وزارة الصحة
Ministry of Health

The Saudi Center for
Evidence Based Health Care

Saudi Arabian Handbook for Healthcare Guideline Development

Version 1.0

April 2014

Contents

Acronyms and Terms

A full glossary of terms and their definitions may be found at the end of this handbook.

AMSTAR	Assessment of Multiple Systematic Reviews
AGREE	Appraisal of Guidelines for Research and Evaluation
AHRQ	Agency for Healthcare Research and Quality
CADTH	Canadian Agency for Drugs and Technologies in Health
COI	Conflict of interest
DOI	Declaration of Interest
GAB	Guideline Advisory Board
GRADE	Grading of Recommendations Assessment, Development and Evaluation
GDT	Guideline Development Tool
GIN	Guideline International Network
GL	Guideline
GP	Guideline Panel
ICER	Incremental cost-effectiveness ratio
IOM	Institute of Medicine
MeSH	Medical Subject Headings
MoH	Ministry of Health
NGC	National Guideline Center
NICE	National Institute for Health and Care Excellence (UK)
PICO	Patient/Population-Intervention-Comparison-Outcome
QALY	Quality-adjusted life years
WHO	World Health Organization

1. Preamble

Guidelines, if based on the best available evidence for the decision criteria that determine the direction and strength of a recommendation, are an ideal tool to support health care decision makers. The handbook for guideline development in Saudi Arabia establishes a framework for guideline development in the Kingdom of Saudi Arabia that addresses organizational and practical issues to ensure practice is evidence based and addresses the need of the population.

The aim of Healthcare Guidelines is to reduce unnecessary variation in practice through involvement of all relevant groups including health care professionals, such as nurses, physicians, allied health workers and patients in the development of health care recommendations based on the best available evidence. Such guidelines must provide support rather than dictate care; they will not be cookbooks. For that reason, much of the suggested methodology focuses on approaches, such as the Grading of Recommendations Assessment, Development and Recommendations (GRADE) approach with its rationale approach to decision determinants.

The handbook places emphasis on using existing evidence syntheses, sometimes from existing guidelines, as a means to develop recommendations for Saudi Arabia. Emphasis is also placed on using decision determinants (so called evidence to decision frameworks) to ensure that information that is relevant for the target populations is sought for and integrated rather than placing undue emphasis on existing guidelines that may be outdated or not considerate of the context.

The approach and methodology are also based on the work done for the World Health Organization in 2006, modelled on country specific advice given to the country of Estonia and advice given to numerous professional societies and other organizations. It is based on research in the field of evidence to decisions using existing highly credible systematic reviews rather than de novo reviews and placing emphasis on transparency, including conflict of interest management. Existing standards by the Guideline International Network and authorities such as the Institute of Medicine are carefully integrated. Finally, the experience collected during a large guideline development effort in Saudi Arabia in December 2013 helped fine tune the approach described here.

This handbook has been developed by a team of researchers at McMaster University, Hamilton Canada (headed by Dr. Holger Schünemann with collaboration from Drs. Reem Mustafa, Alonso Carrasco, Romina Brignardello-Petersen, Jan Brozek and Wojtek wiercioch) with involvement of key informants in Saudi Arabia who commented on an early draft and provided invaluable advice during in-depth interviews. In particular the authors would like to thank Prof. Lubna Al-Ansary, Dr. Noha Dashash, Dr. Sohail Bajammal, Prof. Hassan Baaqeel, Dr. Zulfa Al Rayess, Dr. Yaser Adi and Dr. Rajaa Alraddadi who provided important insights.

The ministry of health welcomes the opportunity to introduce this new era of supporting care providers in the Kingdom.

Preamble written by Holger Schünemann, Chair of the Department of Clinical Epidemiology and Biostatistics at McMaster University, Hamilton, Canada.

2. Introduction

2.1 Guideline definition

A guideline (GL) is a product that contains recommendations about health interventions, including interventions focusing on health care related tests and test strategies.[1-3] As defined by the World Health Organization, the recommendation is a statement that assists health care providers and recipients of health care in making the best possible health care decision. A recommendation implies a choice between different interventions that may have an impact on health and that have influence on resource use as well as other consequences. The direction and strength of a health care recommendation depends not only on the magnitude of anticipated benefits and harms or burden, but also on the certainty in the intervention effects, the value the population places on these associated outcomes and interventions and the impact on resources, including considerations around feasibility, acceptability and equity.[4,5] Such considerations may be highly context or setting specific and may require local information or evidence.

The main difference between a guideline and a typical textbook is that a guideline provides answers as actionable statements to foreground questions; advice about "what to do" rather than background questions which deal with "how or why does it work". There is broad agreement that these statements should be based on systematic reviews.[6-8] Systematic reviews are transparent syntheses of the available *best* evidence for a given question following established methodology. Given the *best available* evidence should be used, the synthesis of evidence may be derived from different types of studies, such as randomized trials and various types of observational studies, depending on the type of questions and availability of evidence. Detailed processes on guideline development are available through various resources, such as the guideline development checklist on: http://cebgrade.mcmaster.ca/guidecheck.html.[9]

2.2 When is a guideline the right approach?

Before beginning the process of guideline development, prioritization should include considerations if and what type of a guideline is the correct approach to solving the problem.[10,11] The need for rapid responses may lead to providing interim or preliminary advice or guidelines that will later be supported by a fully developed guideline.

2.3 Context for guideline development in Saudi Arabia

Several groups have supported or carried out the development of guidelines in Saudi Arabia with limited co-ordination. There has been no uniformly accepted approach to guideline development and this has resulted in a wide array of different guideline formats and compilation processes. The Ministry of Health of Saudi Arabia (MoH) has embarked on standardizing and coordinating guideline development nationally.

This handbook has two main goals: 1) to summarize the internationally accepted methods and approaches related to the health care guideline enterprise; and 2) to provide an approach on how to successfully implement and sustain the guideline enterprise in Saudi Arabia. It intends to

cover all aspects of the guideline enterprise, starting with assessing the need for one (prioritization) and finishing with the distribution and implementation covered in the following 18 topics found to be relevant for guideline development.[9]

1. Organization, Budget, Planning and Training
2. Priority Setting
3. Guideline Group Membership
4. Establishing Guideline Group Processes
5. Identifying Target Audience and Topic Selection
6. Consumer and Stakeholder Involvement
7. Conflict of Interest Considerations
8. (PICO) Question Generation
9. Considering importance of outcomes and interventions, values, preferences and utilities
10. Deciding what Evidence to Include and Searching for Evidence
11. Summarizing Evidence and Considering Additional Information
12. Judging Quality, Strength or Certainty of a Body of Evidence
13. Developing Recommendations and Determining their Strength
14. Wording of Recommendations and of Considerations of Implementation, Feasibility and Equity
15. Reporting and Peer Review
16. Dissemination and Implementation
17. Evaluation and Use
18. Updating

Although the need for country-specific guidelines is envisaged in most areas of health care due to the need to consider costs and values in addition to the health care evidence, the use of international resources is encouraged.

2.4 Information reviewed for this handbook

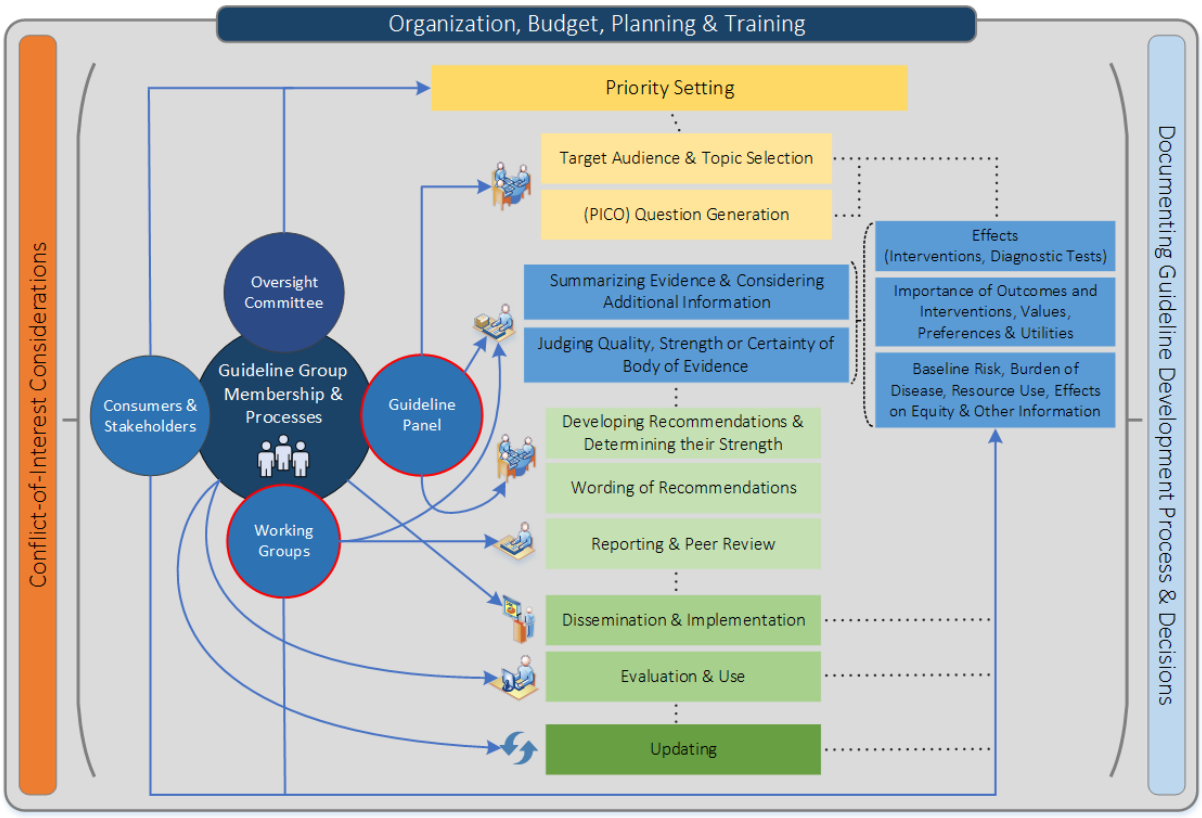
Writing this handbook involved multiple steps and a mixed methods approach. First, critically reviewing different sources to develop a first draft of this handbook. A list of these sources include: A series of articles on “Improving the use of research evidence in guideline development (16-article series)”,[12] another series entitled “Integrating and Coordinating Efforts in COPD Guideline Development (14-article series)”,[13] articles published in Implementation Science on developing CPGs (3-article series),[14] Estonian Handbook for guideline Development (work done by Dr. Schünemann and colleagues on that handbook served as partial template for this handbook),[11] AGREE II: Advancing guideline development, reporting and evaluation in health care,[15] Conference on Guideline Standardization (COGS): Standardized Reporting of Clinical Practice Guidelines, Institute of Medicine: Clinical Practice Guidelines We Can Trust,[2] and Guideline International Network (GIN): Toward International Standards for CPGs.[16] Additionally, multiple manuals of guideline developers were reviewed. These manuals include: Argentina National Academy of Medicine, Colombia Ministry of Health and Social Security, Peru Ministry of Health, Spain Ministry of Health, American College of Cardiology-American Heart Association, Cancer Care Ontario, Canadian Task Force on Preventative Health Care, Kaiser Permanente, National Health and Medical Research Council (NHMRC), National Institute for Health and Clinical Excellence (NICE), New Zealand Guidelines Group, Scottish Intercollegiate Guidelines Network

(SIGN), United States Preventive Services Task Force (USPSTF), Centers for Disease Control and Prevention (CDC), and the World Health Organization (WHO). Second, using information from interviews with key Saudi stakeholders to obtain feedback on local needs and requirements that are relevant for the adaptation. Third, critical revision of initial draft based on comments, written response, and development work based on the KSA context following a workshop in KSA.

3. Overview of the guideline enterprise

A formal process to support guideline development under the auspices of the MoH is a prerequisite for consistent application of the guideline development process. Figure 3.1 describes this process (from reference[9]). A national guideline center (NGC) will be established. A 10 person advisory board to oversee organization, budget, planning and training will include representation of various stakeholders (e.g. MoH, medical societies, care providers, National Guard, Military hospitals) and act as advisory board to support organization, planning and training. Members of the advisory board will be selected based on qualification and background. The advisory board is chaired by an elected member. The NGC is chaired by a member who is nominated by the MoH and approved by the advisory board. The NGC will nominate one or more oversight committees with representation from various stakeholders for guideline projects depending on the type of guidelines. Further roles of the NGC will be described below and are represented by the blue circles in Figure 3.1. [9]

Figure 3.1:



Legend: Flow diagram of the guideline development process. The steps and involvement of various members of the guideline development group are interrelated and not necessarily sequential. The guideline panel and supporting groups (e.g. methodologist, health economist, systematic review team, a secretariat for administrative support) work collaboratively, informed through consumer and stakeholder involvement. They report to the oversight committee. While deciding how to involve stakeholders early for priority setting and topic selection, the guideline group must also consider how developing formal relationships with the stakeholders will enable effective dissemination and implementation to support uptake of the guideline. Furthermore, considerations for organization, planning and training encompass the entire guideline development project, and steps such as documenting the methodology used and decisions made, as well as considering conflict-of-interest occur throughout the entire process.

The need for a guideline can be identified by any body (e.g., professional medical society, patient group, academic institution etc.), but requires coordination by the NCG. This need will be described in a topic proposal (see Chapter 4) and a draft scope (see chapter 5), and a proposal for the guideline panel membership needs to be made, bearing in mind the panel requirements for multidisciplinary.

A guideline topic proposal is submitted to the NGC advisory board. Guidelines intended to be financed by the MoH must be submitted to the NGC for approval. Despite the effort of centralizing guideline development through the NGC, other entities in Saudi Arabia may wish to develop guidelines independently. Such guideline developers using other financing mechanisms and not developing guidelines as part of the activity of the NGC are encouraged to submit their proposals as well to benefit from methodological advice of the NGC.

The roles of the NGC are:

- developing an oversight committee for guideline topics
- proposing and evaluating guideline topic(s) to be financed;
- consulting on and approving the composition of guideline panel;
- evaluating conflict of interests of panel members;
- overseeing and acting as an advisory resource for the work of the guideline panel;
- finalizing the initial scope with the panel;
- signing off on the final scope;
- being an arbiter in situations of lack of agreement on issues other than recommendations (e.g. authorship issues);
- approving the final guideline (note, this function is not to alter recommendations but to ensure that the guidelines are methodologically sound).

3.1 Guideline panel

The guideline is drafted by a guideline panel. The panel should be multidisciplinary and should incorporate representatives of specialities involved in the relevant guideline.[15,17-22] The panel should include representatives of patient and/or consumer groups.[2,17,19,23,24] Patients may be familiar with the topic and its treatments based on personal experience and may be able to

provide information and evidence relative to the guideline. Conflicts of interests of all panel members must be managed appropriately as described below.

The initiator of the guideline presents the potential composition of the Panel and the name of the proposed chair to the NGC for approval. The NGC may deliberate on the composition of the Panel.

The Panel should include:

- medical experts;
- methodologists;
- health economist;
- representatives from key stakeholders and organizations involved in implementation, including:
- representatives from consumer or patient associations;
- representatives from the medical faculty of a university;
- representatives of organizations involved in the health-care process and who are likely to be end-users of the guideline;

The size of the panel depends on the topic of the guideline, but is generally up to 20 persons. The size of a guidelines panel should be small enough for effective group interaction, but large enough to ensure adequate representation of relevant views.[11,16,18,19,23,25-27]

The roles of the Guideline Panel Members include:[11,16-18,25,28,29]

- Comment on the initial scope selected by the NGC and finalize it (including the formulation of clinical questions and choosing outcomes), taking into account the views of stakeholders. During the development of the questions for the guideline, the guideline panel has to consider which clinical questions may require information from existing guidelines or from systematic reviews.
- Review draft recommendations based on the presented evidence, with explicit consideration of the overall balance of risks and benefits. The assumption for the Panel is that the research evidence to support a particular recommendation is global, whereas costs, values and preferences, feasibility, acceptability and equity of recommendations are local considerations, and therefore should be the basis of adaptation of international recommendations for local situations.
- Approve recommendations according to the GRADE approach, taking into account values and preferences and resource implications.
- Decide on consultation needed for the draft guideline.
- Plan and agree on the primary methods for implementation and indicators for measuring the use of the guideline.
- After guideline finalization, facilitate the process of implementation (i.e. to act as opinion leaders for and advocates of the guideline).
- Work closely with the guideline support groups (see below).

Specifically, the backgrounds of various panel members would be: [2,10,11,17,19,23,25,30,31]

- experts who should represent the perspective(s) of health-care professionals, as well as social care and other professionals, where relevant.
- involved in the care of patients affected by the guideline topic; detailed evidence research expertise is not necessary, although an understanding of evidence-based medicine is essential.
- methodologists in assessing clinical evidence and developing guidelines, should be included as appropriate, ideally as a panel co-chair. Inclusion of a methodologist in a leading role, particularly one with experience in the guideline development process, is recommended to explain to the panel the evidence retrieval process and to guide the process of formulating recommendations.
- patient representatives, e.g. from patient organizations (or a representative of the patient with the relevant chronic condition) – they will represent the view of the patient(s) with the relevant condition.
- managers and other health professionals may represent the view of the health-care services and provide expert opinion on the implementation of guidelines.
- health economists and/or bio-statisticians can provide an analysis or explanation of the costs of health services, cost-effectiveness, data on the provision of health care services and medicines.

