abou-Setta (University of Manitoba, Winnipeg, Canada); Yaolong Chen (Lanzhou University, Lanzhou, China); Mauricio Ferri (University of Calgary, Calgary, Canada); Lisa Hartling (University of Alberta, Edmonton, Canada); Joerg Meerpohl (University of Freiburg, Freiburg, Germany); Rose Relevo (Oregon Health & Science University, Portland, United States of America); and Vivien Welch (University of Ottawa, Ottawa, Canada).

This handbook was produced under the overall direction of Najeeb Al-Shorbaji and Hooman Momen.

Funder: The Bill & Melinda Gates Foundation provided funding to support development of this handbook.

Production editor: Melanie Lauckner
Design & layout: Sophie Guetaneh Aguettant

Contents

Foreword

This handbook provides step-by-step guidance on how to plan, develop and publish a World Health Organization (WHO) guideline. It covers the methods, processes and procedures for producing a document that meets WHO standards. It does not provide detailed technical guidance on many of the steps: this can be obtained from the references in the handbook and through references to the published scientific literature listed on WHO's Guideline Review Committee (GRC) intranet site (available to WHO staff). Additional chapters of this handbook containing detailed guidance on selected topics are also available and new chapters will continue to be added to the GRC intranet site in response to the needs of WHO guideline developers.

The first edition of this handbook, published in 2012, provided general guidance on the steps involved in guideline development and on GRC and WHO processes and procedures. This, the second edition, provides additional detailed guidance on each step in guideline development, as well as two additional chapters. Chapter 5 deals with the importance of considering equity, human rights, gender and the social determinants of health in formulating recommendations. The chapter provides specific entry points for integrating these issues into each and every guideline developed by WHO. Chapter 11 describes the development of rapid advice guidelines in the context of a public health emergency, with a focus on how the standard methods of evidence review and synthesis need to be modified to produce such guidelines.

The science underpinning evidence identification and synthesis and the translation of a body of evidence into recommendations continues to evolve. Because of this, any manual on how to produce a guideline requires frequent reassessment and updating. The GRC Secretariat is committed to providing up-to-date guidance that reflects the latest methods and approaches in the peer reviewed literature and the best practices internationally. Feedback on this handbook is welcome: please contact the GRC Secretariat at WHO with comments and suggestions for additional topics for future updates of the handbook or for online dissemination.

GRC Secretariat

1. Introduction

1.1 What is a WHO guideline?

A WHO guideline is any document developed by the World Health Organization containing recommendations for clinical practice or public health policy. A recommendation tells the intended end-user of the guideline what he or she can or should do in specific situations to achieve the best health outcomes possible, individually or collectively. It offers a choice among different interventions or measures having an anticipated positive impact on health and implications for the use of resources.

Recommendations help the user of the guideline to make informed decisions on whether to undertake specific interventions, clinical tests or public health measures, and on where and when to do so. Recommendations also help the user to select and prioritize across a range of potential interventions.

1.2 Why does WHO develop guidelines?

WHO develops guidelines whenever Member States, WHO country offices, external experts or other stakeholders ask for guidance on a clinical or public health problem or policy area. This generally happens when they are uncertain about what to do or how to choose among a range of potential policies or interventions. Uncertainty can be triggered by a new public health problem or emergency; the uncovering of new evidence; an absence of good-quality evidence (or of any evidence at all); or a change in resource availability or access to services.

1.3 What principles underlie WHO guidelines?

The Twelfth General Programme of Work (GPW) of WHO (2014) specifies that in its normative and standard-setting work, WHO is and will remain a science- and evidence-based organization with a focus on public health. Guidelines are the fundamental means through which the Organization fulfils its technical leadership in health, as identified in the GPW. The GPW states that "WHO's legitimacy and technical authority lie in its rigorous adherence to the systematic use of evidence as the basis for all policies" (*1*).

WHO has adopted internationally recognized methods and standards for guideline development to ensure that its guidelines are of the highest quality. WHO guidelines must be developed in observance of the following principles:

- Guidelines address an area of uncertainty and an unmet need for guidance.
- Guidelines reflect the core WHO value of the "right to health" (*2*).
- The process of developing recommendations is explicit and transparent: the user can see how and why a recommendation was developed, by whom, and on what basis.
- The process of developing guidelines is multidisciplinary and includes all relevant expertise and perspectives, including input from stakeholders.
- The processes and methods used in each step of guideline development aim to minimize the risk of bias in the recommendations.
- Recommendations are based on a systematic and comprehensive assessment of the balance of a policy's or intervention's potential benefits and harms and explicit consideration of other relevant factors.
- The evidence used to develop WHO guidelines is publicly available.
- Recommendations can be implemented in, and adapted to, local settings and contexts.
- Guidelines should be tailored to a specific audience. (The audiences that WHO guidelines can target include public health policy-makers, health programme managers, health-care providers, patients, caregivers, the general public and other stakeholders.)

1.4 What is the aim of this handbook?

This handbook presents the important principles of guideline development and provides stepwise instruction on the technical and procedural aspects of developing a WHO guideline. It aims to steer readers through the process to ensure that the resulting guideline is credible; the recommendations accurately reflect the balance of potential benefits and harms; the guideline meets WHO's publication standards in terms of content, methods and presentation; the underlying evidence is accessible; and the recommendations are implementable and positive in their impact on population health.

1.5 For whom is this handbook intended?

This handbook is intended for:

- any WHO department, programme or staff member wishing to produce a guideline;
- members of a WHO guideline steering group;
- members of a WHO guideline development group (GDG);
- members of a WHO guideline external review group; and
- anyone interested in understanding how WHO develops guidelines.

1.6 What is in this handbook?

This handbook follows the entire development of a WHO guideline – from determining if a guideline is needed through to eventual publication. The guideline development process and its primary contributors are summarized in Table 1.1.

1.7 Types of WHO guidelines

It is important to consider what type of guideline will best fit the intended purpose, as this will determine the methods, resources and time frame for development, finalization and dissemination (see Table 1.2). WHO guidelines comprise a broad spectrum of products that vary mainly in terms of the following features:

- purpose;
- scope;
- the point in time at which the guideline is being developed relative to the life-span of an intervention;
- the organizations or entities developing the guideline;
- the presence in the guideline of new versus previously published recommendations; and
- the timeline.

WHO develops four main types of guidelines, defined in terms of the above characteristics: standard, consolidated and interim guidelines, and guidelines produced in response to an emergency or urgent need. In addition, there are other, less frequent types of guidelines.

Table 1.1. The guideline development process at WHO

Stage/primary contributor	Step	Chapter
Planning		
WHO Member State, WHO country office or public/private entity	Request guidance on a topic	1
WHO technical unit	Determine if a guideline is needed; review existing WHO and external guidelines	2
	Obtain approval for guideline development from the director of the relevant technical unit at WHO	2
	Discuss the process with the GRC Secretariat and with other WHO staff with experience in developing guidelines	2
	Form the WHO guideline steering group	3
	Identify sufficient resources; determine the timeline	2
WHO guideline steering group	Draft the scope of the guideline; begin preparing the planning proposal	2,4
	Identify potential members of the GDG and its chair	3
	Obtain declaration of interests and manage any conflicts of interest among potential GDG members	6
WHO guideline steering group and GDG	Formulate key questions in PICO format; prioritize outcomes	5, 7
WHO guideline steering group	Finalize the planning proposal and submit it to the GRC for review	4
GRC	Review and approve the planning proposal	4
Development		
Systematic review team	Perform systematic reviews of the evidence for each key question	8
	Evaluate the quality of the evidence for each important outcome, using GRADE as appropriate	9
WHO guideline steering group	Convene a meeting of the GDG	10,11
GDG	Formulate recommendations using the GRADE framework	10,11
WHO steering group	Draft the guideline document	10,11
External review group	Conduct external peer review	12
Publishing and updating		
WHO guideline steering group and editors	Finalize the guideline document; perform copy-editing and technical editing; submit the final guideline to the GRC for review and approval	12
GRC	Review and approve the final guideline	12
WHO guideline steering group and editors	Finalize the layout; proofread	12
	Publish (online and in print as appropriate)	12
WHO technical unit and programme manager	Disseminate, adapt, implement, evaluate	13
WHO technical unit	Update	12

GDG: guideline development group; GRADE: Grading of Recommendations Assessment, Development and Evaluation; GRC: Guideline Review Committee; PICO: population, intervention, comparator, and outcome.

Table 1.2. Characteristics of the types of guidelines produced by WHO

Type of guideline	Purpose	Scope	Developer	New or existing recommendations	Development period
Primary types of guidelines					
Standard	To provide recommendations on a specific topic or condition	Focused or comprehensive	WHO technical staff	Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate	6 months to 2 years
Consolidated	To aggregate all the existing guidance on a disease or condition	Comprehensive	WHO technical staff	Existing recommendations that have been evaluated and found to be up to date; may contain some new recommendations	1 to 2 years
Interim	To provide guidance when new interventions, exposures or diseases arise or when new evidence becomes available or data are likely to be incomplete	Focused	WHO technical staff	New	6 to 9 months
Guidelines produced in response to an emergency or urgent need	To meet an emergent or urgent public health need when the short timeline mandates a modified process	Focused	WHO technical staff	Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate	1 to 3 months
Other types of guidelines					
Developed in collaboration with (an) external organization(s)	To provide recommendations on a specific topic or condition when organizations have a shared interest or remit	Focused or comprehensive	WHO technical staff and staff from the external organization(s)	Usually new; may contain existing recommendations if they have been evaluated and updated as appropriate	1 to 2 years
Developed by (an) external organization(s)	To provide recommendations on a specific topic or condition when a guideline produced by an external organization already exists	Focused or comprehensive	External organization(s)	Existing recommendations; may be updated	1 to 3 months
Adaptation of existing WHO guidelines	To develop recommendations specific to the local context where they will be implemented	Focused or comprehensive	Policy-makers and programme managers in WHO Member States	Reflect the content of the original guideline	1 to 3 months

1.7.1 Standard guidelines

A standard guideline covers a clinical or policy area (e.g. the treatment of postpartum haemorrhage or the minimum requirements for the safe delivery of care to people with HIV infection). Such guidelines vary greatly in scope and focus: they might address the use of a single drug for a disease or condition, such as naloxone injection by lay persons for suspected opioid overdose, or they might encompass the full scope of a condition or public health problem, such as the diagnosis, screening and treatment of type 2 diabetes mellitus. Recommendations in a standard guideline are either developed de novo or by updating previous WHO guidelines.

Standard guidelines generally focus on one or more of the following:

- clinical interventions (e.g. the management of severe acute malnutrition in infants and children);
- health-care system or policy approaches (e.g. country pharmaceutical pricing policies);
- public health interventions or exposures (e.g. optimal intake of dietary folate in pregnant women);
- diagnostic tests (e.g. fluorescent light-emitting diode [LED] microscopy for the diagnosis of tuberculosis), or
- surveillance and monitoring (e.g. surveillance guidelines for measles, rubella and congenital rubella syndrome in the WHO European Region).

Standard guidelines usually take between 9 and 24 months to complete, depending on their scope, and should be prepared after wide consultation on their need, scope and rationale. They should be supported by one or more systematic reviews of the evidence and finalized after one or two meetings of the GDG. A standard guideline has to be reviewed by a specified date that will depend on how fast the evidence in the topic area is expected to change. Most WHO guidelines fall into this category.

1.7.2 Consolidated guidelines

A consolidated guideline (also known as a compilation of guidelines) contains recommendations from existing WHO guidelines, or from guidelines produced by other organizations that have followed processes consistent

with those used by WHO (see Section 1.8.2). Producing consolidated guidelines is complex because existing guidelines may need to be updated and new recommendations may have to be added to address important gaps in the existing guidance. In addition, maintaining the document is difficult, since individual recommendations may become outdated at different times. Production times for consolidated guidelines vary widely.

All recommendations contained in a consolidated guideline should be up to date and approved by the WHO Guideline Review Committee (GRC): existing recommendations may need to be updated during preparation of the consolidated guideline. During the updating process, all the standard procedures as outlined in this handbook should be followed. Existing recommendations must be thoroughly, clearly and explicitly cross-referenced.

Consolidated guidelines require review by the GRC if any of the included recommendations were initially published without GRC review; the updating process led to changes in any of the existing recommendations; or new recommendations were developed. A compilation of guidelines that includes recommendations developed by organizations external to WHO must also be reviewed by the GRC. Only consolidated guidelines whose recommendations have all been previously approved by the GRC and have remained unchanged during the updating process do not require review by the GRC.

1.7.3 Interim guidelines

Interim guidelines are produced when WHO is asked to provide guidance when the available data and information are most certainly incomplete, especially if additional data are anticipated in the near future. This can occur when an intervention – a drug, a medical device or a health practice, for example – has just appeared or is about to be made available; a new disease or condition has emerged; or new data on an existing intervention, exposure, disease or condition have been reported. Interim guidelines usually have a very focused scope and a short shelf-life. They should always clearly indicate when additional evidence affecting the interim recommendation(s) is expected to be reported, and thus when an update is anticipated. Although the target audience or other stakeholders may demand that interim guidance be generated quickly, this type of guideline fully complies with all processes and procedures and meets the standards set out in this handbook.

1.7.4 Guidelines in response to an emergency or urgent need

Certain circumstances preclude the development of a standard guideline. This is the case when WHO must provide guidance in response to public health emergencies. Depending on the type of event or situation, such guidelines may need to be produced within hours, days, weeks or months. Their purpose and the methods used to develop them vary with the time frame in which guidance is required.

Although guidelines that must be produced rapidly cover a wide spectrum, it is convenient to consider two basic types:

- **Emergency (rapid response) guidelines** – Public health emergencies may necessitate a response from WHO within hours to days. Hence, many of the guideline development processes and methods outlined in this handbook are not applicable. WHO staff will need to quickly identify relevant existing guidelines produced by WHO or other entities or may need to issue recommendations based on expert opinion only. The basic principles of high-quality guidelines, as outlined in Section 1.3, must be adhered to: the recommendations should be applicable to the specified population and setting; the sources for the recommendations must be indicated; to the extent possible, efforts must be made to minimize the risk of bias; and the development process should be explicit and transparent. It may not be feasible to perform a systematic review of all available evidence. However, only sources of high-quality evidence should be used. It is important that the decision-making process be documented and that the rationale for each recommendation be stated, even if it is based on indirect or very limited evidence or on expert opinion. The GRC Secretariat is developing further guidance on this type of guideline.
- **Rapid advice guidelines** – If a public health event continues for an extended period, the initial emergency (rapid response) guidelines must be reviewed to take into account both the evidence emerging from the event and a systematic review of the relevant evidence. Such rapid advice guidelines will follow WHO processes more closely and must meet the standards for guideline development at WHO. These guidelines are published with a review-by date that indicates when the guidance may become invalid, or when it will be updated or converted to a standard guideline. Rapid advice guidelines must be reviewed by the GRC, both at the planning proposal and final guideline stages. Chapter 11 provides detailed guidance on when and how to produce this type of guideline.

For a variety of reasons, WHO may wish to issue recommendations quickly, even in the absence of a public health emergency or continuing urgency. It may choose to do so, for example, when a new drug becomes available, when new information on an existing technology is likely to change existing guidance, or because a Member State or external entity has made an urgent request for guidance. Recommendations issued in such situations must be based on a systematic review and assessment of the evidence and meet all the standards outlined in this handbook, including GRC review at the planning proposal and final guideline stages. However, it may be possible to produce such guidelines within a relatively short time, if the appropriate resources and expertise are available.

1.8 Additional types of guidelines produced by WHO

1.8.1 Guidelines developed in collaboration with external organizations

Clinical and public health 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.

Collaborative guidelines must adhere to the basic principles for WHO guidelines. When developing such guidelines, it is particularly important to note that:

- no funding can come from commercial sources, and conflicts of interest must be declared and managed as per WHO procedures;
- a systematic review of the relevant evidence is required;
- the methods used to develop the recommendations must be explicit and transparent;
- the experts who develop the recommendations must adequately represent the guideline's target audience, whether global or regional; and
- WHO Press should be consulted on publishing agreements, including matters relating to copyright and open access.

When collaborating with an external organization, WHO may not be able to insist that all WHO processes and procedures for guideline development be followed. WHO staff planning such guidelines need to submit a planning proposal to the GRC describing in detail how the guideline will be developed and highlighting any deviation from the procedures set out in this handbook. The GRC will assess such proposals on a case-by-case basis.

1.8.2 Guidelines developed by external organizations

Guidelines that serve WHO's purpose may already exist and WHO may consider adopting, adapting or incorporating all or part of a guideline developed by an external organization. In order to adopt or endorse such a guideline, WHO staff in the relevant technical unit must carefully evaluate the guideline's quality and applicability to WHO's needs.

The guideline must meet WHO's standards for guidelines as depicted in this handbook. The two standards that externally developed guidelines can seldom meet are: (a) applicability to a global target audience (most notably low- and middle-income countries); and (b) formulation by a panel of experts with representation from all relevant WHO regions. WHO rarely adopts the recommendations produced by third parties because they fail to comply with these requirements.

A plan for development of a WHO guideline that endorses, adopts or adapts, in whole or in part, a guideline produced by another organization should be submitted to the GRC for assessment, as with any WHO guideline. The planning proposal must include the rationale for selecting that particular guideline; a description of the process used to evaluate its quality, applicability and currency; and the methods and rationale for any revisions (including updating of the evidence and recommendations).

1.8.3 Adaptation of existing WHO guidelines

WHO guidelines originally intended for one setting may be adapted for use in another. For example, guidelines on routine obstetric care can be adapted for emergency settings. Adaptations of WHO guidelines must follow standard GRC procedures, including GRC review of the planning proposal and the final guideline. Guidelines may also be adapted in emergency or urgent situations, when existing guidelines need to be quickly identified and assessed. The GRC Secretariat is developing further guidance on when and how to adapt existing guidelines in such situations.

1.9 Information products that are not considered guidelines

Certain types of documents are not considered guidelines and do not need to be submitted to the GRC for review. They include:

- documents that state established principles (e.g. human rights, WHO constitutional issues);
- WHO Secretariat reports and other papers submitted to the Governing Bodies (Executive Board, Regional Committees, World Health Assembly);
- information documents that report facts, describe evidence, or document or review existing practices and interventions, provided that such documents are not making recommendations or advocating commitment of resources;
- documents containing standards for manufacturing health technologies, such as pharmaceuticals and vaccines;
- "how to" documents such as operational manuals (e.g. how to set up a research project or how to implement a service) or implementation guides or tools based on GRC-approved guidelines; and
- documents that describe standard operating procedures for organizations or systems.

If you are not sure whether your proposed document is a guideline, please consult with the GRC Secretariat.

1.10 The Guideline Review Committee

The GRC was established by WHO's Director-General in 2007 to ensure that WHO guidelines are of high quality, that they are developed using a transparent and explicit process, and that, to the extent possible, recommendations are based on evidence (*3*).

The GRC is composed of approximately 30 individuals, including representatives from all WHO regions as well as external members, and meets monthly to review submitted documents.

1.10.1 Terms of reference of the GRC

All WHO publications containing recommendations must be approved by the GRC according to WHO policies and procedures (*4*). The GRC reviews every WHO guideline twice during its development – once at the initial planning stage and again after the recommendations have been developed and the guideline document has been finalized and edited.

The review of a planning proposal includes an assessment of whether the proposed guideline development process is consistent with the steps described in this handbook. The review of final submissions ensures that the approved process has been followed and that the final guideline document meets WHO standards as laid out in this handbook, contains clear and actionable recommendations, and meets all WHO reporting requirements. The GRC also provides suggestions and advice on guideline development processes and procedures and on the methods for developing high-quality guidelines at any stage of the development process. The GRC does not provide guidance on the technical content of the guidelines, but rather on how to develop guidelines that will meet WHO standards.

GRC approval is part of WHO's internal clearance processes for the publication of guidelines. The GRC evaluates planning proposals and final guidelines that have been approved by the relevant directors of the technical unit(s) responsible for the document or their regional counterparts, and before clearance of the final guideline by the assistant director-general (ADG) or regional director.

The specific terms of reference for the GRC are to (*5*):

- define appropriate and standardized processes for guideline development at WHO;
- ensure that all guidelines published by WHO are of high quality and comply with explicit standards;
- develop and implement a plan to ensure that GRC members have the necessary knowledge of the approved methods for guideline development and to identify opportunities to build capacity in guideline development among WHO staff; and
- develop collaborations with other organizations and international networks that have methodological expertise and skills in guideline development, adaptation and implementation.

1.10.2 The GRC Secretariat

The GRC Secretariat has several important roles. In addition to supporting the GRC, it provides WHO staff with technical advice on guideline development, sets benchmarks, and evaluates guideline development processes, documents and standards to achieve quality assurance and improvement for WHO guidelines. More specifically, the GRC Secretariat:

- coordinates and provides technical support on guideline development to WHO technical units at headquarters and in the regional offices;
- organizes training on guideline development methods and procedures for WHO staff;
- provides administrative and technical support for the work of the GRC;
- collaborates with other organizations and international networks that provide methodological expertise in guideline development, adaptation and implementation; and
- maintains a database of GRC submissions.

Further details on the processes and procedures of the GRC and its Secretariat are found on the WHO GRC intranet site (*6*).

1.11 References

1. Twelfth General Programme of Work: not merely the absence of disease. Geneva: World Health Organization; 2014 (http://www.who.int/about/resources_planning/twelfth-gpw/en, accessed 11 November 2014).
2. Constitution of the World Health Organization. Geneva: World Health Organization; 1946 (http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf, accessed 11 November 2014).
3. Information Note - Establishment of a WHO Guidelines Review Committee [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2007 (http://intranet.who.int/admin/infonotes/documents/16_2007_en.pdf, accessed 11 November 2014).
4. WHO eManual/VIII - Information Products/VIII.2 Planning and Content Development/VIII.2.3 Guidelines [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2013 (http://emanual.who.int/p08/s02/Pages/VIII23Guidelines.aspx, accessed 11 November 2014).
5. WHO eManual/VIII - Information Products/VIII.1 General/VIII.1.3 Publishing policy groups [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://emanual.who.int/p08/s01/Pages/VIII13Publishingpolicygroups.aspx, accessed 11 November 2014).
6. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/about/, accessed 11 November 2014).

2. Planning guidelines

The more planning and thought that goes into guideline development at the beginning, the more efficient the entire process will be and the better the end product. The principles of good project management and high-quality research apply to guideline development. Careful consideration of the issues raised in this chapter and efforts put into developing a thorough and reasoned planning proposal (see Chapter 4) will reduce the time and resources expended on the guideline and result in a superior, more useful final product.

2.1 Is this guideline really needed?

When planning to develop a guideline, you must begin by asking a critical question: *Is this guideline really needed?* Guideline development is time-consuming and expensive. You must carefully consider several issues before embarking on the process.

Priority-setting in the normative work at WHO needs continuous evaluation and revision as priorities evolve within and across technical units. Requests for guidelines exceed the number of guidelines that can be produced with available resources. In keeping with the need to prioritize, be sure to ask certain key questions before deciding to develop a guideline.

2.1.1 Who wants the guideline?

Have one or more WHO Member States requested it? WHO guidelines must generally meet global or regional needs and have a public health perspective.

2.1.2 Should WHO be the organization to produce this guideline?

Have WHO's governing bodies requested the guideline? Is WHO in the best position to issue guidance on this topic? Does the topic fall within the scope of WHO's remit? Is another organization better suited to produce this guide-

line owing to its resources, implementation skills and experience, or its local and regional knowledge?

2.1.3 Do guidelines on the topic already exist?

Have guidelines on the same topic already been issued by other departments at WHO or by other organizations? Do they address the needs of WHO's Member States? Avoid duplicating existing work by consulting the relevant departments, the GRC Secretariat and the WHO library. If existing guidelines are identified, examine their quality and applicability to the current need. Consider updating existing WHO recommendations if they are out of date or were developed before the GRC was formed.

2.2 What purpose and audience does the guideline serve?

WHO staff developing guidelines need to have a clear goal and audience in mind when they set about developing a guideline.

2.2.1 What will the guideline achieve?

Will the guideline address poor practice or change programme approaches or health policy while aiming to fulfil the right to health? Will the guideline provide direction for users in areas in which they feel uncertain about what to do? Resolving uncertainty should be the focus of most guidelines.

2.2.2 What are the specific objectives of the guideline?

What health indicators or other measures do you hope to improve? What test or approach do you wish to evaluate and provide guidance on? Clear, achievable objectives will govern and facilitate the development of your guideline.

2.2.3 Who is the target audience?

Guidelines must have a clearly defined target audience (end-user) which is identified early in the guideline development process, and the recommendations need to be tailored to that audience. (Note that the end-users are those who will apply the guidelines, not the recipients of the recommended interventions.) Writing documents to meet the needs of policy-makers, health-sector managers, clinicians and other health professionals simultaneously is not straightforward and should be avoided.

2.2.4 Who are the recipients of the interventions?

Who are the service users and other individuals whose health and well-being will be affected by the recommended interventions? Attention to their perspectives and needs must be built into the guideline development process from the beginning.

2.3 When is the guideline needed?

The optimal time to start developing a guideline and the time frame for completing the task must be carefully considered at the outset.

2.3.1 Why now?

Is this the best time to develop recommendations, or are new data expected to emerge in the near future? Will the existing infrastructure or policies change in a manner likely to affect guideline implementation?

2.3.2 When is the guideline needed?

The projected timeline for guideline development needs to be matched with the time when the guidance is needed. Realistically, a high-quality standard guideline will take at least 9 to 12 months to produce if it has a fairly narrow scope. On the other hand, if the guideline aims to cover a large number of questions and involves many de novo systematic reviews, it may take as long as 2 or 3 years to produce.

2.3.3 Is the guideline a response to a situation calling for urgent advice?

If so, the guideline development process and methods may differ from those used for standard guidelines. Under a compressed timeline, the scope of a guideline may have to be greatly reduced. In addition, rapid review and guideline development processes and methods can result in an end product with a high risk of bias. WHO staff must balance this risk with the demand to produce recommendations rapidly (see Chapter 11).

2.4 Will the recommendations in the guideline be implemented?

The purpose of WHO guidelines is to improve the health and well-being of individuals and populations. To accomplish that, guidelines need to be disseminated, adopted or adapted, and their recommendations implemented.

2.4.1 Are the guideline and its implementation within the relevant programme's work area and budget?

Under the WHO reform (*1*) all work at WHO should fall within the scope of the General Programme of Work and Programme Budget, as agreed with WHO's Governing Bodies. If the guideline does not fall within this remit, it will be difficult to justify the resources and budget needed to complete the work.

2.4.2 Is a process for disseminating and implementing the recommendations in place?

Who is likely to implement the recommendations? What level of guidance is needed and in what format? What infrastructure and services are needed for implementation? Is access to the service an important consideration? These issues should be considered from the start, as they affect who should be involved, what resources are required and how the recommendations are crafted and presented.

2.5 Who should be involved in developing a guideline?

Guideline development involves collaboration among a large number of groups and people (see Chapter 3). These should be engaged from the outset.

2.5.1 Do you have your director's agreement?

You will need your director's formal approval before your planning proposal or your final guideline can be considered by the GRC. In addition, your ADG will need to approve the final guideline after its review and approval by the GRC.

2.5.2 What WHO departments must be involved?

Do other departments at headquarters or in WHO's regional offices, including those with cross-cutting mandates, have an interest in the guideline, have relevant expertise or are producing similar products? Prepare a list of technical units at WHO and of key external organizations, experts and stakeholders that will need to be consulted or involved in the guideline development process.

2.5.3 Have you identified experts in guideline methods?

Generally WHO staff commission systematic review and guideline development experts to assist in formulating the key questions; identifying, assessing and synthesizing the evidence; and translating the evidence into recommendations. Such experts should be identified early in the guideline development process.

2.6 Additional considerations

Once you are certain that a guideline is needed by its intended end-users and will benefit the service users or other recipients of the recommended measures or interventions, you need to consider several other factors before embarking on the project.

2.6.1 Do you have the necessary resources to develop a guideline?

To produce a high-quality guideline, you will need resources to identify or perform systematic reviews, convene meetings and cover production costs. Note that WHO cannot develop guidelines using money from commercial entities (see Chapter 6). Guideline development also involves long hours of dedication by WHO staff.

2.6.2 What publication types and formats are you considering?

What level of detail and format will your target audience find most useful? Guideline users are not generally interested in a detailed description of how guidelines are developed to conform to WHO standards, so you should plan and budget for derivative products such as summaries, algorithms or wall charts specifically tailored for the end-users of the guideline. Electronic versions, perhaps accompanied by short paper publications, may be more practical and cheaper than print versions.

2.6.3 What translations are you planning?

What languages are spoken by the people most in need of the advice in your guideline? Consider the implications for your budget and time frame and carefully choose the language(s) for the translation of your guideline.

2.7 Scoping the guideline

Scoping is the process of defining what the guideline will and will not include. To establish the scope, you need to determine:

- the areas of practice or policy to which the guideline applies;
- the interventions, approaches or exposures of interest (i.e. the priority topics);

- the individuals and/or populations (including subpopulations) that the recommendations are intended to affect; and
- the important outcomes – both benefits and harms – that may result.

The process of scoping a guideline will establish the focus for the recommendations, as well as the key questions that will govern the search for evidence to inform the recommendations. This process should ensure that the guideline is of manageable size, adequately focused, and capable of being executed within the allocated time frame and with the available resources.

Scoping is one of the most difficult but important aspects of guideline development. If the scope is right, the guideline process should be manageable and the end product relevant to the end-user. If the scope is wrong, resources are wasted and the end-user will remain uncertain about the optimal intervention or approach.

2.7.1 How to scope the guideline

Scoping a guideline involves a series of steps, generally performed or coordinated by the responsible technical officer.

- **Draft the potential scope** – Draft a proposed scope and a list of potential priority topics.
- **Set up the WHO guideline steering group** ("steering group" for the remainder of the text) – Convene a small group of WHO staff, including representatives of all relevant departments, to provide feedback on the proposed scope of the guideline.
- **Refine the list of priority topics** – Identify the key issues: 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 in which a change in practice is desired and feasible, and on areas where inequity, controversy or uncertainty exist. Although some background information may be useful, avoid extensive exploration of background questions (e.g. epidemiology, pathology, pharmacology) unless they directly relate to the area of controversy that the guideline is intended to resolve.

- **Search the literature** – Do a preliminary search of the literature to identify relevant information, including existing guidelines and systematic reviews, health technology assessment reports and economic evaluations of 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 conducted to retrieve the appropriate evidence.
- **Sharpen the focus** – Step back and ask if you need to include all these topics. For now try to restrict the final list to a minimum, as it tends to expand during the development of the guideline.
- **Formulate key questions** – Use the topic list to formulate the key questions to be answered in the guideline. These questions will guide the search for evidence and are best developed using the "population, intervention, comparator and outcome" (PICO) format (see Chapter 7).
- **•Equity, human rights and gender** – Will the recommendations affect equity, human rights and gender? Are the relevant subpopulations considered in the key questions? These issues should be considered early and throughout the guideline development process (see Chapter 5).
- **Review** – Once the steering group has established the guideline's scope, a summary of the scope should be circulated to the GDG and sometimes other stakeholders for comments. These groups should be reminded that WHO is producing a guideline, not a textbook, as the reviewers will almost invariably suggest expanding the planned scope.
- **Reconsider** – Once you have incorporated the external feedback as appropriate, review the scope and key questions again. Is the scope feasible? Is the time frame reasonable? Are sufficient financial and human resources available?

2.8 References

1. WHO reform [website]. Geneva: World Health Organization; 2014 (http://www.who.int/about/who_reform/en, accessed 11 November 2014).

3. Contributors and their role in guideline development

When developing WHO guidelines, four groups need to be established (Table 3.1):

- the steering group;
- the GDG;
- the external review group; and
- the systematic review team.

These groups have different skills, perspectives, roles, responsibilities and tasks. They are established at different times, but all work to produce a high-quality guideline.

3.1 The steering group

Once the technical unit has decided to proceed with developing a guideline, the steering group should be formulated, led by the responsible technical officer. The steering group includes members from all WHO departments and regional offices whose work deals directly with the topic of the guideline. It is wise, however, to limit the group to fewer than 8 or 10 members to maximize efficiency, although some guidelines require a larger steering group to encompass representatives from all relevant departments and regions.

Steering group members must be prepared to allocate a lot of time to this work: senior WHO staff who cannot do so should not be listed as members. Instead, they should be consulted as appropriate during the development process. If the guideline is being developed jointly with another organization, individuals from that organization will also be members of the steering group. Otherwise, the steering group is composed exclusively of WHO staff from headquarters and the regional offices.

The role of the steering group is to:

- provide administrative support for guideline development;
- draft the scope of the guideline and key questions in PICO format (see Chapters 2 and 7);
- identify the systematic review team and guideline methodologist(s);
- develop and finalize the planning proposal for submission to the GRC (see Chapter 4);

- oversee evidence retrieval, assessment and synthesis;
- select members of the GDG and the external review group;
- collect and assess disclosures of interest and manage conflicts in collaboration with the director of the technical unit and in consultation with the Office of Compliance, Risk Management and Ethics (CRE), as needed (see Chapter 6);
- organize GDG meetings;
- draft recommendations based on the decisions of the GDG;
- draft the final guideline, in collaboration with the technical writer;
- oversee peer review, review comments and revise the draft guideline as appropriate;
- submit the final guideline to the GRC and revise as indicated to meet GRC requirements;
- oversee publication and dissemination of the guideline; and
- monitor new information, user needs and requests that inform when an update may be needed.

The responsible technical officer is responsible for the efficient and effective function of this group and for liaising and consulting with departments and experts internal to WHO, and with the chair and members of the GDG as needed.