Panel members are asked to make a commitment to attend meetings for the guideline development process, in order to ensure continuity and effective participation in the process.[11,18,26,27,32]

3.2 Guideline panel chair

The choice of the co-chairs (ideally one content expert and a methodologist) of the panel is important to ensure that the panel will be able to work effectively. In most situations, groups work most effectively if the chairs have knowledge of the content, but there must be particular expertise in facilitating groups, interpreting evidence and developing guidelines. People who are experts in the content area of the guideline and who have strong views about interventions or aspects that may be included should not chair a guidelines panel. A panel would be chaired jointly by a methodologist and a content expert with appropriate division of tasks and labor.[2,16,17,26,32-36]

3.3 Guideline support unit

The panel is supported in its work by a guideline support unit within the NGC, consisting of experts in methodological aspects of guideline development, evidence retrieval and assessment and health economics. If the guideline is financed by the MoH, the MoH provides the guideline support unit for the panel[2,17,18,27,29,32,36]. The guideline support unit reports to the NGC advisory board and works with the guideline oversight committee forming various working groups.

The roles of the guideline support unit are: [11,18,25,32,37,38]

- provide technical and administrative support for developing the guideline;
- make panel members aware of items on the guideline checklist;
- preparation of materials, evidence retrieval and summary for recommendations;
- organizing the panel meetings;
- use and manage the guideline development tool (www.guidelinedevelopment.org);
- prepare draft records (minutes) of all panel meetings, taking special care to document areas of controversy and dissent.

Once finalised by the panel, the guideline is endorsed by the NGC, making sure the appropriate methodology has been followed while developing it.

The principles for the production and dissemination of the guideline are described in the chapter on implementation and dissemination of this handbook.

4. Topic proposal and selection

4.1 By whom and how are topics proposed?

General topics for guideline development can be proposed by medical societies, the medical faculty of a university or MoH (the proposer is subsequently called "the initiator"). Topics together with initial scope must be presented by the initiator to NGC on October 31 of each year.

4.2 What is the process for making a proposal?

Topics can be triggered by many different inputs: regular audits, feedback from practitioners, variations in care, guidelines being issued by other entities that need to be adapted, introduction of new interventions, emerging health problems, etc.[10,18,21,23,36,39]
Topic proposal will need an active communication between the initiator and other potential stakeholders including the MoH and the guideline support unit to provide background information and statistical data for the proposals.

4.3 Who selects the topic?

The selection of topics for guidelines intended to be financed by the MoH has to be made by the NGC taking into account the initial scope (see Chapter 3). In process of choosing topic(s) to be financed and approved applicability of further guideline should be taken into account (including organisational and potential resource implications of a guideline). This would help to avoid situation when the NGC chooses to finance a guideline topic which implementation is organizationally not feasible and evidently not affordable to the health system.

4.4 How is topic selection done?

The NGC will assess the topics together with draft scope documents presented annually based upon the criteria listed below. The NGC advisory board will evaluate all topics. A nine point Likert-type scale will be used; were 9 is the score of the most important and useful topic and 1 is the score of the topics that are not important or useful to address. Following resolution of possible misunderstandings, the average will be used to determine priority topic. Raters (members of the NGC) will be encouraged to use the whole range of the scale to allow for differentiation between topics' importance.

In general, the NGC will evaluate topics based on an assessment of:

- ✓ **Burden of disease**
  - the population suffering the disease/condition in Saudi Arabia (incidence, prevalence, mortality,)
  - the resource impact of the disease/condition in Saudi Arabia

- ✓ **Variations**
  - **practice variation** and variations in **health outcome** by different
    - regions in Saudi Arabia
    - providers in Saudi Arabia
    - level of care (primary care, specialist services)
    - patient populations (here and after most critical subgroups under the disease/condition can be identified if necessary)
    - international practice compared with Saudi Arabia
  - variation in **treatment cost**s (regions, providers, level of care, patient populations). Treatment costs analyses can be conducted using data from databases.
    - service treatment (all treatment costs in certain period)
    - pharmaceuticals
    - hospitalization (rate, length of stay)

- ✓ **Potential**
  - **potential for modernization** of current practice
    - availability of new interventions (including diagnostic tests and strategies)
    - availability of new evidence that will likely change practice
    - availability of new service delivery
  - **potential result** of successfully implemented guideline
    - measurable impact on health (indicators)
    - more cost-effective use of resources

- ✓ **Problem statement and the purpose of the guideline**
  - **problem statement** is completed by the initiator based on the information listed above eg, *"persons having condition X in the Riyadh area are hospitalized more frequently and their average prescription cost for drug B is different from other regions in Saudi Arabia."* and the **purpose of the guideline**: eg, "*to guarantee up-to-date treatment with equitable costs for persons with condition X irrespective of region*"

- ✓ **Initial scope prepared by initiator** (See template below)

- ✓ **Relationship of topics and scope to health related government priorities.**

The NGC may exclude proposed topics if the topics proposed are not potential subjects for guidelines. The MoH may propose topics of special importance that receive financing through purpose directed channels. Topics receiving directed financing should be prioritized.

5. Defining the scope of the guideline

5.1 What is the scope of the GL?

The scope of a guideline provides a framework within which to conduct the guideline development work, that is the topics that should generally be addressed.[17-19,25,27,36] Proposing topics and the scope of the guideline will be influenced by existing guidelines and systematic reviews. The NGC will make a list of preferred topics available. These topics will result from surveying existing systematic reviews and guidelines from international organizations to allow for adaptation of guidelines.

Considering the resources for possible guideline topics, scoping should be conducted in stages:

1. Drafting the scope
2. Consulting with stakeholders about the draft scope
3. Finalizing the scope

5.2 Who prepares the scope?

The initial scope, with questions and preliminary outcomes, is prepared by the initiator of the clinical guideline. The scope is finalized by the oversight committee for an approved topic, in cooperation with the NGC, and signed off by NGC, e.g. the advisory board.

5.3 Drafting the initial scope

After the general topic is defined by the group proposing the topic, the aspects of care that the guideline will cover should also be defined:[17-19,25,27,36]

- population to be included or excluded (e.g., specific age groups or people with certain types of disease);
- healthcare settings (primary or specialized care);
- the different types of interventions and treatments to be included or excluded (diagnostic tests, surgery, rehabilitation, lifestyle advice). Does the potential guideline complement other programs or interventions in the particular therapeutic area?
- information and support for patients and carers;
- the preliminary outcomes that will be considered (benefits and potential harms to patients, impact on health insurance, society perspective); this list will be completed by the guideline panel;
- links with other relevant guidance. Are there any similar guidelines available in Saudi Arabia in this particular therapeutic area? If so, will the new guideline replace or supplement the existing one(s)?

On the basis of these aspects, **formulate two-page document with a scope that:**

- provides an overview of what the clinical guideline will include and what will not be covered;
- identifies the key questions (clinical, as well as organizational, regulatory, etc). It is useful to formulate the questions using the PICO format (see below for formulating questions);
- chooses and rates the outcomes in the PICO (see below for choosing and rating outcomes)
- sets the boundaries of the development and provides a clear framework to enable the work to stay within the agreed priorities;
- informs the development of the detailed review questions from the key clinical issues and the search strategy;
- provides information about the expected content of guideline;
- ensures that a minimum set of essential aspects, questions and recommendations is covered;
- ensures that the guideline will be of reasonable size (no more than 20 key questions are suggested) and can be developed within a specified time period.

evaluates, if:

- any existing guideline in Saudi Arabia covers this topic?

- up-to-date evidence is likely to be available on the topic (see list of preferred topics)?

finalizes:

- the title of guideline;

- who should be key stakeholders for implementation for further consultation on the scope, if they have not already been involved in preparing it.

Table 5.1 Submission of topic and scope

Domain	Description
Describe the general topic:	
Does the potential guideline complement other programs or interventions in the particular therapeutic area?	
Population to be included or excluded (e.g., specific age groups or people with certain types of disease):	
Healthcare settings (e.g. primary or specialized care):	
The different types of interventions and treatments to be included or	

excluded (diagnostic tests, surgery, rehabilitation, lifestyle advice).	
Information and support for patients and carers to be provided:	
The preliminary outcomes that will be considered (benefits and potential harms to patients, impact on health insurance, society perspective)	
Links with other relevant guidance. Are there any similar guidelines available in Saudi Arabia in this particular therapeutic area? If so, will the new guideline replace or supplement the existing one(s)?	
Provide an overview of what the clinical guideline will include and what will not be covered:	
Identify some of the key questions (clinical, as well as organizational, regulatory, etc) following PICO format:	
Describe the up-to-date evidence that is available on the topic (see list of preferred topics)?	
Who are the key stakeholders for implementation and for further consultation on the scope, if they have not already been involved in preparing it.	

5.4 Informing other stakeholders about the initial scope

When the initial scope is prepared the initiator must decide who else should be consulted and involved[2,18,27,29,33]. This can be an informal process, the main purpose of which is to check that the initial scope is clearly understood.

5.5 Formulating questions for the scope

The selection of questions (and their components) that are to be addressed in the guideline has major consequences for the scope of the guideline. The questions will drive the direction (inclusion and exclusion of data) and determine the type of information that will be searched for and assessed. The questions are also the starting point for formulating the recommendations. It is very important that the questions are clear and well defined, and that there is agreement

about them among panel members. The guideline development tool can be used to complete this task.[10,19,23,26,35,38-40]

Updating a guideline may include a change of scope; not only the questions but also the selection of critical outcomes may differ from the original guideline.

It is helpful to start by dividing the types of information and questions into three main categories:

Definition/background questions

e.g., What is chronic obstructive pulmonary disease (COPD)?
What are the causes of caries?

Facts/foreground questions

e.g. What is the effect of inhaled steroids in COPD?
What types of public health interventions reduce the incidence of caries?

Recommendation/decision

e.g. Should inhaled steroids in COPD be used?
Should regular dental hygiene visits be included in dental care?

Guidelines should focus on the recommendation and decision and minimize the description of the definition and background to what is needed to put the recommendations in context.

The questions to be covered by the guideline should be identified on the basis of clinical, public health or policy needs and input from clinicians and other experts. Input from consumer or patient groups may also be helpful. Questions should focus on areas where changes in policy or practice are needed and/or controversy may exist.[41,42]

During the development of the questions for the guideline, the guideline panel should consider which questions may need information from systematic reviews or from existing guidelines.[7,11,18,19,23,27,33] Questions that may require new systematic reviews will have the greatest impact on the time taken to complete the guideline. Thus, the preferred approach is one in which existing, highly credible systematic reviews can be used.

The foreground questions are the most important ones for a guideline and they are used to inform the recommendation/decision and they may require a systematic review and quality assessment of the evidence about effects using the GRADE approach. Ideally systematic reviews inform adaptation issues, values and preferences, clinical needs and baseline risks.

For priority setting and defining the scope, the initial list of types of question may be probably be a long one. Some examples could be:

- What is the frequency of the condition or issue of interest? (background)
- What causes the condition or issue of interest? (etiology)
- Who has the condition or issue of interest? (diagnosis)
- What happens if someone gets the condition or issue of interest? (prognosis)
- How can we treat the condition or issue of interest? (interventions)

- What policies should we introduce to alleviate the condition or issue of interest? (policy intervention)

To formulate these general questions in a way that they can be answered, the PICO framework is useful: [10,19,23,43,44]

Population *In patients with cancer*	(What factors are essential?)
Intervention *what is the impact of blood thinning with heparin*	(Specific intervention or class?)
Comparator *compared with no heparin*	(Compared with nothing or with standard treatment)
Outcome *on…*	(Patient-relevant outcomes, including both benefits and potential side effects and burden and over what time, e.g. mortality at 2 years)

This format can also be used, with slight modifications, for questions on prevalence and incidence, etiology (exposure-outcome) and diagnosis. For instance:

- *In women in Saudi Arabia (P), what is the frequency of breast cancer (O)?*
- *In men over 40 years of age (P), what is the rate of lung cancer (O) in smokers versus non-smokers (C)?*
- *In babies born (P), does screening with a new rapid diagnostic test (I, C) accurately detect disease?*

5.6 Choosing and rating outcomes

Once the clinical questions for the guideline have been defined, identify the key outcomes that need to be considered in making the recommendations. Specially define the outcomes for foreground questions and for the outcomes that will be *critical* for making decisions and recommendations. These outcomes will also be used to guide the evidence retrieval and synthesis. It is important to focus on the outcomes that are important to patients, and to avoid the temptation to focus on those outcomes that are easy to measure and are often reported (unless those are also the important outcomes).[11,44,45]

Step 1. Create an initial, comprehensive list of possibly relevant outcomes for each question, including both desirable and undesirable outcomes from the interventions that will be considered in the recommendations.[45]

Step 2. Score the relative importance of each outcome from 1–9. Rating an outcome 7–9 indicates that the outcome is critical for a decision to recommend or not recommend a particular intervention or diagnostic test, 4–6 indicates that it is important, and 1–3 indicates that it is not important. The average score for each outcome can be used to determine the relative importance of each outcome, although it is helpful to provide the range of results as well. Sometimes people with different perspectives (patients, physicians, researchers, policy-makers)

have different opinions about which outcomes are important.[44,45] Therefore all these stakeholders should have an opportunity to contribute to the discussion on the selection of critical outcomes either by participation in the panel or by consultation.

Please note that meetings are not required to accomplish this task. These ratings can be conveniently completed using electronic tools.

5.7 Identifying resource implications

Once the key questions are formulated, the initiator should list the resource implications for the potential interventions that may be recommended. This might include for example, possible costs of new medicines or diagnostic tests, or possible outcomes, such as admission time to hospital, that might be associated with costs.[40,46] This step will inform any budget-impact assessment that will be carried out by one of the working groups of the guideline support unit.

5.8 Finalising the scope

Topics together with initial scope must be presented to NGC according to template (see Table 5.1). The NGC will assess the topics together with initial scope documents and will approve topics to be proper for guideline development. The NGC will consult and approve the composition of the guideline panel.

The panel may revise the initial scope based on the clinical importance of some the questions and outcomes, the potential evidence available or the potential for recommendations that will be useful in the Saudi Arabian health care context.[11,18,29,43] **It is critical to maintain the scope as narrow as sensible to ensure feasibility of completing the guideline in a timely manner.**

Final scope will be approved by NGC advisory board.

6. Panel meetings

The purpose of the meetings and expected tasks must be clearly laid out at the start, including:[11,18,32]

- what is expected from meeting participants;
- what needs to be achieved during the meeting;
- what can be done afterwards;
- what follow-up will take place with meeting participants;
- what the ground rules and processes to be followed (there should be no discussion about the process, i.e. members of the panel agree to the process when they agree to become a member).

Decisions are made based on consensus and voting as a form of forced consensus is used only in exceptional situations when consensus cannot be reached through discussion.[47] If voting takes place, existing models can be used.[48]

The panel will usually benefit from 2 to 3 face-to-face meetings with a minimum of one face-to-face meeting (the meeting to agree on recommendations). The purpose of the first meeting is generally to finalize the scope of the proposed guideline. At the (essential in person) meeting when recommendations are formulated, the panel reviews recommendations based on evidence prepared by guideline support unit. Another meeting might include finalizing plans for dissemination and for assessment of implementation of the guideline. Additional consultations (outside group meetings) may be held through electronic communication.

- If the purpose of the meeting is to formulate recommendations:
  - distribute the evidence profiles (see Appendix 4.1 for an example based on the GDT) prepared by the methodologist before the meeting, ideally two weeks before the meeting;
  - distribute the evidence to decision tables (see Appendix 4.2 for an example based on the GDT) prepared by the methodologist in consultation with the guideline panel before the meeting, ideally two weeks before the meeting;
  - at the meeting, present draft recommendations that have been prepared by the guideline support unit (meeting participants will comment on these and refine them).

6.1 Management of conflict of interests

- Nominated Panel members should declare to the NGC their conflict of interests, for example according to the declaration used by the World Health Organization.[22,41,49,50] The NGC oversight committee will decide whether any declared interest are such that a proposed panel member should not be included, for example due to significant financial or personal ties with a company who has an interest in a product that is the subject of the guideline; the NGC advisory board will resolve conflicts.

- Once the Guideline Panel is approved by the NGC, the guideline support unit collects declarations of interest (can be done electronically with the GDT) before the first meeting and later if there are any changes so that there is enough time to intervene if necessary (e.g. if any invited participant needs to be excluded owing to major conflicts or to prevent there being too many participants with potential conflicts of interest).
- At each panel meeting each participant reports verbally potential conflicts of interest (with actions taken if necessary); all panel members and any individuals who have direct input into the guideline should update their declaration of interests form before each panel meeting. Any changes to a panel member's declaration of interests should be recorded in the minutes of the panel meeting.
- Declarations of interests will be published in the final full guideline. [10,11,19,22,41,49-52]
- Recusal or excusal from certain decisions or recommendations is appropriate. If guideline panels involve members with (limited) conflict of interest, Chairs and group members on a guideline group should ensure that committees are reminded of the specific COI before discussion of individual conclusions or recommendations on which those COI bear. This will allow recusal from recommendations of those with important COI. Group chairs can play an active role and excuse group members from discussions or decision-making on particular recommendations.
- Procedures for handling disputes in conflict of interest resolution: Final decision about inclusion and exclusion from a panel with rest with the NGC. Chairs of guideline panels will have to review individuals' conflicts before each panel meeting and evaluate if the NGC should be involved based on new or changing conflict of interest. Once a member is approved for participation in a guideline panel, the panel chair will determine if specific panel members should be excused from individual recommendations or part of the discussion.

7. Evidence retrieval

7.1 What is 'evidence' for guideline development?

A summary of all relevant research evidence is essential when developing a recommendation, ideally based on systematic reviews.[49,53-56] In contrast to narrative reviews, systematic reviews address a specific question and apply a rigorous scientific approach to the selection, appraisal and synthesis of relevant studies. Systematic reviews, if conducted properly, reduce the risk of selective citation (the 'my favourite study' approach) and improve decisions.