3.2 The guideline development group

The GDG is made up of external experts whose central task is to develop evidence-based recommendations. The GDG also performs the important task of finalizing the scope and key questions of the guideline in PICO format. This group should be established early in the guideline development process, once the steering group has defined the guideline's general scope and target audience and begun drafting the key questions.

Potential members of the GDG are identified by the steering group and are selected to encompass the technical skills, diverse perspectives and geographic representation needed. The group should be small enough for effective group interaction and decision-making, but large enough to ensure that all relevant expertise and perspectives are represented. A group of 10 to 20 is usually feasible and effective, although some GDGs are larger if the scope of the guideline is broad. The group can hold online or teleconference meetings but will usually need to have at least one face-to-face meeting to formulate the recommendations based on the systematic reviews of the evidence and other information (see Chapter 10).

The members of the GDG are not commissioned and do not receive any financial compensation other than for direct expenses associated with their work on the guideline. The responsible technical officer may develop terms of reference so potential GDG members clearly understand their roles and responsibilities before committing themselves. Members of the GDG participate in the guideline development process and at meetings as individuals and not as representatives of the institutions or organizations with which they are affiliated.

The role of the GDG is to:

- provide input into the scope of the guideline;
- assist the steering group in developing the key questions in PICO format;
- choose and rank priority outcomes that will guide the evidence reviews and focus the recommendations;
- examine the Grading of Recommendations Assessment, Development and Evaluation (GRADE) evidence profiles or other assessments of the quality of the evidence used to inform the recommendations and provide input;
- interpret the evidence, with explicit consideration of the overall balance of benefits and harms;
- formulate recommendations taking into account benefits, harms, values and preferences, feasibility, equity, acceptability, resource requirements and other factors, as appropriate; and
- review and approve the final guideline document before submission to the GRC.

3.2.1 Composition of the guideline development group

The GDG is multidisciplinary and composed of individuals from all WHO regions likely to use the guideline, except for employees of WHO or other United Nations organizations. Its membership should be balanced in terms of gender and geography. Possible conflict of interest is also an important consideration when selecting and confirming GDG members (see Chapter 6). There are several ways to identify, nominate and select members of the GDG. In addition to drawing members from established technical networks and WHO collaborating centres, consider publishing an open call for nominees. The aim is to have a diverse group that includes:

- relevant technical experts;
- end-users, such as programme managers and health professionals, who will adopt, adapt, and implement the guideline;

- representatives of groups most affected by the recommendations in the guideline, such as service users and representatives of disadvantaged groups;
- experts in assessing evidence and developing guidelines informed by evidence; and
- other technical experts as required (e.g. a health economist or an expert on equity, human rights and gender).

3.2.1.1 The chair of the guideline development group

The selection of the chair of the GDG is a key decision. The steering group usually selects the chair, but the choice should generally be agreed upon by members of the GDG. The chair should be an expert in facilitating groups that reach decisions based on consensus; be experienced at critically appraising and interpreting evidence and developing evidence-informed recommendations; and have no financial interests related to the guideline's topic. Although the chair should have a general knowledge of the topic of the guideline, no one with strong views about the interventions under consideration should chair the GDG. The chair should have experience engaging in consensus-based processes involving people with different opinions. The chair can be a guideline methodologist with expertise in evidence synthesis and in formulating recommendations based on evidence. A vice-chair should also be identified by the steering group to stand in if the chair is absent and to share in the chair's tasks and responsibilities. The expertise of the chair and vice-chair should be complementary, especially with regard to expertise in the content area versus guideline development methods or implementation. Complementary skills and perspectives will also help to balance the influence of a chair who is a content expert and has opinions on specific recommendations. Another acceptable option is to have two co-chairs with equal responsibilities and complementary expertise and perspectives. For instance, one co-chair might be a guideline methodologist and the other an expert on the subject at hand.

3.2.1.2 Technical experts

Individuals selected for their technical expertise in a guideline's subject area are critically important to GDGs but should not dominate the group. A balanced group includes a range of expertise and institutional and professional affiliations.

3.2.1.3 End-users of the guideline

People with direct experience in managing the condition or problem addressed by the guideline and who will have a role in implementing the new recom-

mendations – members of governmental and nongovernmental organizations, programme managers, health-care workers and other end-users of the guideline – should participate in the GDG. For example, palliative care nurses would participate in developing a guideline about pain management; hospital administrators would help to develop a guideline on infection prevention and control in health-care settings. The aim is to ensure that the final guideline document is useful to its end-users and readily understood by them.

3.2.1.4 Representatives of the people affected by the recommendations

Individuals who are likely to be affected by the intervention(s) or approach(es) under consideration in the guideline – or their representatives – bring invaluable perspectives to the guideline development process. They can help to ensure that the guideline reflects the needs of its intended beneficiaries and can be effectively implemented, and they can assist the GDG in understanding the impact of the recommendations in real life. For example, guidelines on the management of diabetes would benefit from input by people with diabetes; similarly, guidelines on human resources for health would benefit from input by labour union representatives. Although finding such individuals with the necessary background is not easy when developing global guidelines, an increasing number of groups are operating at the international level. Many countries have nongovernmental organizations whose members may be able to participate in the GDG in an individual capacity, or attend meetings as observers on behalf of their organization (see Section 3.5.3).

Involving service users in groups developing guidelines helps to ensure that:

- the questions addressed are relevant to service users;
- important aspects of the experience of illness are considered;
- critical outcomes are identified and prioritized; and
- the balance of benefits and harms of the intervention is appropriately considered when recommendations are formulated.

Certain barriers can stand in the way of service user participation in guideline development. They include:

- the lack of organized service-user groups, which makes it difficult to identify individuals able to participate in GDGs;
- the fact that an individual cannot represent the varied perspectives and experiences of all persons affected by a disease or condition; and
- the complex scientific terminology used by guideline developers and topic experts.

Experiences from organizations such as the United Kingdom's National Institute for Health and Care Excellence (NICE) suggest that service users make critical contributions to guideline development when provided with training and support (*2*, *3*).

3.2.1.5 Experts in guideline development

Ideally, at least one of the technical experts in the GDG should also have expertise in the processes and methods for developing evidence-based guidelines.

3.2.1.6 An economist

An economist can be an important contributor to a GDG if resource-related issues are at play in the formulation of recommendations. This GDG member can advise on matters of economic efficiency, such as cost–effectiveness, and on any other resource implications of the interventions under consideration. The economist can also advise on how to search for and interpret relevant economic data and the evidence on resource use. If modelling of economic data is used to inform one or more recommendations, it is essential that the GDG include one or more individuals with expertise in economic modelling or that an expert is commissioned to attend the GDG meeting.

3.2.1.7 An expert on equity, human rights and gender

Depending on the topic of the guideline, a GDG member with expertise in matters of equity, gender and human rights can contribute to the analysis and interpretation of evidence and determine how the intervention might affect certain subpopulations. For example, they could bring insights into how women and men – in all of their diversity and across the life-course, subject to different gender norms, and belonging to different income and education groups – could be affected differently by the recommendations in the guideline.

3.3 The external review group

The external review group is composed of persons interested in the subject of the guideline as well as individuals who will be affected by the recommendations (often referred to as "stakeholders"). Thus, the external review group may include technical experts, end-users, programme managers, advocacy groups and individuals affected by the condition addressed in the guideline, among other stakeholders. This group is generally established by the steering group after the GDG is identified and once the guideline's scope and

key questions have been drafted. Methods for recruitment vary. The steering group and GDG can suggest names or issue an open call for interested persons and organizations. Like the GDG, the external review group should be balanced in terms of geography and gender and should provide diverse perspectives. If important perspectives and stakeholders are missing from the GDG, these should be represented in the external review group.

Members of the external review group can be asked to participate in different stages of the guideline development process, depending on the nature of the topic and the needs of the steering group. The external review group may review the guideline's scope and key questions (in PICO format) in the early stages of the guideline development process, and the final guideline document at the end. When the external review group reviews the final guideline, its role is to identify any errors or missing data and to comment on clarity, setting-specific issues, and implications for implementation – not to change the recommendations formulated by the GDG. If external review group members have major concerns regarding one or more recommendations, the GDG should meet to discuss and address them. Review of the final guideline by the external review group is often referred to as peer review.

3.4 The systematic review team

Systematic reviews of the evidence are the basis for most types of recommendations (see Chapter 8). Because WHO staff usually lack the time to perform these reviews, they normally commission them from external contractors. These contractors should be identified very early in the guideline development process because they have expertise in the development of key questions and can help the steering group to establish a reasonable scope that conforms to the available budget and timeline.

Systematic reviews can be commissioned from any group with the necessary expertise and no financial conflicts of interest. The Cochrane Collaboration (*4*) and the Campbell Collaboration (*5*) have editorial teams whose expertise covers a broad range of topics relevant to WHO guidelines. These teams may be interested in updating an existing review or in performing a review de novo. They can be located via their organizational websites, or the GRC Secretariat can help to identify the appropriate contact person. Systematic review teams that are interested in working with WHO are listed on the GRC intranet site (*6*).

Table 3.1. Contributors and their role in guideline development

Name	Primary responsibilities	Main functions	When is the group established?	How is the group established?	Leader(s)	Essential expertise	Other characteristics	Number of members	Require DOI?
Steering group	Oversee the guideline development process	Administration; draft the scope; identify members of the GDG and external review group; oversee conduct of the systematic review; draft the final guideline	At the first step, as soon as a decision is made to develop a guideline	The WHO technical unit that initiated the guideline identifies members	Responsible technical officer	Administration; technical expertise in the guideline topic; systematic review and guideline methods	Members include representatives from all relevant WHO technical units at headquarters and in the regional offices	4–10	As dictated by WHO rules and regulations (*1*)
Guideline development group	Formulate recommendations; the general scope and content of the guideline	Provide input into the scope and key (PICO) questions; attend GDG meeting(s); formulate recommendations; critically review the final guideline document	Once the steering group has drafted the scope and a first draft of the key questions in PICO format	Selected by the steering group	Chair and perhaps a co-chair or vice-chair	Technical expertise; needs of the end-users; programme managers; personal experience with the condition or will be affected by the recommendations; methods for developing evidence-based guidelines	Gender balanced and broad geographic representation from all WHO regions	10–20, occasionally more	Yes
External review group	Provide diverse and real-world perspectives	Peer review of the draft final guideline; may provide input into the scope and key questions	Once the scope, key questions and the GDG have been established	Selected by the steering group with input from the GDG	None	As for the GDG	As for the GDG; can complement any missing diversity or perspectives in the GDG	5–20, depending on their role and the guideline topic	Yes for individuals, no for peer reviewers representing their organization

continues ...

... continued

Name	Primary responsibilities	Main functions	When is the group established?	How is the group established?	Leader(s)	Essential expertise	Other characteristics	Number of members	Require DOI?
Systematic review team	Provide a comprehensive, objective synthesis of the evidence to inform each recommendation	Provide input into the key questions; perform systematic reviews of the evidence; assess the quality of the body of evidence and develop GRADE evidence profiles	Once the steering group has drafted the scope and the key questions	Selected by the steering group	Team leader	Systematic review methods; the GRADE approach to guideline development	Experience developing high-quality systematic reviews on public health topics; ability to understand the guideline topic area	2–6 or more depending on the scope	Yes
Guideline methodologist	Oversee the process of developing recommendations based on evidence	Review GRADE evidence profiles developed by the systematic review team; attend the GDG meeting and assist the group in developing recommendations using the GRADE framework	Once the steering group has drafted the scope and the key questions	Selected by the steering group	None	Development of recommendations based on evidence; systematic review methods; GRADE	Specific topic expertise may be helpful but is not essential; ability to understand the guideline topic area	One	Yes
Observers at the GDG meeting	Observe	Ensure the transparency of the processes; engage partners; facilitate implementation	When the GDG meeting is being planned	Selected by the steering group with input from the GDG	None	Various related to the guideline topic	Variable	Variable	No

continues ...

... continued

Name	Primary responsibilities	Main functions	When is the group established?	How is the group established?	Leader(s)	Essential expertise	Other characteristics	Number of members	Require DOI?
Funder	None with respect to process of development or content of the guideline	Provide funding	After the technical unit decides to proceed with guideline development	Identified by the steering group	None	None	WHO does not accept any commercial support for guideline development	One or more depending on the scope	No (it must be determined that they do not receive any funds from commercial entities)

DOI: declaration of interests; GDG: guideline development group; GRADE: Grading of Recommendations Assessment, Development and Evaluation; PICO: population, intervention, comparator and outcome.

3.5 Other individuals and groups involved in guideline development

3.5.1 The guideline methodologist

At least one methodologist – an expert in guideline development processes and methods – should be involved in the development of WHO guidelines. The methodologist complements the technical expertise of the subject matter experts, programme managers and other members of the GDG. The guideline methodologist is an expert in systematic reviews, GRADE, and the translation of evidence into recommendations. They generally have experience formulating public health recommendations. Methodologists should be identified early in the guideline process so that they can participate in planning, scoping and the development of key questions. The methodologist also plays a critical role in GDG meetings by helping the GDG to formulate recommendations informed by the evidence in a transparent and explicit manner (see Chapter 10).

3.5.2 Consultants with additional technical expertise

Additional commissioned experts are sometimes involved in the development of a guideline. These individuals have expertise in other essential areas, such as decision analysis (modelling), economics or epidemiology. They may play a variety of roles in the guideline development process and may be invited to attend and present at GDG meetings. They do not contribute to the formulation of recommendations, however, unless they are official members of the GDG, in which case they would not have a contract with WHO (see Section 3.2).

3.5.3 Observers at the meeting of the guideline development group

Stakeholders such as nongovernmental organizations, advocacy groups, funders, target audiences, and service-users may be invited to send a member or representative to GDG meetings. Their role is to observe, although the meeting chair may ask them for information or an opinion. They do not participate in the formulation of recommendations.

3.5.4 Designated writer/editor

To ensure coherence, clarity and accuracy, one person should be responsible for drafting the guideline, incorporating comments from the GDG and peer review, and finalizing the document before publication. A clearly written guideline and a well-documented process are critical to the final clearance and usability of the guideline, so the writer needs to be involved in the planning and development stages, attend GDG meetings, and work closely with the steering group and the GDG chair. The writer, who may be a WHO staff member or can be commissioned, must be familiar with the *WHO style guide (7)* and should consult with the equity, human rights and gender experts at WHO if they have questions about how persons should be referred to and presented in the guideline.

3.6 Funders of the guideline

The funder(s) of a guideline can play no role in the guideline development process and should not influence the recommendations. Private funders, including industry or foundations, may wish to observe GDG meetings, but they must not contribute in any way to the systematic review and evidence appraisal process and, most particularly, to the formulation of recommendations. Governmental and intergovernmental agencies (e.g. the Centers for Disease Control and Prevention [United States of America] or the Joint United Nations Programme on HIV/AIDS [UNAIDS]) may fund and/or partner in guideline development. Employees of such funders or partners cannot be members of the GDG but should be included in the steering group. An employee of a governmental or United Nations organization who provides specified voluntary funding to support development of the guideline should not be a member of the GDG. However, if the membership of such a person is considered essential in light of the expertise required within the GDG, then this dual relationship must be identified as a conflict of interest and managed accordingly.

3.7 Identifying and managing conflicts of interest

Chapter 6 and the website of the Office of Compliance, Risk Management and Ethics (*8*) explain the steps followed to identify, manage and report conflicts of interest involving GDG members and commissioned experts. Declara-

tions of interests (DOI) are collected and reviewed before making appointments to the GDG. Any subsequent changes in GDG members' declaration of interests must be reported to the steering group. Potential candidates for membership who have major conflicts of interest, be they financial or non-financial, cannot be appointed to the GDG. Minor conflicts of interest can be managed at the individual level (e.g. by restricting participation in parts of the GDG meeting) or at the group level (e.g. by balancing GDG members' perspectives and experiences). See Chapter 6 for details.

3.8 Managing an effective guideline development group meeting

Along with the GDG chair, the steering group is responsible for seeing to it that the GDG works effectively and efficiently to develop recommendations that meet WHO standards.

3.8.1 The role of the chair and vice-chair

During meetings, the chair must ensure that GDG members can present their viewpoints and that all relevant issues are discussed in a respectful and efficient manner. In addition, the chair should keep the group focused on the agenda; reflect on and summarize the opinions of GDG members; raise issues that could inform the decision process; and manage the group so as to achieve consensus. The chair and vice-chair should not impose their own opinions on the group, however they may put forth their views in a purely personal capacity.

3.8.2 Decision-making in the guideline development group

The steering group must formulate a plan delineating how decisions will be made by the GDG. This plan must be formulated and agreed upon by the group's chair and vice-chair before the first GDG meeting. It should be presented to GDG members at the beginning of this first meeting, with the opportunity for questions and discussion. GDGs sponsored by WHO should generally make their recommendations through consensus, which can be achieved by various means. Acceptable approaches for group decision-making must be defined and made explicit before any recommenda-

tions are formulated, and a plan as to how to proceed if consensus cannot be achieved must also be in place. Such decisions cannot be made when the GDG encounters serious internal dissention (*6*).

3.9 Planning an effective meeting

Most guidelines involve at least one face-to-face meeting of the GDG to formulate recommendations. Some groups hold an initial meeting to finalize a guideline's scope and key questions (see Chapter 7). GDG meetings generally cover a lot of material and in them a number of important decisions are made within a short time. The steering group must ensure that all in attendance understand their role and the expected outputs of the meeting by providing GDG members with clear information about how the meeting will be run and about the tasks and roles of those present.

3.9.1 Goals of the meeting

The steering group should consider the following questions when planning a meeting of the GDG:

- What is the purpose of the meeting?
- What are the specific objectives for the meeting?
- What decisions do not require deliberation by the GDG (and can therefore be made by the steering group)?
- What specific decisions and outputs are expected from the meeting?

3.9.2 Preparing for the meeting

- How are meeting participants expected to prepare before the meeting?
- What materials need to be sent to GDG members ahead of time and when should they be sent to allow adequate time for review?
- Have all DOI forms been collected and assessed and have appropriate management plans been agreed upon with the director of the responsible technical officer's technical unit?

3.9.3 Roles and process at the meeting

- What roles are played by the steering group, GDG members, the methodologist, the systematic review team's representative and observers?
- How will declarations of interest and conflicts of interest be presented? What will be the level of detail? How will new disclosures be assessed and managed at the GDG meeting?

If the purpose of the GDG meeting is to finalize the scope and the key questions for guideline development:

- Who will draft the scope and key questions and how will these questions be identified?
- Will GDG members have the opportunity to review the questions before the meeting?

If the purpose of the meeting is to develop recommendations:

- Will a summary of the evidence be presented? If so, by whom and in what format?
- Will the GRADE evidence profiles (summaries of the quality of the evidence for each outcome, see Chapter 9), be finalized before the meeting, or will a draft be discussed and revisions made during the meeting?
- Will the steering group draft recommendations before the GDG meeting? How detailed or explicit will those draft recommendations be?
- What process will the GDG follow to finalize its recommendations?
- Will a standardized template be used to translate evidence into recommendations (see Chapter 10)? If so, will these be partially completed before the GDG meeting. If so, by whom?
- Who will record the meeting deliberations? Will a meeting report be prepared, separate from the guideline document? (If the meeting report contains recommendations, it must be submitted to the GRC with the final guideline document, and the report cannot be publicly released before GRC approval.)

3.9.4 Follow-up after the meeting

- What follow-up will take place with meeting participants?
- If recommendations are not finalized at the GDG meeting, how will they be finalized after the meeting?

3.10 References

1. WHO eManual /III - Human Resources/III.1 Duties, Obligations and Privileges/III.1.2 Declaration of Interests [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://emanual.who.int/p03/s01/Pages/DeclarationofInterests.aspx, accessed 27 November 2014).
2. National Institute for Health and Care Excellence (NICE) [website]. London: National Institute for Health and Care Excellence; 2014 (http://www.nice.org.uk/, accessed 12 November 2014).
3. NICE patient and public involvement policy. London: National Institute for Health and Care Excellence; 2014 (http://www.nice.org.uk/about/nice-communities/public-involvement/patient-and-public-involvement-policy, accessed 14 November 2014).
4. The Cochrane Collaboration [website]. Oxford: The Cochrane Collaboration; 2014 (http//www.cochrane.org, accessed 12 November 2014).
5. The Campbell Collaboration [website]. Oslo: The Campbell Collaboration; 2014 (http://www.campbellcollaboration.org/, accessed 12 November 2014).
6. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/, accessed 12 November 2014).
7. WHO style guide, second edition [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2013 (http://intranet.who.int/homes/whp/documents/sg13_web_v4%20pdf%20-%20adobe%20reader.pdf, accessed 12 November 2014).
8. About CRE (Office of Compliance, Risk Management and Ethics) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/cre/about/, accessed 12 November 2014).

4. Preparing the planning proposal

4.1 What is a guideline planning proposal?

A guideline planning proposal is a document that outlines the rationale and scope of a proposed guideline and the methods and resources involved in developing it. The proposal includes many of the components of a protocol for a systematic review and guideline, plus additional information on the groups contributing to guideline development and the administrative steps required for WHO guidelines. The responsible technical officer must submit a planning proposal to the GRC for each guideline to be developed and the GRC's approval is required for guideline development to proceed to completion.

4.2 What is the purpose of the planning proposal?

To embark on any project or produce any document, including guidelines, its developers need an explicit, thoughtful and comprehensive plan or roadmap. The planning proposal serves several important purposes.

First, a planning proposal helps to ensure that a high-quality and credible guideline is produced. The GRC and the Secretariat can provide constructive feedback, which helps to ensure that the final guideline will meet WHO standards. In addition, many steps in guideline development must be performed correctly early in the process, since they cannot generally be remedied later, and a planning proposal facilitates the use of appropriate approaches. For example, individuals with diverse perspectives and expertise are needed to delineate the scope and priority questions: it is not possible to incorporate missing viewpoints late in the guideline development process.

Second, the planning proposal codifies the relationship between the proposed guideline and existing and planned WHO guidelines and products. One of the important roles of the GRC is to ensure that recommendations are consistent across WHO guidelines, and that any overlap is purposeful. This helps to avoid duplication of efforts and to identify opportunities for the appropriate integration of related guidance across technical units at WHO.

Third, a carefully constructed planning proposal makes it more likely that the guideline development process will be smooth and efficient. Resources will closely match the proposed scope; relevant expertise will be identified and engaged early and throughout the development process; and potential difficulties, both procedural and technical, can be anticipated and planned for.

Fourth, the planning proposal promotes accountability. When the scope, methods and timeline of the guideline are laid out and approved by the steering group a priori, each member of the steering group and the director of the technical unit leading the process are more accountable for the final guideline. Adherence to the original timeline may also be improved.

Finally, and perhaps most importantly, the planning proposal promotes the integrity of the development process and of the final guideline. By deciding in advance on the steps and methods to be followed, arbitrariness in decision-making can be minimized. In addition, thoughtful development and detailed documentation of the planned methods and key questions and outcomes may reduce the risk of bias in the evidence synthesis and in the formulation of recommendations. Changes in methods mid-stream based on the findings of the systematic review or the interests of dominant GDG members must be avoided.

4.3 Who develops the planning proposal?

The responsible technical officer in the technical unit leading the guideline development process is responsible for preparing the planning proposal, in close collaboration with other members of the steering group. Broad input must be obtained, however, to ensure a useful product. The GDG, or at least a subset of that group including the chair, should provide input into the guideline's scope and key questions in PICO format. The systematic review team and the guideline methodologist should also provide input into the key questions before they are finalized. It may also be useful to have members of the external review group provide input.

If resources allow, the GDG may meet in person before the planning proposal is finalized to define the scope and discuss the methods to be used. Otherwise, a virtual meeting can be held, or discussions can occur via email.

4.4 When should the planning proposal be developed and finalized?

The planning proposal is the first major product of the guideline development process, and its preparation starts as soon as the technical unit decides to develop a guideline. The responsible technical officer should strive to develop and submit the proposal to the GRC for review as quickly as possible, but with sufficient consideration and preparation of the many elements in the proposal. The most difficult and time-consuming part of the planning proposal is formulating the key questions (in PICO format). These questions need to be well developed in the planning proposal, and the proposal is submitted to the GRC only when this has been completed.

4.5 Preparing the planning proposal

Table 4.1 provides a list of all the topics that should be considered when writing a planning proposal, and the GRC intranet site (*1*) provides detailed instructions for reporting on each topic. The purpose of these tables is not to provide instructions for each and every step in the guideline development process, but rather to indicate what should be *reported* in the planning proposal; specific instructions on how to develop a guideline are contained in this handbook.

The planning proposal is not a lengthy document, but it must contain all items in the instructions described below, presented in a succinct and cogent manner. Detailed background explanations and materials are not necessary. The planning proposal and annexes should be submitted via email directly to the GRC Secretariat before the deadline for submission of documents for the next GRC meeting. A template for a planning proposal and a checklist derived from this template are available on the GRC intranet site (*1*). The responsible technical officer must complete and submit the checklist with the planning proposal for review by the GRC.

Planning proposals will only be reviewed by the GRC once they have been approved in the WHO electronic clearance system (ePub) by the director of the responsible technical unit.

Table 4.1. Essential components of the guideline planning proposal

Topic	Subtopic	Responsible person or group/collaborators[a]
1. Background and context	1.1 Disease burden and distribution across subgroups	RTO/None
	1.2 Background on the intervention or topic	RTO/SG
	1.3 History of this guideline	RTO/None
2. Rationale	2.1 Why this guideline is needed	RTO/SG
3. Target audience	3.1 The end-user(s) of the guideline	SG/GDG
4. Persons affected by the recommendations	4.1 Service users, patients, populations affected by the recommendations in the guideline	SG/GDG
5. Related guidelines	5.1 WHO guidelines related to the current guideline	RTO/SG
	5.2 Relevant guidelines produced by external organizations	RTO/SG
6. Goal and objectives	6.1 Goal of this guideline	SG/None
	6.2 Specific objectives	SG/None
7. Contributors to guideline development	7.1 Steering group	RTO/TU
	7.2 Guideline development group	SG/None
	7.3 Systematic review team	RTO/SG
	7.4 External review group	SG/GDG
	7.5 Guideline methodologist(s)	RTO/SG
	7.6 Stakeholders, including service users	SG/GDG
	7.7 External partners	SG/GDG
8. Management of the guideline development group	8.1 Selection of the chair, vice-chair	SG/None
	8.2 Group processes and decision-making	SG/None
9. Conflict of interest	9.1 Collecting disclosures of interest	RTO/None
	9.2 Assessing disclosures of interest	RTO, director TU/SG, CRE
	9.3 Managing conflicts of interest	RTO, director TU/SG, CRE
	9.4 Confidentiality agreement	RTO/None
10. Formulating key questions	10.1 Background questions	SG/GDG
	10.2 Foreground (key/PICO) questions	SG/GDG
	10.3 Important and critical outcomes	SG/GDG
11. Systematic review methods	11.1 Need for new systematic review	SG/SRT, GM
	11.2 Study inclusion and exclusion criteria	SG/GDG, SRT, GM
	11.3 Evidence identification and retrieval	RTO/information scientist, SRT, GM
	11.4 Quality assessment of the primary studies	SRT/None
	11.5 Synthesis of the body of evidence for each outcome	SRT/None
	11.6 Quality assessment of the body of evidence for each outcome	SRT/RTO, GM
12. Evidence to recommendations	12.1 Use of the GRADE framework	SG/SRT, GM
	12.2 Factors to consider, e.g. values and preferences; resource use; equity, human rights and gender	SG/SRT, GM, GDG
	12.3 Tools for formulating recommendations	RTO/GM

continues ...

... continued

Topic	Subtopic	Responsible person or group/collaborators[a]
13. Writing the guideline document	13.1 Writer, editor	RTO/None
14. Peer review	14.1 Process	RTO/SG
15. Logistics and resources	15.1 Funding	RTO/SG
	15.2 Budget	RTO/TU
	15.3 Timeline	RTO/SG, SRT, GM
	15.4 Other logistics	RTO/None
16. Implementation and evaluation	16.1 Publication formats	SG/None
	16.2 Derivative products	SG/GDG
	16.3 Implementation	SG/GDG
	16.4 Adaptation	SG/GDG
	16.5 Evaluation	SG/GDG
17. Updating	17.1 Plans for when and how to update	SG/None
	17.2 Strategies for identifying new information	RTO/SG

CRE: Office of Compliance, Risk Management and Ethics; GDG: guideline development group; GM: guideline methodologist; GRADE: Grading of Recommendations Assessment, Development and Evaluation; PICO: population, intervention, comparator and outcome; RTO: responsible technical officer; SG: steering group; SRT: systematic review team; TU: technical unit.

[a] This column indicates who or which group is primarily responsible for defining the methods and approach to the topic and who else also collaborates in determining the approach.

4.6 What if there are major changes in scope or approach after the GRC has approved the planning proposal?

It is not unusual for changes to occur after the planning proposal is developed and approved by the GRC. New or unanticipated evidence, or new ideas from the steering group, the GDG or other interested parties, can engender discussions and the decision to change course. Such changes in planning proposals generally fall into one of the following categories:

- **A major increase or decrease in scope** – Example: new key questions might be added or, less commonly, questions might be removed.
- **A change in the packaging of the guideline** – Example: a subset of the recommendations is published before the entire set, due to an urgent request from Member States.
- **A major change in the proposed methods** – Example: the planning proposal might indicate that long-term follow-up data from comparative cohort studies will be used to inform cost–effectiveness analyses,

but after exploration of the available data, modelling will be needed to examine economic efficiency.

- **A change in a key question without a major change in scope** – Example: after further discussion and input from the guideline methodologist, a key question might be deemed poorly formulated and therefore unanswerable, and the question is therefore reformulated.
- **Minor changes in process or methods** – Example: the original timeline is extended, new members are added to the GDG, or the search strategies or bibliographic databases to be examined are changed.

Major changes to the methods and/or scope presented in the planning proposal (the first three categories listed) need to be reviewed by the GRC Secretariat and assessed, in some cases, by the GRC. On the other hand, minor changes (such as the fourth and fifth categories listed) do not require discussion with the GRC or the Secretariat. If you are uncertain as to whether a change is major or minor, please discuss it with the Secretariat. Proposed major changes, together with their rationale and potential impact on the integrity of the guideline development process, should be documented as an addendum to the original planning proposals and submitted to the Secretariat for review and assessment.

4.7 For how long is an approved planning proposal in effect?

The GRC approves planning proposals for a period of six months beyond the publication date indicated in the planning proposal. If the final guideline is not submitted to the GRC before the end of the six-month period, the responsible technical officer should discuss the delay with the GRC Secretariat. If the extended timeline may compromise the original methods and bibliographic database searches, a revised planning proposal may have to be submitted to the GRC. Significant delays in the production of the final guideline should be discussed with the Secretariat.

4.8 References

1. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/, accessed 14 November 2014).

5. Incorporating equity, human rights, gender and social determinants into guidelines

5.1 Reflecting WHO's values in its guidelines

The aim of all WHO's work is to improve population health and decrease health inequities. Sustained improvements to physical, mental and social well-being require actions in which careful attention must be paid to equity, human rights principles, gender and other social determinants of health. Attention to these areas enables WHO to more effectively deliver on its commitments in its Twelfth General Programme of Work *(1)*. This chapter describes how these important considerations can be integrated into each step of the guideline development process and suggests eight entry points for doing so.

5.2 Background

The WHO Constitution (*2*) states that:

> "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity."
>
> and that
>
> "The enjoyment of the highest attainable standard of health is one of the fundamental rights of every human being without distinction of race, religion, political belief, economic or social condition."

Guideline developers need to consider how their guidelines will contribute to the realization of the right to health, as we shall henceforth refer to the concept embodied in the second of these statements. In so doing, they would be wise to draw on the fields dealing with equity, human rights, gender and other social determinants of health.

5.2.1 Equity

The progressive realization of the right to health involves a concerted and sustained effort to improve health across all populations and to reduce inequities in the distribution of health. Inequities are inequalities that are judged to be unfair (i.e. both unacceptable and avoidable) *(3)*. Several measures of health inequalities have been developed. They generally rely on health indicators and equity stratifiers obtained from a variety of sources, including population-based sources, such as household surveys, censuses and vital registration systems; institutional sources, including resource records, service records and patient records; trial or project reports; and surveillance systems (which comprise both population-based data and institutional data). To choose the most appropriate set of measures in a given situation, one needs to understand the distinctions among these measures and their implications, such as the use of absolute versus relative measures of inequity.

5.2.2 Human rights

To be fulfilled, the right to the highest attainable standard of health requires a set of social arrangements – norms, institutions, laws, and enabling environments – that are conducive to the enjoyment of this right. Realization of the right to health is closely related to that of other human rights, including the right to food, housing, work, education and non-discrimination; equality; access to information; and participation (*4*). Specifically, human rights dimensions are integrated into the provision of health services by paying due attention to non-discrimination; to the availability, accessibility, acceptability and quality of information and services; to informed decision-making; to privacy and confidentiality; and to participation and accountability (*5*).

The right to health comprises both freedoms and entitlements. Freedoms include the right to control one's health and body (e.g. sexual and reproductive rights) and to be free from interference (e.g. free from torture and from non-consensual medical treatment and experimentation) (*4*). Entitlements include the right to a system of health protection that gives everyone an equal opportunity to enjoy the highest attainable level of health (*4*).

5.2.3 Social determinants of health and health inequities

The social determinants of health are the conditions in which people grow, live, work and age (*6*). Health inequities are differences in health that are unfair, avoidable and remediable because they depend largely on the social determinants of health, which are amenable to improvement. WHO's normative work, including guideline development, must thus take these determinants into consideration (*7*).