Many guideline organizations rely on groups such as the Cochrane Collaboration for systematic reviews for use in guideline development. In countries or organizations with limited resources (including staff and expertise), however, it may be more practical and efficient to use existing systematic reviews including those used for already existing guidelines as the basis for local guideline development or to adapt recommendations from existing guidelines, and only occasionally develop recommendations *de novo.* The assumption here is that the research evidence to support a particular recommendation is 'global' whereas costs, values and

preferences and the feasibility of recommendations are 'local' considerations, and therefore should be the basis of adaptation of existing, sometimes international, recommendations.

Guidelines and recommendations will therefore be developed from a variety of sources, where the emphasis in existing guidelines is on the possibility of extracting information from highly credible systematic reviews from these guidelines:

1. recommendations and systematic reviews developed from published clinical guidelines that were created by independent organizations or groups that meet specified criteria (see “Retrieving and assessing existing guidelines”);
2. recommendations developed from existing systematic reviews;
3. recommendations developed from new systematic reviews;

Existing guidelines could be assessed for their credibility using validated tools such as the AGREE tool.[15] However, the emphasis is on finding systematic reviews that include the information of interest. If a guideline recommendation is required when there is truly little evidence to support a decision, then the panel will need to document the reasons for developing the recommendation based on little evidence and the basis for their judgement.[6,20,42,57,58] Such a recommendation may also be the basis for a proposal for research.

7.2 Evidence to decision tables

Regardless of the source of a recommendation or a systematic review, evidence to decision tables should be completed by a guideline panel, ideally for each recommendation or set of related recommendations based on the obtained information (see appendix 6.2).[57,58] Existing evidence to decision tables can be used and checked for relevance and possibility of adaptation.

7.3 Prioritizing evidence retrieval

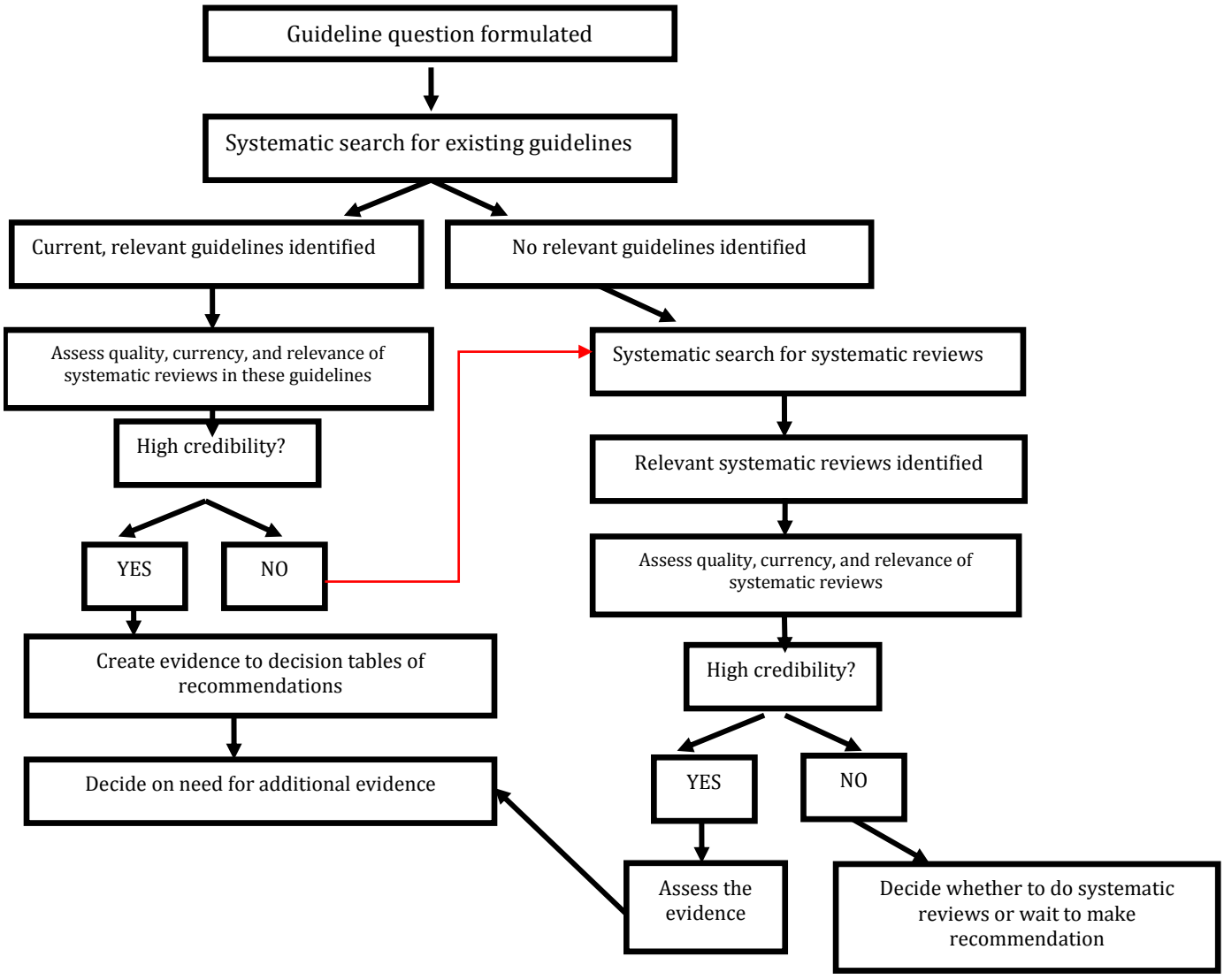
Whatever the source of the evidence, retrieving evidence to support every recommendation in a guideline may simply not be feasible. This is where it becomes important to identify priority questions or issues that the guideline should address (see section on defining the scope).[11]

To avoid duplication, the process outlined below starts by using existing guideline recommendations, and checking the evidence that relates to the recommendations (i.e. availability of systematic reviews supporting them), then describes the full process of developing recommendations based on systematic reviews, and includes a process for undertaking systematic reviews. This third option should be carried out only when there is no existing basis for recommendations and when the question is a major issue for the guideline to cover.[2,10,11,16,18-21,26,27,29,32,36] The methodology of development of systematic reviews is not covered in this handbook. Preparation of systematic reviews should follow the Cochrane Handbook for Systematic Reviews of Interventions (available at: http://www.cochrane.org/training/cochrane-handbook).

7.4 The process of evidence retrieval

The process of evidence retrieval, assessment and synthesis is described in further detail below and is summarized in the figure below (adapted from Estonian Handbook for guideline development)

Figure 7.1



7.5 Retrieving and assessing existing guidelines

Start by conducting a systematic search for existing guidelines (generally, those published in the last 5 years to ensure currency) on the same topic(s).[11] Guidelines can be difficult to find through electronic databases, so the following sources, in addition to Medline, may be helpful:

- the National Guideline Clearinghouse - http://www.guideline.gov/
- websites of guideline-producing agencies, such as NICE
- the guidelines international network (GIN) database of guidelines

It is strongly recommended to consult with an expert in information retrieval to ensure the use of a sound search strategy.

The search strategy should include key words for target population, intervention and comparators if relevant, etc. MeSH terms could be used. See http://www.ncbi.nlm.nih.gov/mesh.

The search strategy should be clearly documented and should specify:[19,23,36]

- the details of the sources (including web sites) searched, and the search used in each database including the date when the search was performed;

If relevant guidelines are identified the following aspects should be assessed:

1) are the guidelines based on systematic reviews?
    - if not, they should not be used.
    - If they are evidence based, are evidence summaries provided and are evidence to decision tables provided? (including GRADE evidence profiles, summary of findings tables, or references to systematic reviews)

2) who funded the guideline development?
    - what processes were used to manage conflicts of interest? If these are not described, the guidelines should not be used further, but there may be relevant systematic reviews or evidence profiles incorporated into them that can be helpful.

For the assessment tool it is suggested to use AGREE instrument questions 8-11 and 22-23.[15]

Once it is decided if the guidelines can be used as the basis for development of local recommendations, one will need to identify the recommendations in them that are relevant to the scope of the guideline. One approach that has been used is to make a table that includes all similar recommendations from different guidelines on the same topic, compare their sources (who produced them and if there are systematic reviews).

If the recommendations and the sources are the same, the main task is to develop or use evidence to decision tables (see section on developing recommendations). It is also possible that new evidence may be available that might need to be considered. Pragmatic decisions will have to be made about how to supplement the evidence in existing guidelines with new evidence, if

necessary. Advice on this should also be obtained from the content experts on the guidelines panel.

If an existing guideline has used GRADE evidence profiles as the basis for evidence presentation (based on systematic reviews), it may be possible to update the evidence profile and then reassess the recommendation, adding in considerations such as costs, local values and preferences, feasibility and other factors in the evidence to decision tables.

7.6 Retrieving existing systematic reviews

Rationale

Systematic reviews, if conducted properly, reduce the risk of selective citation and improve the reliability and accuracy of decisions. Systematic reviews should be assessed for their quality (see below "Adequacy of systematic reviews").[20,27,28,33,35,36]

Each systematic review under consideration should have a protocol that describes:[11,55]

- the search strategy used to identify all relevant published – and unpublished – studies;
- the eligibility criteria for the selection of studies;
- how studies will be critically appraised for risk of bias;
- an explicit method of synthesis of results and, if feasible, a quantitative synthesis of the results of studies to estimate the overall effect of an intervention (meta-analysis).

The first step is to identify relevant systematic reviews for each of your questions. The most readily accessible biomedical database is PubMed. The PubMed "Clinical Queries" or "Special Queries" options permit specific searches to be set up to identify systematic reviews of different types of studies identified with MeSH terms (see http://www.ncbi.nlm.nih.gov/mesh). This includes searches of the Cochrane Database of Systematic Reviews. A expert in information retrieval should be consulted through the NGC.

7.7 Adequacy of systematic reviews

Once the reviews are retrieved, they should be checked for:[11,18,23,59]

- relevance (to the questions to be addressed in the recommendations);
- timeliness (assessed by date of last update);
- quality (assessed by a standard critical appraisal instrument).

There are multiple checklists available for critical appraisal of systematic reviews, such as the one developed by the Oxford Centre for Evidence-Based Medicine (http://www.cebm.net/index.aspx?o=1157)_or, better, the validated AMSTAR toolto assess the credibility of a systematic review. [60,61] An update of the AMSTAR tool is being prepared (under the leadership of one of the authors of this handbook) and table 7.1 shows the current draft version which could be used for the assessment of systematic reviews.

If there are several relevant systematic reviews, use the most recent one that is of high quality. If the review is of high quality but more than two years old, consider updating the review to include more recent evidence and compare if other reviews include more or additional studies.

Table 7.1 Credibility of the Systematic Review Process

Did the Review Explicitly Address a Sensible Clinical Question?
Was the Search for Relevant Studies Exhaustive?
Was the Risk of Bias of the Primary Studies Assessed?
Did the review address possible explanations of between-study differences in results?
Did the review present results that are ready for clinical application?
Were Selection and Assessments of Studies Reproducible?
Did the Review Address Confidence in Effect Estimates (i.e, quality of evidence)?

8. Grading the quality of evidence

Assessing the retrieved evidence is a crucial step that enables the guideline panel to formulate recommendations.[62,63] The GRADE system for preparing evidence profiles, assessing quality of evidence and developing recommendations should be used.[5,63,64] This approach allows for a structured and transparent assessment of the quality of evidence for each outcome. For each question, there should be relevant data (from the systematic review) for all the outcomes (benefits and harms) that were rated as important.

If GRADE tables have already been prepared for the published guidelines they should be used as the basis for formulating recommendations using the evidence to decision frameworks as described above.[54,65] If there are no GRADE evidence summaries (evidence profiles or summary of findings), the guideline panel will have to decide whether to retrieve the systematic reviews on which the recommendations are based, and to prepare evidence summaries, or simply to use the existing recommendations, and apply considerations of cost, local values and preferences and feasibility. For potentially high-cost interventions it is strongly suggested that the systematic review be retrieved and evidence summaries prepared.

The GRADE handbook contains (electronic help file in the GDT) all the instructions for developing GRADE evidence profiles, and the software for the entire guideline development process can be accessed on http://www.guidelinedevelopment.org/.

In the GRADE system, the quality of evidence in the context of clinical practice guidelines reflects "the extent to which confidence in an estimate of the effect is adequate to support recommendations". [63] It is implicit in the definition that guidelines panels have to judge the quality of the evidence along with the specific context in which the evidence is being used.

Table 8.1 explains the quality assessment according to GRADE. Although RCTs start as high quality evidence, they can be downgraded depending on whether issues of risk of bias, indirectness, imprecision, inconsistency, and publication bias are detected in the body of evidence. On the other hand, observational studies start as low quality evidence; however, they can be upgraded to moderate or high quality evidence if they are methodologically sound and evidence of a large magnitude of effect, dose-response gradient or plausible confounding, which would reduce a demonstrated effect, are identified. It is important to highlight that the assessment of the quality of the evidence should be conducted at an outcome level, across studies.
This handbook will not provide details on grading as this is described in the GDT and in the series of articles cited below. An overview is provided here.

Table 8.1 GRADE's approach to rating quality of evidence (aka confidence in effect estimates)

For each outcome based on a systematic review and across outcomes (lowest quality across the outcomes critical for decision making)

1. Establish initial level of confidence		**2. Consider lowering or raising level of confidence**		**3. Final level of confidence rating**
Study design	***Initial confidence in an estimate of effect***	***Reasons for considering lowering or raising confidence***		***Confidence in an estimate of effect across those considerations***
		⍰ Lower if	**⍰ Higher if***	
Randomized trials⍰	**High confidence**	**Risk of Bias** **Inconsistency** **Indirectness** **Imprecision** **Publication bias**	**Large effect** **Dose response** **All plausible confounding & bias** • would reduce a demonstrated effect **or** • would suggest a spurious effect if no effect was observed	**High** ⨁⨁⨁⨁
				Moderate ⨁⨁⨁◯
Observational studies⍰	**Low confidence**			**Low** ⨁⨁◯◯
				Very low ⨁◯◯◯

*upgrading criteria are usually applicable to observational studies only.

8.1 Limitations that can reduce the quality of the evidence:

1. Risk of bias or limitations in the detailed study design and execution: RCTs and observational studies may suffer from limitations in the study design that could increase the risk of misleading results. Although this assessment is conducted at a study level, the risk of bias can differ across outcomes. [66] Some of the reasons for downgrading by one or two levels the quality of the evidence of RCTs are:

- Lack of allocation concealment
- Lack of blinding (particularly if outcomes are subjective and their assessment highly susceptible to bias)
- Large loss to follow-up
- Failure to adhere to an analysis according to intention-to-treat principle
- Selective reporting of events: investigators neglect to report outcomes that they have measured (typically those for which they observed no effect).

Consider the following example from the GRADE series in the Journal of Clinical Epidemiology. The table below was extracted from a systematic review summarizing the evidence on the use of flavonoids for treating haemorrhoids (Figure 8.1) [67]. The table describes the risk of bias assessment of all the included studies providing evidence for the outcome of persisting symptoms. Most of the trials did not provide enough information to determine which method was used to generate the randomization sequence nor the appropriateness of the allocation concealment. Most of the studies described blinding using the terms double blinding, with no clear specification of who was blinded. Finally, the majority of the trials failed to conduct an intention-to-treat analysis and did not report enough data to allow readers to conduct it.

After conducting an assessment of the risk of bias at a study level, it is necessary to obtain an overall estimate of the risk of bias for the body of evidence informing a particular outcome. Since we are using RCTs to inform this outcome, the quality of the evidence started as high quality; however, due to issues of risk of bias it has to be downloaded at least one level going from high to moderate.