5.2.4 Gender

"Gender" refers to the roles, behaviours, activities and attributes that a given society considers appropriate for men and women and boys and girls. Thus, gender is socially constructed and in many contexts it is the strongest social determinant of health. "Sex", on the other hand, refers to the biological and physiological characteristics that define men and women. That is, "male" and "female" are sex categories, while "masculine" and "feminine" are gender categories (*8*). It is important to be sensitive to the different gender identities that do not fit into the binary models of masculine versus feminine.

Gender norms, roles and relations influence people's risk of contracting disease and their susceptibility to different health conditions. They also have a bearing on people's access to and uptake of health services and on the health outcomes they experience throughout the life-course. It is important to consider how WHO guidelines can contribute to gender equality.

5.2.5 Social determinants of health

The effects of the social determinants of health are manifest in the form of differences in health outcomes across subpopulations and in a gradient in health status in accordance with income and education. These social determinants contribute to inequity in health because they are responsible for differences in physical and social environments; vulnerability; access to services and products; ability to benefit from services; and ability to cope with ill health and disability. These differentials can all be reduced through the interventions and actions of one or several health programmes and services, or through intersectoral collaboration (*7*). WHO's normative work, including guideline development, must thus take these determinants into consideration.

Box 5.1. Eight entry points for integrating equity, human rights and gender into WHO guidelines

1. **Analysis of evidence on inequities and their determinants.** As appropriate to the health topic of a guideline and as there are reasons to expect differences in the impact of a recommendation across subpopulations, you might examine the stratifiers used in primary data sources and analyse disaggregated data; review additional studies (including qualitative studies) conducted in subpopulations experiencing different health outcomes; explore the related evidence on access barriers and the unintentional consequences of service usage; explore the determinants that require intersectoral action; and review reports on the fulfilment of the right to health that are issued by human rights bodies. PROGRESS-Plus provides a useful framework for considering stratifiers (*9*) and WHO has developed guidance on health inequality monitoring that guideline developers can draw from (*10*). Guidance also exists for reporting on systematic reviews with a focus on health equity (*11*).
2. **Analysis of laws, policies, standards, protocols and guidelines.** Analysing the coherence and implementation of these with respect to equity, human rights, gender and other social determinants is essential to understanding the effect that societies and services can have on the realization of the right to health for the health topic in question. When guideline developers review the evidence base, they can consider how equity, human rights, gender and social determinants are affected by the existing policy, legal, normative, programmatic and monitoring and evaluation frameworks in national contexts, and how a new guideline might serve to better address these issues.
3. **Analysis of the social determinants at play.** To effectively contribute to the reduction of health inequities, guideline developers can identify and take into account those social determinants most relevant for each guideline, including those that influence the following (*7*):
   - *differential exposure to the physical environment*, e.g. adverse workplaces and community settings, poor infrastructures, unhealthy and harmful consumables, etc.;
   - *differential exposure to the social environment,* e.g. social norms that can undermine health, gender expectations and repression, ethnic and racial discrimination, unregulated marketing, etc.;
   - *differential community and individual vulnerability*, e.g. poverty and unemployment, family and community dysfunction, poor knowledge, low levels of health literacy and care-seeking, alcohol abuse, malnutrition, etc.;
   - *differential access to health products and services*, e.g. skewed availability, financial barriers, products and services with poor acceptability, etc.;
   - *differential benefit from health services*, e.g. poor quality health services, discriminatory treatment and care, biased referral systems, services insensitive to needs, limited patient–provider interaction, poor adherence to advice and recommended treatments, etc.; and

continues ...

... continued

- *differential consequences of illness and disability,* e.g. loss of income, impoverishment/catastrophic health expenditure, stigmatization or other forms of discrimination.

4. **Analysis and allocation of resources.** Fulfilment of the right to health for all is not just a matter of the absolute level of resources. It also has to do with how resources are prioritized and allocated to reduce inequities. Guideline developers should reflect on ways that an eventual skewed distribution of attention, resources and efforts might influence the implementation of the final guidelines. They should also include guidance on how to plan or prioritize services and interventions related to guideline roll-out in ways that reduce rather than increase inequity.
5. **Analysis and strategies to address gender issues.** Guideline developers can use certain tools, such as the Gender Analysis Matrix and the Gender Responsive Assessment Scale (*12*), to examine the extent to which the interventions covered in the guideline address gender issues. Guidelines should promote gender-responsiveness in all processes and in the organization of programmes and services.
6. **Analysis and provision of means for civil society and individuals to participate in decision-making for health.** The right to health is best protected when individuals and concerned populations, including those marginalized or otherwise disadvantaged, are actively involved in decision-making on policy and their individual health. Guideline developers can consider how to make the processes and services they recommend more participatory and how to involve the relevant population groups (i.e. so that participation is not the exclusive privilege of the affluent and powerful).
7. **Transparency, accountability and keeping health and other sectoral managers and services to task.** These are essential for reducing health inequities, together with safe mechanisms for reporting and addressing complaints whenever the right to health is threatened or violated, individually or collectively. As relevant, guideline developers can try to proactively build in transparency, define accountability and recommend mechanisms for enforcement of the right to health.
8. **Ensuring gender balance and sensitivity to other equity considerations in the guideline development and production process**. Steering groups should try to ensure sex parity and appropriate gender representation, as well as balance in terms of regional and country income levels among members of the guideline development group and the external review group.

5.3 Eight entry points for integrating equity, human rights, gender and the social determinants of health into WHO guidelines

WHO's normative work, including its recommendations and guidelines, must integrate equity, human rights, gender and the social determinants of health. Eight entry points (Box 5.1) provide a practical approach for integrating these aspects into WHO guidelines. The relevance of specific strategies varies across guideline topic areas, but all strategies should be systematically considered in the guideline development process.

5.4 Incorporating equity, human rights, gender and social determinants into each step in guideline development

Equity, human rights, gender and the social determinants of health must be considered during all three phases of guideline development (Table 1.1) – planning, development, and publishing and updating – and during most steps within each phase.

5.4.1 Phase 1: planning

5.4.1.1 Request for guidance on a topic (Chapter 1)

In their request for guidance, Member States or other stakeholders might explicitly ask that equity, human rights, gender and the social determinants of health be addressed during guideline development. Whether or not this request is made explicit, examining any health issue through the lens of equity, human rights, gender and the influence of social determinants is important. Doing so will help to better understand the needs and gaps to be addressed and may lead to interventions that are more effective in the longer term and that will evoke a feeling of "ownership" in the targeted group or community. This will be particularly relevant when the request for a guideline relates to the accomplishment of global and national public health goals, or if it responds to a specific recommendation emerging from the monitoring mechanisms of the international human rights framework ("treaty bodies") (*13*).

5.4.1.2 Planning to develop a guideline (Chapter 2)

Equity, human rights, gender and the social determinants of health need to be taken into account from the moment the decision is made to begin the guideline development process. The planned achievements should focus not only on the average level of health, but also on how health is distributed within populations and across groups. The idea is to ensure that those of lower social position and with greater needs can benefit more than more advantaged persons. Through this progressive realization of the right to health, a levelling-up of health status is achieved across the population.

Early in the preparatory stage of the guideline development process, it is important to ensure the use of appropriate stratifiers in the systematic reviews that are carried out. PROGRESS is a useful framework for stratifiers. It includes place of residence, race or ethnicity, occupation, gender, religion, education, socioeconomic position, and social capital (*14*, *15*). This list is not exhaustive, however, and PROGRESS-Plus (*10*) adds age, disability and sexual orientation. In addition, it is worth considering other variables that might constitute potential barriers to the desired outcomes, such as legal and policy frameworks that could marginalize or exclude certain populations.

5.4.1.3 Establishing groups involved in developing guidelines (Chapter 3)

The composition of the various groups involved in developing guidelines and the views of their members influence the contents and perspectives of the final guideline, especially the extent to which WHO's values, and the needs of the populations whose health WHO seeks to protect, are reflected in the recommendations. GDGs should include individuals who understand how to take equity, human rights, gender and social determinants into account in efforts to promote better health and who have expertise in doing so. In addition, in line with the human rights principle of greater participation, it is important to ensure that the views and voices of groups that are marginalized or otherwise disadvantaged are also reflected in the GDGs as relevant. It is particularly important to ensure that GDGs are chaired by people who understand and support WHO's emphasis on equity, human rights, gender and the social determinants of health.

5.4.1.4 Declaration and management of interests (Chapter 6)

In reflecting WHO's values, mitigating conflicts of interest in the intellectual and ideological realms is just as important as doing so in the financial realm. Explicit attention should be given to conflicts of interest that can lead to a weakened stance on equity, human rights, gender and social determinants in the final guideline.

5.4.1.5 Formulating questions and selecting outcomes (Chapter 7)

Careful selection of key questions is pivotal for integrating equity, human rights, gender and social determinants into the guideline development process and the content of the guideline. Background questions should encompass both the average effects resulting from an intervention and the distribution of effects across subpopulations. These questions should also address specific human rights and other issues related to laws, policies, standards, protocols and guidelines (see also Box 5.1, item 2). Key (foreground) questions concerning the effectiveness of interventions must take into account the potential for differences in uptake and benefits as a function of social position. In applying PICO to develop key questions, the following must be considered: under P (population), social position and other PROGRESS-Plus stratifiers; under I (intervention), ways to address differentials in exposure, vulnerability, access, benefit and consequences (see Box 5.1, item 3); under C (comparator), the potential effects of alternative choices of action on equity and on people's ability to claim their right to health; and under O (outcome), the desired effects of the recommendation in ensuring the right to health.

5.4.2 Phase 2: development

5.4.2.1 Evidence retrieval and synthesis (Chapter 8)

Existing systematic reviews seldom take equity, human rights, gender and social determinants into account, and a new review or an update of an existing review may be needed to analyse information about equity. Fortunately, the volume of primary research in this field is increasing and the methods and tools for systematic reviews of evidence on equity and for combining qualitative and quantitative research findings are rapidly evolving. Looking beyond averages and disaggregating data sometimes calls for an adjustment of traditional data collection practices and statistical analyses. Furthermore, some research questions pertaining to the realization of the right to health and to public health and policy do not lend themselves to quantitative research methods, and qualitative research and innovative study designs may be needed. The steering group should carefully consider – in consultation with the GDG – when a systematic review focused on important subpopulations, or on the impact of laws and policies on health (see Box 5.1, item 2), may inform the recommendations. The group must assess the probability that the intervention under consideration will have a bearing on equity, human rights, and gender equality, and the likelihood that a systematic review will provide useful data.

The Campbell and Cochrane Equity Methods Group (*16*) encourages authors of systematic reviews to explicitly describe the potential effect of interventions not only on the population as a whole, but across the social gradient (including disadvantaged and marginalized subpopulations) and by sex. Relevant tools include the Equity Checklist for Systematic Review Authors (*17*) and PRISMA-Equity 2012 reporting guidelines (*11*).

5.4.2.2 Evidence assessment (Chapter 9)

Systematic assessment of the evidence on the benefits and harms of interventions needs to be expanded to include attention to the quality of the evidence on equity, human rights, gender and social determinants. Unfortunately, the evidence on important subpopulations and the effects of an intervention on equity will often be sparse and any available evidence may be limited in statistical power to demonstrate a significant effect. Thus the quality of the body of evidence will have to be assessed without being certain as to whether an intervention does, or does not have an important differential effect across subpopulations. If a differential effect is strongly suspected based on plausible biological mechanisms, social theory, or indirect evidence, the quality of the evidence applicable to the subpopulations under consideration may be rated down because of indirectness when applying the GRADE framework (see Chapter 9).

5.4.2.3 Developing recommendations (Chapter 10)

The effect of an intervention on equity is one of the factors that determine the strength of a recommendation: if an intervention is likely to reduce health inequities, a strong recommendation may be warranted, provided it is justifiable in light of other factors. On the other hand, if equity is a key consideration and evidence on the intervention's potential effects on equity is sparse, a conditional recommendation may be appropriate. In the latter case particularly, the guideline should outline the key gaps in knowledge and provide an agenda for future research. Evaluation and monitoring of the impact of recommendations that potentially affect inequities are also critically important and should be articulated in the guideline document.

5.4.3 Phase 3: publishing and updating

5.4.3.1 Producing and publishing the guideline (Chapter 12)

The final guideline and in its derivative products must convey clear messages with respect to equity, human rights, gender and social determinants.

For example, the process of developing a gender-sensitive guideline starts before the first words are put on paper and only ends when the last illustration has been chosen. Throughout the process careful attention should be paid to the use of language, the selection of case examples, the expressions used to refer to population groups, the choice of photographs, etc. Setting the ground rules up front is much easier than trying to modify a document that is ready for layout. In translations of guidelines, gender-sensitive and non-stigmatizing language should always be employed.

5.4.3.2 Adaptation, implementation and evaluation (Chapter 13)

When developing guidelines, clear references to health equity, the social determinants of health, gender equality and relevant international human rights standards and principles, should be included in the sections on implementation, monitoring and evaluation. The interventions recommended in a guideline pose the risk of increasing rather than reducing inequities, depending on how they are implemented. For example, marginalized population groups might be ignored or strategies that make them the last to benefit might be used. Important gender differences might not be considered or unintended effects on certain subpopulations or population strata might occur. Such situations arise because decisions governing implementation are not explicitly related to equity. For instance, vaccination coverage targets might be achieved without benefit to the poorest and most disadvantaged groups. WHO staff have a responsibility to monitor and evaluate how recommendations affect equity, human rights, gender equality and the advancement of the highest attainable standard of health, and how health services and programmes implement the recommended interventions.

5.5 References

1. Twelfth General Programme of Work: not merely the absence of disease. Geneva: World Health Organization; 2014 (http://www.who.int/about/resources_planning/twelfth-gpw/en/, accessed 12 November 2014).
2. Constitution of the World Health Organization. Geneva: World Health Organization; 1946 (http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf, accessed 12 November 2014).
3. Whitehead M. The concepts and principles of equity and health. Int J Health Serv. 1992;22(3):429–45. doi: http://dx.doi.org/10.2190/986L-LHQ6-2VTE-YRRN PMID: 1644507
4. 25 Questions & Answers on Health & Human Rights. Geneva: World Health Organization; 2002. (http://whqlibdoc.who.int/hq/2002/a76549.pdf, accessed 12 November 2014).
5. Ensuring human rights in the provision of contraceptive information and services: guidance and recommendations. Geneva: World Health Organization; 2014 (http://www.who.int/reproductivehealth/publications/family_planning/human-rights-contraception/en/, accessed 12 November 2014).

6. Closing the gap in a generation. Geneva: World Health Organization; 2008 (http://whqlibdoc.who.int/hq/2008/WHO_IER_CSDH_08.1_eng.pdf, accessed 12 November 2014).
7. Blas E, Kurup AS, editors. Equity, social determinants and public health programmes. Geneva: World Health Organization; 2010 (http://whqlibdoc.who.int/publications/2010/9789241563970_eng.pdf, accessed 12 November 2014).
8. Gender, women and health [website]. Geneva: World Health Organization; 2014 (http://www.who.int/gender/whatisgender/en/, accessed 25 November 2014)
9. Handbook on health inequality monitoring with a special focus on low- and middle-income countries. Geneva: World Health Organization; 2013 (http://www.who.int/iris/handle/10665/85345, accessed 12 November 2014).
10. Welch V, Tugwell P, Petticrew M, de Montigny J, Ueffing E, Kristjansson B, et al. How effects on health equity are assessed in systematic reviews of interventions. Cochrane Database Syst Rev. 2010;12(12):MR000028. PMID: 21154402
11. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al.; PRISMA-Equity Bellagio group. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. PLoS Med. 2012;9(10):e1001333. doi: http://dx.doi.org/10.1371/journal.pmed.1001333 PMID: 23222917
12. Gender mainstreaming for health managers – a practical approach. Geneva: World Health Organization; 2011 (http://whqlibdoc.who.int/publications/2011/9789241501071_eng.pdf, accessed 14 November 2014).
13. United Nations' Human Rights [website]. New York: United Nations; 2014 (http://www.un.org/en/rights/index.shtml, accessed 14 November 2014).
14. Evans T, Brown H. Road traffic crashes: operationalizing equity in the context of health sector reform. Inj Control Saf Promot. 2003;10(1-2):11–2. doi: http://dx.doi.org/10.1076/icsp.10.1.11.14117 PMID: 12772480
15. Armstrong RWE, Doyle J, editors. Chapter 21: Reviews in public health and health promotion. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions, version 5.1.0 [updated March 2011]. Oxford: The Cochrane Collaboration; 2011 (http://www.cochrane-handbook.org, accessed 12 November 2014).
16. Campbell & Cochrane Equity Methods Group [website]. Ottawa: The Campbell and the Cochrane Collaborations; 2014 (http://equity.cochrane.org/, accessed 12 November 2014).
17. Ueffing E, Tugwell P, Welch V, Petticrew M, Kristjansson E for the Cochrane Health Equity Field. 2009. C1, C2 Equity Checklist for Systematic Review Authors. Version 2009–05–28. (http://www.equity.cochrane.org/files/equitychecklist.pdf, accessed 12 November 2014).

6. Declaration and management of interests

The declaration and management of conflicts of interest is essential to the development of unbiased and credible recommendations and guidelines. This chapter defines the basic concepts, describes the principles involved, and outlines how WHO staff who develop guidelines can collect declarations of interest and assess and manage any conflicts.

6.1 What is a conflict of interest?

A conflict of interest is an important potential source of bias and diminished credibility in the development of WHO guidelines. "A conflict of interest is a set of circumstances that creates a risk that professional judgment or actions regarding a primary interest will be unduly influenced by a secondary interest" (*1*). The Declaration of Interests form at WHO (*2*) contains a similar definition: [A conflict of interest is] "any interest declared by an expert that may affect or reasonably be perceived to affect the expert's objectivity and independence in providing advice to WHO."

As noted, "primary interest", "secondary interest" and "undue influence" are key terms employed in the definition of a conflict of interest. In WHO guidelines, the primary interest is to serve WHO's Member States by producing recommendations that improve the health and well-being of populations, globally or in specific areas or countries. In other contexts, the primary interest might be, for instance, to protect the integrity of research, the welfare of patients, the objectivity of educational materials, or the equitable allocation of resources.

Secondary interests are all interests other than the primary one(s) that could be affected or potentially affected by a recommendation or guideline. All individuals and organizations involved in developing a guideline have secondary interests, which in most cases are legitimate and appropriate in their own right. Such interests include, for example, a technical expert's desire to publish or obtain funding for his or her research, or a stakeholder's desire to advocate for a disease or condition.

Problems arise, however, when the primary and secondary interests of the parties involved in guideline development are in conflict and, as a result, an individual's secondary interests put them at risk of making decisions that are detrimental to the primary interest. In other words, a conflict of interest exists when the primary and secondary interests are not aligned and there is a chance that, because of this, the person will make decisions that undermine or subvert the primary interest.

Note that a conflict of interests exists whether or not the secondary interest actually exerts an influence on a person's decision(s). A conflict of interest is not a *potential* event: a conflict is present inasmuch as the primary and secondary interests are different. Thus, the phrase "potential conflict of interest" should be avoided. What does lie in the realm of possibility is that this conflict will lead to biased decisions. This may happen or it may not, but the risk is there. For someone involved in guideline development, a conflict of interest entails the risk that the recommendations in the guideline will not reflect the evidence on the benefits and harms of a given intervention. The phrase "perceived conflict of interests" is also not helpful because it implies that the conflict is not "real" and hence not serious. The conflict of interest is real, although its effect on the validity and credibility of a recommendation varies with the severity of the conflict and with the perceptions of the guideline user.

6.2 What types of conflicts of interest exist?

Conflicts of interest are of two basic types: financial and nonfinancial. Both are relevant to guideline development. Financial interests can be directly measured in monetary units, such as for stocks or patents owned, money received for commissioned work, or an honorarium for a speaking engagement. Nonfinancial interests, on the other hand, cannot readily be measured in monetary units and are less tangible and thus more difficult to identify, measure and manage. They include any interest that could be reasonably perceived to affect an individual's objectivity and independence while working with WHO. Examples include a desire for professional advancement or prestige or a drive to publish, to obtain research funding, or to improve one's personal standing in the scientific community. One type of nonfinancial interest is referred to as an "intellectual conflict of interest", defined by one group (*3*) in the context of guidelines as "academic activities that create the potential for an attachment to a specific point of view that could unduly affect an individual's judgment about a specific recommendation".

The declaration of one or more secondary interests (declaration of interests) does not automatically mean that a conflict of interest is present: there is a distinction between the two. A transparent and objective process for assessing a declaration of interests is required to determine if a conflict exists and what its effects might be, and to manage any significant conflict of interest. This chapter describes the processes and procedures for identifying and managing conflicts of interest during guideline development at WHO.

6.3 Why is a conflict of interest important in guideline development at WHO?

Certain data point to an association between the secondary interests of authors and funders and study outcomes (*4*), and between conflicts of interest and the conclusions drawn from systematic reviews (*5*). Data on the effects of industry relationships and the intellectual interests of guideline sponsors and authors are confined to case studies (*6*). Conflicts of interest can introduce bias into almost every step of the guideline development process, from scoping and the development of key questions, to the formulation and presentation of recommendations. In addition, conflicts of interest – with or without bias – can undermine the credibility of specific recommendations, the entire published guideline, and WHO as a global leader in normative and other work. It is thus critically important that WHO guideline developers attend to declarations and conflicts of interest in a proactive, reasoned, transparent and defensible manner.

6.4 What are WHO's policies on conflicts of interest?

Guideline developers at WHO must be familiar with WHO's policies on obtaining declarations of interests and assessing and managing conflicts of interest. An overview of the process for identifying, managing and reporting such conflicts during guideline development at WHO is presented in Box 6.1, with further details provided below. Various websites provide additional information about WHO policies and procedures related to conflicts of interest and guideline development (*2*, *7*, *8*). Additional information can be obtained from the Office of Compliance, Risk Management and Ethics.

Box 6.1. Overview of the process for identifying, managing and reporting secondary interests

1. All potential external contributors to guideline development complete the standard WHO declaration of interests form before invitations to participate in a guideline development group (GDG) or attend a meeting are finalized and before contracts are issued. Potential contributors should submit a curriculum vitae also.

2. The responsible technical officer, in consultation with the director of their department and with input from the steering group, assesses the declaration of interests and curriculum vitae and determines if a conflict of interest exists.

3. Conflicts of interest are assessed for severity (risk of adversely affecting the guideline development process) and a management plan is formulated.

4. At each GDG meeting, declarations of interests are summarized and presented, with the opportunity for GDG members to update and/or amend their declaration. The management strategy for each member with a conflict of interest is also presented.

5. A summary of all disclosed interests and the management plan for any conflict of interest are reported in the final guideline document.

6.5 Who should complete a declaration of interests form?

According to WHO rules and regulations, a declaration of interests form and an analysis of all declarations must be performed whenever an individual provides independent advice to WHO (*7*) – that is, whenever an expert is asked to serve in an advisory role in a personal capacity. In guideline development, this means that anyone invited to participate in the development of a WHO guideline must complete the declaration of interests form and agree to the publication of a summary of that declaration in the guideline (see Table 6.1 and Fig. 6.1). This includes all members of the GDG, individuals who prepare systematic reviews and evidence profiles, guideline methodologists, the technical writer and any other expert who participates in the process in an individual capacity.

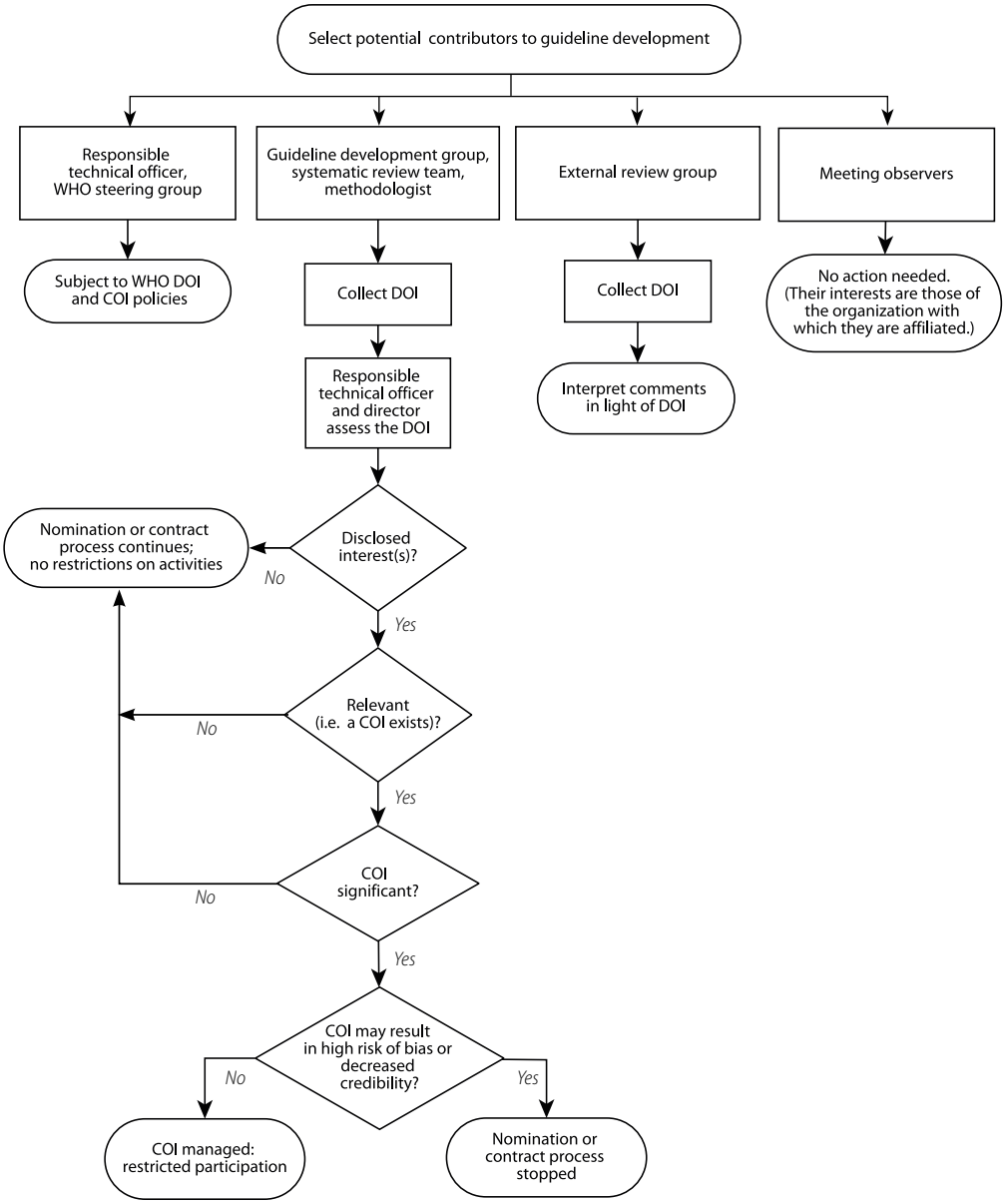
Observers at GDG meetings, including persons representing a governmental or nongovernmental organization (rather than present in an individual capacity) do not need to complete a declaration of interests form. Observers' affiliations as representatives of an organization should be made transparent.

Table 6.1. Declaration and management of conflicts of interest in guideline contributors

Group or individual	DOI required?	Who assesses the DOI?	Management of COI	Comments
Steering group, including RTO	As required by WHO staff rules	WHO staff	This assessment is independent of the guideline process	–
Guideline development group	Yes	RTO and their director, with input from CRE as needed	According to the processes outlined in this chapter	Chair must be free of all financial COI, and free of non-financial interests to the extent possible.
External review group	Yes if they are representing themselves; No if they represent an institution	Steering group with input from CRE as needed	By interpretation of their comments in the context of their COI	These individuals or organizations may be conflicted. Their DOI form (if they represent themselves as an individual) or institutional affiliation (when they represent an entity) should make their interests clear.
Systematic review team	Yes	Steering group with input from the CRE as needed	According to the processes outlined in this chapter	The systematic review team members should have no financial or nonfinancial COI. They may have performed systematic reviews on the topic previously.
Methodologist	Yes	Steering group with input from the CRE as needed	According to the processes outlined in this chapter	The methodologist should have no financial or nonfinancial COI.
Meeting observers	No	Not applicable	By interpretation of their comments in the context of their affiliation(s)	The role and interests of representatives of organizations must be made clear at GDG meetings if the observers participate in any way in the discussions.

COI: conflict of interest; CRE: Office of Compliance, Risk Management and Ethics; DOI: declaration of interests; GDG: guideline development group; RTO: responsible technical officer.

Fig. 6.1. Assessment and management of declarations of interest for groups and individuals involved in guideline development



COI: conflict of interest; DOI: declarations of interest.

6.6 What interests need to be disclosed?

Relevant financial as well as nonfinancial interests should be disclosed and subsequently assessed and managed in order to minimize bias in guideline development and to ensure credible recommendations. A financial conflict of interest arises when an individual or organization receives income or monetary 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 financial interests and the interests of the individual's immediate family members (defined as the spouse, or partner with whom one has a close personal relationship, and the children). Financial interests include, for example:

- personal financial gain such as paid work, consulting income or honoraria and travel stipends;
- support for research, including direct monetary contributions or donations of equipment, laboratory space, etc.;
- proprietary interests and patents;
- grants, fellowships or other financial support to the individual or to their institution; and
- stock shares or bonds in a commercial entity.

The threshold for reporting relevant personal investment interests in the WHO declaration of interests form is 5000 United States dollars (US$): when the value of a single interest is below that amount, it does not have to be reported. Research support must also be disclosed, including non-monetary support valued at more than 1000 US$. Note, however, that certain evidence suggests that individuals can be influenced by financial rewards amounting to far less than these arbitrary thresholds. However, a pragmatic approach to declarations of interest is needed and most organizations that develop guidelines have similar thresholds.

Nonfinancial interests include academic, professional, and personal interests. Financial and nonfinancial interests can overlap. For example, intellectual interests related to career advancement obviously have a monetary component. Examples of roles or positions that might interfere with the objective assessment of a body of evidence include:

- prior publication of a study or systematic review that is part of the evidence base under consideration in the guideline;
- prior public declaration of a firm opinion or position, as in public testimony during a regulatory or judicial process, or in an editorial in a journal; or
- professional or personal affiliation with an organization advocating for products or services related to the subject of the guideline.

Nonfinancial conflicts of interest are ubiquitous and not inherently bad. However, they must be identified and appropriately managed if an unbiased, credible guideline is to be produced.

6.7 Collecting declarations of interests

Declarations of interests are collected using the standard WHO form for experts. Such a form should be completed, signed by the expert or potential contractor, and reviewed and assessed before firm invitations to participate in a guideline development process are issued or contracts are signed. External experts and contractors should be instructed to notify the responsible technical officer of any change in relevant interests, and declaration of interests forms should be updated on a regular basis. Declaration of interests forms are confidential documents that should be securely stored at WHO (whether in hard copy or electronic format). Experts should also be asked to submit a curriculum vitae at the time they submit the declaration of interests form. Table 6.2 lists the principles for collecting, assessing and managing conflicts of interest among persons contributing to guideline development at WHO.

6.8 Assessing declarations of interests

Little is known about the effect of declarations and conflicts of interest on someone's behaviour, how group participants and GDG members and chairs weigh or discount information or opinions presented by conflicted individuals, and how guideline end-users interpret information on conflicts of interest. The following guidance is based on: the scant existing evidence; standard and best practices among leading guideline development groups internationally; sound legal and ethical principles; and expert opinion. In practical terms, three basic steps are involved in assessing declarations of interest.

As a first step, WHO staff review the declaration of interests form in conjunction with the expert's curriculum vitae. If the information in the form is incomplete or unclear, the expert should be contacted for clarification. If the responsible technical officer has any remaining concerns regarding the completeness or accuracy of the information in the form, he or she should discuss this with the WHO Office of Compliance, Risk Management and Ethics or with the Office of the Legal Counsel.

The second step is to determine whether a conflict of interest actually exists: might the primary interest (to develop evidence-based recommendations with objective consideration of all relevant factors) differ from the

Table 6.2. Principles for collecting, assessing and managing declarations of interest for WHO guidelines[a]

Principle	Details	Operational issues and comments
Report all funding sources for the guideline	The funder(s) should be detailed explicitly and made publicly accessible.	All funding sources should be reported in the planning proposal and in the final guideline. The sources of revenue should be summarized to ensure that no funds are received from commercial entities. When the nature of the funder is not apparent from the name, the type of entity should be briefly described, together with its primary purpose and main sources of revenue.
Private sector entities have no role in guideline development	Private sector entities cannot fund WHO guidelines.	The steering group should ensure that potential funders do not receive money from commercial entities.
Complete and accurate DOI is required	Prior to final confirmation of GDG members and contractors, individuals under consideration should declare in writing all interests and activities potentially resulting in a COI.	DOIs should reflect all current and planned commercial, non-commercial, intellectual, institutional, and private/public activities pertinent to the potential scope of the guideline.
Obtain and assess DOI	Disclosures of relevant interests should be assessed and managed according to explicit and transparent criteria developed a priori.	See Table 6.3.
Avoid COI when at all possible	Whenever possible, COIs, both financial and nonfinancial, should be avoided in GDG members.	Significant financial COIs should always be avoided.
Minimize and appropriately manage COIs	Members with COIs should represent not more than a minority of the GDG. Minimize and appropriately manage COIs in order to minimize the risk of bias in decision-making and to maximize the guideline's credibility.	It may be necessary to balance strong opinions if it is not possible to eliminate intellectual COIs among GDG members. When GDG members with intellectual COIs are deemed essential, members with diverse perspectives and experiences should be included in the GDG.
The chair, co-chairs or vice-chair should not have financial COIs	The chair, vice-chair and co-chairs should not have financial COIs and nonfinancial COIs should be avoided if possible.	If nonfinancial COIs in the GDG chairs are inevitable, then the co-chairs should have disparate interests. Consider having as chair or co-chair a methodologist with no intellectual or financial COI.

COI: conflict of interest; DOI: declarations of interest; GDG: guideline development group.

[a] Adapted from Institute of Medicine, National Academies Press; (United States) 2011 standards (*1*).

Table 6.3. Criteria for assessing the severity of a conflict of interest[a]

Criteria	Definition	Considerations
Likelihood of undue influence by the conflicted individual		
Value of the secondary interest	Monetary or other value to the recipient	The higher the monetary value, the greater the effect may be on the recipient's decisions. However, even small gifts can change the behaviour of the recipient. Effect may depend on whether the payment is direct to the individual (e.g. an honorarium) or indirect (e.g. to an academic institution).
Scope of the relationship	Duration and depth of the relationship between the individual and the secondary interest	Relationships that are longer and/or closer may increase the risk of bias in decision-making. For example, a one-time relationship, such as a presentation for industry, is less concerning that a long-term relationship, such as employment by a relevant company.
Extent of discretion	The amount of discretion or authority the conflicted individual has in making important decisions	The role of the conflicted individual in decision-making and the amount of oversight of that individual and of the guideline process in general affect the risk of bias. For example, the chair generally has more discretion, so her or his COIs may have a greater effect than those of other GDG members.
Seriousness of the possible harm from the conflict of interest		
Value of the primary interest	The potential effect of the recommendations on individuals and populations	Requires an assessment of the intended outcomes of the recommendations in the guideline. Recommendations that affect health outcomes such as survival or quality of life, can cause significant harm to the recipients of an intervention if the balance of benefits and harms has not been assessed in an unbiased manner.
Scope of the consequences	The potential for harms caused by an invalid (biased) recommendation across populations	The more individuals that are potentially effected by a recommendation or guideline, the greater is the potential effect of any COI. For example, recommendations that impact the health or well-being of large populations can have tremendous health consequences. The consequences of a COI in the individuals developing a guideline also include diminished trust in these individuals and in the organization sponsoring the guideline (i.e. WHO).
Extent of accountability	The degree of accountability for, and oversight of, the guideline development process	Increased accountability or oversight may lead to lower levels of risk from a COI. The GRC process of independent review and assessment of draft guidelines promotes accountability and helps to ensure that WHO policies on COI are enforced. The external peer review process also increases accountability.

COI: conflict of interest; GDG: guideline development group; GRC: Guideline Review Committee.

[a] Adapted from the Institute of Medicine, National Academies Press; (United States) 2009 (*9*).

expert's other (i.e. secondary) interests? This involves an assessment of the relevance of the secondary interest to the guideline and is usually relatively easy to answer for financial interests.

The third step is much more difficult to address: Does the conflict of interest pose a risk to the guideline? Might the declared interest interfere, or be perceived as interfering, with the individual's ability to objectively evaluate the available evidence and formulate an unbiased recommendation?

Conflicts of interest represent a spectrum; they are not absolute situations. Conflicts of interest vary in severity and hence in the risk that they pose to the guideline development process. The severity of a conflict depends on two factors: the likelihood that decisions in respect of the primary interest will be unduly influenced by a secondary interest, and the seriousness of the adverse outcomes that could result from that influence.

The factors that should be considered when evaluating the severity of a conflict disclosed by an external expert are described in Table 6.3. Note that the assessment focuses on the primary and secondary interests and has nothing at all to do with the characteristics (or character!) of the individual who has disclosed a conflict of interest.

6.9 Managing conflicts of interest at the individual level

All declarations of interests should be reviewed and assessed by the responsible technical officer in consultation with the director of the technical unit. These individuals must decide whether a conflict of interest is severe enough to require a management strategy. In other words, they must decide to what extent the individual can participate in development of the guideline. Decisions are made on a case-by-case basis depending on the severity of the conflict of interest. Merely disclosing the conflict does not constitute an adequate management strategy: careful consideration must be given to each declaration of interests, with formulation of an appropriate strategy in each case.

6.9.1 Guideline development group members

At the level of the individual potential GDG member, management options represent a spectrum that includes:

- **No action** – the conflict of interest requires no action beyond declaration at the GDG meeting and reporting in the published guideline.

- **Participation is restricted** – the individual is excluded from parts of meetings and of the guideline development process.
  - The conflicted individual can be excluded from the formulation of specific recommendations but allowed to participate in all discussions.
  - The conflicted individual can be barred from participating in discussions as well as in the formulation of the recommendations. He or she can be asked to leave the meeting during the development and ratification of any recommendations related to his or her conflict of interest.
- **No participation is allowed** – the conflict of interest is deemed serious enough to preclude membership in the GDG or participation as a contractor for WHO in a specific guideline development process.

In general, certain individuals should not participate at all in the development of a guideline. This includes those who declare significant personal financial interests in one or more companies with a commercial interest in the outcome of the guideline, and those who have intellectual conflicts of interest that are severe and/or cannot be adequately managed at the group level (see Section 6.10).

Conflicts of interest that clearly preclude participation of an individual in a GDG or as a contractor during guideline development include:

- owning more than US$ 5000 worth of shares in a company that manufactures or sells a product or technology under consideration in the guideline;
- holding a patent on a product or technology that may be recommended in the guideline; and
- having a spouse (or partner with whom the individual has a close personal relationship) or children who work for the manufacturer of a product or technology under consideration in the guideline.

In certain situations, the responsible technical officer and director of the relevant technical unit should consider limiting an expert's participation. This applies when the expert:

- is or has been recently employed by a company with an interest in a product related to the guideline, or has been hired by such a company as a consultant, adviser, paid speaker, or opinion leader;
- has financial associations with multiple companies that have commercial interests in the outcome of the guideline;

- has received research funding from companies that have commercial interests in the outcome of the guideline, whether the funding was directed to the expert's institution or to the individual;
- is an author or co-author of one or more key studies within the body of evidence underpinning a recommendation, particularly if the body of evidence is limited; and
- is or has been involved in a major academic programme of work that concerns the intervention, approach or exposure under consideration in the guideline, including conducting trials or systematic reviews and publishing conclusions or opinions on the benefits and/or harms.

There is no evidence to support the use of a particular monetary threshold beyond which an individual must provide a declaration of interests. However, for practical reasons, a threshold is commonly used. (WHO's present threshold is US$ 5000.) There is also no evidence to suggest that an individual who has received money from multiple companies is any less biased than an individual who has received funding from only one. Receipt of funding from multiple for-profit companies may be indicative of a general pro-industry approach (*10*). There are also no data to support distinguishing between funding paid directly to the individual or to an institution on the person's behalf. The lack of evidence to guide the management of conflicts of interest makes it difficult and the decisions somewhat arbitrary. Therefore it is critical that the processes, procedures and rules that are implemented for each group developing a guideline must be explicit, transparent, consistent, thoughtfully constructed and defensible. The Office of Compliance, Risk Management and Ethics provides guidance on the assessment and management of declarations and conflicts of interest (*7*) and will provide advice on how to handle individual cases.

6.9.2 Peer reviewers

Declarations of interests should be collected from all peer reviewers. When reviewers are affiliated with organizations that clearly have a vested interest in the outcome of specific recommendations, a declaration of interests is not usually required; reporting the reviewer's affiliation or institution in the guideline is considered sufficient disclosure.

The external review group, including peer reviewers, may have significant conflicts of interest. In fact, these individuals may have been selected

because of their perspectives and secondary interests. For example, a pharmaceutical company representative may be asked to comment on a new drug from the industry perspective, or developers of a guideline on sugar intake might solicit peer review from a representative of the soft-drink industry. Reviews from such individuals or organizations on a draft guideline may be helpful in anticipating and dealing with controversy, improving the clarity of the final document, and promoting engagement with all stakeholders.

The declarations of interests of the peer reviewers must be transparent to the readers of the guideline, and if significant conflicts of interest exist, they must be appropriately managed. In the process of examining and acting on the comments made by these reviewers, the steering group and the GDG need to assess each comment for validity on a case-by-case basis, in the context of the reviewer's conflicts of interest. Further information is found on the GRC Secretariat website.

6.10 Managing conflicts of interest at the group level

While financial conflicts of interest are usually managed at the individual level, nonfinancial conflicts of interest (including intellectual ones) may be managed at both the individual and group level. Group composition is influential in any decision-making process and observational evidence shows a correlation between the composition of GDGs and specific recommendations. For example, physicians tend to recommend procedures that they personally deliver, whereas multidisciplinary groups tend to be more conservative in their recommendations (*11*, *12*). If individuals with a particular viewpoint have a dominant role in selecting GDG members, this can lead to the selection of members that support a particular recommendation (so-called "committee stacking") (*13*).

6.10.1 Minimizing the risk posed by conflicts of interest when constituting guideline development groups

There is scant evidence upon which to base recommendations on the optimal composition of GDGs to minimize the risk that intellectual conflicts of interest will bias their decisions. The following suggestions are based largely

on expert opinion and standards recommended by other organizations (*1*). (See also the guidance in Chapter 3.)

- GDG chair, co-chairs and vice-chair
  - The chair, co-chair or vice-chair of a GDG should not have any financial conflicts of interest.
  - If the co-chairs or the chair and/or vice-chair are topic experts with intellectual conflicts of interest, then they should have diverse opinions and perspectives.
  - Rather than a topic expert, the chair, vice-chair, or one of the co-chairs can be a guideline methodologist, since such an individual is not likely to have any intellectual conflicts of interest.
- General members of the GDG
  - The majority of members of the GDG should have no conflicts of interest, either financial or nonfinancial.
  - Individuals with financial conflicts of interest should generally not be members of GDGs. This applies especially to individuals with substantial financial interests in an intervention under consideration in the guideline.
  - If the GDG must include some members with financial and/or intellectual conflicts of interest, every effort should be made to balance the perspectives of these individuals in the group. This can be achieved by selecting people whose opinions are known to differ, including a variety of stakeholders.
  - GDG members who are involved in either primary research or systematic reviews relating to the recommendations in question should disclose these activities. The GDG should be composed of individuals with diverse perspectives, training and experiences to keep the recommendations from reflecting a single viewpoint that was conceived before examining and discussing the systematic review of the evidence.
  - Members of the GDG should generally not perform the systematic review, develop the GRADE evidence profiles, or write the final guideline document. These tasks are best done by a team that is separate and independent from the GDG.

- Methodologists
  - Ideally several of the topic experts will also have experience with systematic reviews and guideline development methods. However, GDG members do not often have such expertise, so a methodologist is commissioned to guide and inform the guideline development process, referring the group back to the evidence and to other explicit factors, helping to mitigate the effects of intellectual conflicts of interest.
- Stakeholders
  - In addition to content experts, the GDG should include a variety of stakeholders, including experts in implementation, programme managers, health-care workers, and people who will be affected by the recommendations.

6.10.2 Minimizing the risk posed by conflicts of interest during the guideline development process

Appropriate management of disclosures and conflicts of interest does not end with the assessment of the declarations and the formulation of a plan for the management of conflicts. The responsible technical officer and the steering group have additional and continuous responsibilities.

- Although GDG members participate in guideline development at WHO in an individual capacity, they must all provide their institutional affiliations, along with the goals and purpose of their institutions at the beginning of each GDG meeting.
- The declarations of interests made by GDG members should be provided to all participants at meetings (whether an in-person or a virtual meeting) as one of the first items on the agenda. In addition, all members should be given the opportunity to update or otherwise revise their declarations of interest at that time. If there are any changes to previously declared interests, WHO staff should record those changes and request that a new declaration of interests form be completed. WHO staff will then need to judge whether any new declarations are relevant to the guideline topic and, if so, what the best way to manage any conflicts of interest might be. All decisions should be clearly documented and shared with the GDG.

6.11 Reporting declarations of interest in the guideline

A summary of how declarations of interest were collected, assessed and managed must be included in the main guideline document. If no conflicts were disclosed, this needs to be stated as well. The interests disclosed and the way any conflicts of interest were managed can be presented in either tabular or narrative form. The GRC will not approve a guideline document that does not contain this information.

The following are examples of how conflicts of interest should be reported in the guideline:

- Dr N.C. reported being an investigator on trials for GlaxoSmith-Kline, Quintiles, Uriach and Biomarin, but not for any products or products related to those being considered in this guideline. She also reported holding shares in Biota (valued at more than US$ 5000), which makes antiviral drug X. She was therefore excluded from all discussion of the use of antivirals in prevention and treatment and from the formulation of recommendations on antiviral agents.
- 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 National Institutes of Health (United States) from the use of gossypol for cancer; and is a consultant to several start-up companies, none of which has products on the market. Since no products related to any of these items were under consideration in this guideline, no action was taken.
- Dr A.F. reported that his spouse is an employee of Merck, Sharpe and Dohme, Brazil. He was therefore excused from the room during the review of the evidence and discussion of recommendations on drug X, made by Merck.

The completed declaration of interests forms are kept confidential at WHO in the custody of a senior officer for at least 10 years, and must not be distributed or made public. The responsible technical officer can make available to people outside WHO a summary of the forms' contents, but the forms themselves can only be released into the public domain under the explicit direction of WHO's Director-General.

6.12 What sources of funding are appropriate for WHO guidelines?

WHO guidelines cannot be funded, in whole or in part, by private sector entities or by non-state actors that are not at "arm's length" of a private sector entity (*2*). Private sector entities are commercial enterprises. That is to say, they are businesses that are intended to make a profit for their owners. The term also refers to entities that represent private sector entities, or are governed or controlled by private entities. This includes (but is not limited to) business associations representing commercial enterprises; entities not at an arm's length of their commercial sponsors, and enterprises that are partially or fully state-owned but act like private sector entities. Sometimes it is difficult to determine where a potential funder obtains its revenues. The responsible technical officer and the steering group must make every reasonable effort to elicit this information before accepting funding. Questions regarding the appropriateness of specific funders for guideline development should be directed to the Office of Compliance, Risk Management and Ethics.

6.13 References

1. Graham RMM, Wolman DM, Greenfield S, Steinberg E, editors. Clinical practice guidelines we can trust. Washington (DC): The National Academies Press; 2011 (http://www.nap.edu/openbook.php?record_id=13058, accessed 11 November 2014).
2. Declaration of Interests (WHO Experts) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/cre/ethics/doiexperts/, accessed 11 November 2014).
3. Guyatt G, Akl EA, Hirsh J, Kearon C, Crowther M, Gutterman D, et al. The vexing problem of guidelines and conflict of interest: a potential solution. Ann Intern Med. 2010;152(11):738–41. doi: http://dx.doi.org/10.7326/0003-4819-152-11-201006010-00254 PMID: 20479011
4. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev. 2012;12:MR000033. PMID: 23235689
5. Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA. 1998;279(19):1566–70. doi: http://dx.doi.org/10.1001/jama.279.19.1566 PMID: 9605902
6. Norris SL, Holmer HK, Ogden LA, Burda BU. Conflict of interest in clinical practice guideline development: a systematic review. PLoS ONE. 2011;6(10):e25153. doi: http://dx.doi.org/10.1371/journal.pone.0025153 PMID: 22039406
7. WHO eManual/VIII - Information Products/VIII.7 Ethical Issues/VIII.7.1 Conflicts of interest [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2011 (http://emanual.who.int/p08/s07/Pages/VIII71Conflictsofinterest.aspx, accessed 11 November 2014).

8. WHO eManual/VIII - Information Products/VIII.2 Planning and Content Development/ VIII.2.3 Guidelines [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2013 (http://emanual.who.int/p08/s02/Pages/VIII23Guidelines.aspx, accessed 11 November 2014).
9. Lo BFM, editor. Conflict of interest in medical research, education, and practice. Washington (DC): The National Academies Press; 2009 (http://www.ncbi.nlm.nih.gov/books/NBK22935/, accessed 11 November 2014).
10. Wang AT, McCoy CP, Murad MH, Montori VM. Association between industry affiliation and position on cardiovascular risk with rosiglitazone: cross sectional systematic review. BMJ. 2010;340:c1344. doi: http://dx.doi.org/10.1136/bmj.c1344 PMID: 20299696
11. Hutchings A, Raine R, Sanderson C, Black N. A comparison of formal consensus methods used for developing clinical guidelines. J Health Serv Res Policy. 2006;11(4):218–24. doi: http://dx.doi.org/10.1258/135581906778476553 PMID: 17018195
12. Norris SL, Burda BU, Holmer HK, Ogden LA, Fu R, Bero L, et al. Author's specialty and conflicts of interest contribute to conflicting guidelines for screening mammography. J Clin Epidemiol. 2012;65(7):725–33. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.12.011 PMID: 22498428
13. Lenzer J, Hoffman JR, Furberg CD, Ioannidis JPA; Guideline Panel Review Working Group. Ensuring the integrity of clinical practice guidelines: a tool for protecting patients. BMJ. 2013;347:f5535. doi: http://dx.doi.org/10.1136/bmj.f5535 PMID: 24046286

7. Formulating questions and selecting outcomes

The formulation of the questions that the guideline needs to address strongly influences the final recommendations, so getting this right is crucial. The questions are used to systematically search the evidence base for answers in the areas of uncertainty or controversy that the guideline seeks to clarify. In the course of exploring the scope of the guideline, identifying potential topics, and discussing areas of uncertainty and controversy, the questions that need to be asked should become clear. Because these questions drive the evidence search and form the basis of the recommendations, they should be carefully crafted and precise.

7.1 Types of questions

There are two basic types of questions: background and foreground questions. Background questions provide context and rationale for the guideline; foreground questions directly inform and underpin recommendations.

7.1.1 Background questions

These questions pertain to important background information on the issues under consideration and their context. They do not relate to evidence of the type that directly informs recommendations, but they do lead to information which helps to frame and formulate the foreground questions. Background questions may pertain, for instance, to definitions; the prevalence, burden and distribution of the problem or disease at hand; or the pathophysiologic mechanisms underlying the effects of exposures or potential interventions. Here are some examples:

- What are the potential mechanisms by which human papilloma virus infection causes cervical cancer? What is the relative risk of cervical cancer among women with laboratory evidence of human papilloma virus infection? Does the risk vary by age group, socioeconomic status, comorbidities or other factors?
- What is the prevalence of HIV infection among individuals in the African Region with multi-drug resistant tuberculosis?

- How does the physical and social environment affect the prevalence of obesity?

The answers to background questions may be found in a wide range of information sources, ranging from basic scientific or pharmacokinetic data from animal studies, to surveillance data on incident cases, to theoretical frameworks explaining behaviour change in connection with community educational programmes. These questions help to establish the context for the recommendations and a systematic review may be unnecessary or an ineffective use of time and resources. Nevertheless, the answers to these questions must be based on relevant and objective evidence in order to generate a high level of confidence in the results. For example, data on the incidence or prevalence of a disease might be obtained from a single valid source, which should be duly cited and justified. (Note that it is possible, in some situations, for a question on prevalence to be central to prioritizing a set of interventions or establishing baseline risk. In such cases it is a foreground question requiring a systematic review.)

7.1.2 Foreground or key questions

These questions form the basis of the search for the evidence that will underpin the recommendations. They will be of various types, depending on the nature of the guideline and the specific topics it will cover. When a potential recommendation focuses on an intervention, the key questions pertain to the efficacy, effectiveness and potential harms of the intervention, as well as to factors such as acceptability, feasibility, cost and cost–effectiveness, and to the values and preferences of the people who will be affected by the recommendations. When the topic focuses on a diagnostic test, the key questions will revolve around sensitivity, specificity, positive and negative predictive value and other related parameters. For recommendations related to environmental, genetic and other exposures, key questions will relate to baseline risk or prognosis and to the relative risk of specific outcomes from the exposure. Examples:

- What is the effect of vaccination against human papilloma virus on the incidence of invasive cervical cancer? What are the possible adverse effects of such vaccination?
- What is the risk of type 2 diabetes mellitus among females aged 45 to 65 years with a body mass index between 25.0 and 30.0 kg/m? Does the risk vary across different population groups and different socioeconomic circumstances?

- In multi-drug-resistant tuberculosis patients, does the addition of bedaquiline to WHO-recommended second-line drug therapy improve outcomes, as reflected by culture conversion and the time to such conversion?

Foreground questions are the most important ones for guideline development. The evidence identified on these questions will be used by the GDG to underpin the recommendations and these questions usually require a systematic review and assessment of the quality of the evidence for selected outcomes using the GRADE approach (see Chapter 9) (*1*). Foreground questions should therefore be framed in a way that enables a systematic search of the literature. The PICO format is an effective way to do this.

7.2 Formulating key questions in PICO format

PICO is an acronym for population, intervention (or exposure), comparator and outcome – four elements that should be considered in any question governing a systematic search of the evidence. Sometimes T is added for the timing of the measurement of outcomes or for the duration of the intervention or exposure, and S for the setting where the recommendations will be implemented (PICOTS). For simplicity we use PICO in this handbook, wrapping the T into the I (intervention or exposure) component, and the S into either the P (population) or the I. The PICO format also provides a useful structure for delineating inclusion and exclusion criteria for the body of evidence (see Chapter 8) and for formulating recommendations (see Chapter 10).

7.2.1 Population

What group or population is targeted by the intervention or exposure under consideration?

- How can the members of the target population be most accurately described? What are the relevant demographic characteristics? Consider age, sex and other social, geographic and environmental characteristics that are of relevance to the guideline topic.
- What is the setting? Is it, for example, hospitals, communities, workplaces or schools?

- Should any subpopulations be considered? These may include, for example, subpopulations whose gender, education, income, geographical location or other characteristics makes them subject to increased environmental exposures or to reduced access to health services, or less able to benefit from such services.
- Should certain subpopulations be specifically considered and analysed separately? Consider for example, pregnant women, children or people living with HIV infection.

7.2.2 Intervention or exposure

What intervention or exposure is under consideration?

- What treatment, procedure, diagnostic test, prognostic factor, risk factor, behavioural change, social activity, screening test, or preventive measure is being evaluated?
- What health policy or health-care system is being evaluated? Is it, for example, pharmaceutical pricing policies at the country level, or the integration of an opportunistic screening programme into rural primary care clinics?
- What social, economic, or environmental determinants are being evaluated? Some possibilities are social protection; regulation of product formulae, marketing and advertisement; occupational health standards; urban planning; or psychosocial support to families and communities.
- Are variations in the level of an intervention or exposure being considered? Examples are drug dosages and frequency, a range of levels of a risk factor or environmental exposure, or the intensity of an educational intervention.

"Intervention" is defined very broadly. It can mean anything from diagnostic tests or other technologies, to complex public health measures, to measures aimed at modifying aspects of the health-care system, to give a few examples. For interventions that are complex, consider which components are of greatest interest to the target audience of the guideline and how they might best be described. An "exposure" is any factor that can influence the risk of a given outcome. Some exposures determine baseline risk or prognosis, whereas others can change baseline risk. One example of the latter is the presence of chemical contaminants in the air or community water supply.

7.2.3 Comparator

What courses of action or exposures stand as alternatives with which to compare those recommended in the guideline?

- The comparator in a recommendation may be standard practice – including no specific treatment – or another intervention relevant to current practice or policy.
- Comparators may be placebo, no intervention, the standard of care, a gold or reference standard diagnostic test, variations of the intervention under consideration (such as different dosages), a different level of exposure or intervention, or a completely different intervention.
- Comparisons may also be established, for example, between individual-level and population-level interventions, or between interventions involving the health sector alone versus the health sector in collaboration with other sectors.

The most important comparators are generally those that are most closely related to current practice because they provide guideline developers with the information needed to formulate recommendations relevant to end-users' needs.

7.2.4 Outcomes

What are the outcomes of the intervention or exposure that matter most to the individuals and populations affected by the guideline? In the context of clinical interventions, these are sometimes referred to as patient-oriented or patient-centred outcomes.

- What are the potential benefits of the intervention or exposure?
- What are its potential harms?
- What impact will it have on equity (distribution of health)?

Outcomes, both positive and negative, need to be carefully selected with input from experts, the programme managers who will implement the recommendations and the individuals who will be affected by them (see Section 7.6).

7.2.5 Examples of key questions in PICO format

The following are well-formulated key 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) within a 5-year period?
- 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) by 12 months of age?
- In an urban population (P), is exposure to an environmental chemical (I), compared with no exposure (C), associated with the risk of cancer (O) at 10 years of follow-up?
- Is breast cancer screening (I) in women 70 years of age or older with an average risk of breast cancer (P) as cost–effective as no screening (C) in preventing death from breast cancer (O)?
- In a national population (P), how does one intervention (I) perform, compared with another (C), in terms of cost per quality-adjusted life years gained over a 5-year period (O)?
- In general populations (P), does gender balance in health facility staffing (I), compared with lack of gender balance (C), improve equity in health outcomes (O)?"

The following is a poorly-formulated key question:

- Is drug X effective for treating multi-drug-resistant tuberculosis?

  - What is the specific population subject to the intervention?
  - Over what time period is drug X given and what is the period of follow-up?
  - What is the exact drug regimen that is being examined? Should drug X be added to standard regimens recommended in existing WHO guidelines, or should it replace other drugs?
  - What is drug X being compared to? Standard treatment regimens?
  - What outcomes are under consideration? Intermediate outcomes, such as culture conversion by 2 months? Or health outcomes, such as death? What about harms?

The key question is better phrased as:

- Among different groups of multi-drug-resistant tuberculosis patients (P), does the addition of drug X for Y months (I) to the WHO-recommended regimen (C) improve the rate of cure at 24 months (O)?

The following key question is also poorly formulated:

- Is screening of adults for diabetes cost-effective?

  - What is the specific population? Although this may not be specified in the key question *per se*, "adults" should be further defined in terms of age, sex, risk factors for cardiovascular disease, for example.
  - What type of diabetes is being considered here? Diabetes mellitus? Type 1 and 2, or just type 2?
  - What are the operating characteristics of the screening test?
  - What is the comparator? No screening?
  - Cost–effectiveness is based on a specific outcome, such as quality-adjusted life years gained: for what outcome is cost–effectiveness being examined in this question?
  - Cost–effectiveness is a relative concept: what is the threshold used to assess whether screening is considered cost-effective?
  - For what time frame is cost–effectiveness being considered?

In general, key questions are poorly crafted if they have not included all four components of the PICO format, or if they do not provide sufficient specificity for one or more of these components, such that a systematic review search strategy cannot be devised without further clarification.

Key questions in PICO format can become unwieldy at times if all four components are encompassed together with additional details on each component. It is therefore usually preferable to craft a simple, clear key question that includes the essential aspects of the PICO format, and follow it with a detailed list of inclusion and exclusion criteria for the body of evidence that will be used to answer the key question (see Chapter 8).

7.3 Types of key questions

Although the PICO framework was originally devised to address the efficacy, effectiveness and potential harms of an intervention, it can be applied to other types of questions, as outlined in Table 7.1. The general types of questions most relevant to WHO guidelines relate to the following:

- intervention efficacy and effectiveness;
- intervention harms;
- diagnostic approaches and test characteristics;
- the values and preferences of the individuals affected by an intervention;
- risk assessment, including baseline risk (prognosis), and additional risk from a given exposure; and

Table 7.1. Types of key questions in PICO format

Type of key question	Syntax	Population	Intervention	Comparator	Outcomes	Example
Intervention efficacy and/or effectiveness	Among individuals with a certain disease or condition (P), how effective is a certain treatment or approach (I) in producing an outcome of benefit (O)?	What populations are of interest? Which subpopulations?	What intervention, treatment or approach is being considered?	What is/are the main alternative treatment(s) or approach(es)?	What outcomes matter most to individuals affected by the disease or condition?	Among children with HIV (P), how effective is isoniazid prophylaxis (I) compared to no prophylaxis in preventing tuberculosis (O)?
Intervention harms	Among individuals with a certain disease or condition (P), what are the unintended consequences (harms) of a certain treatment or approach (I)?	As above	As above	As above	What harms or undesirable effects of the intervention matter most to individuals who experience the intervention?	Among children with HIV (P), what are the harms and burden (O) of isoniazid prophylaxis (I) used to prevent tuberculosis compared to no prophylaxis?
Diagnosis	Among patients with a certain condition (P), how accurate is a certain test (I) in diagnosing a given disease (O) compared with a reference standard (C)?	To which populations would the test be applicable? Which subpopulations?	What test or strategy is being evaluated?	What is the comparison test, (often the reference standard or current diagnostic approaches)?	What is the diagnostic accuracy of the test or strategy for detecting the target condition (usually sensitivity, specificity, predictive value and related parameters)?	Among patients with acute chest pain (P), how accurate is an electrocardiogram (I) in diagnosing acute myocardial infarction (O) compared to serum enzyme testing (C)?
Values and preference	What are the values and preferences (O) of individuals with a certain disease or condition (P) in terms of the potential outcomes of the intervention or exposure (I)?	What populations are of interest? Which subpopulations?	What intervention, treatment or approach is being considered?	What is/are the main alternative treatments or approaches?	Which outcomes matter most to individuals affected by the disease or condition? How do they feel about the possible benefits of the intervention or exposure versus its possible harms? What are the attitudes of people affected by the intervention or exposure regarding the intervention?	How do health-care workers rate outcomes potentially impacted by wearing personal protective equipment when working in Ebola treatment centres? Relevant outcomes might include manual dexterity, ability to communicate with patients, becoming infected with Ebola virus, or death. What are health-care workers' (P) preferences (O) regarding wearing a face shield (I) versus goggles (C)?

continues ...

... continued

Type of key question	Syntax	Population	Intervention	Comparator	Outcomes	Example
Risk or prognosis	Among patients with a certain disease (P), does a prognostic or risk factor (I), change the risk of a certain event (O) relative to baseline risk (C)?	As above	What exposure is of interest? Which individual or environmental factors predict outcomes?	What is the baseline risk (the risk in the unexposed)?	What is the incidence or prevalence of the condition In the exposed?	Among patients with prostate cancer (P), do lumbar metastases (I), increase 5 year mortality (O) compared to no metastatic disease?
Resource considerations	What is the cost of intervention X in setting Y? What is the cost–effectiveness of intervention X in setting Y, for outcome Z?	As above	What intervention, treatment or approach is being considered?	What is/are the main alternative treatments or approaches?	What is the cost of intervention X? Which outcome matters most to individuals affected by the disease or condition and would provide a metric for cost–effectiveness?	What is the cost (O) of latex-free gloves in West Africa (I) for use in managing persons in Ebola treatment units (P)? Is drug X (I) cost–effective in preventing death (O) from cardiovascular disease in adults with type 2 diabetes mellitus (P)?

HIV: human immunodeficiency virus; PICO: population, intervention, comparator and outcome.
This table is based in part on Chapter 4 of the NICE guidelines manual (*2*).

- resource considerations, including cost and measures of economic efficiency, such as cost–effectiveness.

7.4 Broadly- versus narrowly-focused key questions

Questions based on the PICO format may be broad or narrow in scope. A broad question will lead to a comprehensive summary of a larger body of evidence and more generalizable findings, but it may also require significantly more resources to answer. A broad question may also yield very heterogeneous evidence, making interpretation difficult. On the other hand, a narrow question may be easier to manage, but the evidence might be sparse and the 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 several narrow questions.

Example of a broad PICO question:

- Do financial incentives (I), compared with 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 with no housing allowance (C), improve the retention (O) of health workers (P) in rural areas?

For many recommendations, several separate key questions in PICO format will be required to find all the relevant evidence. In other words, several key questions may be needed to inform the "recommendation question". For example, a GDG may wish to develop a recommendation on screening for type 2 diabetes mellitus, in which case the recommendation question might be:

- Should opportunistic screening be performed to identify persons with type 2 diabetes mellitus?

Some of the key questions in PICO format that need to be answered in order to formulate the recommendation include, for example:

- What is the burden of disease from diabetic complications and mortality in terms of quality-adjusted life years lost?
- Does stringent control of glycaemia, serum lipids and blood pressure in people with type 2 diabetes result in better health outcomes than less intensive treatment or no treatment?

- Do interventions in people with type 2 diabetes detected through screening improve health outcomes, compared with no intervention or interventions started after clinical diagnosis?
- What are the psychological, sociological and iatrogenic adverse effects (or harms) of screening for type 2 diabetes?

As noted, a seemingly simple question (Is screening effective?) involves a large number of carefully crafted questions, each of which may call for a systematic review. In other words, a one-to-one relationship between recommendations and key questions does not exist (see Section 7.8).

7.5 Number of key questions

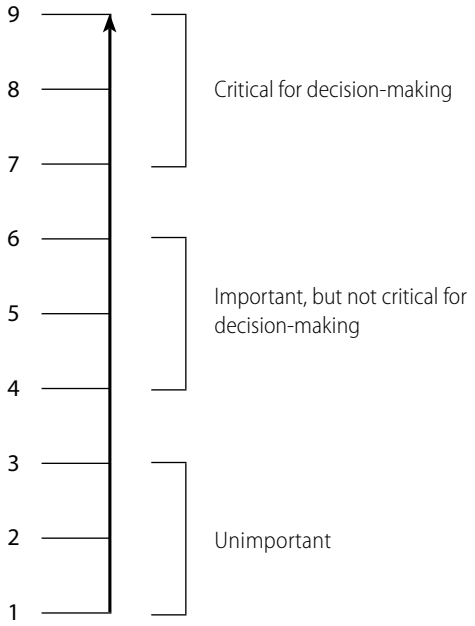
The number of key questions for a guideline varies greatly, depending on the scope of the guideline (the number of areas of uncertainty) and the available resources and timeline. In addition, key questions vary greatly in scope and complexity for each component of PICO. As noted above, for some types of recommendations, several key questions are needed to provide the evidence base; for others, only one key question and one systematic review are required. There is no point in wrapping several key questions into one: the workload and resources required remain the same.