Figure 8.1 Risk of bias example

Risk of bias for measurement of symptoms in studies of flavonoids in patients with hemorrhoids

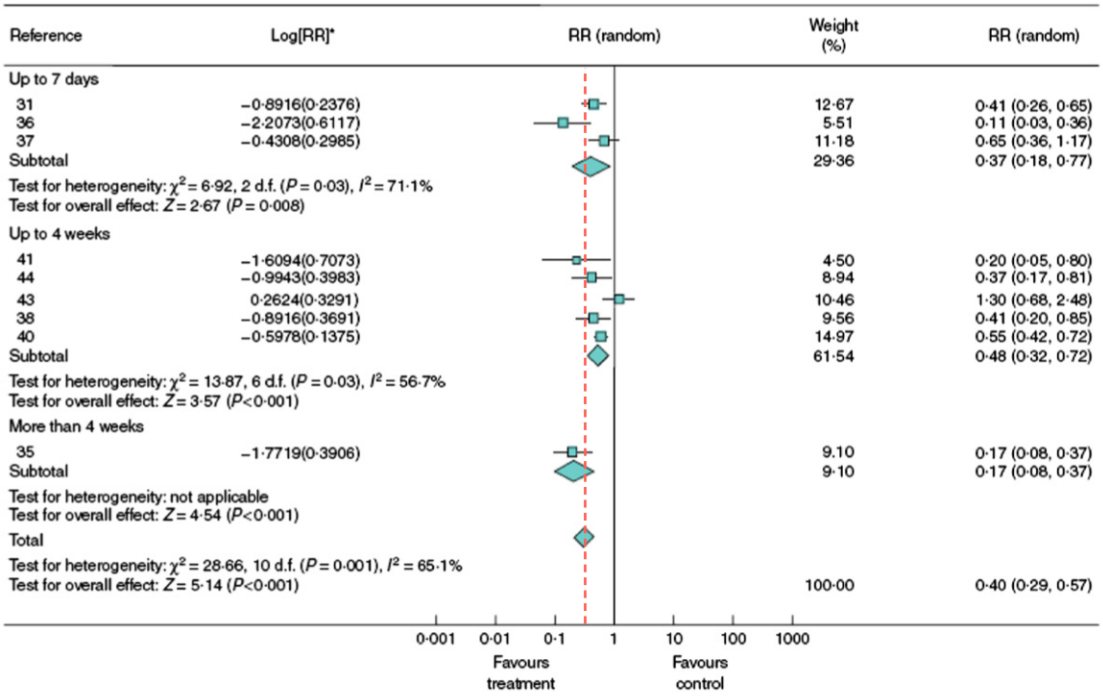
Study[c]	Randomization	Allocation concealment	Blinding	Loss to follow-up[a]/IT principle observed or per protocol analysis	Other
Dimitroulopoulos D, 2005	Adequate[b] Computer-generated random numbers[b]	Sealed opaque envelopes[b]	Described as single blind Care givers, patients, and data collectors blinded[b]	3%/protocol	Unvalidated symptom measure
Misra MC, 2000	Adequate Computer-generated random numbers[b]	Adequate Sealed opaque envelopes[b]	Patients and physicians[b] Described as double blind Placebo identical appearance	2%/protocol	Unvalidated symptom measure
Godeberge P, 1994	Adequate[b]	Adequate Sealed opaque envelopes[b]	Patients, physician-investigator, data manager, statistician, and authors blinded	6%/protocol	
Cospite M, 1994	Unclear	Unclear	Unclear Described as double blind	12%/IT	Unvalidated symptom measure
Chauvenet-M, 1994	Unclear	Unclear	Unclear	11%/protocol	Unvalidated symptom measure
Ho Y-H, 2000	Adequate Drawing of sealed opaque envelopes[b]	Adequate Sealed opaque envelopes	All parties blinded[b]	0%/IT	Unvalidated symptom measure
Thanapongsathorn W, 1992	Unclear	Unclear	Unclear Described as double blind	I2%/protocol	Unvalidated symptom measure
Titapant V, 2001	Unclear	Unclear	Unclear Described as double blind Placebo identical appearance	12%/protocol	Unvalidated symptom measure
Wijayanegara H, 1992	Unclear	Unclear	Unclear Described as double blind	3%/protocol	Unvalidated symptom measure
Annoni F, 1986	Unclear	Unclear	Unclear Described as double blind Placebo identical appearance	Uncertain/unclear	Unvalidated symptom measure
Thorp RH, 1970	Unclear	Unclear	Physicians and patients blinded Described as double blind Placebo identical appearance	20%/protocol	Unvalidated symptom measure
Clyne MB, 1967	Bottles numbered consecutively in accordance to random tables	Unclear	Physicians and patients blinded Described as double blind Placebo identical appearance	Uncertain/protocol	Unvalidated symptom measure
Sinnatamby CS, 1973	Unclear	Unclear	Physicians and patients blinded Described as double blind	53%/protocol	Unvalidated symptom measure
Trochet JP, 1992	Randomized by blocks of three (method unclear)	Unclear	Physicians blinded Placebo identical appearance	Uncertain/IT	Unvalidated symptom measure

Abbreviation: IT, intention-to-treat principle observed.
[a] No important differences in rate of loss to follow-up between flavonoid and control groups in any study.
[b] Data provided by authors.
[c] For full citation of the references cited in this table, see Alonso-Coello et al.[36]

2. Inconsistency: Widely differing estimates of the treatment effect (i.e. heterogeneity or variability in results) across studies suggest true differences in underlying treatment effect. When evidence of heterogeneity exists, but investigators fail to provide a plausible explanation, the quality of evidence should be downgraded by one or two levels, depending on the magnitude of the inconsistency in the results. [66]

A systematic review provides a summary of the data from the results of a number of individual studies. If the results of the individual studies are similar meta-analysis is used to combine the results from the individual studies and an overall summary estimate is calculated. The meta-analysis gives weighted values to each of the individual studies according to their size. The individual results of the studies need to be expressed in a standard way, such as relative risk, odds ratio or mean difference between the groups. Results are traditionally displayed in a figure (see Figure 8.2) called a **forest plot.**

Figure 8.2 Forest plot



The forest plot depicted above represents a meta-analysis of 9 trials that assessed the effects of flavonoids for the treatment of haemorrhoids [67]. Individual studies are represented by a square and a horizontal line, which corresponds to the point estimate and 95% confidence interval of the relative risk. The size of the square reflects the weight of the study in the meta-analysis. The solid vertical line corresponds to 'no effect' of treatment – a relative risk of 1.0. When the confidence interval includes 1 it indicates that the result is not significant at conventional levels (P > 0.05).

The diamond at the bottom represents the combined or pooled relative risk of all 9 trials with its 95% confidence interval. In this case, it shows that the treatment reduces persisting symptoms by 60% (RR 0.40 95% CI 0.29 to 0.57). Notice that the diamond does not overlap the 'no effect' line (the confidence interval doesn't include 1) so we can be assured that the pooled RR is statistically significant. The test for overall effect also indicates statistical significance (p<0.001).

Heterogeneity can be assessed by eyeballing or more formally with statistical tests, such as the Cochran Q test and the I^2 value. With the "eyeball" test one looks for the similarity of the point estimates and the overlap of the confidence intervals of the trials with the summary estimate. In the example above note that the dotted line running vertically through the combined relative risk does not cross the horizontal lines of all the individual studies (3/9) indicating small to moderate degree of heterogeneity among the included studies. If Cochran Q is statistically significant there is an indication for heterogeneity. If Cochran Q is not statistically significant but the ratio of Cochran Q and the degrees of freedom (Q/df) is > 1 there is possible heterogeneity. If Cochran Q is not statistically significant and Q/df is < 1 then heterogeneity is very unlikely. In the example above Q/df is >1 (28.66/10= 2.866) and the p-value is significant (0.001) indicating heterogeneity. Finally, the I^2, which quantifies the proportion of the variation in point estimates due to among-study differences is large (65%). The higher the I^2 the greater the inconsistency (i.e. that differences between studies are not likely due to chance).

3. Indirectness: Two types of indirectness are relevant. First, a review of the evidence comparing the effectiveness of alternative interventions (say A and B) may find that randomized trials are available, but they have compared A with placebo and B with placebo. Thus, the evidence is restricted to indirect comparisons between A and B.

Second, an evidence review may find randomized trials that meet eligibility criteria but which address a restricted version of the main review question in terms of population, intervention, comparator or outcomes. For example, suppose that in a review addressing an intervention for secondary prevention of coronary heart disease, the majority of identified studies happened to be in people who also had diabetes. Then the evidence may be regarded as indirect in relation to the broader question of interest because the population is restricted to people with diabetes. The opposite scenario can equally apply: a review addressing the effect of a preventative strategy for coronary heart disease in people with diabetes may consider trials in people without diabetes to provide relevant, albeit indirect, evidence. This would be particularly likely if investigators had conducted few if any randomized trials in the target population (e.g. people with diabetes). Other sources of indirectness may arise from interventions studied (e.g. if in all included studies a technical intervention was implemented by expert, highly trained specialists in specialist centres, then evidence on the effects of the intervention outside these centres may be indirect), comparators used (e.g. if the control groups received an intervention that is less effective than standard treatment in most settings) and outcomes assessed (e.g. indirectness due to surrogate outcomes when data on patient-important outcomes are not available, or when investigators sought data on quality of life but only symptoms were reported). Review authors should make judgements transparent when they believe downgrading is justified based on differences in anticipated effects in the group of primary interest. Review authors may be aided and increase transparency of their judgments about indirectness if they use 8.2 (available in the GDT software). [68]

Table 8.2 Judgements about indirectness by outcome

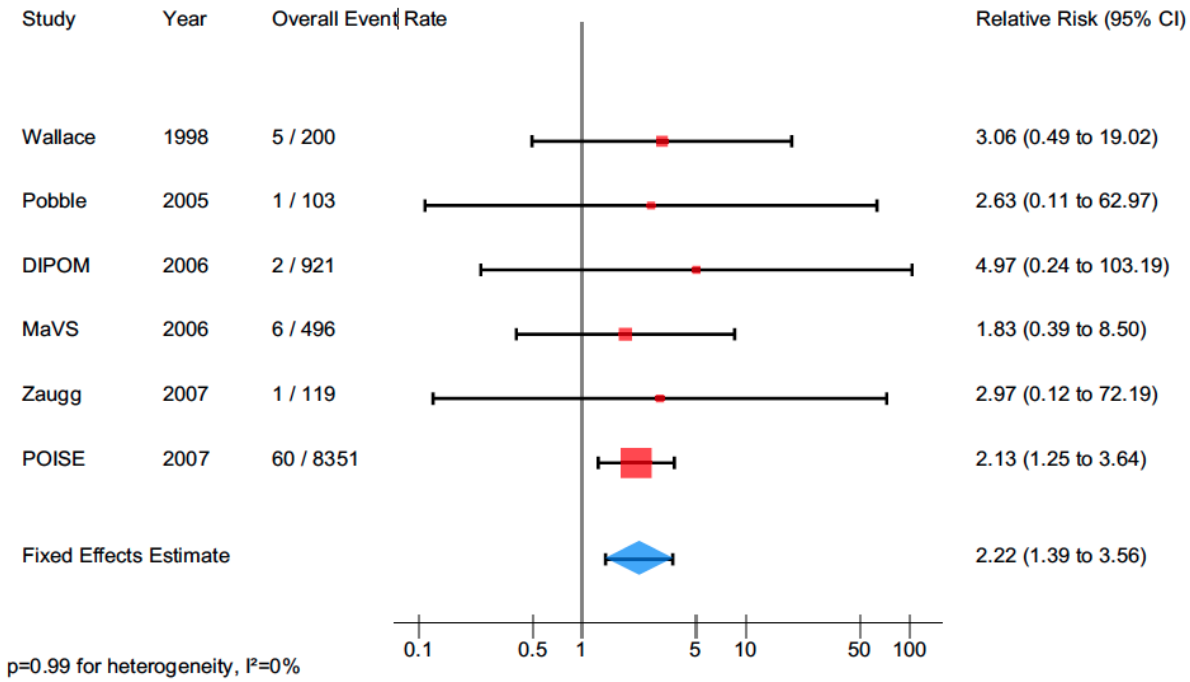
Outcome		
Domain (original question asked)	**Description (evidence found and included, including evidence from other studies) – consider the domains of study design and study execution, inconsistency, imprecision and publication bias**	**Judgment - Is the evidence sufficiently direct?**
Population:		Yes ☐ Probably yes ☐ Probably no ☐ No ☐
Intervention:		Yes ☐ Probably yes ☐ Probably no ☐ No ☐
Comparator:		Yes ☐ Probably yes ☐ Probably no ☐ No ☐
Direct comparison:		Yes ☐ Probably yes ☐ Probably no ☐ No ☐
Outcome:		Yes ☐ Probably yes ☐ Probably no ☐ No ☐
Final judgment about indirectness across domains:		☐ **No indirectness** ☐ **Serious indirectness** ☐ **Very serious indirectness**

For example, a panel of a clinical practice guideline was formulating a recommendation about the use of antiviral therapy for treating avian influenza. The available evidence showed that RCTs of high quality (low risk of bias) have demonstrated that antiviral treatment is effective for managing seasonal influenza; however, the panel had concerns about whether the underlying biology of seasonal influenza is different enough from the avian one to decrease the confidence in the estimates of effect. In this case, the quality of the evidence can be downgraded by one or two levels due to indirectness.

4. Imprecision: Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect. In the context of clinical practice guidelines, if a recommendation or a clinical decision-making would differ if the upper versus the lower boundary of the CI represented the truth, one should consider to rate down for imprecision. [69]

For example, the forest plot below represents a meta-analysis of 6 studies about the use of β-blockers for preventing cardiovascular events in patients undergoing noncardiac surgery. [70] The pooled estimate (blue diamond) suggests a doubling of the risk of stroke using the intervention (RR: 2.22; 95%CI: 1.39-3.56). In this case, two main arguments can support the decision of rating down the quality of the evidence due to imprecision. First, the lower boundary of the 95% confidence interval suggests a 39% increase in the risk of stroke, while the upper limit suggests a 256% reduction on the risk for this cardiovascular event. Since both, appreciable benefit and considerable harm are being included in the 95% confidence interval. Second, Only 75 events (from 10,889 participants) are informing this outcome. If conventional sample size estimation is conducted, 43,586 participants are required to detect a clinically relevant difference (α 0.05, β 0.20, β-blocker group's 1% rate, and an effect size of 0.25). Under these circumstances, a guideline panel may decide to downgrade due to imprecision.

Figure 8.3 Forrest plot showing imprecision

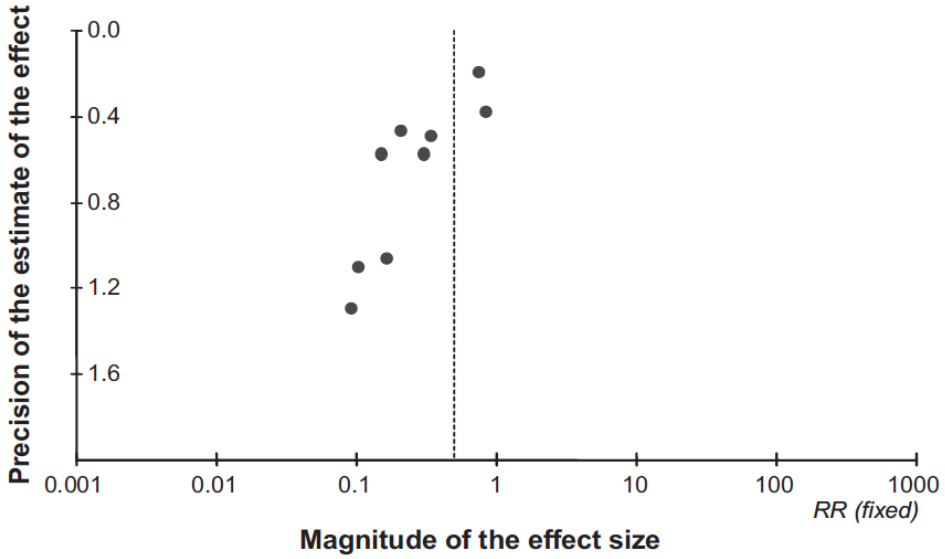


5. Publication bias: Corresponds to a systematic underestimate or an overestimate of the underlying beneficial or harmful effect of an intervention due to the selective publication of studies. That is, investigators fail to report studies they have undertaken (typically those that can be considered as "negative" results) or journals are less likely to accept studies that show no effect for publication.[71] Publication bias should be suspected especially when studies are consistently small, and sponsored by the industry.

Following the example described earlier in this section, a systematic review of the use of flavonoids for treating haemorrhoids included 9 studies reporting on the outcome persistent symptoms[67]. Figure 8.4 depicted below is a **funnel plot.** The x-axis represents the magnitude of the effect size, while the y-axis is the precision of the estimate of the effect. The small dots populating the figure represent the point estimate of each included trial. Larger studies tend to be grouped around the pool estimate (dotted vertical line) and show more precision (located at the top) than smaller studies (located at the bottom). The ideal funnel plot should be symmetric, which means that small "positive" and "negative" trials are equally distributed around the point estimate. Asymmetric funnel plot suggests the presence of publication bias, particularly when studies are missing at the bottom right quadrant.

The figure suggests the presence of publication bias due to the asymmetry of the funnel plot (studies at the bottom right quadrant are missing). Likewise, all the included trials recruited on average no more than 100 participants and all of them were industry sponsored. Thus, collecting these different pieces of information, the suspicious for publication bias cannot be discarded. Under these circumstances, the quality of the evidence should be rated down by one level due to publication bias.

Figure 8.4 Funnel plot to detect publication bias



8.2 Factors that can increase the quality of the evidence:

1. Large magnitude of effect: When methodologically strong observational studies yield large or very large and consistent estimates of the magnitude of a treatment or exposure effect, the quality of the evidence increases. In this particular case, the weaknesses of the observational design is unlikely to explain all of the apparent benefit or harm, even though observational studies are likely to provide an overestimate of the true effect. The larger the magnitude of effect, the stronger becomes the evidence: [72]. A large effect (e.g. RR > 2 or RR < 0.5) in the absence of plausible confounders, or a very large effect (e.g. RR > 5 or RR < 0.2) in studies with no major threats to validity, might qualify for this.

For example, a systematic review studied the association between infant sleeping position and sudden infant death syndrome. [73] An impressive 4-fold increase in the risk of sudden death was found for front compared to back sleep position (OR: 4.1; 95% CI: 3.1-5,5). Subsequent studies supported these findings increasing even more the confidence in the estimates of effect. When large magnitude of effect are reported in the body of evidence for a particular outcome, this allows to rate up the quality of the evidence for observational studies from low to moderate or even high quality of the evidence.

2. Dose-response gradient: The presence of a dose-response gradient may increase the confidence in the findings of observational studies. Only studies with no threats to validity or any other of the criteria for downgrading can be upgraded by one or two levels. [72]

For example, observational data shows that patients with supra-therapeutic anticoagulation therapy levels have an increased risk for bleeding. The strength of this association increases when there is evidence of a dose-response gradient between the two variables. In this particular case, the higher the levels of the international normalized ratio, the higher the risk of bleeding.[74] Under this circumstance, and as long as the quality of the body of evidence from observational studies shows, otherwise, rigorous methodologies, it can be rated up from low to moderate or high confidence in the estimates of effect.

3. Plausible confounding, which would reduce a demonstrated effect: All plausible confounding from observational studies may be working to reduce the demonstrated effect or increase the effect if no effect was observed. [72] This phenomenon increases the confident in the association demonstrated and consequently, increases the quality of the evidence.