7.6 Selecting and rating outcomes

The purpose of any recommendation is to achieve a net benefit. Thus, selecting the most important outcomes is critical to producing a useful guideline. The value attached to a given outcome by different populations and subpopulations varies. For this reason, it is essential to ask members of the GDG – which should contain end-users, implementers, patient representatives, policy-makers, technical experts and perhaps also other stakeholders – to identify the key outcomes that need to be considered in a guideline.

The steering group should compose an initial list of relevant outcomes, including the intervention's or exposure's desirable and undesirable effects and its potential effects on health equity. The steering group should then ask the GDG to identify any other outcomes that have not been listed. Once a workable list of outcomes has been collected, they need to be ranked in order of priority. This can be done effectively by asking GDG group members to rank the outcomes on a scale from 1 to 9, where 7–9 rates the outcome as critical for a decision, 4–6 rates them as important and 1–3 rates them as

Fig. 7.1. Scale for rating outcomes in the guideline development process



unimportant (Fig. 7.1). The individuals doing the rating should be given a particular perspective from which to assess the outcomes; most often this is the perspective of the service users or other individuals who will be directly affected by the recommendations.

The mean score for each outcome – or the median if the scores are skewed – can then be used to determine its relative importance. It may be helpful to provide the range of results as well. An iterative approach may be needed to define the final list of outcomes, ranked in order of priority, upon which the systematic review and recommendations will focus. In-person discussion at the scoping meeting or the use of a Delphi approach via email are suitable options to rank outcomes.

Generally no more than seven outcomes (both beneficial and harmful) considered important or critical to the formulation of the recommendations should be selected (*3*). Selecting too many outcomes makes it difficult to compare across outcomes when balancing the overall benefits and harms of an intervention.

7.7 Finalizing the key questions and priority outcomes

Key questions should be finalized by the steering group after receiving input from all the relevant experts, including end-users (e.g. programme managers, partner agencies, and consumer and patient groups). Because the number of questions that call for systematic reviews will be a major determinant of the time and resources needed to complete the guideline, the steering group should aim to restrict the questions to those dealing with the areas of greatest controversy and uncertainty. This will help to ensure a comprehensive and useful guideline. The following steps should be followed to develop and finalize the key questions (Fig. 7.2):

Step 1: generate an initial list of questions

The steering group develops an initial list of potential questions based on the agreed-upon scope of the guideline. These are then divided into background and foreground questions.

Step 2: draft the key questions in PICO format

The steering group, with input from the GDG, the systematic review team and the guideline methodologist, applies the PICO framework to the foreground questions.

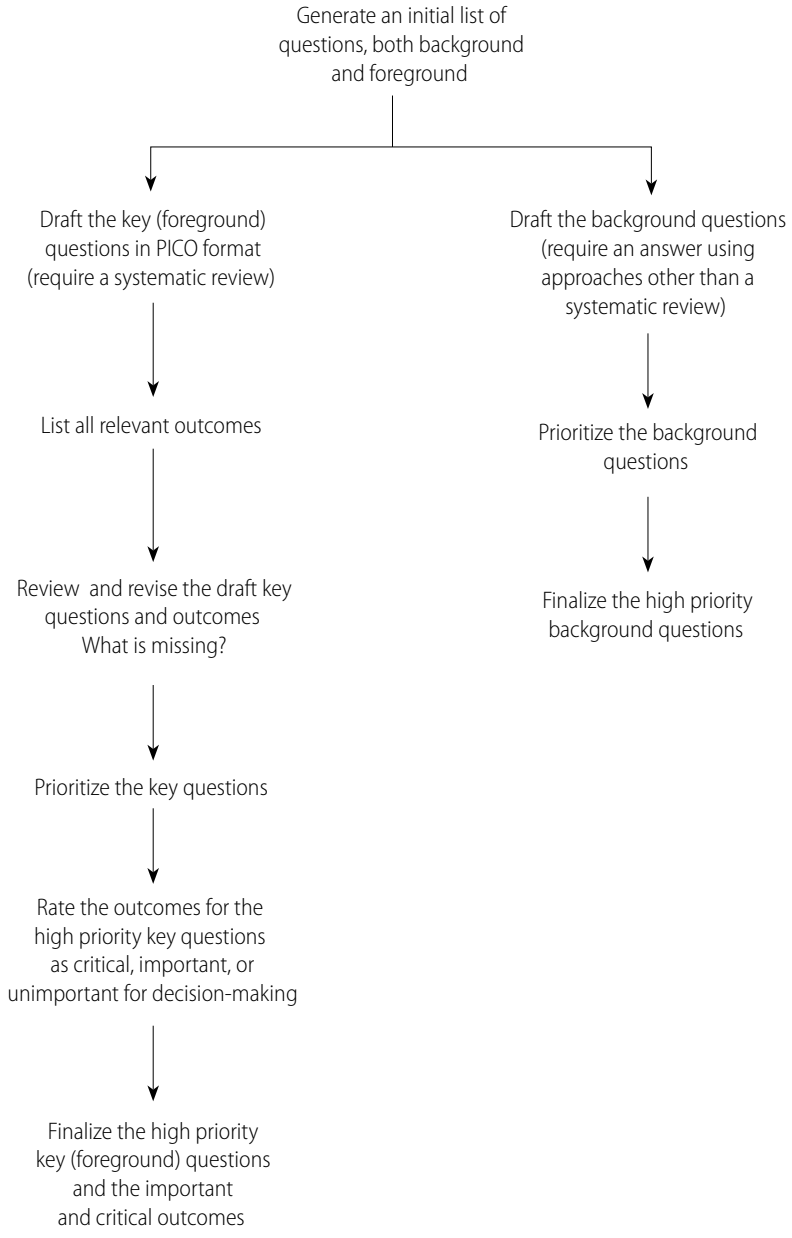
Step 3: list relevant outcomes

The steering group should list relevant outcomes, including both the potential benefits and harms of the intervention or exposure. The GDG then reviews this list and may add additional outcomes considered important.

Step 4: review and revise

The list of questions and outcomes of interest should be sent by the steering group to the GDG and possibly to other stakeholders for review and comment.

Fig. 7.2. Prioritizing key questions and outcomes



PICO: population, intervention, comparator and outcome.

Step 5: prioritize the key questions

The steering group, with input from the GDG, the systematic review team and the methodologist, finalizes the list of background and foreground questions and ranks them in order of priority. Determining which foreground questions need systematic reviews follows next.

Step 6: rate the outcomes

The GDG and/or other stakeholders, such as service users and others directly affected by the recommendations, then rate or rank the listed outcomes using a formal process, such as that described above.

Step 7: finalize the key questions and the important and critical outcomes

The steering group should list the high-priority key questions and the outcomes which the GDG will use to formulate recommendations.

7.8 Analytic frameworks

A diagram illustrating the components of the intervention or exposure, and its relationship to the population, comparator and outcomes, is an extremely useful tool for the systematic review team, the steering group and the GDG. Such diagrams go by different names: analytic frameworks, evidence models, logic models, causal pathways, decision trees, etc. The analytic framework illustrates the various components of the key question, the relationships among these components and the underlying mechanisms and pathways by which interventions or other exposures affect health and other outcomes (*4*).

Formulating an analytic framework helps the steering group to better understand the various factors under consideration and their context, and the relationship between intermediate outcomes (e.g. serum cholesterol levels, rates of participation by community members in an intervention to increase physical activity) and distal health outcomes (e.g. death from cardiovascular disease, quality of life). By depicting all of the elements between an intervention and important health outcomes (such as death), the types of evidence used to inform each linkage and any assumptions become clear. For example, in examining the effectiveness of a community intervention to increase

levels of physical activity, the outcome in studies included in the evidence base may focus on behaviour (i.e. activity), and a single systematic review may be selected to assert the relationship between activity levels and diminished rates of death from cardiovascular disease. Analytic frameworks can also help to illuminate reasons for agreements and disagreements when the GDG formulates recommendations as each linkage is discussed in turn (*5*).

The steering group should consider developing analytic frameworks when drafting and revising the scope and key questions at the beginning of the guideline development process. The GDG, the systematic review team and the guideline methodologist should provide input. The analytic framework may be included with the planning proposal or in the final guideline publication, particularly when recommendations address a series of linked interventions, outcomes and health states.

7.9 References

1. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al.; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924–6. doi: http://dx.doi.org/10.1136/bmj.39489.470347.AD PMID: 18436948
2. National Institute for Health and Clinical Excellence. The guidelines manual – Chapter 4. London: National Institute for Health and Clinical Excellence; 2012 (http://www.nice.org.uk/article/pmg6, accessed 11 November 2014).
3. Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. J Clin Epidemiol. 2013;66(2):158–72. doi: http://dx.doi.org/10.1016/j.jclinepi.2012.01.012 PMID: 22609141
4. Woolf SH, DiGuiseppi CG, Atkins D, Kamerow DB. Developing evidence-based clinical practice guidelines: lessons learned by the US Preventive Services Task Force. Annu Rev Public Health. 1996;17(1):511–38. doi: http://dx.doi.org/10.1146/annurev.pu.17.050196.002455 PMID: 8724238
5. Helfand M, Balshem H. AHRQ series paper 2: principles for developing guidance: AHRQ and the effective health-care program. J Clin Epidemiol. 2010;63(5):484–90. doi: http://dx.doi.org/10.1016/j.jclinepi.2009.05.005 PMID: 19716268

8. Evidence retrieval and synthesis

WHO recommendations must be based on the best available evidence. All relevant evidence should be identified, synthesized and presented in a comprehensive and unbiased manner. This is challenging, yet it is essential to developing valid recommendations and high-quality guidelines.

Recommendations in WHO guidelines should be based on a systematic review of the scientific literature guided by specific key questions about the intervention, exposure or approach under consideration (see Chapter 7).

8.1 What is a systematic review?

A systematic review is "a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to extract and analyse data from the studies that are included in the review" (*1*). If conducted properly, systematic reviews reduce the risk of bias and improve the reliability and accuracy of conclusions based on evidence. Systematic reviews should underpin all WHO recommendations on the efficacy, effectiveness, and harms of interventions; the use of diagnostic tests; exposure limits or safety thresholds for various substances; and all other topics for which WHO issues recommendations.

The characteristics of a systematic review are:

- specific, objective and clearly focused key questions;
- explicit, transparent and reproducible methods;
- pre-defined eligibility criteria for included studies;
- a comprehensive and systematic search for all studies that meet eligibility criteria;
- an assessment of the risk of bias of the included studies;
- a description and synthesis of the characteristics and findings of the individual studies and of the body of evidence; and
- valid and clearly presented conclusions, with information on their applicability to the key question.

Not all reviews are systematic and not all systematic reviews are of high quality. So-called "narrative" or non-systematic reviews are missing one or more of the essential characteristics noted above. Non-systematic reviews

and low-quality systematic reviews should not inform WHO guidelines and WHO staff must be able to recognize these and understand their limitations.

A systematic review is different from a meta-analysis. The latter term refers to the quantitative synthesis (pooling) of an outcome across comparable studies to achieve a pooled estimate of effect (see Section 8.3.5). A systematic review may contain one or more meta-analyses, but this is not a requirement. Likewise, a meta-analysis may be a pooled estimate of an outcome from a cohort of selected studies, and not from studies identified through a systematic review. This latter approach is to be avoided in most situations.

8.2 Identifying and evaluating existing systematic reviews

After scoping the proposed guideline (see Chapter 2) and crafting the key questions (see Chapter 7), the next step in guideline development is to identify one or more systematic reviews to address each key question. Existing systematic reviews will inform the guideline development process, whether or not a new systematic review or an update of an existing review will be required.

8.2.1 Sources for existing systematic reviews

The search for existing systematic reviews in bibliographic databases, such as Medline and EMBASE, should be comprehensive and conducted on the basis of the predefined key question(s). This search can be performed by the steering group or can be commissioned to an external team. Search filters specific for reviews are available in search engines such as PubMed or OVID and are useful for preliminary scoping. The characteristics of these search filters vary, however. Furthermore, the filters cannot reliably distinguish between a systematic review and a non-systematic review or identify high-quality reviews in particular (*2*, *3*). Search strategies for identifying systematic reviews should be developed, validated and documented in the same way as search strategies for identifying primary studies (see Section 8.3.2).

The Cochrane Collaboration, a large global network that produces systematic reviews, is a nongovernmental organization in official relations with WHO (*4*). The responsible technical officer should search The Cochrane Library for existing reviews and for the protocols of reviews that are under development. In addition, the Cochrane Collaboration may be able to iden-

tify forthcoming reviews or planned updates of existing reviews that have not yet been published in the Cochrane Library. The GRC Secretariat can refer WHO staff to the Cochrane Collaboration focal point, who will liaise with the relevant Cochrane groups to help identify reviews that are in progress.

The Database of Abstracts of Reviews of Effects (DARE), developed and maintained by the Centre for Reviews and Dissemination at the University of York, in the United Kingdom of Great Britain and Northern Ireland, is also part of the Cochrane Library (*5*). This database contains details of systematic reviews that evaluate the effects of health-care interventions and of the delivery and organization of health services, as well as reviews of how the wider determinants of health, such as housing, transport and social care, may affect health. Other excellent sources for systematic reviews are Epistemonikos (*6*) and Health Systems Evidence (*7*).

The Campbell Collaboration produces systematic reviews of research evidence on the effects of interventions in the fields of crime and justice, education, international development and social welfare (*8*). Like Cochrane systematic reviews, Campbell reviews are structured and meet certain quality standards.

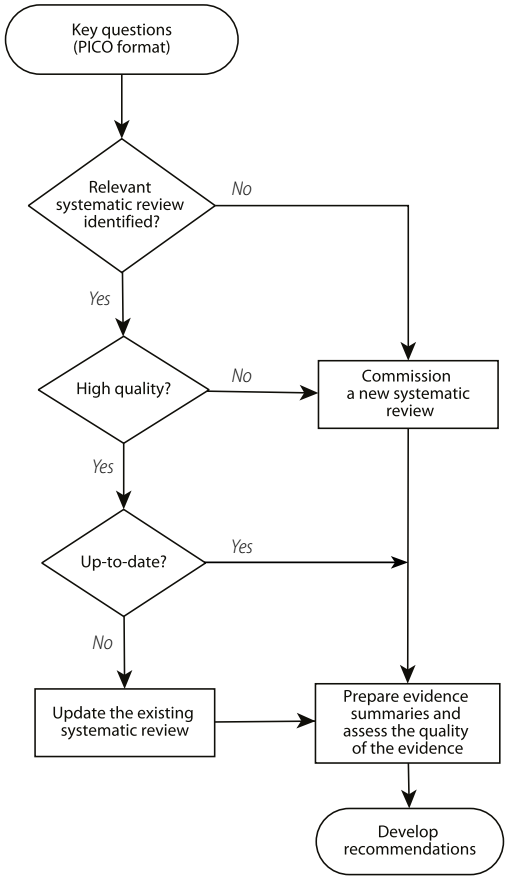
PROSPERO, an international database of systematic reviews in health and social care (*9*), is another valuable resource. It includes Cochrane reviews as well as other systematic reviews, many of which were prospectively registered. Before commissioning a new review, check this database for relevant reviews that may be under way or near completion.

Searching for existing guidelines directly applicable to the guideline topic under development may yield high-quality systematic reviews. The National Guideline Clearinghouse of the US Department of Health and Human Services (*10*) is an excellent source for international guidelines.

8.2.2 Is a new systematic review needed?

It is not always necessary to commission a new systematic review (see Fig. 8.1). If one or more relevant, current and high-quality systematic reviews exist, these should be used. Updates, if needed, are usually less expensive and time-consuming than new reviews.

Fig. 8.1. Is a new systematic review needed?



PICO: population, intervention, comparator and outcome.

8.2.3 Evaluating the relevance and quality of existing systematic reviews

Once existing systematic reviews are retrieved, the steering group needs to assess their relevance, quality, and timeliness before deciding whether a new review needs to be commissioned. To assess relevance, compare the key question of the existing systematic review to the key questions that were developed during the guideline scoping exercise, considering each component of PICO. Most frequently, the existing review does not entirely match

the current key question. Nonetheless, the review may address one facet of WHO's key question or provide useful background information, and the list of included studies may inform WHO's systematic review.

If the existing systematic review addresses one of the key questions of the guideline, then its quality should be assessed. The following checklists may be used to assess the quality of systematic reviews:

- assessment of multiple systematic reviews (AMSTAR) (*11*), or R-AMSTAR (*12*);
- Oxman and Guyatt index for the quality of review articles (1991) (*13*).

Note that a checklist merely provides a list of items that should be appraised. In deciding whether a systematic review is of sufficient quality to inform a WHO guideline, all items need to be considered in combination. None of the checklists mentioned above should be used as a scoring tool because the items they contain are not equally weighted and summary scores are therefore misleading.

Once a systematic review has been found relevant and of high quality, whether it is up to date or not must be determined. There is no rule for dismissing a review on the basis of the time since publication (for example, two years); it depends on the topic and on the availability and rate of production of new information. In some topic areas, such as drug development for multidrug-resistant tuberculosis, data on new drugs are published monthly. However, in areas of public health in which little research is being conducted, an older review may still be considered "current". If a Cochrane review is available, the relevant review group should be contacted to determine if an update is planned or is already under way.

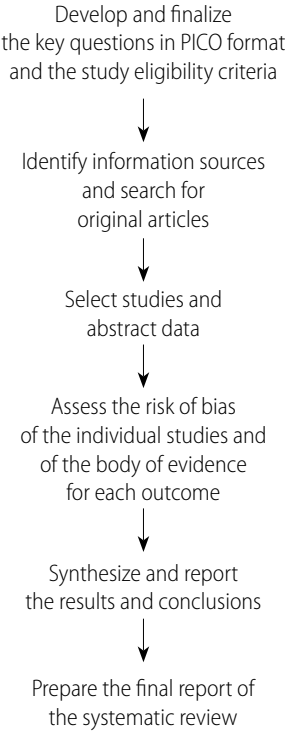
If existing up-to-date, high-quality systematic reviews provide inconsistent results, further exploration is necessary. Do the reviews address slightly different questions or have different inclusion/exclusion criteria? How do the methods for data synthesis compare? Are the data the same but interpreted differently? If so, why? Such an examination can be very informative, and a rationale is needed as to why a specific review was selected to inform a WHO guideline.

8.3 Steps in performing a systematic review

There are six basic steps in the systematic review process (see Fig. 8.2) (as distinct from the guideline development process described in Chapter 1). Step 1, which involves developing and finalizing the key questions, should be performed by the steering group with input from the systematic review

team; the remaining steps are usually performed by the systematic review team in collaboration with the steering group.

Fig. 8.2. Steps in developing a systematic review



PICO: population, intervention, comparator and outcome.

Step 1: develop and finalize the key questions and the study eligibility criteria

In addition to specifying a guideline's key questions in PICO format, it is essential to develop explicit and thoughtful criteria for deciding which studies to include or not include in the systematic review, be they primary studies or existing reviews. The inclusion and exclusion criteria, as they are known, are drafted by either the steering group or the systematic review team, reviewed

by both groups, and adopted by consensus agreement between them. The responsible technical officer may also ask the GDG to provide input. Inclusion/exclusion criteria can be conceptualized using the PICO format, with additional inclusion of: (1) the time frame for outcome measures or the duration of the intervention; (2) further specification of setting and subpopulations; (3) study design (i.e. randomized controlled trials only, or various types of nonrandomized experimental or observational studies); (4) publication language; (5) a date range for publications; and additional specifications that are relevant to a given topic.

Step 2: identify information sources and search for original articles

The development and finalization of the strategies for searching for evidence that meets the established inclusion criteria is a fundamental step in developing guidelines. It must be done correctly or the quality of the guideline may be compromised. Engaging an expert in information science with specific expertise in bibliographic database searching related to systematic reviews is essential. Several individuals, including topic experts and one or more systematic review and guideline methodologists, should review the draft search strategies before they are finalized. The systematic review team must consult with a WHO information scientist before finalizing the search to ensure that the search strategy includes all relevant databases and the appropriate search terms. There are tools and templates for planning, implementing and documenting a literature search (*14*).

Searching the literature involves a balance between sensitivity (i.e. the number of relevant records identified by a search strategy as a proportion of the total number of relevant records) and precision (i.e. the number of relevant records retrieved by a search strategy as a proportion of the total number of records retrieved). A very sensitive search may not miss any relevant articles, but it may involve screening an infeasible number of records. An information specialist with expertise in bibliographic database searching for systematic reviews can advise on the trade-offs among alternative approaches for investigating each key question.

Although the general concepts and approaches to searching are similar for any systematic review, the methods employed for WHO guidelines involve unique aspects. It is important to search for studies from low- and middle-income countries in all WHO regions, as well as from more standard literature sources. Some journals are not well represented in Medline and

other commercial databases such as EMBASE and CAB Abstracts. Regional databases grouped under the general heading of the Global Index Medicus 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. A collection of bibliographic databases and other information sources relevant to low- and middle-income countries can be obtained on the Norwegian Satellite of the Cochrane Effective Practice and Organisation of Care Group website (*15*).

Validated search filters may be useful in limiting the number of records (i.e. increase precision). Filters are available for study designs and for specific databases. An example is the Cochrane Highly Sensitive Search Strategy for identifying randomized controlled trials in Medline (*16*). The use of filters requires caution, as each filter has limitations in terms of sensitivity and precision. An information specialist familiar with these tools can provide advice.

If resources allow, the systematic review should, in most cases, include a search for evidence in WHO's six official languages. The approach used to identify and synthesize these studies should be specified in the terms of reference for the contracted systematic review team, and adequate resources should be available for the translation of studies published in all relevant languages other than English. Many databases exist in other languages – the China Academic Journals Full-text Database is one example (*17*). However, a native speaker should generally perform the searches of these databases.

A search for grey literature (studies not indexed in commercial bibliographic databases) should be carefully considered for all systematic reviews underpinning WHO guidelines. Unpublished data or data that are difficult to retrieve can be critically important to inform a WHO guideline. Other important sources of information are study registries, such as those in the international Clinical Trials Registry Platform (ICTRP) sponsored by WHO (*18*), which includes ClincialTrials.gov (*19*).

In addition to searching the relevant bibliographic databases and the grey literature, the systematic review team should review the reference lists of existing guidelines and of high-quality reviews on the topic of interest.

It is essential that the systematic review team use reference management software to organize the records obtained from searching. WHO staff should be familiar with such software and use it to formulate the reference lists in the final guideline document. The systematic review contractor should supply a copy of the reference library in electronic format (without attached pdfs of included publications) as a final deliverable.

Step 3: select studies and abstract data

As the first stage in selecting relevant studies, records retrieved from the bibliographic databases and from other sources are recorded and assessed for eligibility by examining their titles and abstracts only. This assessment is performed in accordance with the inclusion and exclusion criteria developed a priori. The full text of articles found to be potentially relevant on the basis of their titles and abstracts is retrieved and examined in light of the same inclusion criteria in the second stage of study selection. Data from eligible studies are then extracted into pre-defined templates that generally include the characteristics of the study design and of the population, intervention, comparator and outcomes. To ensure accuracy, at least two people should independently assess the eligibility of the studies identified and extract data from study reports.

The search strategy and results should be carefully documented. This involves reporting the databases searched, the strategy used to search each database, the total number of citations retrieved from each database, and the reasons for having excluded some publications after reviewing the full text.

The flow of articles throughout the search and up to the final cohort of included studies should be depicted with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (*20*), which includes the number of excluded articles and the reasons for any exclusions at the full-text screening stage (*21*). The PRISMA diagram is included in an annex to the report of the systematic review or within the text of the report.

Step 4: assess risk of bias of the individual studies

Each study included in a systematic review should be assessed for risk of bias. The main types of bias for intervention studies include selection, attrition, performance and detection bias (*22*). Nonrandomized studies have similar sources of bias, the most important being selection bias, which results in the presence of systematic differences in the baseline characteristics of the groups being compared. Chapter 9 provides additional details on how to assess the risk of bias at the individual study level.

The following resources may be helpful when determining which instrument to use to assess the risk of bias of studies included in a systematic review:

- Cochrane risk of bias tool for randomized controlled trials (*23*);
- Cochrane risk of bias tool for nonrandomized studies of interventions (*24*);

- Quality assessment tools project report, Canadian Agency for Drugs and Technologies in Health, 2012 (*25*);
- Assessing the risk of bias of individual studies in systematic reviews of health care interventions, the Agency for Healthcare Research and Quality, USA, 2012 (*26*).

Step 5: synthesize and report the results and conclusions

The findings of the systematic review may be synthesized in a narrative manner or quantitatively with a pooled estimate of effect for certain outcomes (in other words, a meta-analysis). The review should describe how data were handled and why a given approach to synthesis was taken for each outcome.

The results of the systematic reviews are presented to the GDG, along with an assessment of the confidence in the estimates of effect for the critical and important outcomes (i.e. quality assessment using GRADE [see Chapter 9]). This presentation generally occurs at an in-person meeting during which recommendations are formulated (see Chapter 10). The two most common ways of presenting the evidence are briefly described here, but for more details, please consult the Cochrane handbook (*22*).

If the data extracted from the systematic review meet certain requirements (the most important one being a high level of homogeneity in study design and in population, intervention, comparator and outcomes across studies), then the data can be combined across comparable studies in a meta-analysis. A meta-analysis is a statistical method used to summarize the quantitative results of independent studies, providing a summary estimate of effect with a confidence interval. By combining information from all relevant studies, meta-analyses can provide more precise estimates of the effects of an intervention than estimates derived from the individual studies included within a review. The results of a meta-analysis are usually displayed in tables and in a figure called a forest plot.

If a meta-analysis is not feasible or appropriate due to heterogeneity or to the qualitative nature of the data, the evidence should be presented in the form of a narrative synthesis. The method used to produce this synthesis needs to be specified before the work begins and must be followed rigorously to avoid introducing bias. Studies can be stratified if many of them fulfil pre-specified criteria. Patterns of effect can be elicited and described in a narrative fashion. In addition, evidence on important subgroups (e.g. poor countries) should be reviewed separately from the overall body of the evidence to help guide decision-making. Subgroup and sensitivity analyses

should be specified a priori in the review protocol to prevent the introduction of bias with post hoc analyses.

Once the data are synthesized across studies, the systematic review team uses GRADE to assess the quality of the body of evidence across studies for each outcome (see Chapter 9).

Step 6: prepare the final report of the systematic review

Systematic reviews conducted in developing WHO guidelines should be reported in a standard format using the PRISMA reporting guidelines (*20*). For some purposes, the methods and results section of the review may be all that WHO will require; for other purposes it may require a full report with detailed introduction and discussion sections. GRADE evidence profiles should generally be part of the final report of the systematic review (see Chapter 9). This report should be sufficiently detailed for WHO staff and the GDG to know exactly what was done and what the findings are. However, excessive detail in the text of the characteristics and findings at the individual study level is unnecessary.

8.4 Including qualitative research

Qualitative data can address certain types of key questions that cannot be answered by quantitative research methods, such as "how" and "why" a given intervention produces its effects. Qualitative evidence can help explain, interpret and apply the quantitative results of a systematic review. A synthesis of the findings from qualitative research can be done as part of the scoping of the guideline and can help to define and refine the key questions. Qualitative data may also inform the interpretation of studies on both benefits and harm, address questions on the contextual barriers and facilitators to an effective intervention, and examine the values and preferences of the people receiving the intervention or experiencing the outcomes the intervention can affect. Qualitative research can also help bring to light what some subpopulations experience when they access and use services, including unintended social consequences, such as stigmatization or impoverishment. The GRC intranet site (*27*) provides additional guidance on when and how to use qualitative data to inform WHO guidelines. The Special Sections of the Cochrane handbook (*22*) also includes a chapter on qualitative research in Cochrane reviews (*28*) and the Cochrane Qualitative and Implementation Methods Group website provides additional information (*29*).

8.5 Including other types of data in the systematic review

Many other types and sources of data may be relevant to WHO guidelines. For example, for questions about substances that are potentially toxic to humans or the environment, mechanistic and animal studies may be relevant. Questions about cost and measures of economic efficiency, such as cost–effectiveness, may require information on prices from manufacturers or data from programme evaluations, or may involve modelling based on evidence-informed assumptions. The values and preferences of the people affected by a guideline may be explored by conducting a literature search or by generating primary data through surveys conducted as part of guideline development. Additional data on an intervention's feasibility or potential effects on equity, or on the infrastructural requirements it entails, may also be needed.

8.6 Updating existing systematic reviews

Updating an existing systematic review can be a less expensive and more efficient way to obtain the synthesis of the evidence required to underpin a recommendation. Updating a systematic review is a complex process, however. Like all the methods outlined in this chapter, it calls for significant experience and training if it is to be done well. Once a high-quality and applicable systematic review is identified, its authors should be contacted by the responsible technical officer to see if they are planning to update the review or are doing so already. This is particularly relevant for Cochrane reviews, which are periodically updated.

Updating is particularly difficult when the original authors are not available to update their review. The published report seldom contains all the information needed to perform an update, and the original authors may not respond to requests for clarification or additional (unpublished) information. Older reviews may not be up to current standards in the use of quality assessment tools and they are often missing an assessment of the confidence in the effect estimate for each outcome (see Chapter 9). Some of the data required to update a pooled estimate (meta-analysis) may not be in the public domain, a problem that further complicates the work of the team performing the update.

In spite of the difficulties mentioned, updating an existing systematic review is a reasonable option provided a high quality, fully reported review exists that uses current methods for identifying, appraising and synthesizing the evidence.

8.7 How to commission a systematic review

A new systematic review is needed if a relevant, high-quality systematic review is not identified. Or if a high-quality review is not up to date, it will need to be updated. If WHO lacks the appropriate expertise and resources (i.e. staff time) for these tasks, then it will have to commission them from an external source.

Performing systematic reviews correctly takes time, expertise and resources and the systematic review team needs to be carefully selected. Members of the Cochrane Collaboration may be commissioned to update an existing Cochrane review or to perform a review de novo. The cost of a systematic review varies widely and depends on the level of expertise required, the daily rate of the contractor and his or her team, and the estimate of the total number of days of work. The expertise and number of days depend, in turn, on the number and complexity of the key questions, the volume of literature on the questions of interest, the number of languages involved and the need for translation, among other factors. An estimate of the extent of the work involved can be made by WHO staff through an initial scan of the evidence, or potential contractors can be asked to include the results of a scan when responding to the request for proposals issued by WHO's responsible technical officer.

In commissioning a systematic review, the 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 supplier's systematic review protocol before the evidence search is started;
- request regular updates from the supplier on the progress of the review; and
- assess the quality of the deliverable(s).

Systematic review teams commissioned by WHO must have expertise in using GRADE and should produce GRADE evidence profiles (see Chapter 9). Separate individuals or teams should not be commissioned to perform the systematic review and the GRADE evidence profiles.

The terms of reference for a systematic review should be detailed and explicit to ensure that the deliverables meet the needs of the GDG. Detailed instructions and a template for developing terms of reference are available on the WHO GRC intranet site (*27*).

8.8 Systematic review quality standards

Regardless of the supplier selected, systematic reviews used to inform WHO recommendations must be developed according to the standards outlined by the Cochrane Collaboration in the Cochrane handbook (*22*). This handbook has a chapter dealing with reviews in public health and health promotion. The evidence-informed standards developed by the Institute of Medicine of the United States also provide a useful list of elements for high-quality reviews (*30*).

8.9 Systematic review reporting standards

PRISMA contains reporting standards for systematic reviews (*20*) and these standards should be the basis for reviews performed to develop WHO guidelines. The PRISMA-Equity 2012 extension (*31*) includes additional reporting standards related to health equity. The terms of reference for systematic review contractors should include a reference to these standards.

8.10 References

1. Glossary of terms in The Cochrane Collaboration. Version 4.2.5. Oxford: The Cochrane Collaboration; 2005. (http://www.cochrane.org/sites/default/files/uploads/glossary.pdf, accessed 11 November 2014).
2. Montori VM, Wilczynski NL, Morgan D, Haynes RB; Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ. 2005;330(7482):68. doi: http://dx.doi.org/10.1136/bmj.38336.804167.47 PMID: 15619601
3. Lee E, Dobbins M, Decorby K, McRae L, Tirilis D, Husson H. An optimal search filter for retrieving systematic reviews and meta-analyses. BMC Med Res Methodol. 2012;12(1):51. doi: http://dx.doi.org/10.1186/1471-2288-12-51 PMID: 22512835
4. The Cochrane Collaboration [website]. Oxford: The Cochrane Collaboration; 2014 (http://www.cochrane.org, accessed 11 November 2014).
5. Centre for Reviews and Dissemination [website]. York: University of York; 2014 (http://www.crd.york.ac.uk/CRDWeb/, accessed 11 November 2014).
6. Epistemonikos [website]. (http://www.epistemonikos.org/, accessed 11 November 2014).
7. McMaster Health Forum [website]. Hamilton: McMaster University; 2014 (http://www.healthsystemsevidence.org/, accessed 11 November 2014).
8. The Campbell Collaboration [website]. Oslo: The Campbell Collaboration; 2014 (http://www.campbellcollaboration.org/, accessed 11 November 2014).
9. Centre for Reviews and Dissemination [website]. York: University of York; 2013 (www.crd.york.ac.uk/prospero/prospero.asp, accessed 11 November 2014).
10. National Guideline Clearinghouse [website]. Plymouth: ECRI Institute; 2014 (http://www.guideline.gov, accessed 11 November 2014).

11. AMSTAR [website]. Ottawa: AMSTAR; 2012 (http://amstar.ca/Amstar_Checklist.php, accessed 11 November 2014).
12. Kung J, Chiappelli F, Cajulis OO, Avezova R, Kossan G, Chew L et al. From Systematic Reviews to Clinical Recommendations for Evidence-Based Health Care: Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance. Open Dent J. 2010;4:84–91. doi: http://dx.doi.org/10.2174/1874210601004020084 PMID: 21088686
13. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol. 1991;44(11):1271–8. doi: http://dx.doi.org/10.1016/0895-4356(91)90160-B PMID: 1834807
14. South CHL. The Literature Search Process: Guidance for NHS Researchers. Version 5. Shrewsbury and Telford Health Libraries; 2013 (http://library.sath.nhs.uk/wp-content/uploads/2013/09/researchguidance.pdf, accessed 11 November 2014).
15. Norwegian Satellite of the Cochrane Effective Practice and Organisation of Care Group. LMIC Databases [website]. Oslo: Norwegian Knowledge Centre for the Health Services; 2013 (https://epocoslo.cochrane.org/lmic-databases, accessed 11 November 2014).
16. Lefebvre CME, Glanville J. Chapter 6: Searching for studies. In: Higgins JTP, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.1.0. [updated March 2011]. Oxford: The Cochrane Collaboration; 2011. (http://handbook.cochrane.org, accessed 25 November 2014).
17. Database CKRI. KNS. [website]. Beijing: Tsinghua Tongfang Knowledge Network Technology Co.; 2014.(http://oversea.cnki.net/kns55/brief/result.aspx?dbPrefix=CJFD, accessed 11 November 2014).
18. International Clinical Trials Registry Platform (ICTRP) [website]. Geneva: World Health Organization; 2014 (http://www.who.int/ictrp/en/, accessed 11 November 2014).
19. ClinicalTrials.gov [website]. Maryland: U.S. National Institutes of Health; 2014 (https://clinicaltrials.gov/, accessed 11 November 2014).
20. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097. doi: http://dx.doi.org/10.1371/journal.pmed.1000097 PMID: 19621072
21. The PRISMA Statement [website]. Ottawa: Ottawa Hospital Research Institute; 2014 (http://www.prisma-statement.org/index.htm, accessed 11 November 2014).
22. Higgins JTP, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011 (http://handbook.cochrane.org, accessed 11 November 2014).
23. The Cochrane Collaboration. The Cochrane Collaboration's tool for assessing risk of bias. Manchester: Cochrane Oral Health Group; 2014 (http://ohg.cochrane.org/sites/ohg.cochrane.org/files/uploads/Risk%20of%20bias%20assessment%20tool.pdf, accessed 11 November 2014).
24. Sterne JACHJ. Reeves BC on behalf of the development group for ACROBAT-NRSI. A Cochrane Risk Of Bias Assessment Tool: for Non-Randomized Studies of Interventions (ACROBAT-NRSI), Version 1.0.0. Oxford: The Cochrane Collaboration; 2014 (http://www.riskofbias.info, accessed 11 November 2014).
25. Bai A, Shukla VK, Bak G, Wells G. Quality Assessment Tools Project Report. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2012 (http://www.cadth.ca/media/pdf/QAT_final.pdf, accessed 11 November 2014).

26. Viswanathan MBN, Berkman ND, Dryden DM, Hartling L. Assessing risk of bias and confounding in observational studies of interventions or exposures: further development of the RTI Item Bank. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 (http://www.ncbi.nlm.nih.gov/books/NBK154461/, accessed 11 November 2014).
27. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/, accessed 14 November 2014).
28. Noyes JPJ, Pearson A, Hannes K, Booth A. Chapter 20: Qualitative research and Cochrane reviews. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 5.0.1. [updated September 2008]. Oxford: The Cochrane Collaboration; 2008 (http://handbook.cochrane.org/, accessed 11 November 2014).doi: http://dx.doi.org/10.1002/9780470712184.ch20
29. Qualitative and Implementation Methods Group [website]. Oxford: The Cochrane Collaboration; 2014 (http://cqim.cochrane.org/, accessed 11 November 2014).
30. Institute of Medicine. Finding what works in health care: standards for systematic review. Washington (DC): The National Academies Press; 2011.
31. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al.; PRISMA-Equity Bellagio group. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. PLoS Med. 2012;9(10):e1001333. doi: http://dx.doi.org/10.1371/journal.pmed.1001333 PMID: 23222917

9. Evidence assessment

This chapter describes the steps involved in assessing the evidence gathered in the systematic review. Specifically, it explains how to quantify, for each key question (in PICO format), the risk of bias in the outcomes of the individual studies found; how to determine the quality of the evidence for each outcome relevant to decision-making across included studies and, finally, how to assess the quality of the evidence for all outcomes needed to formulate a recommendation. Chapter 10 will explain how the GDG will then use the entire body of evidence identified through the systematic review and other sources, including the assessments of its quality, to formulate recommendations. In this chapter and the next, we focus on key questions and evidence pertaining to interventions. WHO staff can find additional guidance on key questions related to diagnosis and diagnostic tests on the GRC intranet site (*1*).

WHO uses the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the quality of a body of evidence, and to develop and report recommendations. The GRADE working group began its work in 2000 with the goal of developing a common, sensible and transparent approach to grading the quality of evidence in support of recommendations in health care and assessing the strength of the recommendations. Currently many national and international guideline development groups, including WHO, use the GRADE approach. The GRADE working group continues to develop new methods, update and evolve existing methods, and monitor and evaluate the quality and utility of its approaches. As such, the methods continue to evolve as the evidence underpinning the approaches grows and experience with the methods expands.

Detailed information on GRADE is available to WHO staff on the WHO GRC intranet site (*1*) and to the public on the following sites:

- GRADE working group (*2*);
- GRADE online learning modules (*3*);
- GRADE series in the *Journal of Clinical Epidemiology (4*); and
- GRADE profiler software (GRADEpro) and GRADE applications in the GRADEpro Guideline Development Tool (G_2DT) (*5*).

9.1 What is the "quality of the evidence"?

Once the evidence has been retrieved and synthesized through a well conducted systematic review, its quality needs to be assessed. In the context of evidence syntheses, the GRADE working group defines the quality of the evidence as the "extent to which one can be confident that an estimate of the effect or association is correct" (*6*). In the context of guideline development, the quality of the evidence reflects the confidence that the estimates of an effect are adequate to support a particular decision or recommendation. Alternative terms to describe the quality of the evidence are "certainty of the evidence" or "confidence in the estimates of effect".

9.2 What are GRADE evidence profiles?

GRADE evidence profiles contain the assessment of the quality of the evidence and a summary of findings across studies for each important or critical outcome and each key question (in PICO format). The GDG uses these summaries as the basis for its discussions and to formulate recommendations (Table 9.1) Outcomes are listed in rows and the judgements made about the factors that determine the quality of the body of evidence are described briefly for each outcome, along with a summary of the effect estimates for each. Additional details are provided in explanatory footnotes. Further examples of evidence profiles are available to WHO staff on the GRC intranet site (*1*).

9.3 Who performs the quality assessment?

Assessing the evidence and developing evidence summaries are specialized tasks that should be performed by a methodological expert with experience using GRADE. Systematic review teams commissioned to support guideline development at WHO must have expertise in using GRADE to conduct evidence assessments.

Table 9.1. Example of a GRADE evidence profile: is cryotherapy preferable to no treatment in women with histologically confirmed cervical intraepithelial neoplasia?

Number of studies	Design	Quality assessment					Number of patients		Effect		Quality
		Limitations	Inconsistency	Indirectness	Imprecision	Other	Cryotherapy (%)	No treatment (%)	Relative (95% CI)	Absolute at 1 year (95% CI)	
Recurrence of CIN II-III[a]											
1	Randomized trials	No serious	No serious	Serious[b]	Very serious[c]	None	1/29 (3.4)	2/31 (6.5)	OR 0.52 (0.04 to 6.04)	30 fewer per 1000 (from 62 fewer to 230 more)	Very low
Cervical cancer[d]											
3	Observational studies	No serious	No serious	No serious	Serious[c]	None	3/222 (1.4)	9/285 (3.2)	–	20 more per 1000 (from 40 fewer to 70 more)	Very low
29	Observational studies	Serious[e]	No serious	No serious	Serious[f]	None	–	1.0[f]	RR 0.39[f]	6 fewer per 1000[f]	Very low

CI: confidence interval; CIN: cervical intraepithelial neoplasia; GRADE: Grading of Recommendations Assessment, Development and Evaluation; GRC: Guideline Review Committee; OR: odds ratio; RR: relative risk.

[a] Follow-up 12 months for randomized trials; 6—16 months for observational studies.

[b] Recurrence rates from pooled analysis of observational studies providing cryotherapy with no controls (with 30 000, 7200, and 21 000 women, respectively) show: 6% recurrence of all CIN types, 2% recurrence of CIN I, 4% recurrence of CIN II—III after cryotherapy. Heterogeneity among studies was high.

[b] All women with a diagnosis of CIN I.

[c] Few events with wide confidence intervals, including appreciable harm with cryotherapy.

[d] Follow-up mean: 6 months to 16 months.

[e] Based on studies with no control.

[f] In observational studies with no independent control, the relative risk reduction with cryotherapy is 86%; considering spontaneous regression of 28%, the relative risk reduction with cryotherapy is approximately 61% [86% − (28% × 86%)] corresponding to a relative risk of 0.39. Using 1% baseline risk without cryotherapy, the absolute risk reduction with cryotherapy is 0.61% over 1 year.

Note: The complete evidence profile is available to WHO staff on the GRC intranet site (*1*).

9.4 What is the role of WHO staff in assessing the quality of the evidence?

WHO staff responsible for developing guidelines need to understand the basic concepts governing the assessment of the quality of the evidence – and how to apply them in GRADE evidence profiles – to supervise commissioned work and advise GDG members. WHO staff who perform systematic reviews in house must have an in-depth knowledge of GRADE.

GDG members should be briefed on how to assess the quality of the evidence before meeting to formulate recommendations. Depending on their familiarity with guideline development and GRADE, this can be done with a combination of the online training modules, publications and presentations available through the sites listed earlier in this chapter. Additionally, guideline development meetings can start with an introduction to GRADE presented by the methodologist or the GRC Secretariat. This is particularly effective if the GDG have previously completed the introductory online module (*3*).

9.5 How is the quality of the body of evidence assessed for intervention studies?

GRADE categorizes the quality of the evidence as high, moderate, low or very low (Table 9.2). These quality ratings apply to the body of evidence for each outcome assessed for each key question and not to individual studies. A judgement on the risk of bias of each individual study included in the body of evidence is needed, however, to assess the quality domain of study limitations (see Section 9.5.1.1).

For key questions that address interventions (*6–8*), the starting point or baseline for rating the quality of the evidence is always the study design, broadly classified into two types:

- randomized controlled trials (RCTs); and
- nonrandomized trials and observational studies, including interrupted time-series analyses, cohort and case–control studies, cross-sectional studies and other types of studies, such as case series and case reports.

Although RCTs are the preferred source of evidence for measuring the effects of interventions, in many instances guideline developers must rely on information from nonrandomized trials or observational studies. This is particularly so when the guideline developers evaluate the potential harms of an intervention and the feasibility of implementing it in a real-world setting, or

Table 9.2. Quality of evidence in GRADE

Quality level	Definition
High	We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate	We are 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.
Low	Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low	We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

GRADE: Grading of Recommendations Assessment, Development and Evaluation.

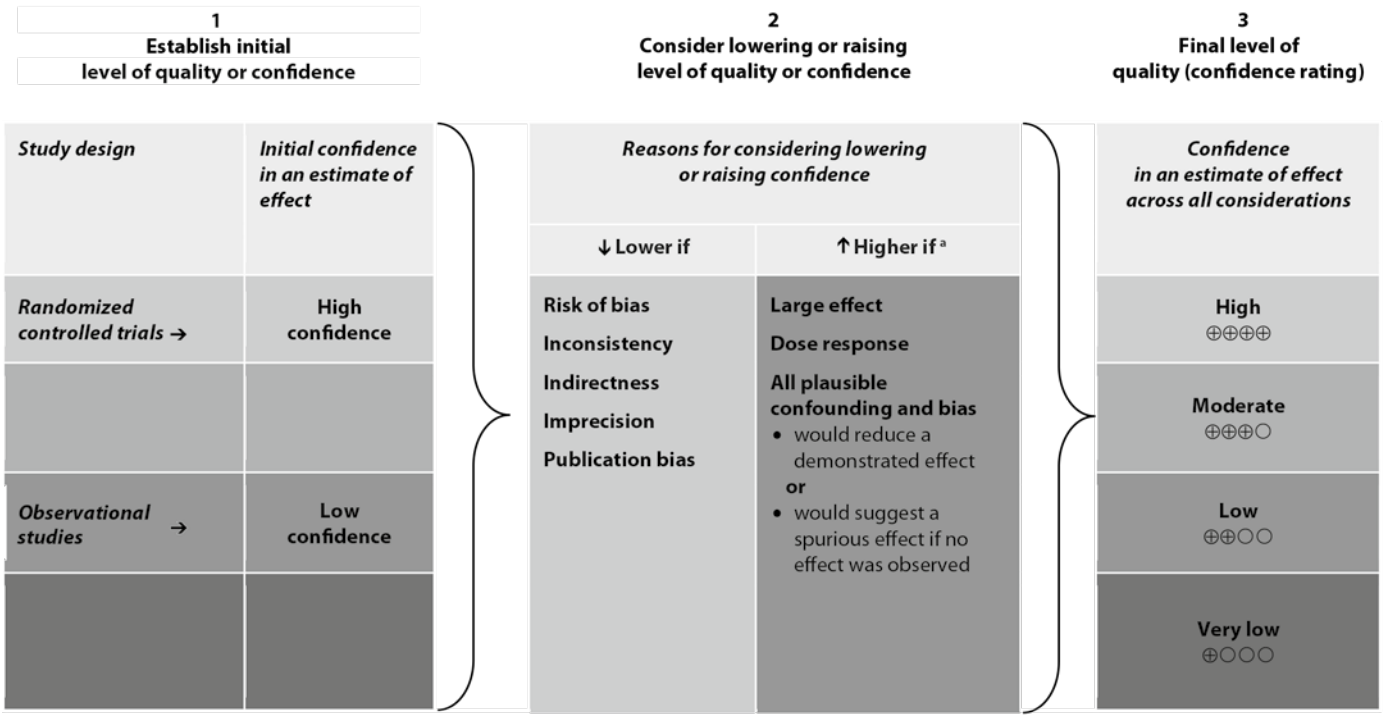
the barriers and facilitators to doing so. Thus, relevant data can be obtained from both RCTs and observational studies, and each type of evidence complements the other. Generally speaking, meta-analyses should not be used to pool estimates of effect from studies having different study designs. Rather, the results for the various types of studies should be presented separately and then synthesized across study designs in a narrative manner, with a focus on the highest quality evidence of relevance to the key question.

9.5.1 Five factors can lower the quality of the body of evidence for each outcome

A body of evidence based on RCTs is rated as being of high quality at the outset; evidence from nonrandomized trials or observational studies is rated as being of low quality. For both types of studies, these initial ratings can be adjusted in light of five factors: limitations in study design and execution; indirectness; imprecision; inconsistency; and publication bias (see Fig. 9.1). For a given body of evidence, the ratings are conducted for each outcome. This requires detailed knowledge of the individual studies included in the body of evidence. The factor "limitations in study design and execution" is assessed initially as the risk of bias at the level of the individual study and then across studies, while the other four factors that can lower the quality of the evidence are assessed for each outcome across all the included studies.

The criteria used to downgrade the quality of the evidence and the reason for applying these criteria should be provided in explanatory footnotes in the GRADE evidence profile, so that the reader can understand exactly how each

Fig. 9.1. The GRADE approach to rating quality of evidence for each outcome



GRADE: Grading of Recommendations Assessment, Development and Evaluation.

[a] Criteria for upgrading the quality are only applicable to observational studies without any reason for downgrading.

quality domain was assessed and why. The quality of evidence should not be rated in a mechanistic way. Although GRADE suggests criteria for rating each factor up or down, the overall quality of the evidence for each outcome is rated after considering all criteria.

9.5.1.1 Limitations in study design and execution

For RCTs, several criteria are used to assess any limitations in study design and execution (i.e. risk of bias). The following characteristics are the distinguishing features of the studies that yield the best quality evidence (*9*, *10*):

- random sequence generation;
- concealment of allocation to treatment group;
- blinding of participants and investigators, particularly if the outcomes were measured subjectively and thus may be subject to bias;
- reporting of data on all study participants, including attrition and exclusions from analysis; and
- complete reporting of all study outcomes that were specified a priori.

For observational studies, the main criteria depend on the study design and can be categorized as follows:

- application of appropriate eligibility criteria;
- use of an unbiased approach to measurement of exposure and outcomes;
- adequate control for confounding; and
- documentation and consideration of differential withdrawals of study participants across treatment groups.

Once the risk of bias has been assessed for each individual study, it is then summarized across studies for each outcome. Study limitations across the body of evidence for each outcome can be categorized as follows:

- No serious limitations – the majority of the studies in the review meet all the minimum quality criteria for the particular study design.
- Serious limitations – one of the minimum criteria for quality is not met by the majority of studies in the review. This results in a lowering of the overall quality rating by one level (e.g. "high" becomes "moderate" for RCTs or "low" becomes "very low" for observational studies).
- Very serious limitations – the risk of bias may have a strong influence on the estimate of effect and study limitations are present in the majority of studies contributing data on a given outcome in the review. This typically results in a lowering of the quality by two levels.

9.5.1.2 Inconsistency

Inconsistency is present when the results for a given outcome are not similar across studies (*11*). Some inconsistency will always be present. The magnitude of the differences in the direction and size of the effect observed in different studies, the significance of such differences, and whether any of these differences can be explained guide the decision as to whether important inconsistency exists. Inconsistency may arise from random variation or from differences across studies in the populations, interventions, comparators or outcomes. To explore the sources of inconsistency one may have to conduct sensitivity or subgroup analyses.

Important inconsistency is present if:

- the point estimates vary widely across studies;
- confidence intervals show minimal overlap or none; or
- in the case of a meta-analysis, the test for heterogeneity yields a statistically significant result and the I^2 value is high (*12*).

If all the results of the studies on a single outcome show overlapping confidence intervals, important inconsistency is not likely to exist. If the results are inconsistent – for example, when the results of the largest trial contradict those of the smaller trials – the overall quality of the evidence may be lowered by one level. If the results are very inconsistent, the evidence may be downgraded for the particular outcome involved by two levels. If only one study exists for a given outcome, inconsistency is not present and the quality of the evidence should not be lowered on this account. However, when there is only one study, the quality of the overall body of evidence may be lowered for reasons such as publication bias, imprecision or indirectness.

9.5.1.3 Indirectness

The characteristics of study results known as directness, generalizability, external validity, transferability and applicability refer to similar concepts that are encompassed in GRADE by the term "directness" (*13*, *14*).

Generally, there are two types of indirectness.

- Indirectness arises when the identified evidence differs in terms of population, intervention, comparator or outcome (PICO), from that pertaining to the question of interest to the GDG or to the authors of a systematic review. While all evidence is indirect to some degree, serious or very serious indirectness will cause the quality of the body of evidence for a given outcome to be rated down by one or two levels. If the measured effect applies to a population that is likely to differ from the target population of the guideline, the quality of the evidence should be rated down. Depending on the framing of the key question,

intermediate and surrogate outcomes are generally rated down because of indirectness, as they do not provide direct evidence on the health outcomes that ultimately matter to individuals and populations.

- Indirectness also occurs when no direct comparison of the intervention of interest with an alternative approach of interest (comparator) is available. For example, if the guideline panel is interested in comparing intervention A with intervention B but only studies comparing A with C and B with C are found, A can only be compared with B from indirect evidence. Such indirect evidence would be of lower quality than the evidence derived from a direct comparison of A and B and would be rated down.

9.5.1.4 Imprecision

In general, results are imprecise when studies include relatively few participants and few events and thus large uncertainty (i.e. wide confidence intervals) surrounds the estimate of effect (*8*, *15*). For GDGs, if the confidence interval for the pooled estimate of effect crosses the threshold established for making one decision versus another, then the body of evidence is imprecise for the particular outcome in question and the quality of the evidence is lower than it would be otherwise owing to uncertainty in the results.

Systematic review teams can use the 95% confidence interval for the pooled estimate of effect as the primary criterion for judging the presence of imprecision. Alternatively, they can use the optimal information size which is determined using a conventional calculation of the sample size needed for a single adequately powered trial to detect the minimum effect of interest (*15*). If the total number of patients included in a systematic review is less than the number of patients generated by a conventional sample size calculation, one should consider downgrading the quality of the evidence for imprecision.

GRADE defines the quality of the evidence somewhat differently for systematic reviews and for guidelines, especially in terms of the criteria for downgrading for imprecision. This is because GDGs need to consider the context when making a recommendation, whereas in systematic reviews judgements in connection with specific outcomes are usually made without regard to context. Thus, GDGs should carefully examine the systematic review authors' judgments about imprecision.

In formulating a recommendation, all outcomes are considered together, with attention to whether they are critical, or important but not critical for decision-making. The decision to downgrade the quality of the evidence for imprecision depends on the threshold established as the basis for a decision or a recommendation and on the trade-off between desir-

able and undesirable consequences. Determining the acceptable threshold involves an explicit judgement.

GDGs should use the following steps to decide whether to downgrade the quality of the body of evidence on a given dichotomous outcome due to imprecision:

- First, consider whether the boundaries of the confidence interval are on one and the same side of the decision-making threshold. Does the confidence interval cross the threshold for deciding to recommend or to not recommend an intervention? If the answer is yes, one would downgrade the evidence for imprecision, irrespective of where the point estimate lies.
- If the confidence interval does not cross the threshold, one needs to determine if the criteria for the optimal information size are met. If the answer is yes, one would not downgrade the evidence for imprecision.
- Alternatively, if the event rate is very low and the sample size is very large across studies for the outcome of interest (at least 2000 participants), one would not downgrade the evidence for imprecision.

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

The same logic for downgrading the quality of the evidence because of imprecision applies to continuous variables, where the optimal information size will require a sample size calculation for the continuous variable. If the sample size exceeds 400, imprecision is unlikely to be present.

9.5.1.5 Publication bias

Publication bias is the systematic underestimate or overestimate of the underlying beneficial or harmful effect of an intervention or exposure resulting from the selective publication of studies based on the study results; studies in which no effect is found are less likely to be published. Searches of trial registries and the grey literature can help to identify unpublished studies and thus minimize the risk of bias (see Chapter 7). The risk of publication bias may be assessed using funnel plots and appropriate statistical tests. The limitations of such tests should be noted, however: the existence or nonexistence of publication bias cannot be confirmed – it can only be suspected. When publication bias is suspected, the quality of the evidence should be downgraded by one level (*16*).

9.5.2 Three factors can increase the quality of the evidence from observational studies

Observational studies are considered low quality evidence from the outset by virtue of their design. If, and only if, there are no further limitations, such as additional risk of bias (if, in other words, there is no reason to downgrade their quality), then the assessor can consider upgrading the quality of the evidence according to three criteria: dose–response gradient; direction of plausible bias; and the magnitude of the effect. (It is important to note that factors that could reduce the quality of the evidence are always considered before those that could serve to upgrade it) (*17*).

9.5.2.1 Dose–response gradient

The presence of a dose–response gradient may generate greater confidence in the findings of observational studies and may thereby increase the quality of the evidence by one level.

9.5.2.2 Direction of plausible bias

On occasion, lack of adjustment for all plausible confounders (possible sources of bias) in the analysis of a rigorous observational study is likely to have resulted in an underestimate of an apparent treatment effect or an overestimate of the effect, if no effect was observed. In such situations, the quality of the evidence for the relevant outcome can be upgraded by one level. For example, if only sicker patients receive an experimental intervention or exposure, yet they still fare better, the actual effect of the intervention or exposure is likely to be even greater than the data suggest.

9.5.2.3 Magnitude of the effect

When a body of evidence from observational studies yields large or very large, precise and consistent estimates of the magnitude of a treatment or exposure effect, one can have greater confidence in the results. In such situations, the observational study design is unlikely to explain the apparent benefit or harm in its entirety. The larger the magnitude of an effect, the stronger the evidence becomes.

Decisions to upgrade the quality of the evidence because of large or very large effects should be based not only on the point estimate, but also on the precision surrounding that effect (i.e. the width of the confidence interval): one should rarely upgrade the quality of the evidence on the basis of a large effect when the confidence interval overlaps substantially with effects below

the threshold of clinical importance. One is more likely to upgrade the quality of the evidence because of a large or very large effect when:

- the effect is rapid;
- the effect is consistent across populations;
- the previous trajectory over time of an outcome is reversed; or
- the large magnitude of an effect is supported by indirect evidence.

The final assessment of the quality of the evidence for each outcome is determined by considering how the three factors covered in this section contribute to the overall certainty of the effect estimate (Fig. 9.1). A body of evidence from observational studies for which no reason to downgrade exists can be upgraded by one or two levels, depending on the overall assessment of the other factors described above.

9.6 How is the overall quality of the evidence determined?

Guideline developers review all the information from the systematic review and, if needed, reassess and make a final decision about which outcomes are critical and which ones are important in light of the recommendations that they aim to formulate. The GDG assigns an overall quality to the evidence, based on a combined rating of the quality of the evidence across all outcomes that are considered critical for answering the key question (i.e. for making a decision or a recommendation).

GDGs must determine the overall quality of the evidence across all the critical outcomes for each recommendation. Because quality of evidence is rated separately for each outcome, the quality frequently differs across outcomes. If the quality of the evidence is the same for all critical outcomes, then this is the level of quality that applies to all of the evidence supporting the answer to the key question. If the quality of the evidence differs across critical outcomes, the overall confidence in effect estimates cannot be higher than the lowest level of confidence in the effect estimates for an individual outcome. Therefore, the lowest quality of the evidence for any single critical outcome determines the overall quality of the evidence.

The judgement about which outcomes are critical for decision-making may depend on the evidence. Although it happens rarely, the overall quality of the evidence is sometimes not based on the outcomes judged critical at the beginning of the guideline development process. There can be two reasons for this:

- An outcome turns out not to be critical for decision-making (e.g. a particular adverse event considered critical at the outset of the guideline

development process turns out to be very infrequent and of questionable relevance to the intervention).

- If there is higher quality of evidence for some critical outcomes that is sufficient to support a recommendation, then there is no reason to downgrade the overall quality of evidence because of lower quality for another critical outcome when the recommendation would not change.

9.7 Is GRADE applicable in all situations?

Confusion and concerns exist as to whether GRADE is applicable to all situations in which a GDG issues a recommendation regarding an intervention. The strength of the GRADE approach rests on the use of a structured framework for the assessment of the quality of the evidence and on the requirements that processes be explicit and judgements transparent. GRADE has been applied to a wide range of interventions in clinical medicine, public health and health policy (*18*). While the ease of applying the GRADE approach will vary according to the type of evidence being assessed, the circumstances in which GRADE cannot be applied are rare. In such cases, the rationale and an alternative approach must be provided in the planning proposal and approved by the GRC. The GRADE approach is also applicable to questions regarding diagnostic tests (WHO staff can consult the GRC intranet site (*1*) for additional guidance) and for questions of prognosis or baseline risk. The GRADE working group has not yet finalized methods for assessing other types of questions and evidence, however. For example, the approach for assessing the quality of evidence required to answer questions regarding environmental exposures and economic efficiency is under development, as are the methods for assessing the quality of evidence based on network meta-analyses and modelling.

9.8 References

1. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/, accessed 14 November 2014).
2. GRADE working group [website]. The GRADE working group; 2014 (http://www.gradeworkinggroup.org/, accessed 12 November 2014).
3. GRADE Online Learning Modules [website]. Ontario: McMaster University; 2014 (http://cebgrade.mcmaster.ca/, accessed 12 November 2014).

4. GRADE series in the Journal of Clinical Epidemiology [website]. Maryland Heights: Elsevier Inc.; 2014 (http://www.jclinepi.com/content/jce-GRADE-Series, accessed 12 November 2014).
5. GRADEpro Guideline Development Tool [website]. Ontario: McMaster University and Evidence Prime Inc.; 2014 (http://www.guidelinedevelopment.org, accessed 12 November 2014).
6. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401–6.
7. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al.; GRADE Working Group. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ. 2008;336(7653):1106–10. doi: http://dx.doi.org/10.1136/bmj.39500.677199.AE PMID: 18483053
8. Schünemann HBJ, Brozek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. [Updated October 2013]. The GRADE Working Group (http://www.guidelinedevelopment.org/handbook/, accessed 12 November 2014).
9. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). J Clin Epidemiol. 2011;64(4):407–15. doi: http://dx.doi.org/10.1016/j.jclinepi.2010.07.017 PMID: 21247734
10. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al.; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011 Oct 18;343:d5928. doi: http://dx.doi.org/10.1136/bmj.d5928 PMID: 22008217
11. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al.; GRADE Working Group. GRADE guidelines: 7. Rating the quality of evidence–inconsistency. J Clin Epidemiol. 2011;64(12):1294–302. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.03.017 PMID: 21803546
12. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557–60. doi: http://dx.doi.org/10.1136/bmj.327.7414.557 PMID: 12958120
13. Schünemann HJ, Tugwell P, Reeves BC, Akl EA, Santesso N, Spencer FA, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Res Synth Methods. 2013;4(1):49–62. doi: http://dx.doi.org/10.1002/jrsm.1078
14. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al.; GRADE Working Group. GRADE guidelines: 8. Rating the quality of evidence–indirectness. J Clin Epidemiol. 2011;64(12):1303–10. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.04.014 PMID: 21802903
15. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence–imprecision. J Clin Epidemiol. 2011;64(12):1283–93. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.01.012 PMID: 21839614
16. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence–publication bias. J Clin Epidemiol. 2011;64(12):1277–82. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.01.011 PMID: 21802904
17. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al.; GRADE Working Group. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol. 2011;64(12):1311–6. doi: http://dx.doi.org/10.1016/j.jclinepi.2011.06.004 PMID: 21802902
18. Akl EA, Kennedy C, Konda K, Caceres CF, Horvath T, Ayala G, et al. Using GRADE methodology for the development of public health guidelines for the prevention and treatment of HIV and other STIs among men who have sex with men and transgender people. BMC Public Health. 2012;12(1):386. doi: http://dx.doi.org/10.1186/1471-2458-12-386 PMID: 22640260

10. Developing recommendations

Once the evidence has been identified and synthesized and its quality assessed, the GDG, with the support of the steering group, has the critically important task of formulating recommendations based on the evidence. GRADE provides a framework to accomplish this task, with explicit consideration of specific factors that may affect the direction and strength of each recommendation. This chapter outlines those factors and provides guidance on how to assist the GDG in formulating clear, actionable recommendations.

While formulating recommendations, the GDG needs to take a particular perspective – the health system's or the health service user's, for example – about which it needs to be clear, since the perspective adopted influences, among other things, the impact of an intervention on resources. In the context of WHO as a global public health agency, most recommendations are based on a health systems perspective (as opposed to an individual patient perspective). The particular perspective of each recommendation should be agreed upon and communicated to the GDG by the steering group, before recommendations are formulated.

10.1 Factors that determine the direction and strength of recommendations

Four main factors determine the direction and strength of a recommendation in public health (*1–5*):

- the confidence in the estimates of effect of the evaluated evidence (i.e. the quality of the evidence) (see Chapter 9);
- values and preferences related to the outcomes of an intervention or exposure;
- the balance of benefits and harms; and
- resource implications.

In addition to these four main factors, several other considerations are important when formulating public health, health system and health policy recommendations. These factors overlap with the main four factors listed above, particularly with values and preferences. However, some GDGs find it useful to discuss the following as distinct considerations:

- the importance or priority of the problem being addressed;

- equity and human rights;
- acceptability; and
- feasibility.

Table 10.1 lists the detailed GRADE criteria that determine the direction and strength of a recommendation and describes how they influence the recommendation. These criteria are also included in the GRADE guideline development tool (*6*). To formulate a recommendation, the GDG considers each factor in turn and judges its importance and effect on the recommendation.

Table 10.1. Factors that determine the direction and strength of a recommendation

Factor	How the factor influences the direction and strength of a recommendation
Quality of the evidence	The quality of the evidence across outcomes critical to decision-making will inform the strength of the recommendation. The higher the quality of the evidence, the greater the likelihood of a strong recommendation.
Values and preferences	This describes the relative importance assigned to health outcomes by those affected by them; how such importance varies within and across populations; and whether this importance or variability is surrounded by uncertainty. The less uncertainty or variability there is about the values and preferences of people experiencing the critical or important outcomes, the greater the likelihood of a strong recommendation.
Balance of benefits and harms	This requires an evaluation of the absolute effects of both benefits and harms (or downsides) of the intervention and their importance. The greater the net benefit or net harm associated with an intervention or exposure, the greater the likelihood of a strong recommendation in favour or against the intervention.
Resource implications	This pertains to how resource-intense an intervention is, whether it is cost–effective and whether it offers any incremental benefit. The more advantageous or clearly disadvantageous the resource implications are, the greater the likelihood of a strong recommendation either for or against the intervention.
Priority of the problem	The problem's priority is determined by its importance and frequency (i.e. burden of disease, disease prevalence or baseline risk). The greater the importance of the problem, the greater the likelihood of a strong recommendation.
Equity and human rights	The greater the likelihood that the intervention will reduce inequities, improve equity or contribute to the realization of one or several human rights as defined under the international legal framework, the greater the likelihood of a strong recommendation.
Acceptability	The greater the acceptability of an option to all or most stakeholders, the greater the likelihood of a strong recommendation.
Feasibility	The greater the feasibility of an option from the standpoint of all or most stakeholders, the greater the likelihood of a strong recommendation. Feasibility overlaps with values and preferences, resource considerations, existing infrastructures, equity, cultural norms, legal frameworks, and many other considerations.