For example, a systematic review of observational studies summarized the evidence about the contrast between private for-profit vs. private not-for-profit hospitals and death rates.[75] The review showed that there was a higher death rate in private for-profit institutions than private non-for profit ones. One potential source of bias is the severity of the diseases between patients in these hospitals. Since it is expected that patients in the non-for-profit hospital were sicker than the ones in for-profit hospitals, this potential bias would work against the review findings, providing more confidence for the association found. In this case, the quality of the body of the evidence from observational data informing about death rates and profit hospital characteristic can be upgraded by one or two levels from low quality to moderate or even high quality evidence.

If further information is required about the process of assessing the quality of the evidence or grading the recommendations, please visit the GRADE guideline series published in the Journal of Clinical Epidemiology.[40,44,65,66,68,69,71,72,76-84]

8.3 Presenting the evidence to the Guideline Panel

Draft evidence summaries, along with a draft assessment of values, preferences, and costs, should be sent to the members of the panel before the meeting. Panel members should be asked to identify any relevant evidence that is missing from the summaries. The final summaries are then used as the basis for drafting recommendations.[54,65]

9. Assessing cost implications

When developing guideline recommendations the cost implications of alternative strategies have to be taken into account by the guideline panel.[85,86] Guideline panels should evaluate the potential consequences of alternative scenarios and consider cost implications in addition to health outcomes. In the KSA context cost implication analysis are divided into budget impact and economic evaluation analysis. Economic evaluation of interventions not included in the health care services list financed by MoH and the reimbursed pharmaceuticals list should be done according to procedures set in legislation.

It is important to assess the cost implications related to potential changes in current clinical practice standard related to each developed recommendation. This has to be done in parallel with developing recommendation enabling to consider cost information as one input into the process when moving from evidence to recommendations.[10,40,85] Generally, all important resource use associated with the recommendation (suggested intervention and, if available, the comparator) – are assessed.

Consideration of the economic consequences of potential guideline recommendations has to be taken after defining the final scope of the guideline. In this stage it is also suggested to decide which of the recommendations probably will need economic evaluation in addition to budget impact analysis. A summary of budget impact analysis should be done for all initial recommendations by describing also alternatives. A full economic evaluation might be worthwhile if an unbiased effectiveness measure is available, and a review of country relevant existing economic studies may be useful to inform the definition of resource use (for a costing exercise only or for a full economic evaluation).[40,85]

The description of resource use and costs should be made from the perspective of the health system by identifying the main resources required to implement a specific recommendation. It is important to include resource use associated with the provision of the intervention, subsequent investigations and care, and adverse effects.[40,85] Implications not only for MoH but also for other stakeholders (hospitals etc) should be taken into account. These should be grouped as costs incurred by the patient, the health system and society. Those incurred by the patient and health system should be described (e.g. drug, admissions, visits, examinations). Other resources, such as patient and care-giver time, should generally be considered only when they are considered to be very important in that context as they are difficult to measure and to put a value on reliably. It is also important to define the time horizon for inclusion of resource use – when are important differences in resource use likely to occur (in the short term or the long term).[10,40,46,85,86]

Cost-effectiveness analyses must be done selectively. A full economic evaluation of cost-effectiveness if conducted has to take into account the costs and health outcomes (effects) of an intervention assessed in relation to its comparator, and present incremental cost-effectiveness ratio (ICER). Effectiveness measures can be natural units (e.g. disease episodes or deaths prevented), two-dimensional quality-adjusted life years (QALYs) in a cost–utility analysis, or can be expressed in monetary terms in a cost–benefit analysis. Cost-effectiveness analyses often use decision-analytic methods in order to combine evidence from different sources and to extrapolate from the limited time-horizons of existing studies on health outcomes. Once the cost-effectiveness of an intervention is established, an evaluation should be made as to whether the intervention is affordable and represents value for money.[85]

Costs must be taken into account additionally to clinical evidence in process of approving specific recommendation by guideline panel.[46] Wording recommendation as strong is suggested only in cases if the intervention or pharmaceutical is affordable in Saudi Arabia or accepted for financing by MOH or some other state agency.

Once resource use is measured, a range of monetary values can be estimated for each item of resource use. For reporting on this costing exercise, it is important not just to document the aggregate costs (number of units of resource use x unit costs of resource) associated with an intervention, but also to report as far as possible disaggregated costing information (i.e. all the associated resource use and unit costs separately).

Practical guidance:
Refer to interpretation of strong and weak recommendations.
Guideline panels will take a health systems perspective. Panels will label situations in which resources are a driver of conditional/weak recommendations. When panels feel that they should make a strong recommendation based on the overwhelming benefits, then it will be up to the ministry to implement or make a decision about implementation. If a panel feels that there is benefit but the technology is possibly too costly to implement then they might offer a conditional recommendation indicating that the resource considerations are the key factor for the conditional recommendation. This provides appropriate guidance to clinicians and the ministry (as opposed to misleading guidance if offering strong recommendation for something that is not available) and allows MoH to prioritize across health conditions and identify recommendations that are conditional/weak solely bearing on resources/implementability.

10. Developing recommendations

Draft (neutral) recommendations are prepared by guideline support unit and final recommendations (direction and strength) must be approved by panel.

For each recommendation, the quality of evidence and strength of recommendation should be presented.[13,57,87,88] Recommendations should specify the perspective that is taken (e.g., individual patient, health care system or society) and which outcomes were considered (including costs, if assessed). The language used in recommendations should be clear and direct, indicating an unambiguous action (e.g., all patients with disease A should be offered treatment B by health professionals).[11,57,88] Preferably the language should be consistent across recommendations (e.g., all strong recommendations phrased with "should" – see below for wording of recommendations).[11]

10.1 How a panel decides on recommendations

The panel should reach recommendations based on consensus. Consensus does not necessarily mean unanimity, however, and in some cases, at the discretion of the chair, a vote may need to be taken.[89] Voting can then be used as a tool to work toward consensus. Panel members collaborate with the chair to achieve the wording of final recommendations. The group should discuss and agree on the process at the beginning of the meeting.

It is most effective if the group considers draft recommendations that have been prepared by the guideline support unit. A suggested process is as follows:

- the draft recommendation is presented by the guideline support unit using the evidence to decision framework or tables, with a justification and reference to the relevant evidence (evaluated by GRADE) summary;
- the evidence is reviewed and discussed by the group, considering the balance of evidence for benefits and harms;
- the panel considers costs or resource utilization, ideally presented by health economists of the guideline support unit, budget impact, and possibly cost-effectiveness, along with values and preferences;
- if necessary, the first recommendation is modified;
- final agreement on the recommendation is reached.

10.2 Grading recommendations

The strength of a recommendation reflects the degree of confidence that the desirable effects of adherence to the recommendation outweigh the undesirable effects.[4,62,64,87]

Desirable effects can include beneficial health outcomes, less burden and greater savings. Undesirable effects can include harms, greater burden, and increased costs. Burden here refers to the demands of adhering to a recommendation that patients or care-givers (e.g., family members) may find onerous – such as having to undergo more frequent tests or opting for a treatment that may require a longer time for recovery.[57,77]

Although the degree of confidence is a continuum, the GRADE system defines two categories – strong and conditional (also known as "weak"). A strong recommendation is one for which the guideline development group is confident that the desirable effects of adherence outweigh the undesirable effects. This can be either in favour of or against an intervention. A weak recommendation is one for which the panel concludes that the desirable effects of adherence probably outweigh the undesirable effects, but the group is not confident about the trade-off.[4,5,62,87] Reasons for not being confident may include:

- absence of high-quality evidence;
- presence of imprecise estimates of benefit or harm;
- uncertainty or variation in how different individuals value the outcomes;
- small benefits;
- benefits that are not worth the costs (including the costs of implementing the recommendation).

Despite the lack of a precise threshold for moving from a strong to a conditional (also known as "weak") recommendation, the presence of important concerns about one or more of the above factors make a weak recommendation more likely (see Table 10.1). The Guideline Development Panel should consider all these factors and make the reasons for their judgments explicit.

Implications of a strong recommendation are:[4,5,62,87]

- For patients: most people in your situation would want the recommended course of action and only a small proportion would not.
- For clinicians: most patients should receive the recommended course of action. Adherence to this recommendation is a reasonable measure of good-quality care.
- For policy-makers: the recommendation can be adopted as a policy in most situations. Quality initiatives could use this recommendation to measure variations in quality.

Implications of a conditional recommendation are:[4,5,62,87]

- For patients: the majority of people in your situation would want the recommended course of action, but many would not.
- For clinicians: be prepared to help patients to make a decision that is consistent with their own values.
- For policy-makers: there is a need for substantial debate and involvement of stakeholders.

It is strongly suggested to present the implications in a written copy of each guideline to facilitate interpretation. However, the implications should not be seen as definition of the strength of a recommendation.

Table 10.1 Factors that may influence the strength of recommendations

Factor	Examples of strong recommendations	Examples of conditional (weak) recommendations
Quality of evidence	Many high-quality randomized trials have demonstrated the benefit of inhaled steroids in asthma	Only case series have examined the utility of pleurodesis in pneumothorax
Uncertainty about the balance between desirable and undesirable effects	Aspirin in myocardial infarction reduces mortality with minimal toxicity, inconvenience and cost	Warfarin in low-risk patients with atrial fibrillation results in small stroke reduction but increased risk of bleeding and substantial inconvenience
Uncertainty or variability in values and preferences	Young patients with lymphoma will invariably place a higher value on the life-prolonging effects of chemotherapy over treatment toxicity	Older patients with lymphoma may not place a higher value on the life-prolonging effects of chemotherapy over treatment toxicity
Uncertainty about whether the intervention represents a wise use of resources	The low cost of aspirin as prophylaxis against stroke in patients with transient ischaemic attacks	The high cost of clopidogrel and dipyridamole/aspirin as prophylaxis against stroke in patients with transient ischaemic attacks

Many recommendations are labelled as either strong or conditional. However, because the "conditional" label may sometimes be misinterpreted, other options exist. These include the use of terms such as "strong/weak" or "strong/qualified".

The wording of recommendations is important.[57,77] To ensure that end users will understand the specific linguistic and cultural contexts of the wording, sample text should be validated with them. The key to the wording must always be attached to the guideline. Some examples are in table 10.2 below.

Table 10.2 Wording of recommendations

	Wording 1	Wording 2	Wording 3
Strong recommendation for	We recommend…	Clinicians should…	We recommend…
Weak recommendation for	We suggest	Clinicians might…	We conditionally recommend…
Weak recommendation Against	We suggest...not	Clinicians might not…	We conditionally recommend...not
Strong recommendation Against	We recommend …not	Clinicians should not…	We recommend …not

Example of a conditional recommendation: The KSA MoH panel suggests sublingual immunotherapy for treatment of adults with seasonal or intermittent allergic rhinitis (conditional recommendation; Moderate-quality evidence).

Example of a strong recommendation: The KSA MoH panel recommends intranasal corticosteroids rather than intranasal H1-antihistamines for treatment of adults with seasonal or intermittent allergic rhinitis (Strong recommendation; High-quality evidence).

Please refer to appendix 6.2 for detailed examples.

10.3 Indicators for implementation

The Guideline Development Panel should suggest indicators for monitoring the implementation of the guideline and its impact, based on the final recommendations as part of the evidence to decision tables.[10,11,18,21,23,59] When conditional recommendations are selected (ideally only those based on high quality evidence) the decision making process (a dyad approach between the patient and the clinician) can function as quality indicator.

In general, indicators can be *process* indicators (for example, prescription rates for specific medicines; length of hospital stay) or *outcome* indicators, such as readmission to hospital due to a specific cause, or clinical events (for example, patients experiencing myocardial infarction).

The indicators that are selected by the guideline panel should be events or processes that are expected to be affected as result of the recommendation and in some instances may be the same as the critical outcomes used by the Panel in making recommendations.[10,11,18,21,23,59,90-92] They should also be processes or events that can be measure through routine data collection by the MOH or through audit that can be done as part of the guideline implementation process. There is no pre-specified number of indicators required for a guideline, but if there are several strong recommendations, there may need to be several indicators.

The final selection of indicators should be done in consultation with the MOH and the key societies likely to be involved in implementing the guideline and approved by the NGC.

11. Producing and disseminating the guideline

The person or group in charge of producing the guideline must be identified in the early stages of the guideline development. If a member of the panel group will produce the draft of the guideline, there must be sufficient time in the timeline to do this. If the services of an external group will be used, it is important to consider the experience of the group and what are the resources that must be allocated to this purpose.

The guideline should have three main parts: 1. Executive summary including a summary of the key recommendations; 2. Main text; and 3. The evidence to recommendation tables and evidence profiles.

The key recommendations of the guideline should be easily identifiable[15] and should answer to the guideline questions in a clear and concise manner. The guideline summary must provide an overview of the guideline. It should briefly describe the scope, aim, target audience, methods and recommendations. The main text must describe with details the guideline development process. It must be complemented with appendices that could include the following:

- Information about panel members and external reviewers
- Conflicts of interest of the panel members
- Search strategies and electronic datasets used (with dates and number of hits)
- If recommendations from existing guidelines were adapted, quality assessment of the guidelines from which they were obtained
- Risk of bias assessment of the primary studies included, or methodological quality assessment of the systematic reviews included
- Evidence profiles used by the panel members to formulate the recommendations
- Evidence to decision frameworks used to formulate the recommendations
- Supplementary material for clinicians and patients

The dissemination of the guideline should aim to deliver it to the target audience in order to facilitate its implementation. Not only a printed version, but also an online version, publication in peer-reviewed journals, educational interventions such as meeting and conferences, and introduction to electronic medical record systems should be considered.

The plan for dissemination should consider the scale to which the guideline will be implemented, and used strategies according to this.

12. Guideline authorship

The decision about guideline authorship rests with the NGC. Generally, the guidelines will be authored by the "The Saudi Center for EBHC". Panel members will be acknowledged by name as members of the "Saudi Expert Panel" if they participate meaningfully in the guideline development process, e.g. in meetings, question development, developing recommendations, review of the final guideline document. Other contributors, such as systematic review authors or support staff will be listed under appropriate headings, such as working groups. The address for correspondence of the guidelines will be the "The Saudi Center for Evidence Based Health Care".

13. Updating a guideline

The process for updating a guideline must already be planned during the guideline development, and it should be reported in the guideline publication. The aim of updating a guideline is to include any new evidence relevant to the guideline questions, and to reflect any important change in the setting for which the recommendation was developed.

Some of the main reasons for updating a guideline are:[93]

- Availability of new evidence regarding the benefits and harms of the interventions recommended in the guideline
- Availability of new evidence regarding critical and important outcomes, for which there was not high quality evidence when the recommendations were formulated
- Development of new intervention that could be a valid alternative to those recommended in the guideline
- Changes in the health system, which modify the setting for which the guideline was developed
- Changes is issues related to the resources needed to implement the recommendations

The updating plan should specify a time period after a complete revision of the evidence included in the guideline would be done. On average, this revision should be done every 3-5 years.[11,94] The specific time depends on the speed to which the evidence in a specific medical area evolves.

When updating a guideline, the same process for guideline development described in this handbook should be used.

14. Glossary and references

The glossary (from reference[9]) includes definitions of terms and acronyms appearing throughout the handbook to help with interpretation of the items included. Related terms in the list are grouped into categories describing the various aspects of guideline development.

Term	Definition
Groups, individuals, and organizations involved in the guideline development process	
Guideline development panel	The entire group of healthcare and other professionals, stakeholders, patients and carers, research and technical staff who develop a guideline. The guideline development group may consist of several task-specific subgroups or committees such as the oversight committee, guideline panel, stakeholder and consumer consultants, and working group. Certain individuals may be members of more than one subgroup or committee (e.g. a clinician scientist as a member of the working group and guideline panel).[11,95]
Oversight committee	A body overseeing the guideline development process, whose tasks include the priority setting, and selection of potential guidelines for development out of proposed topics, recruitment and appointment of members for the guideline panel, and approval of the final guideline for publication and dissemination. May also be referred to as an executive committee or guideline advisory board.[11]
Guideline panel	Decides on topics to be covered within the guideline, formulates questions, develops and agrees on the recommendations in the guideline using evidence summaries prepared by the working group, and endorses the final guideline document for approval by the oversight committee. Members of the guideline panel may often be referred to as 'panelists'.[11]
Chair (of the guideline panel)	The leading member of the guideline panel. This person is neutral and has an expertise in coordinating groups of healthcare professionals and patients and caregivers. Someone who is qualified and experienced in strategies and facilitation of optimal group processes, ensuring all members of the panel have equal opportunity to contribute and freely express their opinion without feeling intimidated. This individual is not necessarily an expert of any specific clinical domain.[2,25]
Co-chair (of the guideline panel)	Should be appointed when the guideline panel is especially large or the task particularly complex. Co-chairs should also have experience leading groups but should represent a

	different discipline (clinical or methodological) than the Chair.[2,25]
Working group	A group of individuals tasked with the preparation and technical aspects of guideline development such as assisting the guideline panel in formulating PICO questions, conducting systematic reviews, rating quality of evidence, preparing evidence summaries and background documents for guideline panel discussions, writing the guideline, and reviewing comments from stakeholders and public consultation. Works closely with the guideline panel to ensure the work to achieve goals and objectives for the guideline is completed.
Secretariat	A group of individuals tasked with supporting the guideline development group in preparing for the development and writing of the guideline. The Secretariat provides technical support as well as administrative support (e.g. scheduling meetings and teleconferences, distributing documents).[11]
Stakeholder	An individual, group or an organization that has an interest in the organization and delivery of health care and will have an interest in the content of or the outcome of a guideline. This may include health care providers, professional societies and colleges, experts in a disease or condition, research institutions, and policy makers.[11,95]
Consumer	Consumers of healthcare include: (a) individual patients, (b) carers, including patients' family and friends, (c) members of the public (both as potential patients and as funders of healthcare through taxation, insurance or direct payments), (d) voluntary and community organizations that represent the interests of patients, carers and the public, (e) advocates representing the interests of patients, carers and other client groups. They are described collectively as 'consumers' (without implying consumerist assumptions about health services) and are distinct from other consumers of guidelines such as health professionals, commissioners and providers of services.[96]
Carer	Provide non-reimbursed care and/or support to patients (e.g. family members, friends) and have knowledge of the issues that are important to patients and carers. May also be referred to as caregivers.
Advocate	Someone who speaks on behalf of a patient, or a group of patients to help them make their wishes known.[97]
Sponsoring organization	The organization that funds the development of a guideline and will endorse it for publication and dissemination.