10.2 Detailed criteria that should be considered when moving from evidence to recommendations

10.2.1 Quality of the evidence

The quality of the evidence – the degree of confidence in the estimates of effect – is a key factor in determining the strength of a recommendation (see Chapter 9). The higher the quality of the evidence, the more likely a strong recommendation is warranted. Uncertainty can surround both benefits and harms and if the corresponding evidence associated with one or more important or critical outcomes is of low or very low quality, a conditional recommendation is much more likely.

10.2.2 Values and preferences

The values and preferences of individuals and populations affected by the recommendation determine the strength of the recommendation. These values and preferences pertain to the relative importance people assign to the outcomes associated with the intervention or exposure; they have nothing to do with what people think about the intervention itself. For example, if a set of interventions for the prevention and treatment of HIV and other sexually transmitted infections including condom use and HIV testing and counselling, is targeted to men who have sex with men, the relevant values and preferences are those of this group, and their views on the potential benefits and harms (outcomes) of the intervention should be elicited and included in the decision-making process (*7*).

Data on the values and preferences of people affected by the recommendations can be quantitative (e.g. utilities of different health states) or qualitative (e.g. from surveys of patients or other stakeholders). Ideally, a systematic review is performed and data on values and preferences are identified. If relevant data cannot be identified in the published or grey literature, primary data collection may be undertaken if time and resources allow. A survey of the stakeholders who will be most affected by the recommendations may yield valuable insights into the relative importance they attach to the benefits and harms of an intervention.

If primary data are not available and cannot be collected, then the experience of GDG members can inform the discussion. The questions to ask are how important the health outcomes linked with an intervention seem to the people affected and how much variation across populations exists in this respect. When significant uncertainty surrounds people's values and preferences or when these vary substantially among and across populations and subpopulations, a conditional recommendation is more likely to be warranted. GDG members find it very difficult to accurately describe the perspectives of people affected by a recommendation. If GDG members are used as proxies for the people who will ultimately be affected by the recommendation under development, this approach must be performed and reported transparently and the results must be interpreted with caution.

10.2.3 Balance of benefits and harms

When considering the balance between an intervention's or exposure's benefits and harms, the GDG should examine the magnitude of the effects and the relative importance of the outcomes, including any disadvantages or inconveniences associated with the intervention (as informed by data on values and preferences). If the benefits clearly outweigh the harms, a strong recommendation is likely to be issued. On the other hand, a conditional recommendation is more likely if there is uncertainty about the balance of benefits versus harms or when the anticipated net benefits are small.

10.2.4 Resource implications

Resource considerations are generally included in the formulation of recommendations, unless there is a deliberate and explicit decision to omit them. In considering resource implications, the GDG can be informed by a formal economic evaluation based on estimates collected during evidence retrieval and by modelling of cost–benefit and cost–effectiveness. If a full evaluation is not possible, resource implications can be anticipated and described in a qualitative manner. The more advantageous or disadvantageous the resource implications are, the greater the likelihood of a strong recommendation for or against the intervention. A conditional recommendation is more likely to be issued if the resource implications are uncertain (if, for example, data on

costs and how they vary across settings are not available, or if it has not been determined whether the net benefits of the intervention are worth the costs).

10.2.5 Priority of the problem

GDGs should consider – and be provided with – evidence about the burden of disease and the baseline risk, prevalence or incidence of the problem that is addressed. These considerations are particularly important when the GDG needs to prioritize across interventions.

10.2.6 Equity and human rights

Interventions have implications for the progressive realization of the right to health and the options given in a guideline can reduce or increase health inequities. Policies or programmes that increase equity or reduce inequity are generally assigned higher priority by GDGs than those that do not. GDGs must carefully consider how the intervention might affect equity and human rights and strive to formulate recommendations in line with WHO's core principles. The greater the likelihood that the intervention will reduce inequities or increase equity and the greater the accessibility of an option to its intended recipients, the greater the likelihood that a strong recommendation will be issued.

10.2.7 Acceptability

The lower the acceptability of an intervention to the most important stakeholders, the lower the likelihood that it will be recommended. Or, if it is recommended, the more likely it is that a strategy to address concerns about acceptability during implementation will be included in the guideline with the recommendations. Acceptability is affected by several factors, such as who benefits from an intervention and who is harmed by it; who pays for it or saves money on account of it; and when the benefits, harms and costs occur. Lack of acceptability may revolve around the distribution of the benefits, harms and costs of a given intervention; its undesirable short-term effects despite desirable long-term effects (benefits); or the ethical principles or judicial con-

siderations involved. The greater the acceptability of an option to all or most stakeholders, the more likely it is that a strong recommendation will be issued.

10.2.8 Feasibility

The less feasible an option is (i.e. the greater the barriers to its implementation), the lesser the likelihood that the intervention will be strongly recommended. Feasibility is influenced by the resources available, programmatic considerations, the existing and the necessary infrastructure and training, and many other factors.

10.3 Evidence-to-recommendation tables

Evidence-to-recommendation tables depict how the factors that determine the direction and strength of a recommendation inform the process of developing the recommendation. These tables enhance the transparency of the process, focus the discussions of the GDG and permit recording of the judgements made about each factor and how each one contributed to the recommendation. Because these factors overlap to some extent, the GDG should, at a minimum, document its judgements on the four main factors. The systematic review team and the methodologist can provide most of the information required for the evidence-to-decision tables before the GDG meets to formulate recommendations. Examples of these tables are available to WHO staff on the WHO GRC intranet site (*8*).

10.4 The strength of the recommendation

The strength of a recommendation expresses the degree to which the GDG is confident in the balance between the desirable and undesirable consequences of implementing the recommendation. When a GDG is very certain about this balance (i.e. the desirable consequences clearly outweigh the undesirable consequences), it issues a strong recommendation in favour of an intervention. When it is uncertain about this balance, however, it issues a conditional (or "weak") recommendation. Table 10.2 provides an aid to interpreting the strength of a recommendation.

Table 10.2. Interpretation of strong and conditional recommendations for an intervention

Audience	Strong recommendation	Conditional recommendation
Patients	Most individuals in this situation would want the recommended course of action; only a small proportion would not. Formal decision aides are not likely to be needed to help individuals make decisions consistent with their values and preferences.	Most individuals in this situation would want the suggested course of action, but many would not.
Clinicians	Most individuals should receive the intervention. Adherence to the recommendation could be used as a quality criterion or performance indicator.	Different choices will be appropriate for individual patients, who will require assistance in arriving at a management decision consistent with his or her values and preferences. Decision aides may be useful in helping individuals make decisions consistent with their values and preferences.
Policymakers	The recommendation can be adopted as policy in most situations.	Policy-making will require substantial debate and involvement of various stakeholders.

10.4.1 Strong recommendations

Strong recommendations communicate the message that the guideline is based on the confidence that the desirable effects of adherence to the recommendation outweigh the undesirable consequences. Strong recommendations are uncommon because the balance between the benefits and harms of implementing a recommendation is rarely certain. In particular, GDGs need to be cautious when considering making strong recommendations on the basis of evidence whose quality is low or very low.

10.4.2 Conditional or weak recommendations

Recommendations that are conditional or weak are made when a GDG is less certain about the balance between the benefits and harms or disadvantages of implementing a recommendation. Conditional recommendations generally include a description of the conditions under which the end-user should or should not implement the recommendation.

10.5 Formulating recommendations

The evidence-to-recommendation tables depict not only the evidence and judgements leading to a recommendation, but also the justifications for the recommendation's direction and strength. They also describe the subgroups considered, the process used to formulate recommendations (e.g. if voting took place) and key issues surrounding implementation, evaluation and monitoring (see Chapter 13) as well as research gaps.

10.5.1 Reaching agreement on recommendations

Draft recommendations can be prepared either by the steering group before the GDG meets to formulate recommendations, or during the meeting by the GDG itself. Under the leadership of the chair, the GDG reviews and discusses the GRADE evidence profiles presented by a representative of the systematic review team. In addition, the GDG considers other relevant criteria and evidence as discussed above. If the steering group has not drafted the evidence-to-decision tables before the GDG meeting, the GDG assesses each of the factors in Table 10.1 at its meeting and completes the evidence-to-decision tables. If the steering group has drafted these tables, they are reviewed at the meeting and each assessment is confirmed by the GDG.

The GDG needs to agree on the direction and the strength of the recommendations. WHO recommendations should be formulated by consensus (see Chapter 3). If the GDG is in strong disagreement with respect to the strength and direction of a recommendation, a conditional recommendation is appropriate.

10.6 Writing recommendations

Recommendations need to be clear and actionable, reflect the PICO format and contain an indication of their strength and of the quality of the evidence on which they are based. Outcomes should generally not be mentioned in the recommendation to avoid the impression that only single outcomes are relevant or were considered.

The language of each recommendation is critically important. Wherever possible, it should be consistent across all recommendations in a guideline, which should be written in the active voice. GRADE recommends using terms or phrases

such as "should" or "strongly recommend" for strong recommendations, and "suggest" or "consider" for conditional ones. The use of confusing or ambiguous phrases should be avoided. For example, the phrase "not recommended" could be interpreted to mean either that no recommendation for or against the intervention has been made, or that the intervention should not be implemented. In the latter instance, "we recommend against intervention X…" is the preferred wording.

A recommendation should include a justification as to why it is strong or conditional and why it is for or against a given intervention. It should also contain a set of remarks explaining the conditions and context in which the recommendation applies and the points to bear in mind regarding implementation. The quality of the underlying body of evidence (high, moderate, low or very low) should be specified. Each recommendation should be linked to a summary of the evidence (e.g. a published systematic review or the systematic review in an online annex), the GRADE evidence profiles and the evidence-to-decision tables.

10.7 When not to make recommendations

In rare situations, a GDG may decide that the evidence is not sufficient to be able to formulate a recommendation. For instance, it may not be appropriate to make a recommendation when no evidence about the effects of an intervention is found, in which case the following statement could be included: "No recommendation can be made because evidence on the effectiveness (or harms) of intervention X was not identified."

Instead of providing a recommendation, it is sometimes appropriate to publish the results of the systematic review showing a lack of studies, or an overview of the available interventions, without accompanying recommendations. However, these approaches should be rare; not making a recommendation should be considered only in exceptional cases because in most situations guidance from WHO is needed, despite the dearth or lack of evidence. When no evidence is available, this should be made clear and the basis for any recommendations made should be presented. For example, in the absence of rigorous research evidence, case reports and a compilation of national or individual experiences may be the basis for formulating a (conditional) recommendation. This approach and its limitations must be clearly documented.

10.8 Future research

When gaps in the evidence are such that significant uncertainty exists with respect to the balance of an intervention's benefits and harms, such knowledge gaps should be described and questions and methods for addressing the gaps should be suggested. The research agenda can be prioritized, if appropriate. This section of the guideline document will inform future recommendations and guidelines. In formulating questions for future research, the GDG should be as specific as possible about what is needed and why. The PICO framework is also very useful for this purpose.

10.9 References

1. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol. 2013;66(7):719–25. doi: http://dx.doi.org/10.1016/j.jclinepi.2012.03.013 PMID: 23312392
2. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al.; GRADE Working Group. Going from evidence to recommendations. BMJ. 2008;336(7652):1049–51. doi: http://dx.doi.org/10.1136/bmj.39493.646875.AE PMID: 18467413
3. Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, et al.; GRADE Working Group. Incorporating considerations of resources use into grading recommendations. BMJ. 2008;336(7654):1170–3. doi: http://dx.doi.org/10.1136/bmj.39504.506319.80 PMID: 18497416
4. Schünemann HJ, Oxman AD, Akl EA, Brozek JL, Montori VM, Heffner J, et al.; ATS/ERS Ad Hoc Committee on Integrating and Coordinating Efforts in COPD Guideline Development. Moving from evidence to developing recommendations in guidelines: article 11 in Integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. Proc Am Thorac Soc. 2012;9(5):282–92. doi: http://dx.doi.org/10.1513/pats.201208-064ST PMID: 23256172
5. Schünemann HJ, Fretheim A, Oxman AD; WHO Advisory Committee on Health Research. Improving the use of research evidence in guideline development: 1. Guidelines for guidelines. Health Res Policy Syst. 2006;4(1):13. doi: http://dx.doi.org/10.1186/1478-4505-4-13 PMID: 17118181
6. GRADEpro Guideline Development Tool [website]. Ontario: McMaster University and Evidence Prime Inc.; 2014 (http://www.guidelinedevelopment.org, accessed 12 November 2014).
7. Akl EA, Kennedy C, Konda K, Caceres CF, Horvath T, Ayala G, et al. Using GRADE methodology for the development of public health guidelines for the prevention and treatment of HIV and other STIs among men who have sex with men and transgender people. BMC Public Health. 2012;12(1):386. doi: http://dx.doi.org/10.1186/1471-2458-12-386 PMID: 22640260
8. Guidelines Review Committee (GRC) [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2014 (http://intranet.who.int/homes/ker/grc/, accessed 12 November 2014).

11. Rapid advice guidelines in the setting of a public health emergency

WHO develops guidelines in response to emergencies or to an urgent need, In such cases, a rapid response guideline may be issued within one day to several weeks, or a rapid advice guideline may be needed within a somewhat longer timeline (see Chapter 1). This chapter focuses on the second type of guideline; the methods for producing a rapid response guideline are under development by the GRC Secretariat.

The aim of this chapter is to provide detailed guidance on how to produce evidence-informed rapid advice guidelines in the context of a public health emergency. We focus on how guideline development can be accelerated, with particular attention to the planning phase, the execution of the evidence reviews that will inform the recommendations, and the mobilization of a GDG to formulate recommendations.

11.1 What is a rapid advice guideline?

In response to a public health emergency, WHO must provide global leadership and timely guidance in the form of an evidence-informed guideline produced within one to three months. The steering group for such a rapid advice guideline must follow all of the basic steps for guideline development as described in this handbook, but with modifications to meet the accelerated timeline.

11.2 What is a rapid review?

Because recommendations issued by WHO need to be based on the best available evidence, guidelines typically draw on evidence from well conducted and reported systematic reviews, whether they exist already or are undertaken to inform the development of the guideline. Producing systematic reviews can take varying amounts of time – from six months to more than one year – depending on the complexity of the topic and the available resources. In a public health emergency, conducting a conventional or standard systematic review de novo may not be feasible. Rapid reviews have emerged as a streamlined approach to synthesizing evidence quickly, typi-

cally for the purpose of helping decision-makers in health-care settings to make decisions expeditiously.

No universally accepted definition of a rapid review has been formulated yet. For the purposes of this guidance, we define a "rapid review" as a type of evidence review that is produced using accelerated and/or modified systematic review methods (*1*).

A rapid review retains the core values shared by the evidence synthesis community, including thoughtful scoping and formulation of the questions to be addressed by the review; transparency; replicability; careful assessment of the quality of the information incorporated into the review; efforts to minimize bias at every stage, and the clear presentation of information focused on the intended users' needs.

When evidence is needed to inform an emergent issue outside the timeline of a traditional systematic review, efficiencies may be adopted by limiting the scope of the review; limiting the outcomes of interest; adding more resources so that more reviewers work in parallel; and streamlining the processes so they are more efficient. When efficiencies are not enough, however, methods for delivering evidence syntheses in a timely manner to decision-makers should be used.

11.3 What is the evidence on the validity and impact of rapid reviews and rapid advice guidelines?

To date limited guidance has been provided on when and how to conduct a rapid review, and few data exist on how a rapid review performs compared with a standard systematic review. Available data do point to wide and increasing interest – and great variability – in the approaches used and level of reporting of the methods and results across organizations producing rapid reviews and rapid advice guidelines (*2–4*). Rapid reviews may be narrower in scope and more tailored to the needs of the commissioning body and intended users than standard systematic reviews (*2*). Most importantly, little is known about how rapid reviews compare with standard systematic reviews in terms of bias and credibility, the ways in which rapid advice guidelines are developed and implemented, and the impact of such guidelines on health outcomes.

11.4 Is a rapid advice guideline needed?

The first and most important question to consider is whether a rapid advice guideline is appropriate in the setting of the public health emergency at hand.

The need for, and appropriateness of, undertaking a rapid advice guideline, in light of the potential limitations of this approach, warrant consideration in the initial planning stage of a rapid advice guideline. Several issues need to be considered when deciding whether to develop a rapid advice guideline instead of a standard guideline, or to defer development of a guideline altogether.

11.4.1 What is the type of emergency and the risk to public health?

The first step is to examine the public health event that is driving the request for a rapid advice guideline. Emergencies may be classified as natural, technological, or conflict-related and may be of sudden onset (e.g. earthquakes, tsunamis, chemical crises) or more gradual onset (e.g. deteriorating situations in armed conflict, progressive disease outbreaks, drought or food insecurities). All types of emergencies can evolve into prolonged or protracted situations.

WHO and Member States use the manual entitled *Rapid risk assessment of acute public health events* to assess "any outbreak or other rapidly evolving situation that may have negative consequences for human health and requires immediate assessment and action" (*5*). Risk is characterized by level and is based on broad descriptive definitions of likelihood and consequences, represented in the form of risk matrices. The WHO Emergency Response Framework describes WHO's roles and responsibilities between the initial alert of an event and its subsequent classification based upon verification and risk assessment (*6*). WHO categorizes emergencies from Grade 1 (those with minimal expected public health consequences) to Grade 3 (those involving events in one or more countries and having significant public health consequences that call for a substantial regional response and/or international response).

11.4.2 Is the event novel?

WHO staff may consider producing a rapid advice guideline primarily in the face of one of two types of events:

- a new situation (e.g. a new strain of influenza; the Middle East respiratory syndrome [MERS] coronavirus; or an earthquake); or
- an event encountered previously but causing problems in a different context (e.g. a natural disaster; a change in disease pattern such as the Ebola virus disease in West Africa in 2014; or a prolonged armed conflict compounded by a disease outbreak).

If the event is not novel, high-quality guidelines that are relevant and applicable to the emergency at hand may already exist and a new guideline may not be needed.

11.4.3 Does uncertainty need to be urgently addressed?

Guidelines are indicated when there is uncertainty about what to do in a specific situation. Rapid advice guidelines are no different. WHO staff may be uncertain about what advice to provide, or there may be uncertainty in the field, with different players having different viewpoints and approaches. The question here is how quickly the uncertainty needs to be addressed.

11.4.4 What is the anticipated time frame for the event?

The intent of rapid advice guidelines is to provide urgently needed, evidence-informed recommendations that can be implemented within one to three months of the appearance of an unfolding situation or crisis. If from the outset an event is judged likely to persist beyond six months, a rapid advice guideline may not be suitable. It is important to weigh the impact of developing recommendations using standard processes and timelines, rather than produce a guideline that may be prone to serious limitations under an accelerated timeline.

11.4.5 Will the rapid advice guideline be rapidly implemented?

Rapid advice guidelines should only be developed if a mechanism is in place for disseminating and implementing them, and if implementing the recommendations is feasible in the context of the emergency. Various factors need to be carefully considered before embarking on the guideline development process. They include the existence of health systems and other infrastructure, the acceptability of the proposed intervention, the training requirements involved, and resource availability.

In summary, the decision to develop a rapid advice guideline should be informed by the above-mentioned factors, in combination with the overall judgment of the relevant technical units and WHO's senior officers. Before the development of a rapid advice guideline is undertaken, a clearly defined need must exist; the scope of the intended guideline must be manageable;

adequate resources, including expertise and funding, must be available; and a plan for dissemination and implementation must be in place.

11.5 How is a rapid advice guideline developed?

The basic steps for developing a rapid advice guideline are identical to those that apply to standard guidelines, as outlined in Chapter 1. There are some differences and additional considerations, however, when developing a rapid advice guideline: the steps involved are depicted in Table 11.1.

11.5.1 Consult the GRC Secretariat early

It is particularly important for the responsible technical officer of a rapid advice guideline to contact the GRC Secretariat early in the guideline development process. The GRC and the Secretariat will provide guidance as to whether a rapid advice guideline is appropriate in the situation at hand. In addition, the Secretariat can assist in developing the planning proposal and identifying a review team and a guideline methodologist. The GRC needs to approve the development of a rapid advice guideline at the planning proposal stage, or else the final guideline cannot be reviewed and assessed in an accelerated manner.

11.5.2 Formulate the various groups involved in guideline development

To determine the most appropriate scope and key questions, the steering group needs to be quickly assembled and all members fully engaged in the guideline development process. The chair and other potential members of the GDG should also be identified early and approached regarding their interest and availability within the set time frame. Ideally one or more key members of the GDG should be available immediately to provide feedback on the proposed scope and key questions. If the draft final guideline is to undergo peer review, potential reviewers – individuals or organizations – should also be identified early in the process and their declarations of interests should be collected.

Public health emergencies frequently pose ethical, social and legal dilemmas, and it is critically important to include individuals with expertise in these areas on the GDG as well as expertise in issues related to equity, gender

Table 11.1. Steps in the development of rapid advice guidelines

Primary contributor	Step	Key points for rapid advice guidelines
Phase 1. Planning		
Member State, WHO country office or public/private entity	Request(s) for guidance on a topic.	The request is in the context of a public health emergency.
WHO technical unit	Determine if guideline is needed; review existing WHO and external guidelines.	The technical unit must determine if a rapid advice guideline is needed, or if a standard or interim guideline would be more appropriate.
	Obtain approval for guideline development from the director of the technical unit.	The director of the technical unit should be involved in all important decisions.
	Discuss the process with GRC Secretariat and with other WHO staff with experience developing guidelines.	The planned guideline is discussed with the Secretariat when it first becomes a possibility.
	Form the steering group.	All relevant departments and regional offices are involved.
	Identify sufficient resources.	
	Determine the timeline.	Usually 1–3 months.
Steering group	Draft the scope of the guideline. Begin preparing the planning proposal.	The literature is scoped through a brief review. The guideline's scope must be narrow and feasible.
	Identify potential members of the GDG and the chair.	Issue invitations early; involve the GDG in the scope and key questions from the beginning.
	Obtain DOIs and manage any COIs among potential GDG members.	The process for rapid advice guidelines and standard guidelines is identical.
Steering group and the GDG	Formulate key questions in PICO format. Prioritize outcomes.	Key questions (in PICO format) include only those of the highest priority, and must be focused and narrow. Background questions are not addressed in a rapid advice guideline.
Steering group	Finalize the guideline planning proposal.	The process is the same as for a standard guideline.
GRC	Review and approve the planning proposal.	The GRC uses an accelerated process for review and disposition.
Phase 2. Development		
Systematic review team	Perform systematic reviews of the evidence for each key question.	The contractor needs to be identified from the outset and involved in the scoping and development of key questions: they can advise on what is feasible in the given time frame.
	Evaluate evidence quality for each important outcome, using GRADE as appropriate.	The process is the same as for a standard guideline.
Steering group	Convene meeting of GDG.	Meeting place and participants need to be identified at the beginning of the development process. The meeting has a similar format and agenda as for the development of a standard guideline.

continues ...

... continued

Primary contributor	Step	Key points for rapid advice guidelines
GDG	Formulate recommendations using the GRADE framework.	The general methods are the same as for a standard guideline. The evidence may be sparse, so the other factors that inform the recommendations must be transparent and based on evidence and on equity, human rights and gender considerations when possible.
Steering group	Draft the guideline document.	The document should be concise and tailored to the end-user.
External review group	Conduct external peer review.	External peer review is recommended for rapid advice guidelines but may not be feasible in some situations.
Phase 3. Publishing and updating		
Steering group and editors	Finalize the guideline document. Perform copy-editing and technical editing. Submit the final guideline to the GRC for review and approval.	This step will have to be performed in an accelerated manner. Editorial staff need to be identified early in the process.
GRC	Review and approve the final guideline.	The GRC uses an accelerated process for review and disposition.
Steering group and editors	Finalize the layout. Proofread.	This step needs to be accelerated and perhaps abbreviated from the standard processes.
	Publish (online and in print, as appropriate).	
WHO technical unit and programme manager	Disseminate, adapt, implement, evaluate.	
WHO technical unit	Update.	From the outset the technical unit must consider the likely shelf-life of the rapid advice guideline and whether a standard guideline will follow and when.

COI: conflict of interest; DOI: declaration of interest; GDG: guideline development group; GRADE: Grading of Recommendations Assessment, Development and Evaluation; GRC: Guideline Review Committee; PICO: population, intervention, comparator and outcome.

and human rights. Although the responsible technical officer may consider these issues peripheral when dealing with the health problem addressed by the rapid advice guideline (e.g. in the case of a disease outbreak), critical human rights issues often come to light and must be addressed in the initial stages of a response. An awareness of gender-related influences in the affected population groups can help target interventions more strategically.

11.5.3 Scope the rapid advice guideline and define the key questions

Once the need for a rapid advice guideline has been established and at least the core members of the steering group have been identified, work should begin on defining the scope of the guideline and on developing the key questions in PICO format. A rapid advice guideline will most likely provide recommendations on the benefits and harms of interventions: it is less likely to deal with matters concerning diagnosis, prognosis or risk thresholds.

WHO staff, with the assistance of an information specialist, should perform a rapid scoping exercise to provide a general sense of the extent and depth of the relevant literature. This is not a systematic search of all potentially suitable sources, but rather, a focused and pragmatic search for the best available, relevant literature, including high-quality systematic reviews and key primary studies. It is important to examine the resources most applicable to the topic under consideration, such as MEDLINE, The Cochrane Library, Google Scholar, the International Clinical Trial Registry Platform (ICTRP), ClinicalTrials.gov, the National Guideline Clearinghouse, and existing WHO guidelines. If WHO has published information or guidance in the early stages of the public health emergency, this information should also be examined. This scoping exercise, including synthesis of the evidence retrieved, should take no longer than one or two days. A brief summary of the results of the scoping exercise should be included in the planning proposal submitted to the GRC.

11.5.4 Prepare the planning proposal for the GRC

Like standard guidelines, rapid advice guidelines require that a planning proposal be submitted to the GRC for review and disposition. The content, level of detail and format are the same as those outlined in Chapter 4. Although the planning proposal takes time to prepare, it is nonetheless essential for the development of a rapid advice guideline, since it serves as a point of reference for the steering group, the systematic review team and the guideline methodologist. The planning proposal describes both the processes and procedures to be used to develop the guideline and provides a detailed outline of the methods that are planned for the rapid review and for translating the evidence into recommendations.

When operating under compressed timelines and with changing staff, as can occur in a public health emergency, it is paramount that the planning proposal be detailed and kept up to date. The rapid review process allows for post hoc adjustments to be made at certain times during the review process, based on the magnitude, complexity and quality of the evidence retrieved and deemed eligible. As a result, the review process is likely to be more fluid and iterative than in a standard systematic review. The planning proposal should thus be a living document, amended as needed and including the rationale for any changes. Complete and accurate documentation ensures transparency and greatly facilitates the drafting of the final guideline document.

11.6 How do rapid reviews compare with systematic reviews?

The core principles of evidence searching and retrieval, including transparency and reproducible and explicit methods, apply to rapid reviews. There are key differences, however, between searching in the context of a rapid review and of a standard systematic review. The rapid review has more restricted search criteria; looks to existing high-quality systematic reviews as the first line of evidence; involves a more targeted and iterative procedure for screening the literature and for data analysis and synthesis; places less emphasis on meta-analyses; and involves a concise and abbreviated report. In addition, in a rapid review the search process is more iterative and hierarchical, depending on the findings at each step. Thus, the types of publication and study designs included in a rapid review, and the bibliographic databases searched, may change as the evidence base is explored.

11.6.1 Types of rapid reviews

The reviews that underpin rapid advice guidelines may be categorized into four basic types: a full, standard systematic review performed rapidly; a review involving a variety of abbreviated methods and including only systematic reviews and existing guidelines; a review of primary studies and existing systematic reviews, or a review of primary studies only. Fig 11.1 outlines the types of rapid reviews, along with their similarities and distinguishing features.

Fig. 11.1. Types of rapid review and their characteristics

Types of rapid reviews	Traditional systematic review (conducted rapidly)	Rapid review of systematic reviews	Rapid review of systematic reviews plus primary studies	Rapid review of primary studies only
Time frame	Up to 16 weeks	Up to 12 weeks	Up to 12 weeks	Up to 12 weeks
Methods				
Question types	Clinical effectiveness, clinical efficacy; safety/harms; diagnostic or screening test accuracy; cost–effectiveness; health systems, education, public health, policy/programmes, or prevention interventions			
Number of questions	Multiple (targeted and narrow in scope)	1 primary question (targeted)		
Literature search limits	No restrictions	Restrictions (e.g. date, study design, language, setting)		
Number of databases searched	No restrictions (comprehensive)	2-3 databases		
Use of systematic reviews	Systematic reviews and primary studies	Systematic reviews only	Systematic reviews plus primary studies	Primary studies only
Grey literature	Yes, as appropriate	Limited (e.g. key websites)		
Screening	2 reviewers	2 reviewers: second review may only review excluded studies at title/abstract phase of screening		
Types of study designs included	RCTs and observational studies as appropriate	Systematic reviews and guidelines only (highest quality)	Systematic reviews and guidelines plus RCTs or observational studies (highest quality)	RCTs or observational studies only (highest quality)
Data extraction	Complete verification	One reviewer with selected verification		
Outcomes	Restricted to 4 critical outcomes or fewer	2-4 critical outcomes only: more if data are available		
Assessment of risk of bias at the individual study level	Yes (using validated instruments when available)			
Assessment of the quality of the body of evidence	GRADE for critical outcomes as appropriate	Reliance on GRADE as reported in the included systematic review(s); or perform de novo for each systematic review		GRADE for critical outcomes as appropriate

GRADE: Grading of Recommendations Assessment, Development and Evaluation; RCTs: randomized controlled trials.

11.7 Steps in the rapid review process

11.7.1 Select the types of evidence to be collected and identify the appropriate sources

Depending on the nature of the question being asked, the purpose of the rapid review and the magnitude of the literature on the topic, various types of evidence may be targeted. In most cases the emphasis will be placed on locating and summarizing evidence from relevant and high-quality "off-the-shelf' systematic reviews or guidelines. In the absence of systematic reviews, high-quality and/or recent primary studies may also be included. Landmark papers may also be included for reference. Finally, quasi-experimental and/or observational studies of high quality may also be considered, depending on the key question and the volume of the available evidence.

Additional databases, including topic-specific and regional databases, may be examined. However, this will depend on the topic under review and on the accessibility of these additional sources. In most rapid reviews, searching involves no more than two to three of the most relevant databases; additional databases may be considered if the rapid review moves to a full systematic review at a later stage.

11.7.2 Develop search strategies

Concepts need to be well defined with search terms that include both medical subject headings (MeSH) and text words. It is important that the draft search strategy be reviewed by:

- at least one other member of the rapid review team;
- one or more content experts, such as WHO technical staff; and
- a WHO information specialist with expertise in systematic reviews.

Validated search filters may be very useful when performing a rapid review (see Chapter 8). The search filters of most interest are those that comprise index terms relating to study type and design, such as *randomized controlled trial* (*RCT*), *systematic review* or *meta-analysis*. Applying study design filters will facilitate citation screening by identifying the highest quality evidence first.

In a standard systematic review, the aim is to maximize sensitivity (recall) while also attempting to maximize specificity (precision). For the purposes of a rapid review, however, the aim may be to maximize precision

(ability to exclude irrelevant articles) rather than recall (the ability to identify all relevant articles). The approach selected, the rationale for using it and its potential limitations should be reported in the review report and in the guideline document.

11.7.3 Search for grey literature

A search for grey literature should be considered but needs to be limited. For example, the websites of relevant organizations can quickly provide relevant data. An information specialist at WHO's headquarters library should be consulted regarding the proposed strategy for searching the grey literature.

11.7.4 Common search restrictions

Several common eligibility restrictions should be considered when developing search strategies in the context of a rapid review (Box 11.1) (*1*). Potential restrictions should be discussed among steering group members and with the review team information specialist to optimally balance precision and recall of the search strategies.

Search strategies for a rapid review will generally have language restrictions, since translation is time-consuming. The languages of inclusion should be carefully selected based on the guideline topic. For example, a rapid review on personal protective equipment for health workers in Ebola treatment centres (*7*), engendered by the Ebola virus disease outbreak in West Africa in 2014, included only literature in English and French owing to the geographic distribution of the outbreak and the opinion of experts that most of the relevant literature was in those two languages. Citations in non-selected languages are generally included during the study identification phase but may be excluded from further analyses if the full-text article is difficult to access or not enough time or resources are available for translation.

Search restrictions should be noted in the planning proposal, in the methods section of the rapid review report, and in the guideline document. The review team should provide – as an appendix to the rapid review report – a list of potentially relevant titles and abstracts identified during the search but published in languages excluded from the analysis.

Box 11.1. Common search restrictions for rapid reviews

Sources

- Usually search no more than two or three key bibliographic databases.
- If time and resources permit, additional resources may be added.

Language

- Language restrictions are frequently applied, as translation is time-consuming and resource-intensive.
- Limitations by language of publication need to be assessed for each topic, with consideration given to the distribution of the disease or condition being addressed and the likely languages of the relevant publications.

Accessible studies

- Publication status is limited to full text only (abstracts are not usually included).
- To maximize efficiency, articles should be electronically available through ejournal subscriptions available to the rapid review team.
- Articles should be purchased directly from a journal only under special circumstances, namely when the paper is deemed essential and is not available through other means.

Grey literature

- The utility of the grey literature is assessed for each topic.
- Websites of relevant organizations may be examined, depending on the subject under review.

Year (search dates)

- Publication dates are limited (e.g. only the most recent decade is searched).
- When applying a year limit, a rationale for the time frame must be provided.

Region

- Restrictions may be placed on the geographical locations of the included studies.
- A rationale should be provided to explain why citations from certain regions, rather than from the global literature, are targeted.