Professional societies	Not-for-profit organizations whose membership consists of healthcare professionals working in a specific field or specialty and whose work focuses on a specific area or topic in health care (e.g. American College of Chest Physicians, European Society of Cardiology). Professional societies are often involved in the development of guidelines for their members and often take policy stances on medical issues and health promotion. May also be referred to as professional organizations or medical societies or associations.
Third party organizations	Organizations or groups that wish to adopt or adapt a guideline for which they were not directly involved in its development. This may often include government departments or ministries of health that do not have sufficient resources to develop guidelines *de novo*, or whose populations and health care settings are similar to those covered in an existing guideline.
Guidelines and topics	
Guideline	A document that focuses on a disease or condition and includes recommendations for appropriate management of patients with this disease or condition. The guideline should be based on the best available evidence and should help healthcare providers by supplementing their knowledge and skills. Guidelines can be tailored to clinical, health policy, health systems or public health settings, among others.[11]
Target audience	The specific group or range of health care provider for whom the clinical practice guidelines are intended, to inform their work in a health care setting. The target audience will have an influence on the breadth and depth of the guideline content.[24] The ***primary audience*** consists of the intended end users of the guideline. For example, if the guideline is for primary care, then the target audience will comprise of primary care physicians and nurses. ***Secondary audiences*** may include any other groups to whom the guideline content will be applicable, such as health care managers, hospital administrators, and policy makers.[98]
Guideline topic	The guideline topic specifies the disease, condition or overall area that will be covered by the guideline (e.g. chronic obstructive pulmonary disease). Guideline developers must consider prioritizing the guideline topics with the greatest potential to improve health care and health outcomes.[39]
Topics within guidelines	Topics within the guideline encompass the content that the guideline will cover. For example, whether the guideline will cover diagnosis of a condition, treatment of a condition, or both, or whether it will focus on topics where there is most uncertainty or variation in practice. Guideline panels must consider and decide on the many issues that may be addressed

	within a guideline that will be important to the target audience. May also be referred to as the scope of the guideline, and will be interrelated with the PICO questions addressed in the guideline.[39]
Steps and processes in guideline development	
Priority setting	Priority setting is the identification, balancing and ranking of priorities by stakeholders. It ensures that resources and attention are devoted to those general areas (e.g. chronic obstructive pulmonary disease, diabetes, cardiovascular disease, cancer, prevention) where health care recommendations will provide the greatest benefit to the population, a jurisdiction, or a country. A priority-setting approach needs to contribute to future plans while responding to existing potentially difficult circumstances.[99]
Peer review	A process of subjecting scholarly works, research, or ideas to the scrutiny of others. Peer review of a guideline and recommendations by those with similar interests and expertise to the people who produced it is intended to ensure the guideline is accurate and valid. Peer review may be internal, conducted by colleagues from the same organization not directly involved in the production of the guideline, or external, conducted by individuals fully independent and removed from the development of the guideline.[11,97]
Dissemination	The active process of distributing information, such as guidelines, to the target end users to ensure maximum exposure, uptake, and implementation. Various methods for dissemination may be used such as a printed version of the full guideline, online version of the guideline, a quick reference guide, mobile application of the guideline, incorporation of guideline recommendations into clinical decision support systems, consumer version of the guideline, education materials detailing the recommendations, conference meetings with target end users, etc. Products other than the main guideline document that are developed are commonly referred to as ***derivative products***.[92]
Implementation	The uptake and incorporation of guideline recommendations into practice by the target end users. An implementation plan should include the identification of potential barriers, criteria and indicators for success, baseline data for the indicators, required resources, training and education needs, identification of existing mechanisms or networks, methods for monitoring the implementation process, reporting and feedback mechanisms, and milestones with timescales.[11,92]
Guideline Adaptation	A systematic approach to using and adjusting existing guidelines produced in one setting for use in a new setting with

	a different cultural or organizational context. The process of adapting a guideline and its recommendations must ensure that the adapted guideline addresses specific health questions relevant to the context of use and that it is suited to the needs, priorities, legislation, policies, and resources in the new target setting.[10]
Group processes	Group processes encompass how and when members of a group interact. For example, the interaction of guideline panel members during a consensus meeting to formulate recommendations.[25]
Consensus methods	Techniques used in decision-making to reach agreement on a particular issue. Consensus may be informal or formal, with examples of formal consensus methods including the Delphi and nominal group techniques.[95]
Quorum	The smallest number of group members that must be present to constitute a valid meeting or voting or consensus process.[95]
Milestones	When major steps are achieved during the guideline development process. Examples include completing the systematic review, having recommendations developed, and publishing the guideline report.[25]
Considerations in the development of a guideline	
Declaration of interest (or disclosure of interest)	A declaration of interest is the disclosure of any potential or actual conflicts of interest that include financial, professional, intellectual or other interests relevant to the subject of the work or meeting to determine possible conflicts of interest. The declaration of interest must also include any relevant interests of others who may, or may be perceived to, unduly influence the expert's judgment, such as immediate family members, employers, close professional associates, or any others with whom the expert has a substantial common personal, financial, or professional interest.[11]
Conflict of interest	A divergence between or individual's private interests and his or her professional obligations such that an independent observer might reasonably question whether the individual's professional actions or decisions are motivated by personal gain, such as financial, academic advancement, clinical revenue streams or community standing. This definition includes a financial or intellectual relationship that may impact an organization's or individual's ability to approach a scientific question with an open mind.[50]
Commercial sponsorship	May apply to individuals or organizations, including funding for the development of a guideline. Of particular concern is the possibility that guideline developers will feel, or be perceived to be, beholden to or pressured by the commercial sponsor to make recommendations favorable to the sponsor's interests.

	Commercial sponsorship may be in the form of industry-sponsored research, clinical services from which a committee member derives a substantial proportion of his or her income, consulting, board membership for which compensation of any type is received.[100]
Barriers to change	Should be identified and considered prior to developing a guideline where recommendations suggest changes in health care practice(s). Barriers to change can exist at various levels of the health care system and include structural barriers (e.g. lack of resources, financial disincentives), organizational barriers (e.g. inappropriate skill mix, lack of facilities or equipment), peer group barriers (e.g. local standards of care not in line with desired practice), professional-patient interaction barriers (e.g., communication and information-processing issues), and competing priorities. There are diverse methods to identify barriers that vary in their formality. Barriers may vary for given resources, across settings, and for different guidelines.[92]
Equity (in health)	Equity in health, or health equity, is a measure of the degree to which health policies are able to distribute well-being fairly. It is the absence of systematic or potentially remediable differences in health status, access to healthcare and health-enhancing environments, and treatment in one or more aspects of health across populations or population groups defined socially, economically, demographically or geographically. Health inequity results from a gap in health status and in access to health services between different social classes, ethnic groups, and between populations in different geographical areas. Guideline panels must consider whether and the extent to which recommendations will have an impact on health equity. May also be referred to as health inequality.[95,99,101]
Values, preferences and utilities	These include patient and carer knowledge, attitudes, expectations, moral and ethical values and beliefs; patient goals for life and health; prior experience with the intervention and the condition; symptom experience (for example breathlessness, pain, dyspnoea, weight loss); preferences for and importance of desirable and undesirable outcomes; perceived impact of the condition or interventions on quality of life, well-being or satisfaction and interactions between the work of implementing the intervention, the intervention itself, and other contexts the patient may be experiencing; preferences for alternative courses of action; and preferences relating to communication content and styles, information and involvement in decision-making and care. This can be related to what in the economic literature is considered utilities. An intervention itself can be considered a consequence of a

	recommendation (e.g. the burden of taking a medication or undergoing surgery) and a level of importance or value is associated with that. The values and preferences of those who will be affected by the recommendations should be integrated into the process of developing the guideline.[96]
Transparency	Transparency involves clearly documenting and presenting details of the entirety of the methods and process that were used to develop a guideline, including the participants involved, the evidence and information reviewed, and judgements made during any decision-making, especially formulating the recommendations. Transparency would allow others to follow and arrive at the same guideline product if replicating the guideline development process.
Credibility of guidelines	The degree to which a guideline's conclusions and recommendations can be trusted. Determined by the methods and approaches used, including timing and editorial dependence such as described by the AGREE II tool, the Institute of Medicine's report on guidelines and the Guideline International Network. May also be referred to as trustworthiness or quality of guidelines.[2,16,102]
Evidence review and consideration of additional information	
Protocol	A document that outlines the plan or set of steps that defines how a guideline will be produced and the methodology that will be used. Before carrying out a guideline, for example, the protocol sets out what questions to be answered, how information will be collected and analyzed, and the framework and consensus methods to be used to formulate recommendations.
PICO question	Population/Patient-Intervention-Comparison-Outcome**;** a mnemonic used in developing specific health care questions to be answered in a guideline. A question generated using the PICO framework will guide which evidence is reviewed and is meant to elicit information about the patient and their condition, interventions of interest that have been undertaken or should be taken, any comparisons between the current intervention and possible alternatives, and outcomes to be desired or achieved.[11]
Population	A group of people with a common link, such as the same medical condition or living in the same area or sharing the same characteristics. The population identified for a guideline is all the people the recommendations are intended to apply to (e.g. adults with diabetes mellitus).[103]
Comorbidity	A disease or condition that exists in a patient in addition to the principal disease of interest being studied or treated (e.g. chronic obstructive pulmonary disease and diabetes mellitus).

	Comorbidities may influence the clinical manifestations and natural history of a disease. May also be referred to as concomitant conditions.[97,104]
Clinical pathway (or care pathway)	The sequence of practices, procedures, tests, interventions and treatments that should be used to provide care for people with a particular clinical condition.[97]
Outcomes	The impact that a test, treatment, policy, program or other intervention has on a person, group or population. Outcomes from interventions to improve the public's health could include a change in people's health and wellbeing or health status. In clinical terms, outcomes could include the number of patients who fully recover from an illness or the number of hospital admissions, and an improvement or deterioration in someone's health, functional ability, symptoms or situation.[97]
Patient-important outcomes	An outcome defined by answering “yes” to the following question: “If one knew that this outcome was the only thing to change with treatment, would the patient consider receiving this treatment even if it was associated with adverse effects, inconvenience, or cost?” Such outcomes include mortality, morbidity, and outcomes reported by patients.[44,105]
Health-related quality of life	A combination of a person's physical, mental and social well-being; not merely the absence of disease. An example of a patient-important outcome.[95]
Surrogate outcomes	Outcomes that are not themselves important health outcomes but may be correlated with patient-important health outcomes (e.g. bone density as surrogate for fractures as the patient-important outcome). May be referred to as substitute or indirect outcomes.[44]
Importance of outcomes	Ranking the relative importance of desirable (e.g. reduced mortality, improvement in health-related quality of life) and undesirable outcomes (e.g. side effects, costs) for the intervention in question allows a guideline panel to determine how much influence the particular outcomes and the results/estimates of effect for those outcomes will have in formulating a recommendation. The relative importance of outcomes is likely to vary according to different values and preferences or when considered from the perspective of patients, clinicians or policy-makers. In the GRADE framework, outcomes are rated as critical for decision-making, important but not critical for decision-making, or low importance for decision-making.[44]
Magnitude of effect	A measure of the difference or relative effect of an intervention on the outcome in the intervention group compared with that in a control group. Also referred to as the

	effect size.[97]
Systematic review	A comprehensive review of the published literature that focuses on a healthcare topic and answers a specific question. An extensive literature search is conducted based on a search strategy to identify all studies. The studies are reviewed, their quality is assessed, and the results are summarized according to the review question.[11]
Evidence retrieval	In the context of systematic reviews, the process of systematically searching for all scientific studies relevant to a particular question, and obtaining them for review. The process also includes obtaining evidence from other sources that may be unpublished.[11]
Selection criteria	The criteria used to decide which studies and study types should be included and excluded from consideration as potential sources of evidence when retrieving evidence during the development of a guideline. Also referred to as inclusion and exclusion criteria.[97]
Expert opinion	An interpretation of evidence. Sometimes based on high quality evidence, such as from randomized controlled trials or well-done observational studies, and other times based on unsystematically collected information, ideally summarized in writing. Expert opinion is often confused with the notion of evidence that is either not available from systematic research or not systematically summarized. Also often used as excuse for not collecting evidence systematically.
Economic evaluation	A set of formal, quantitative methods used to asses one or more interventions, programs, or strategies with respect to their resource use and their expected outcomes. Economic evaluation may involve different study types such as cost-effectiveness analysis, cost-benefit analysis, and economic models.[11]
Quality of evidence	Describes the level of confidence or certainty in the estimates of the effect of an intervention on a specific outcome in a given population. Also called strength of evidence, confidence in estimates, certainty in evidence, levels of evidence.[79]
Evidence table or profile or summary of findings table	A table summarizing the results/estimate of effect from studies for each outcome of interest and the associated quality of evidence. The table provides a concise summary of the key information that is needed by someone making a decision and, in the context of a guideline, provides a summary of the key information underlying a recommendation.[65,97]
Recommendations and formulation of recommendations	
Analytic framework	A framework outlining the criteria that guideline panels use to review the evidence and analyze relevant information to arrive at a recommendation. The analysis may focus on the balance

	between desirable and undesirable consequences, informed by the quality of evidence, magnitude of the difference between the benefits and harms, the certainty about or variability in values and preferences, resource use, equity and other factors (e.g. GRADE/DECIDE Evidence-to-Recommendation framework).[58]
Recommendation	A course of action recommended by the guideline based on clinical questions, evidence retrieval, and consideration of other information in the analytic framework. Recommendations in guidelines may relate to clinical interventions, public health activities, or government policies.[11]
Conditional recommendation	A recommendation for which a guideline panel rested with more uncertainty about whether implementation of the recommended action leads to more desirable than undesirable consequences. Specific conditions may have to be described. Also known as weak recommendation in the GRADE framework.[77]
Research recommendation	A recommendation resulting from a guideline process for use in the context of research only. Guideline panels should consider making research recommendations when there is important uncertainty about the desirable and undesirable effects of an intervention, further research could reduce that uncertainty, and the potential benefits and savings of reducing the uncertainty outweigh the potential harms of not making the research recommendation. The formulation of recommendations for additional research should be as precise and specific as possible. Defining the population, intervention, comparator and outcomes (PICO) explicitly will make research recommendations more helpful.[58,106]
Strength of recommendation	The strength of a recommendation reflects the extent to which guideline developers are confident that the desirable effects of adherence to the recommendation outweigh the undesirable effects.[58,77]
Performance measures	Performance measures are criteria that can be measured to assess the quality-of-care (e.g. a physician following a specific management option). Management options associated with strong recommendations are particularly good candidates for quality criteria.[58]

15. References

1. Academy of Medicine of Malaysia. Clinical Practice Guidelines. 2013; http://www.acadmed.org.my/index.cfm?&menuid=67. Accessed April 22, 2013.
2. Institute of Medicine Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Clinical Practice Guidelines We Can Trust. 2011; http://www.nap.edu/openbook.php?record_id=13058. Accessed April 22, 2013.
3. Appraisal of Guidelines Research and Evaluation. Introduction to AGREE II. 2013; http://www.agreetrust.org/about-the-agree-enterprise/introduction-to-agree-ii/.
4. Brozek JL, Akl EA, Alonso-Coello P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy.* May 2009;64(5):669-677.
5. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj.* Apr 26 2008;336(7650):924-926.
6. Ahmed F, Temte JL, Campos-Outcalt D, Schunemann HJ. Methods for developing evidence-based recommendations by the Advisory Committee on Immunization Practices (ACIP) of the U.S. Centers for Disease Control and Prevention (CDC). *Vaccine.* 2011;29(49):9171-9176.
7. Oxman A, Schunemann H, Fretheim A. Improving the use of research evidence in guideline development: 7. Deciding what evidence to include. *Health Research Policy and Systems.* 2006;4(1):19.
8. Qaseem A, Snow V, Owens DK, Shekelle P. The development of clinical practice guidelines and guidance statements of the American College of Physicians: summary of methods. *Annals of internal medicine.* Aug 3 2010;153(3):194-199.
9. Schunemann HJ, Wiercioch W, Etxeandia I, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* Dec 16 2013.
10. The ADAPTE Collaboration. The ADAPTE Process: Resource Toolkit for Guideline Adaptation. Version 2.0. 2009; http://www.g-i-n.net/document-store/working-groups-documents/adaptation/adapte-resource-toolkit-guideline-adaptation-2-0.pdf. Accessed July 5, 2013.
11. World Health Organization. Estonian Handbook for Guidelines Development. 2011; http://whqlibdoc.who.int/publications/2011/9789241502429_eng.pdf. Accessed April 22, 2013.
12. Oxman AD, Fretheim A, Schunemann HJ. Improving the use of research evidence in guideline development: introduction. *Health Res Policy Syst.* 2006;4:12.
13. Schunemann HJ, Woodhead M, Anzueto A, et al. A guide to guidelines for professional societies and other developers of recommendations: introduction to integrating and coordinating efforts in COPD guideline development. An official ATS/ERS workshop report. *Proceedings of the American Thoracic Society.* Dec 2012;9(5):215-218.
14. Eccles MP, Grimshaw JM, Shekelle P, Schunemann HJ, Woolf S. Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. *Implementation science : IS.* 2012;7:60.
15. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: Advancing guideline development, reporting and evaluation in health care. *Canadian Medical Association Journal.* 2010;182(18):E839-E842.

16. Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P. Guidelines International Network: Toward international standards for clinical practice guidelines. *Annals of internal medicine.* 2012;156(7):525-531.
17. Fretheim A, Schunemann H, Oxman A. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health Research Policy and Systems.* 2006;4(1):15.
18. National Institute for Health and Clinical Excellence. The guidelines manual. 2012; http://publications.nice.org.uk/the-guidelines-manual-pmg6. Accessed April 22, 2013.
19. World Health Organization. WHO Handbook for Guideline Development. 2012; http://apps.who.int/iris/bitstream/10665/75146/1/9789241548441_eng.pdf. Accessed April 22, 2013.
20. Centers for Disease Control and Prevention. *Guidelines and Recommendations: A CDC Primer.* Atlanta, Georgia: Office of the Associate Director for Science Centers for Disease Control and Prevention;2012.
21. New Zealand Guidelines Group. *Handbook for the preparation of explicit evidence-based clinical practice guidelines.* Wellington: New Zealand Guidelines Group;2001.
22. Eccles M, Grimshaw J, Shekelle P, Schunemann H, Woolf S. Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. *Implementation Science.* 2012;7(1):60.
23. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook - Annex D: Completed considered judgement form. 2013; http://www.sign.ac.uk/guidelines/fulltext/50/compjudgement.html.
24. National Health and Medical Research Council. Procedures and requirements for meeting the 2011 NHMRC standard for clinical practice guidelines. 2011; http://www.nhmrc.gov.au/guidelines/publications/cp133-and-cp133a. Accessed April 22, 2013.
25. Kunz R, Fretheim A, Cluzeau F, et al. Guideline Group Composition and Group Processes: Article 3 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):229-233.
26. Agency for Healthcare Research and Quality. U.S. Preventive Services Task Force Procedure Manual. 2008; http://www.uspreventiveservicestaskforce.org/uspstf08/methods/procmanual.htm. Accessed April 22, 2013.
27. Canadian Task Force on Preventive Health Care. Canadian Task Force on Preventive Health Care Procedure Manual. 2011; http://canadiantaskforce.ca/methods/methods-manual/. Accessed April 22, 2013.
28. Cancer Care Ontario. Program in Evidence-Based Care Handbook. 2012; http://www.cancercare.on.ca/about/programs/pebc/pebc-products/. Accessed April 22, 2013.
29. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. 2011; http://www.sign.ac.uk/guidelines/fulltext/50/. Accessed April 22, 2013.
30. Cancer Care Ontario. Program in Evidence-Based Care Document Assessment and Review Protocol. 2012; http://www.cancercare.on.ca/about/programs/pebc/document__review/. Accessed April 22, 2013.
31. European Society of Cardiology. Recommendations for Guidelines Production. 2010; http://www.escardio.org/guidelines-surveys/esc-guidelines/about/Pages/rules-writing.aspx. Accessed April 22, 2013.
32. American College of Cardiology Foundation and American Heart Association. Methodology Manual and Policies from the ACCF/AHA Task Force on Practice Guidelines. 2010;

http://my.americanheart.org/professional/StatementsGuidelines/PoliciesDevelopment/Development/Methodologies-and-Policies-from-the-ACCAHA-Task-Force-on-Practice-Guidelines_UCM_320470_Article.jsp. Accessed April 22, 2013.

33. Davino-Ramaya C, Krause LK, Robbins CW, et al. Transparency matters: Kaiser Permanente's National Guideline Program methodological processes. *The Permanente journal.* Winter 2012;16(1):55-62.
34. Esandi ME, Luca MD, Chapman E, Schapochnik N, Bernztein R, Otheguy L. Guía para la adaptación de Guías de Práctica Clínica. 2008; http://publicaciones.ops.org.ar/publicaciones/otras pub/GuiadeGuias.pdf. Accessed April 22, 2013.
35. Gutiérrez GC, Bossert T, Espinosa JQ, et al. Guía Metodológica para la elaboración de Guías de Atención Integral en el Sistema General de Seguridad Social en Salud Colombiano. 2010; http://www.minsalud.gov.co/Documentos y Publicaciones/GUIA METODOLOGICA PARA LA ELABORACI%C3%93N DE GU%C3%8DAS DE ATENCI%C3%93N INTEGRAL.pdf. Accessed April 22, 2013.
36. Rosenfeld RM, Shiffman RN, Robertson P. Clinical Practice Guideline Development Manual, Third Edition: A Quality-Driven Approach for Translating Evidence into Action. *Otolaryngology -- Head and Neck Surgery.* 2013;148(1 suppl):S1-S55.
37. American College of Rheumatology. Policy and Procedure Manual for Clinical Practice Guidelines. 2012; http://www.rheumatology.org/Practice/Clinical/Guidelines/Clinical_Practice_Guidelines. Accessed July 5, 2013.
38. Ministerio de Sanidad y Consumo. Elaboración de guías de práctica clínica en el sistema nacional de salud: Manual metodológico. 2007; http://www.guiasalud.es/emanuales/elaboracion/index-02.html. Accessed April 22, 2013.
39. Atkins D, Perez-Padilla R, MacNee W, Buist AS, Cruz AA. Priority Setting in Guideline Development: Article 2 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):225-228.
40. Brunetti M, Shemilt I, Pregno S, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *Journal of clinical epidemiology.* Feb 2013;66(2):140-150.
41. National Health and Medical Research Council. Guideline Development and Conflicts of Interest: Identifying and Managing Conflicts of Interest of Prospective Members and Members of NHMRC Committees and Working Groups Developing Guidelines. 2012; http://www.nhmrc.gov.au/guidelines-and-publications/information-guideline-developers/guideline-development-and-conflicts. Accessed April 22, 2013.
42. Woolf S, Schunemann H, Eccles M, Grimshaw J, Shekelle P. Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations. *Implementation Science.* 2012;7(1):61.
43. Burford BJ, Rehfuess E, Schunemann HJ, et al. Assessing evidence in public health: the added value of GRADE. *J Public Health (Oxf).* Dec 2012;34(4):631-635.
44. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of clinical epidemiology.* Apr 2011;64(4):395-400.
45. Schunemann H, Oxman A, Fretheim A. Improving the use of research evidence in guideline development: 6. Determining which outcomes are important. *Health Research Policy and Systems.* 2006;4(1):18.
46. Guyatt GH, Oxman AD, Kunz R, et al. Incorporating considerations of resources use into grading recommendations. *Bmj.* May 24 2008;336(7654):1170-1173.

47. Carling CL, Kristoffersen DT, Oxman AD, et al. The effect of how outcomes are framed on decisions about whether to take antihypertensive medication: a randomized trial. *PLoS ONE.* 2010;5(3):e9469.
48. Jaeschke R, Guyatt GH, Dellinger P, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *Bmj.* 2008;337:a744.
49. Boyd E, Bero L. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Research Policy and Systems.* 2006;4(1):16.
50. Schunemann HJ, Osborne M, Moss J, et al. An official American Thoracic Society Policy statement: managing conflict of interest in professional societies. *American journal of respiratory and critical care medicine.* Sep 15 2009;180(6):564-580.
51. Cancer Care Ontario. Program in Evidence-Based Care Conflict of Interest Policy. 2011; http://www.cancercare.on.ca/cms/one.aspx?objectId=7582&contextId=1377. Accessed April 22, 2013.
52. Oxman A, Schunemann H, Fretheim A. Improving the use of research evidence in guideline development: 14. Reporting guidelines. *Health Research Policy and Systems.* 2006;4(1):26.
53. Cincinnati Children's Hospital Medical Center Evidence-Based Care Group. Let Evidence Guide Every New Decision (LEGEND) Evidence-Based Care Guideline Development and Update Process. 2006; http://www.cincinnatichildrens.org/service/j/anderson-center/evidence-based-care/legend. Accessed April 22, 2013.
54. Guyatt G, Akl EA, Oxman A, et al. Synthesis, Grading, and Presentation of Evidence in Guidelines: Article 7 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):256-261.
55. National Heart Lung and Blood Institute. Systematic Evidence Reviews and Clinical Practice Guidelines. 2013; http://www.nhlbi.nih.gov/guidelines/index.htm. Accessed April 22, 2013.
56. Oxman A, Schunemann H, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. *Health Research Policy and Systems.* 2006;4(1):20.
57. Guyatt GH, Oxman AD, Kunz R, et al. Going from evidence to recommendations. *Bmj.* May 10 2008;336(7652):1049-1051.
58. Schünemann HJ, Oxman AD, Akl EA, et al. 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. *Proceedings of the American Thoracic Society.* 2012;9(5):282-292.
59. World Health Organization. Declaration of interests for WHO experts. http://www.who.int/ipcs/methods/harmonization/areas/mutagenicity_doi.pdf.
60. AMSTAR. Assessing the methodological quality of systematic reviews. 2012; http://amstar.ca. Accessed July 21, 2013.
61. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of clinical epidemiology.* Oct 2009;62(10):1013-1020.
62. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *Bmj.* Jun 19 2004;328(7454):1490.
63. Guyatt GH, Oxman AD, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? *Bmj.* May 3 2008;336(7651):995-998.
64. Brozek JL, Akl EA, Compalati E, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines part 3 of 3. The GRADE approach to developing recommendations. *Allergy.* May 2011;66(5):588-595.

65. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *Journal of clinical epidemiology.* Apr 2011;64(4):383-394.
66. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence--inconsistency. *Journal of clinical epidemiology.* Dec 2011;64(12):1294-1302.
67. Alonso-Coello P, Zhou Q, Martinez-Zapata MJ, Mills E, Heels-Ansdell D, Johanson JF. Meta-analysis of flavonoids for the treatment of haemorrhoids. *Br J Surgery.* 2006;93:909e920.
68. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence--indirectness. *Journal of clinical epidemiology.* Dec 2011;64(12):1303-1310.
69. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. *Journal of clinical epidemiology.* Dec 2011;64(12):1283-1293.
70. POISE Study Group, Devereaux PJ, Yang H, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet.* 2008;371(9627):1839-1847.
71. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence--publication bias. *Journal of clinical epidemiology.* Dec 2011;64(12):1277-1282.
72. Guyatt GH, Oxman AD, Sultan S, et al. GRADE guidelines: 9. Rating up the quality of evidence. *Journal of clinical epidemiology.* Dec 2011;64(12):1311-1316.
73. Gilbert R, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. *International journal of epidemiology.* Aug 2005;34(4):874-887.
74. Schulman S, Beyth RJ, Kearon C, Levine M. Hemorrhagic complications of anticoagulant and thrombolytic treatment: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest.* 2008;133(257):Se98S.
75. Devereaux PJ, Choi PT, Lacchetti C, et al. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* 2002;166:1399e1406.
76. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *Journal of clinical epidemiology.* Jul 2013;66(7):719-725.
77. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *Journal of clinical epidemiology.* 2013;66(7):719-725.
78. Andrews JC, Schunemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *Journal of clinical epidemiology.* Jul 2013;66(7):726-735.
79. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *Journal of clinical epidemiology.* Apr 2011;64(4):401-406.
80. Guyatt G, Oxman AD, Sultan S, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *Journal of clinical epidemiology.* Feb 2013;66(2):151-157.
81. Guyatt GH, Oxman AD, Santesso N, et al. GRADE guidelines: 12. Preparing summary of findings tables-binary outcomes. *Journal of clinical epidemiology.* Feb 2013;66(2):158-172.
82. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of clinical epidemiology.* Apr 2011;64(4):380-382.
83. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). *Journal of clinical epidemiology.* Apr 2011;64(4):407-415.

84. Guyatt GH, Thorlund K, Oxman AD, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes. *Journal of clinical epidemiology.* Feb 2013;66(2):173-183.

85. Edejer T. Improving the use of research evidence in guideline development: 11. Incorporating considerations of cost-effectiveness, affordability and resource implications. *Health Research Policy and Systems.* 2006;4(1):23.

86. Hill SR, Olson LG, Falck-Ytter Y, et al. Incorporating Considerations of Cost-Effectiveness, Affordability, and Resource Implications in Guideline Development: Article 6 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):251-255.

87. GRADE working group. Grading the quality of evidence and the strength of recommendations. 2013; http://www.gradeworkinggroup.org/intro.htm.

88. Hsu J, Brozek JL, Terracciano L, et al. Application of GRADE: Making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implementation science : IS.* 2011;6:62.

89. Murad MH, Montori VM, Sidawy AN, et al. Guideline methodology of the Society for Vascular Surgery including the experience with the GRADE framework. *Journal of vascular surgery.* May 2011;53(5):1375-1380.

90. Fretheim A, Schunemann H, Oxman A. Improving the use of research evidence in guideline development: 15. Disseminating and implementing guidelines. *Health Research Policy and Systems.* 2006;4(1):27.

91. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. *Implementation science : IS.* 2011;6:26.

92. Grimshaw JM, Schünemann HJ, Burgers J, et al. Disseminating and Implementing Guidelines: Article 13 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):298-303.

93. Canadian Health Services Research Foundation. Reader Frindly Writing- 1:3:25. 2010.

94. Wollersheim H, Burgers J, Grol R. Clinical Guidelines to Improve Patient Care. *Neth J Med.* 2005;63(6):188-192.

95. Shekelle PG, Ortiz E, Rhodes S, Morton SC, et al. Validity Of the Agency For Healthcare Research And Quality Clinical practice guidelines: how quickly do guidelines become outdated? . *JAMA.* 2001;286(12):1461-1467.

96. National Institute for Health and Clinical Excellence. The guidelines manual: Appendix L – Abbreviations and Glossary. 2012; http://publications.nice.org.uk/the-guidelines-manual-appendix-l-abbreviations-and-glossary-pmg6d/l2-glossary. Accessed July 21, 2013.

97. Kelson M, Akl EA, Bastian H, et al. Integrating Values and Consumer Involvement in Guidelines with the Patient at the Center: Article 8 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):262-268.

98. National Institute for Health and Clinical Excellence. Glossary. 2013; http://www.nice.org.uk/website/glossary/glossary.jsp. Accessed July 21, 2013.

99. Yawn BP, Akl EA, Qaseem A, Black P, Campos-Outcalt D. Identifying Target Audiences: Who Are the Guidelines For?: Article 1 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):219-224.

100. World Health Organization. Health Systems Strengthening Glossary. 2013; http://www.who.int/healthsystems/hss_glossary/en/index.html. Accessed July 26, 2013.

101. Boyd EA, Akl EA, Baumann M, et al. Guideline Funding and Conflicts of Interest: Article 4 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):234-242.

102. Oxman A, Schunemann H, Fretheim A. Improving the use of research evidence in guideline development: 12. Incorporating considerations of equity. *Health Research Policy and Systems.* 2006;4(1):24.

103. AGREE Research Trust. The AGREE Enterprise Website. 2013; http://www.agreetrust.org. Accessed July 21, 2013.

104. Wilt TJ, Guyatt G, Kunz R, et al. Deciding What Type of Evidence and Outcomes to Include in Guidelines: Article 5 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):243-250.

105. Fabbri LM, Boyd C, Boschetto P, et al. How to Integrate Multiple Comorbidities in Guideline Development: Article 10 in Integrating and Coordinating Efforts in COPD Guideline Development. An Official ATS/ERS Workshop Report. *Proceedings of the American Thoracic Society.* 2012;9(5):274-281.

106. Guyatt G DP, Montori V, Schünemann HJ, Bhandari M,. Putting the patient first: In our practice, and in our use of language. *ACP Journal Club.* Jan-Feb 2004;140(A11).

107. Brown P BK, Chalkidou K, Chalmers I, Clarke M, Fenton M, et al,. How to formulate research recommendations. *Bmj.* Oct 14 2006;333(7572):804-806.

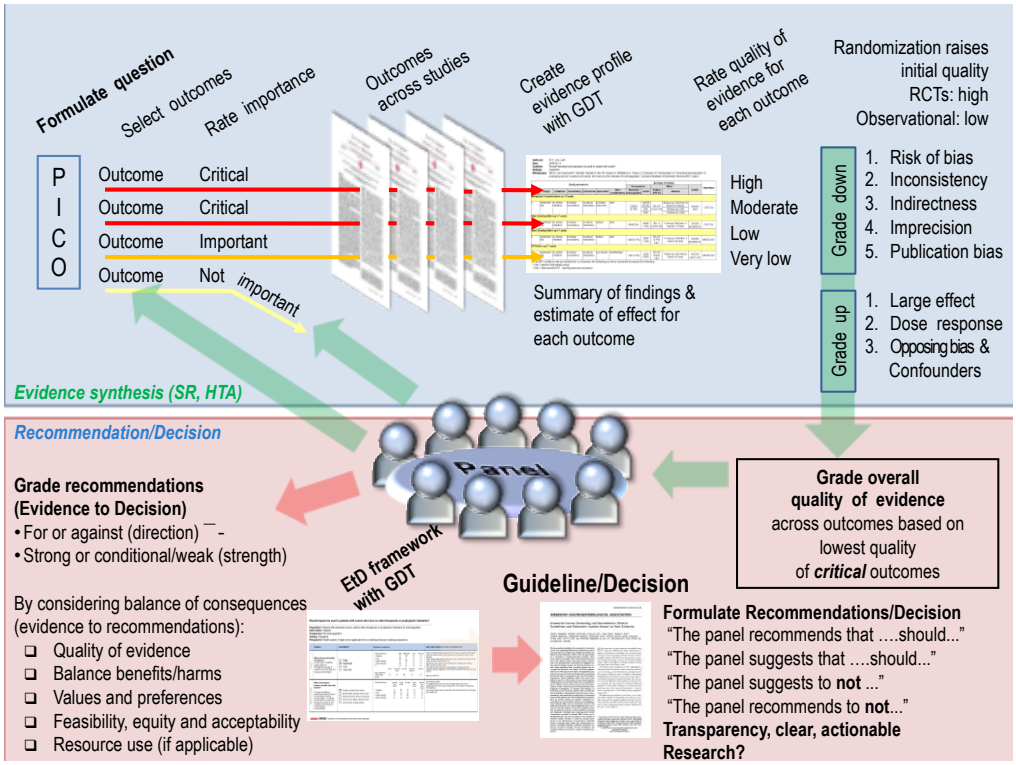
16. Appendices

Appendix 1: Template for topic proposal

Factors	Considerations per each factor	Data provided by
Burden of disease (problem)	mortality (per 1000)	the initiator
	Incidence	the initiator
	Prevalence	the initiator
	resource impact (MOH spending, per year)	the initiator in cooperation with MOH
Variations	Practice variation	the initiator in cooperation with MOH
	Health outcome variation	the initiator in cooperation with MOH
	Variation in treatment costs	the initiator in cooperation with MOH
Potential	Potential for modernization	the initiator
	Potential result on health	the initiator
	Potential impact on resources	the initiator in cooperation with MOH
Problem statement	Based on the information listed above	the initiator
Purpose of the guideline	Based on problem statement	the initiator

Appendix 2: The GRADE process in developing guidelines

This figure describes the GRADE process



The figure demonstrates an optimal process for integrating the GRADE approach into guideline development. It also highlights the relation between the conduct of systematic reviews (evidence synthesis) and guideline development (decision making by a panel). The requirement for a close relationship between systematic reviewers and guideline panels is highlighted. Multidisciplinary guideline panels should be involved in the development of the guideline questions using the PICO framework. The panels are also involved in the process of selecting critical and important outcomes upon which the evidence assessment is focused and prioritizing their importance for decision making.