11.7.5 Other strategies for identifying relevant literature

In the context of a new situation or event, the best (and perhaps only) data might come from the acquisition and analysis of emerging information in real time. In the Ebola virus disease outbreak in West Africa in 2014, essentially no relevant comparative data were obtained through a systematic review of the published literature. WHO staff therefore rapidly devised and implemented a survey of repatriated health-care workers to gather information on these workers' experiences with various types of personal protective equipment.

If time permits, the reference lists of all included studies should be scanned for additional relevant studies to ensure that key publications have not been overlooked.

11.7.6 Screening and study selection

Standard systematic review methods apply to the process of screening the records retrieved from bibliographic databases and other searches. Records should be imported into reference management software in the usual manner to facilitate record management and citation screening and management during the drafting of the guideline document.

As for a standard systematic review, study selection involves a two-step process. First, two people independently screen titles and abstracts and all potentially relevant records, as well as records lacking enough information to determine their eligibility (e.g. no available abstract). Another reasonable approach is for one person to review all titles and abstracts, while a second person examines only the citations that were excluded by the first one. Second, two reviewers examine the full-text publications to determine their eligibility. As for a standard systematic review, consensus on the studies to be included and excluded at this stage should be achieved, with involvement of a third reviewer if necessary.

11.7.7 Use a stepwise approach with emphasis on higher levels of evidence

To keep the scope of the rapid reviews within the bounds dictated by timelines and resources, initially the evidence is often limited to what is found in systematic reviews. If primary studies, such as RCTs, will be included, this needs to be justified in the planning proposal and reflected in the timelines and budget. Further restrictions (e.g. by outcomes or study quality criteria) may be considered to accommodate the inclusion of primary studies. Excluding studies that do not report the outcomes of interest may lead to bias from selective outcome reporting. However, this is a trade-off that must be considered given the condensed timeline and finite resources.

11.7.8 Obtaining publications

Records that are not available electronically are generally excluded because the timeline of a rapid advice guideline is not compatible with the delays

involved in inter-library loans. WHO information scientists should be contacted before concluding that a given publication is not available, since WHO may have ready access to the information. Even if the full text cannot be obtained or translated, the abstract may provide valuable information, particularly when the evidence is sparse.

11.7.9 Data extraction and evidence synthesis

Before starting data extraction, it is important to carefully consider what specific information the GDG will need to make its decisions so that only essential information is extracted. A standard extraction form should be developed to facilitate accurate and reliable data collection, and it should be pilot-tested. One reviewer usually extracts the data and a second one verifies each record. If full verification is not feasible, a random sample of at least 10% of the included studies should be independently checked to provide some measure of quality assurance. If inter-rater agreement is low, all the data extracted will have to be reviewed and verified.

11.7.10 Assess the risk of bias at the individual study level

The review team should assess the risk of bias for individual studies, just as in a standard systematic review. This step is necessary to properly interpret the results of a rapid review and to use GRADE to assess the strength of the body of evidence for each outcome. For rapid reviews particularly, the assessment of the risk of bias may be used to select the studies included in the review, once initial criteria based on study design have been applied.

11.7.11 Evidence synthesis

The rapid review report includes, first, a summary of the overall findings of the literature search and of the general characteristics of the study populations and interventions. A PRISMA flow diagram (*8*) gives the reader an overview of the rapid review process and a snapshot of the evidence identified. All rapid reviews should include a narrative summary of the evidence found, generally organized around the PICO framework.

To begin understanding the data on outcomes, it may be useful to start by classifying the study findings as either favouring the intervention, favouring the control, or inconclusive. Once the cohort of included studies is finalized for each outcome, outcome data can be extracted, together with effect estimates (e.g. odds ratios, relative risks, mean differences, or summary effect [i.e. a meta-analysis], and their corresponding confidence intervals). This implies another round of data extraction to retrieve additional data. Extracting the results in a stepwise and targeted manner can save time when it is unclear at the outset what data will be the most relevant.

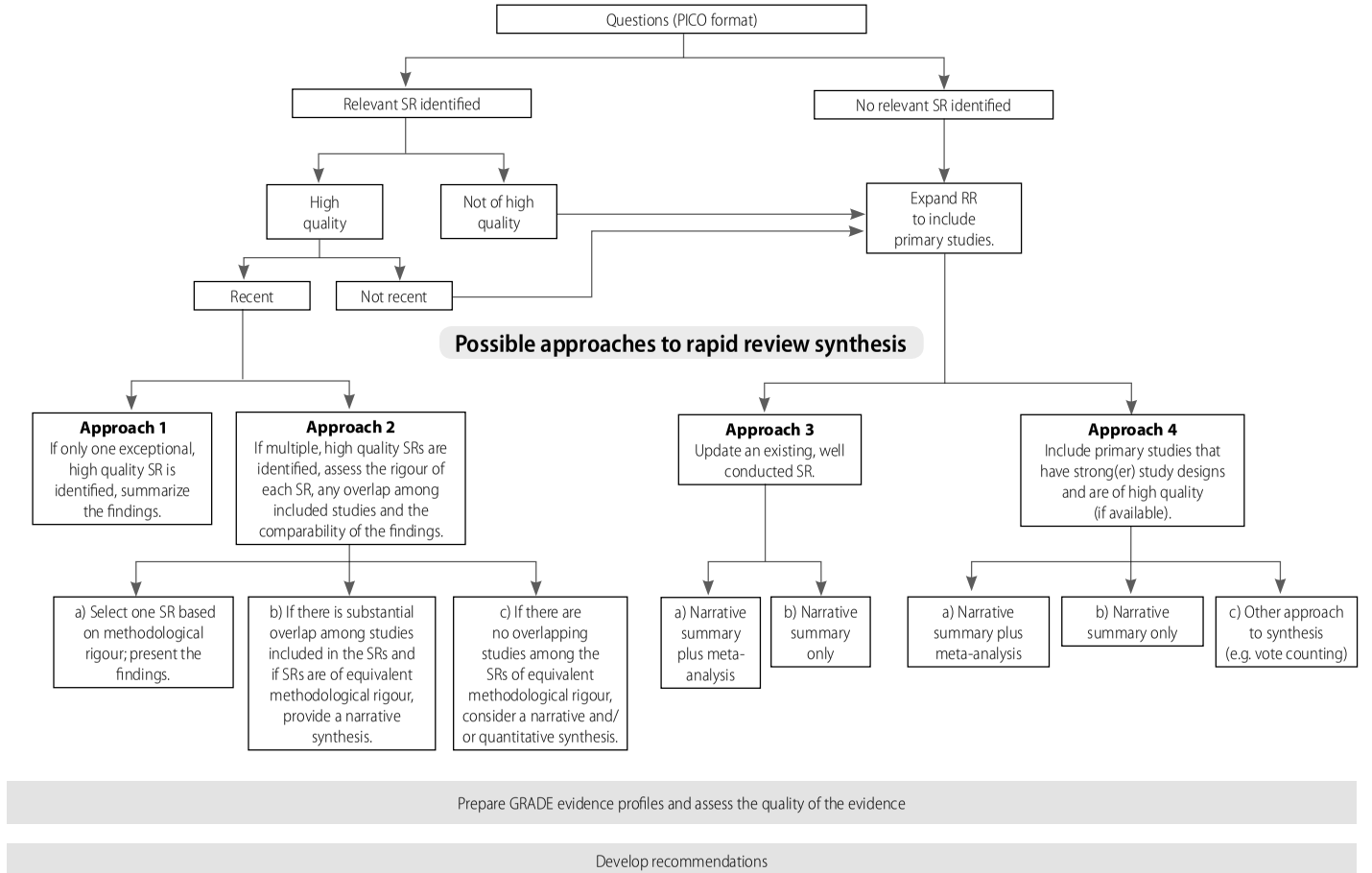
Once the qualitative description of the included studies has been completed and the data on the important and critical outcomes have been synthesized, the rapid review team will finalize the data analysis plan in consultation with the steering group. As per GRADE recommendations, the evidence synthesis should be outcome-specific. A quantitative synthesis of primary studies (i.e. a meta-analysis) may not be feasible for rapid reviews unless time and resources permit. The results of previously published meta-analyses should be reported, however. Fig 11.2 depicts the steps and decisions involved in selecting the type of evidence and the approach to data analysis and synthesis.

11.7.12 Assessing the quality of the body of evidence using GRADE

The quality of the body of evidence pertaining to each outcome should be assessed using GRADE, as appropriate for the type of data involved (see Chapter 9). The focus is on outcomes critical for decision-making in the context of rapid advice guidelines, and not on intermediate, surrogate or other types of outcomes. Exceptions may be made, however, when data are sparse and decisions need to be based on indirect evidence, including intermediate outcomes, for example.

Assessments of the quality of the evidence using GRADE are not restricted to cases in which pooled estimates of effect are available: a narrative synthesis of the evidence can also be assessed. In such situations, the effect estimate will not be a single point estimate with a 95% confidence interval, but rather, a range of observed effects across included studies. The consistency, precision and magnitude of the effect can be assessed, along with other GRADE domains, according to routine guidance. Under the category of "other considerations" in GRADE, the reviewer can list any modifications of standard systematic review methods that could have affected the robustness of the review's conclusions.

Fig. 11.2. Approaches to a rapid review of the evidence



GRADE: Grading of Recommendations Assessment, Development and Evaluation; PICO: population, intervention, comparator and outcome; RR: rapid review; SR: systematic review.

11.8 The rapid review report

The review team needs to produce a concise report that succinctly yet methodically summarizes the methods used and the results of the review. Suggested components of the rapid review report are listed in Box 11.2. All components should be described briefly; excessive detail is not necessary. The rapid review methods should be reported at a level of detail that will allow them to be replicated. A brief section on the gaps in the evidence and future research needs may be very useful and particularly important when data are sparse. A written disclosure should be provided to the effect that the rapid review is not intended to be a gold standard systematic review and that its results should therefore be interpreted with caution and viewed within a specific context.

Box 11.2. Suggested components of the rapid review report

Introduction

- Brief description of the rationale for the rapid review and of the context for the guideline.
- Duration of the rapid review process (with accompanying dates).
- Indication that this is a rapid review and should be interpreted in that light.

Methods

- Final key questions in PICO format.
- How critical and important outcomes were selected.
- Study inclusion and exclusion criteria.
- Search strategies and databases searched.
- Approach to screening citations and identifying the final set of included studies.
- Data extraction process.
- Assessment of the risk of bias at the individual study level.
- Use of GRADE or other approach to assess the quality of the body of evidence for each critical outcome.
- Description of the data synthesis process.

Results

- Complete documentation of the search results, including a PRISMA flow diagram (*8*).
- A summary table of results for each key question.
- GRADE evidence profiles (or modified versions thereof) for each key question.

continues ...

... continued

Discussion

- The strengths and limitations of the review process, focusing particularly on how the methods differed from those of a standard systematic review and the potential risk of bias introduced by the rapid review process.
- Future research needs.

Information page

- Acknowledgements.
- List of authors and collaborators.
- How the rapid review should be cited.
- Declaration of interests of the report authors.
- Sources of funding of the rapid review.
- Disclosure statement regarding the limitations of the rapid review process.

Reference list

Appendices (as appropriate)

- List of studies fulfilling inclusion criteria, with citations.
- List of publications excluded at the full-text screening stage, with citations.
- List of non-English-language or selected foreign language studies that may fulfil inclusion criteria.
- Data extraction tables.
- Risk of bias summary tables.
- GRADE evidence profiles.

GRADE: Grading of Recommendations Assessment, Development and Evaluation; PICO: population, intervention, comparator and outcome; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

11.9 Formulate recommendations

11.9.1 Convene the meeting of the guideline development group

At the very beginning of the guideline development process, the responsible technical officer and the steering group need to plan for the GDG meeting where recommendations will be formulated. Because of the compressed timeline, administrative details need to be addressed early. This includes identifying a meeting space, issuing invitations to potential GDG members, collecting declarations of interest and managing conflicts of interest, and

arranging travel. These tasks take significant time and dedicated administrative support is essential.

Recommendations can be developed via a virtual meeting if time or resource constraints preclude an in-person meeting. In general, however, an in-person meeting is preferable for formulating recommendations for WHO guidelines.

11.9.2 Evidence to recommendations

The GRADE approach for formulating recommendations should be followed when developing rapid advice guidelines (see Chapter 10). The tools used to translate evidence into recommendations are equally applicable to rapid advice and standard guidelines. In the context of a rapid advice guideline, it will seldom be feasible to collect primary data or to perform a review of the values and preferences surrounding the outcomes of interest in the guideline, nor will there be time to collect detailed information on resource use. However, data that can be readily obtained should be collected (e.g. the cost of gloves as part of personal protective equipment in the 2014 Ebola outbreak in West Africa).

11.10 Draft the guideline document

The process and resources needed to draft the final rapid advice guideline document are essentially the same as for standard guidelines. The writer should be identified early and in most situations this will be the responsible technical officer or other member(s) of the steering group. If an external contractor is used, he or she should be identified near the beginning of the guideline development process and they must be involved in all stages of the process.

11.11 External peer review

Peer review of the draft of the final guideline document by key individuals, both internal and external to WHO, is important for all guidelines produced by WHO, including those produced rapidly. In the context of a rapid

advice guideline, peer review will be accelerated and abbreviated. Early in the guideline development process, the steering group should identify three to six key individuals and solicit their interest, availability and commitment to a very short turn-around time for their reviews. Leading external organizations that are involved in the public health emergency should also be asked to review the draft document. Doing so promotes engagement and buy-in during dissemination and implementation and provides the opportunity to raise and address issues before publication. Obviously the peer review period will be markedly shorter than for a standard guideline; reviewers might be given 48 to 72 hours to complete their review.

On very rare occasions, time constraints preclude any form of peer review during the development of rapid advice guidelines. At an absolute minimum, all relevant departments at WHO must be given the opportunity to provide substantive input into the final document.

11.12 Implementation and the importance of context

Most evidence stemming from research is generated in settings and populations that differ from those affected by a public health emergency. Thus, the degree to which such evidence may be directly applied to the current context may be limited. It is important to consider how contextual factors can modify the benefits and harms of an intervention, and how various barriers and facilitators can affect implementation and impact. Therefore, in a rapid review that relies heavily on evidence from systematic reviews, the synthesis should be tailored to the local context throughout all stages of the guideline development process and decision-makers at the national and subnational levels should be provided with the information they need to apply the evidence and recommendations in their setting (*9*).

11.13 Publishing

Preparing the final guideline document for publication involves the same steps as described for a standard guideline. Electronic means will usually be used for initial dissemination, followed by print circulation as required in the local context (see Chapter 12 and Chapter 13).

11.14 Updating

All guidelines need to be kept up to date and consistent with the best available evidence. This is particularly important and difficult to achieve in the context of a public health emergency, when new data are constantly emerging and experience is continually accruing. The technical unit with primary responsibility for the rapid advice guideline must keep abreast of new information and continually assess how such information might affect the recommendations given in the guideline. If data emerge to suggest that the current recommendations need to be revised, WHO needs to be prepared to undertake such a revision.

In light of the above, it is important for the guideline document and any summaries to indicate that the rapid advice guideline may have a short lifespan, particularly if it was developed during the initial phases of a new or evolving situation and data are still being collected. This date by which the rapid guideline will need to be reviewed should be prominently featured in the guideline document and in any summaries and derivative products or implementation tools.

The steering group also needs to decide if and when the rapid advice guideline should be replaced with a standard guideline. This will depend on how closely the development of the rapid advice guideline resembled that of a standard guideline based on systematic reviews, and the rate of emergence of new information. Other important considerations include the extent to which the scope was narrowed for the rapid advice guideline, the demand for recommendations based on additional key questions, and the needs of various subpopulations. For example, the interim guidance produced by WHO during the 2009 H1N1 pandemic was noncommittal on the use of antiviral agents in pregnant women. However, as the pandemic unfolded it became clear that pregnant women were vulnerable to severe disease; thus, greater attention was paid to this population in revised guidance (*10*). In certain situations, WHO may decide to sponsor systematic reviews and to apply the standard guideline development process to update and confirm high-impact or controversial recommendations, even if no substantive new evidence has been reported.

11.15 Guideline Review Committee processes for rapid advice guidelines

The GRC needs to approve the development of a rapid advice guideline. The appropriateness of developing a guideline in the context of a specific public

health emergency is considered when the technical unit contacts the GRC Secretariat; it is formally addressed when the GRC reviews the planning proposal. The planning proposal must provide adequate justification for applying an accelerated and abbreviated process; prolonged emergencies that call for a rapid advance guideline are uncommon.

The GRC must review and ultimately approve both the rapid advice guideline planning proposal and the final guideline, as is the case for standard guidelines. The basic principles and standards outlined in this handbook apply to rapid advice guidelines. Exceptions will be considered on a case-by-case basis, as dictated by the circumstances and needs associated with the public health emergency.

The GRC and its Secretariat will implement abbreviated and accelerated processes when a rapid advice guideline is undertaken and agreed to by the GRC. The GRC will provide expedited reviews of the planning proposal and the final guideline and will hold ad hoc meetings or implement other processes and procedures to provide timely comments and the committee's assessment.

11.16 Conclusions

This chapter outlines the processes and methods used to develop a rapid advice guideline within one to three months in the context of a public health emergency. The development of a rapid advance guideline differs in important ways from the development of a standard guideline. Unlike a standard guideline, a rapid advice guideline must have a very narrow scope to make development feasible within the given time frame. Moreover, WHO staff and external experts must be identified and engaged from the very beginning of the guideline development process, and the GRC Secretariat must be contacted to be able to put in place the required expedited processes and technical support. The evidence review methods used to develop a rapid advice guideline may differ from those used for standard systematic reviews because of constraints in searching bibliographic databases and other sources of information; the need for a more fluid and iterative approach to establishing study inclusion and exclusion criteria, data abstraction and evidence synthesis; and the abbreviated nature of the review report.

The core principles and standards for WHO guidelines apply nonetheless: minimize bias; apply transparent processes and explicit, reproducible methods; and attend to the target audience's needs and to the interests of the individuals and populations affected by the recommendations. Applying

these principles and meeting these standards in the face of an emergency and compressed timelines involves trade-offs, and expertise in guideline development methods and in the topic area of the guideline are required. The guideline's limitations, including its short lifespan, need to be stated in a transparent manner and the responsible technical unit must make a commitment to update the recommendations as needed.

11.17 References

1. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev. 2012;1(1):10.
2. Watt A, Cameron A, Sturm L, Lathlean T, Babidge W, Blamey S, et al. Rapid reviews versus full systematic reviews: an inventory of current methods and practice in health technology assessment. Int J Technol Assess Health Care. 2008;24(2):133–9. doi: http://dx.doi.org/10.1017/S0266462308080185 PMID: 18400114
3. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci. 2010;5(1):56. doi: http://dx.doi.org/10.1186/1748-5908-5-56 PMID: 20642853
4. Harker J, Kleijnen J. What is a rapid review? A methodological exploration of rapid reviews in Health Technology Assessments. Int J Evid-Based Healthc. 2012;10(4):397–410. doi: http://dx.doi.org/10.1111/j.1744-1609.2012.00290.x PMID: 23173665
5. Rapid risk assessment of acute public health events. Geneva: World Health Organization; 2012. (http://www.who.int/csr/resources/publications/HSE_GAR_ARO_2012_1/en/, accessed 12 November 2014).
6. Emergency Response Framework (ERF). Geneva: World Health Organization; 2013. (http://www.who.int/hac/about/erf/en/, accessed 12 November 2014).
7. Personal protective equipment in the context of filovirus disease outbreak. [Rapid advice guideline]. Geneva: World Health Organization; 2014 (http://apps.who.int/iris/bitstream/10665/137410/1/WHO_EVD_Guidance_PPE_14.1_eng.pdf?ua=1&ua=1, accessed 25 November 20140.
8. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097. doi: http://dx.doi.org/10.1371/journal.pmed.1000097
9. English M, Opiyo N. Getting to grips with GRADE-perspective from a low-income setting. J Clin Epidemiol. 2011;64(7):708–10. doi: http://dx.doi.org/10.1016/j.jclinepi.2010.07.016 PMID: 21316192
10. WHO guidelines for pharmacological management of pandemic (H1N1) 2009 influenza and other influenza viruses. Geneva: World Health Organization; 2009. (http://www.who.int/csr/resources/publications/swineflu/h1n1_use_antivirals_20090820/en/, accessed 12 November 2014).

12. Producing and publishing the guideline

This phase of guideline development is critical. To reach the target audience in a timely manner, the guideline's production and publication require careful planning. It is necessary to coordinate the input of many contributors when:

- structuring the guideline;
- managing peer review;
- following WHO's production and publication processes;
- disseminating the guideline; and
- updating its contents.

12.1 Guideline format

All guidelines should have an executive summary, a main body and appendices. The executive summary should contain the key recommendations of the guideline. As executive summaries are often read as stand-alone documents, the strength of the recommendation and the quality of the evidence for each recommendation should be specified in the executive summary, as well as in the main body of the guideline. Authors may wish to note contextual issues for each recommendation and include a brief methods section and citations, although these are not required.

The main text of the guideline should include a table of contents, introduction, methods, recommendations and conclusions. All participants, roles and affiliations should be listed, with their conflicts of interests and how these were managed. The gender parity of the participant groups can be noted.

The systematic review(s), outcome ratings, summaries of findings, GRADE evidence profiles, evidence-to-decision frameworks and tables, and any other relevant documentation can be supplied in appendices and/or may be published electronically, thereby reducing printing and distribution costs for the main guideline. The systematic reviews need to be reported in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) *(1)*. All online material needs to be cited in the guideline itself and supplied to the GRC when final approval is sought.

Prior to submission for clearance, the AGREE-II appraisal instrument *(2)* should be used to check whether the guideline meets international quality standards and reporting criteria.

12.2 Peer review

WHO guidelines must undergo peer review before the draft is finalized for publication. The external review group (see Chapter 3) is primarily responsible for peer review, along with the relevant departments at WHO headquarters and in the regional offices. The final draft guideline with recommendations should be circulated for review before it is submitted into the WHO clearance process and to the GRC. Peer reviewers acting in their individual capacity need to complete a declaration of interests form, while reviewers representing organizations do not need to complete this form (see Chapter 6).

The request to the external review group for comments on the draft guideline must be clear about what changes can be made. Changes at this stage should be restricted to errors of fact, clarifications, and considerations related to implementation, adaptation, and the conditions in which the recommendations apply. If the peer reviewers have major concerns about the wording of the recommendations, the GDG will need to be involved in addressing these concerns and agreeing to any changes.

The process of reviewing comments and responding to them should be specified in the planning proposal (see Chapter 4). It is not necessary to respond to every single comment individually. However, it is important to let your reviewers know how comments were addressed by providing them with a copy of the published guideline (which includes an acknowledgement of their contribution). The responsible technical officer can also send the peer reviewers a version of the guideline with the changes marked, or a separate summary of points and responses.

Peer review and external comment may be sought at various other stages, depending on the timeline, the nature of the controversies around the guideline topic, and the relationships among the different groups contributing to the development of the guideline.

- During the planning stage, drafts of the key questions (in PICO format) can be circulated for comments to technical experts and experts on equity, gender and human rights at WHO headquarters and in the regional offices, as well as to members of the external review group and potential end-users of the guideline.

- Systematic review protocols outlining the search strategy, study eligibility criteria and research synthesis methods can be circulated to members of the external review group for comments.
- The list of included studies identified by searching bibliographic databases can be sent to technical experts to look for studies that have been missed.
- Draft evidence profiles can be circulated to experts, again for the identification of missing data.

All WHO guidelines require some form of external peer review of the draft final guideline. Because of the time constraints in the development of rapid advice guidelines, a small number of external reviewers may be selected and given a very short turn-around time.

12.3 The production process

Production of WHO guidelines should follow the same process as for other WHO publications. Detailed information on each step is available to WHO staff on the WHO intranet (*3*). A brief overview is provided here.

12.3.1 Writing

Identify a writer early in the process and make sure this person is engaged throughout the guideline's development. The writer can be a WHO staff member or contracted on a freelance basis. In either case, 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 roughly calculate the time needed and can begin negotiations with an external writer if necessary. WHO does not have a standard pay scale for writing, 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 in the eManual). When negotiating fees and schedules, calculate a minimum of one week of full time work to produce 5000 words.

Avoid using multiple authors when 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, this approach invariably creates more work than it eliminates and will require an experienced writer or editor to synthesize the different sections.

12.3.2 Legal advice on proprietary products

WHO guidelines should avoid specifying proprietary products when at all possible. Devices and diagnostics used in interventions should be described generically to avoid identification of specific products and trademarks. If in doubt, please contact WHO's legal counsel.

12.3.3 Editing and proofreading

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

12.3.4 Executive clearance and approval by the Guideline Review Committee

GRC review of final guideline documents occurs as part of the final executive clearance. In headquarters, submission to the GRC is done after approval by the relevant director and before submission to the assistant director-general. Documents should be in a final edited form ready for layout, proofreading and printing when they are submitted for final clearance.

12.3.5 Layout

Once you have an edited and carefully checked manuscript that has received full executive clearance from the relevant assistant director-general, 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 using existing publication templates if your department has established these for other publications. After receiving clearance of the full text, you will need a cover

design, an international standard book number (ISBN) and a barcode; the latter two are issued by WHO Press.

12.3.6 Printing

The printing procurement team in WHO Press will provide quotes and arrange for your files to be sent to the printer. You will need to supply an estimate of the number of print copies you require. The sales team in WHO Press can assist with these estimates. You must have the printers' proofs checked again by your proofreader, 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.

12.4 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. Priced distribution is done by WHO Press through sales agents in all regions and by the WHO bookshop. The extent of mandatory free distribution depends on the type of publication but can include depository libraries, schools of public health, schools of medicine, WHO country offices and missions in Geneva. Please consult with WHO Press on the number of print copies required for mandatory free distribution. When thinking about further dissemination of your guideline, consider the options described in the sections that follow.

12.4.1 Online publication

Your guideline can appear on the Internet in a variety of formats. At a minimum, you should contract your designer or typesetter to produce a web-ready portable document format (PDF) – 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 hypertext mark-up language (HTML) and additional materials, both electronic and printed. The WHO web team is a good source of advice.

12.4.2 Archiving

At the time of publication, the department should ensure that archiving requirements are met. In headquarters, departments should send the final electronic file of their guideline to the WHO Library for inclusion in the Institutional Repository for Information Sharing (IRIS) (*4*). In regional offices, departments should send the final electronic file of their information products to their regional office library. If printed copies are produced, one should also be sent for inclusion in the print collection of the respective library.

12.4.3 Translations

Because WHO guidelines target a global audience, it is often necessary to provide 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 their timing discussed with the translation suppliers or regional office involved. Do not forget to budget for translation costs. To reduce such 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 and its strength should not change in translation.

12.4.4 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 Library. To increase awareness of the guideline, the guideline development process and/or recommendations may also be published in peer-reviewed journals, in compliance with WHO's open access and copyright policies. All external publications with WHO authors require internal clearance.

12.4.5 Other forms of dissemination

You should carefully plan a dissemination strategy early in the development process. A variety of approaches should be considered. These can include an offi-

cial launch, a press release and/or conference, an announcement on the WHO website, distribution through regional offices or at meetings, and endorsement by stakeholders and interest groups. Mobile phone applications for guideline dissemination and decision support can also be commissioned. Because the guideline document itself may be quite lengthy, more succinct derivative products should be considered and planned early in the process. These may include charts, manuals and other products designed to facilitate implementation.

12.5 Updating guidelines

12.5.1 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 *(5, 6)*. In deciding on the date by which a guideline should be reviewed, take account of the rate of change of research on the topic, questions for which no evidence has been found, and the potential need for new advice.

12.5.2 Updating recommendations

All WHO recommendations that are not based on the processes and standards outlined in this handbook (particularly those published before the GRC was established in 2007 and those not considering equity, human rights and gender) should be updated to meet current WHO guideline standards. Updating guidelines is 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. If recommendations will be updated incrementally, the planned approach should be discussed with the GRC Secretariat and outlined in the planning proposal.

12.5.3 Guidelines that are near or beyond their "review-by" date

Technical units are responsible for keeping their guidelines up to date. As guidelines near their "review-by" date, they should be carefully examined for currency. If there is reason to believe one or more recommendations need updating, plans should be made to start that process.

12.5.4 Recommendations that may be out of date

Occasionally guideline developers may want to update guidelines before the "review-by" date, particularly if new evidence is published. This new evidence should always be seen in the context of the total body of evidence supporting the recommendations and thus should be part of a new or updated systematic review. Any update that involves changing recommendations needs to be reviewed by the GRC. 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.

If there are concerns that one or more recommendations in a guideline may no longer be valid, the department should make every effort to ensure that the guideline implementers and other stakeholders are aware of the uncertainty and of plans to update the recommendations. Such announcements can be placed on the relevant pages of the WHO website, linked to the online copies of the guideline, circulated directly to the known stakeholders and published in journals.

12.6 References

1. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097. doi: http://dx.doi.org/10.1371/journal.pmed.1000097 PMID: 19621072
2. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al.; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839–42. doi: http://dx.doi.org/10.1503/cmaj.090449 PMID: 20603348
3. WHO eManual/VIII - Information Products [intranet site: available to WHO staff only]. Geneva: World Health Organization; 2011 (http://emanual.who.int/p08/Pages/home.aspx, accessed 11 November 2014).
4. Institutional Repository for Information Sharing [online database]. Geneva: World Health Organization; 2014 (http://apps.who.int/iris/, accessed 10 November 2014).
5. Shekelle P, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? BMJ. 2001;323(7305):155–7. doi: http://dx.doi.org/10.1136/bmj.323.7305.155 PMID: 11463690
6. Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. Ann Intern Med. 2007;147(4):224–33. doi: http://dx.doi.org/10.7326/0003-4819-147-4-200708210-00179 PMID: 17638714

13. Adaptation, implementation and evaluation

Adaptation, implementation and evaluation are the steps that complete the guideline development process. The implementation of new guidance gives rise to new evidence of impact and new research questions. Practice needs to be continually re-evaluated, and guidance updated in light of new evidence. Although implementation and evaluation plans do not need to be described in the guideline, they should be considered by the steering group, made available to all guideline contributors and actively supported by the department authoring the guideline. The steering group and GDG should discuss and document a list of the tools and resources that will need to be made available to countries, such as implementation checklists, costing models and the data that inform assumptions used in economic models. The group may wish to consider which partner(s) will eventually lead guideline adaptation and implementation, the steps they will be expected to take, and any regulatory or licensing implications of specific products.

Research on how best to implement guidelines shows that training, ease of use, financial and professional development incentives and feedback of results to health-care providers all need to be considered in implementation plans. Multicountry evaluations of long-standing WHO guidelines, such as those on the Integrated Management of Childhood Illness (*1*), have been funded and published over the years.

13.1 Adaptation

Adaptation of WHO guidelines, taking into account local circumstances and resource considerations, can be done at the regional, national or subnational level. Translation is one element in the adaptation process. The manual and toolkit for guideline adaptation published by the Guidelines International Network (*2*) describes the approaches used when deciding whether to accept or reject a guideline in a particular setting, whether to endorse its evidence summary and/or recommendations, or whether to modify specific recommendations.

13.2 Implementation

Implementation of a guideline should be taken into account right from the beginning of the guideline development. Implementation is generally the responsibility of national or subnational groups, which explains why their participation in guideline development is critical. WHO headquarters and regional and country offices can support implementation activities by promoting new guidelines at international conferences and providing guideline dissemination workshops, tools, resources and overall coordination.

Implementation strategies are context-specific. The basic steps for implementing a guideline are:

- convene a multidisciplinary working group to analyse local needs and priorities (looking for additional data on actual practice);
- identify potential barriers and facilitating factors;
- determine available resources and the political support required to implement recommendations;
- inform relevant implementing partners at all levels; and
- design an implementation strategy (considering how to encourage the adoption of the recommendations and how to make the overall context favourable to the proposed changes). Implementation or operational research can help inform field testing and rollout strategies to promote the uptake of recommendations.

There is a range of derivative documents or tools that can be developed to facilitate implementation. These can be distributed with the guideline, or local guideline implementers can develop them. Such documents or tools may include a slide set reflecting the guideline content; a "how to" manual or handbook; a flowchart, decision aide or algorithm; fact sheets; quality indicators; checklists; computerized applications; templates, etc.

13.3 Monitoring and evaluation

Monitoring and evaluation systems are used to collect and analyse data to assess the effectiveness and impact of the guideline. The guideline should include outcome or performance measures that can be monitored for the main recommendations. Performance measures may be related to:

- guideline dissemination;
- adaptation and endorsement in the national context;

- policy changes;
- changes in end-user knowledge and understanding;
- changes in practice performance;
- changes in health outcomes and inequities (both by level and distribution); and
- economic or other social consequences.

Ideally, there should be baseline measures against which to assess performance in relation to the potential change induced by the guideline. Operational and implementation research can be performed to assess service providers' and end-users' perceptions, and the values and preferences related to guideline implementation. The guideline should propose a specific set of indicators to be monitored and evaluated, including relevant disaggregation of data.

13.4 References

1. Bryce J, Victora CG, Habicht JP, Vaughan JP, Black RE. The multi-country evaluation of the integrated management of childhood illness strategy: lessons for the evaluation of public health interventions. Am J Public Health. 2004;94(3):406–15. doi: http://dx.doi.org/10.2105/AJPH.94.3.406 PMID: 14998804
2. ADAPTE Collaboration. The ADAPTE Manual for Guideline Adaptation. Version 1.0. Guidelines International Network; 2007 (http://www.g-i-n.net/document-store/working-groups-documents/adaptation/adapte-manual-for-guideline.pdf, accessed 12 November 2014).