Outcomes that are considered critical and important for decision-making should be evaluated in a systematic review. Outcomes rated as less important do not serve as the substrate for recommendations and generally do not need to be considered further for making a recommendation. One of the innovations inherent to the GRADE approach is the evaluation of outcomes across, rather than within, studies. As a consequence, different bodies of evidence may contribute information towards the different critical and important outcomes that are being considered. Those who review the evidence will then assign a degree of confidence or certainty in the estimates of effect of a body of evidence (i.e. the quality of evidence) for each outcome. The degree of certainty can be one of four categories: high, moderate, low or very low on the basis of 8 factors that either increase or decrease the initial quality assignment.

With regards to the initial assignment, randomization is considered the best method to protect against bias and confounding and as such the initial quality rating for a body of evidence from randomized control trials starts as high, but there are 5 factors that lower the quality, and for observational studies 3 factors that can increase the quality. Therefore, an alternative way to conceptualize the use of the factors that influence the rating of the evidence begins with

acknowledging that any confident in estimates starts as low and randomization is one of four factors that lead to an increase in the confidence of effects.

Both the statistical information and the assessment of the quality of evidence are presented in standardized evidence tables organized by outcome and ideally created using the GRADE profiler Guideline Development Tool (GDT) software. Information can either be presented in typical detailed 'evidence profiles' or in the more succinct 'Summary of Findings (SoF)'.

Once all of the outcomes that are critical for decision-making have been evaluated in terms of the quality of the supporting evidence an overall grade of the quality of evidence is assigned. This overall grade is equivalent to the rating for the lowest quality evidence base linked to a critical outcome. This information is then provided to the panel for the purpose of formulating recommendations based on the evidence profiles.

A guideline panel then moves to developing recommendations, the endproduct of the guideline development exercise. Development of recommendations requires considering the following factors: the quality of the evidence, the balance between benefits and harms, a consideration of patient values and preferences, equity, feasibility, acceptability and resource implications. These factors should be seen as broad domains with many items being considered within. Guideline panels will determine and express the fundamental nature of the recommendation either for or against an intervention or diagnostic test or strategy on the basis of evaluating the desirable and undesirable consequences that result across these domains. Panels will then formulate recommendations in a clear and unambiguous way categorising the strength of the recommendation as either a strong or a conditional (weak) recommendation. They will give implementation and research due consideration as part of this process.

Appendix 3: WHO Conflict of Interest Form

DECLARATION OF INTERESTS FOR WHO EXPERTS

WHO's work on global health issues requires the assistance of external experts who **may have interests related to their expertise.** To ensure the highest integrity and public confidence in its activities, WHO requires that experts serving in an advisory role disclose any circumstances that could give rise to a potential conflict of interest related to the subject of the activity in which they will be involved.

All experts serving in an advisory role must disclose any circumstances that could represent a **potential conflict of interest** (i.e., any interest that may affect, or may reasonably be perceived to affect, the expert's objectivity and independence). You must disclose on this Declaration of Interest (DOI) form any financial, professional or other interest relevant to the subject of the work or meeting in which you have been asked to participate in or contribute towards and any interest that could be affected by the outcome of the meeting or work. You must also declare relevant interests of your immediate family members (see definition below) and, if you are aware of it, relevant interests of other parties with whom you have substantial common interests and which may be perceived as unduly influencing your judgement (e.g. employer, close professional associates, administrative unit or department).

Please complete this form and submit it to WHO Secretariat if possible at least 4 weeks but no later than 2 weeks before the meeting or work. You must also promptly inform the Secretariat if there is any change in this information prior to, or during the course of, the meeting or work. All experts must complete this form before participation in a WHO activity can be confirmed.

Answering "Yes" to a question on this form does not automatically disqualify you or limit your participation in a WHO activity. Your answers will be reviewed by the Secretariat to determine whether you have a conflict of interest relevant to the subject at hand. One of the outcomes listed in the next paragraph can occur depending on the circumstances (e.g, nature and magnitude of the interest, timeframe and duration of the interest).

The Secretariat may conclude that no potential conflict exists or that the interest is irrelevant or insignificant. If, however, a declared interest is determined to be potentially or clearly significant, one or more of the following three measures for managing the conflict of interest may be applied. The Secretariat (i) allows full participation, with public disclosure of your interest; (ii) mandates partial exclusion (i.e., you will be excluded from that portion of the meeting or work related to the declared interest and from the corresponding decision making process); or (iii) mandates total exclusion (i.e., you will not be able to participate in any part of the meeting or work).

All potentially significant interests will be **disclosed** to the other participants at the start of the activity and you will be asked if there have been any changes. A summary of all declarations and actions taken to manage any declared interests will be **published** in resulting reports and work products. Furthermore, if the objectivity of the work or meeting in which you are involved is subsequently questioned, the contents of your DOI form may be made available by the Secretariat to persons outside WHO if the Director-General considers such disclosure to be in the best interest of the Organization, after consulting with you. Completing this DOI form means that you agree to these conditions.

If you are unable or unwilling to disclose the details of an interest that may pose a real or perceived conflict, you must disclose that a conflict of interest may exist and the Secretariat may decide that you be totally recused from the meeting or work concerned, after consulting with you.

Name:
Institution:
Email:

Date and title of meeting or work, including description of subject matter to be considered (if a number of substances or processes are to be evaluated, a list should be attached by the organizer of the activity):

_____

_____

Please answer each of the questions below. If the answer to any of the questions is "yes", briefly describe the circumstances on the last page of the form.

The term "you" refers to yourself and your immediate family members (i.e., spouse (or partner with whom you have a similar close personal relationship) and your children). "Commercial entity" includes any commercial business, an industry association, research institution or other enterprise whose funding is significantly derived from commercial sources with an interest related to the subject of the meeting or work. "Organization" includes a governmental, international or non-profit organization. "Meeting" includes a series or cycle of meetings.

EMPLOYMENT AND CONSULTING

Within the past 4 years, have you received remuneration from a commercial entity or other organization with an interest related to the subject of the meeting or work?

1a Employment Yes | No |

1b Consulting, including service as a technical or other advisor Yes | No |

RESEARCH SUPPORT

Within the past 4 years, have you or has your research unit received support from a commercial entity or other organization with an interest related to the subject of the meeting or work?

2a Research support, including grants, collaborations, sponsorships, and other funding Yes| No|

2b Non-monetary support valued at more than US $1000 overall (include equipment, facilities, research assistants, paid travel to meetings, etc.)

Support (including honoraria) for being on a speakers bureau, giving speeches or training for a commercial entity or other organization with an interest related to the subject of the meeting or work? Yes |No|

INVESTMENT INTERESTS

Do you have current investments (valued at more than US $10 000 overall) in a commercial entity with an interest related to the subject of the meeting or work? Please also include indirect investments such as a trust or holding company. You may exclude mutual funds, pension funds or similar investments that are broadly diversified and on which you exercise no control.

3a Stocks, bonds, stock options, other securities (e.g., short sales) Yes| No |

3b Commercial business interests (e.g., proprietorships, partnerships, joint ventures, board memberships, controlling interest in a company) Yes| No |

INTELLECTUAL PROPERTY

Do you have any intellectual property rights that might be enhanced or diminished by the outcome of the meeting or work?

4a Patents, trademarks, or copyrights (including pending applications) Yes | No|

4b Proprietary know-how in a substance, technology or process Yes |No |

PUBLIC STATEMENTS AND POSITIONS (during the past 3 years)

5a As part of a regulatory, legislative or judicial process, have you provided an expert opinion or Yes |No |

testimony, related to the subject of the meeting or work, for a commercial entity or other organization?

5b Have you held an office or other position, paid or unpaid, where you represented interests or defended a position related to the subject of the meeting or work? Yes | No|

ADDITIONAL INFORMATION

6a If not already disclosed above, have you worked for the competitor of a product that is the subject of the meeting or work, or will your participation in the meeting or work enable you to obtain access to a competitor's confidential proprietary information, or create for you a personal, professional, financial or business competitive advantage? Yes |No |

6b To your knowledge, would the outcome of the meeting or work benefit or adversely affect interests of others with whom you have substantial common personal, professional, financial or business interests (such as your adult children or siblings, close professional colleagues, administrative unit or department)? Yes| No|

6c Excluding WHO, has any person or entity paid or contributed towards your travel costs in connection with this WHO meeting or work? Yes| No|

6d Have your received any payments (other than for travel costs) or honoraria for speaking publicly on the subject of this WHO meeting or work? Yes| No|

6e Is there any other aspect of your background or present circumstances not addressed above that might be perceived as affecting your objectivity or independence? Yes| No|

7. TOBACCO OR TOBACCO PRODUCTS (answer without regard to relevance to the subject of the meeting or work)
Within the past 4 years, have you had employment or received research support or other funding from, or had any other professional relationship with, an entity directly involved in the production, manufacture, distribution or sale of tobacco or tobacco products or representing the interests of any such entity? Yes| No|

EXPLANATION OF "YES" RESPONSES: If the answer to any of the above questions is **"yes", check above and briefly describe the circumstances on this page. <u>If you do not describe the nature of an interest or if you do not provide the amount or value involved where relevant, the conflict will be assumed to be significant.</u>**

Nos. 1 - 4: Type of interest, question number and category (e.g., Intellectual Property 4.a copyrights) <u>and</u> basic descriptive details.	**Name of company, organization, or institution**	**Belongs to you, a family member, employer, research unit or other?**	**Amount of income or value of interest (if not disclosed, is assumed to be significant)**	**Current interest (or year ceased)**

Nos. 5-6: Describe the subject, specific circumstances, parties involved, time frame and other relevant details

CONSENT TO DISCLOSURE. By completing and signing this form, you consent to the disclosure of any relevant conflicts to other meeting participants and in the resulting report or work product.

DECLARATION. I hereby declare on my honour that the disclosed information is true and complete to the best of my knowledge.

Should there be any change to the above information, I will promptly notify the responsible staff of WHO and complete a new declaration of interest form that describes the changes. This includes any change that occurs before or during the meeting or work itself and through the period up to the publication of the final results or completion of the activity concerned.

Date: _____ Signature_____

Appendix 4.1: Example of an evidence profile

Evidence profile: Sublingual immunotherapy vs usual care in adults with seasonal/intermittent AR

Author(s): Itziar Etxeandia **Date:** 2013-11-16

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	SLIT	Control	Relative (95% CI)	Absolute		
Allergic rhinitis symptom scores (SS) (follow-up median 7 months[1]) (Better indicated by lower values)												
33	randomised trials	No serious[2]	Serious[3]	no serious indirectness	no serious imprecision	none	1768	1708	-	SMD 0.38 lower (0.49 to 0.27 lower) [4]	⨁⨁⨁◯ MODERATE	CRITICAL
Ocular symptoms (follow-up median 7 months[5]; Better indicated by lower values)												
8	randomised trials	serious[6]	no serious inconsistency[7]	no serious indirectness	serious	none	597	616	-	SMD 0.26 lower (0.06 to 0.46 lower)	⨁⨁◯◯ LOW	IMPORTANT
Medication scores (MS) (follow-up median 7 months[1]) (Better indicated by lower values)												
27	randomised trials	No serious[2]	Serious[3]	no serious indirectness	no serious imprecision	none	1353	1438	-	SMD 0.35 lower (0.47 to 0.23 lower) [9]	⨁⨁⨁◯ MODERATE	IMPORTANT
Combined SS and MS (SMS) (follow-up median 7 months[10]) (Better indicated by lower values)												
5	randomised trials	No serious	Serious[11]	no serious indirectness	no serious imprecision	none	541	546	-	SMD 0.44 lower (0.62 to 0.27 lower) [12]	⨁⨁⨁◯ MODERATE	IMPORTANT
QoL (disease specific RQLQ) (follow-up median 7 months[10]) (Better indicated by lower values)												
6	randomised trials	No serious	Serious[13]	no serious indirectness	no serious imprecision	none	818	840	-	SMD 0.36 lower (0.46 to 0.26 lower) [14]	⨁⨁⨁◯ MODERATE	CRITICAL
Serious adverse effects (follow-up median 7 months[1])												
36	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/2253 (0%)	0/1906 (0%)	not pooled[15]	not pooled	⨁⨁⨁⨁ HIGH	IMPORTANT
Withdrawal due to adverse effect (follow-up median 7 months[1])												
25	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious[16]	none	70/1691 (4.1%)	16/1430 (1.1%)	RR 2.91 (1.72 to 4.92)	21 more per 1000 (from 8 more to 44 more)	⨁⨁⨁◯ MODERATE	CRITICAL
Oral pruritus or burning (follow-up median 7 months[1]) [17]												
19	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	strong association[18]	481/1304 (36.9%)	73/1152 (6.3%)	RR 4.92 (3.16 to 7.67)	248 more per 1000x (from 137 more to 423 more)	⨁⨁⨁⨁ HIGH	CRITICAL
Oral oedema (follow-up median 8 months[1,19])												
7	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	Serious[20]	very strong association[21]	113/763 (14.8%)	4/702 (0.6%)	RR 11.47 (4.66 to 28.24)	60 more per 1000 (from 21 more to 155 more)	⨁⨁⨁⨁ HIGH	CRITICAL
Gastrointestinal adverse effects (follow-up median 7 months[1]; nausea, vomiting, stomach upset, diarrhoea)												
9	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious[22]	none	40/482 (8.3%)	10/413 (2.4%)	RR 2.85 (1.44 to 5.65)	45 more per 1000 (from 11 more to 113 more)	⨁⨁⨁◯ MODERATE	CRITICAL

[1] The duration of maintenance treatment and the period of follow up varied considerably between studies, largely reflecting pre-seasonal, co-seasonal and perennial administration. Range of follow-up was 1 to 48 months

[2] Most studies were at low or unclear risk of bias, mostly because they did not report the sequence generation and in some cases allocation concealment. Majority of studies did not report following intention-to-treat principle and was analysed per-protocol.

[3] There was some inconsistency in the results with I2= -48%49%.

[4] Moderate effect sizes favouring active SLIT in the adults subgroup analysis, and these did not differ significantly in the subgroups analysis of the 42 studies with age (children and adults together (SMD: -0.33 (95%IC -0.42 to-0.25)) , study duration (42 studies) (<6 months, 6-12 months,>12monts), major allergen content (31 studies) (5µg, 5-20 µg, >20 µg) or type of allergen (42 studies) (Grass, Ragweed, Parietaria, tree).

[5] Range: 3.5 to 18 months.

[6] In all studies but one between 10% and 20% of patients withdrew from the study. Majority of studies did not report following intention-to-treat principle and was analysed per-protocol.

[7] There was some inconsistency in results, but removing the studies with extreme results did not substantially change the estimate of effect.

[9] Combined SMD of the 35 studies which included Children and adults was –0.27 (95% CI –0.37 to –0.17) but MSs in children were not significantly better than with placebo treatment (see GRADE profile in the next question).On the other hand small to moderate effect sizes favouring active SLIT were found in all subgroup analyses of the 35 studies, study duration (<6 months, 6-12 months,>12monts), MAC (5µg, 5-20 µg, >20 µg) and type of allergen (Grass, Ragweed, Parietaria, tree).

[10] Range of follow-up was 3 to 10 months

[11] Some heterogeneity between Studies I2: 41%.

[12] When all 6 studies of Children and adults are taking together the combined SMD was similar (–0.40 (95% CI –0.55 to –0.25)), furthermore moderate effect sizes favouring active SLIT were found in all subgroup analyses of those 6 studies conducted in children and adults [study duration (6 studies) (<6 months, 6-12 months,>12monts), MAC (3 studies) (5-20 µg) or type of allergen (4 studies) (Grass)], and these were similar between studies.

[13] Some heterogeneity between Studies I2: 69%. Four of the included studies used the full version of the disease-specific RQLQ to measure QoL, the others an alternative version. Nevertheless the subgroup analysis of those four studies showed a similar combined SMD – 0.34 (95%IC -0.49 to -0.18).

[14] When all 7 studies of Children and adults are taking together the combined SMD was similar -0.37 (95%IC -0.52 to -0.22), moderate effect sizes favouring active SLIT were found in all subgroup analyses of those 7 studies conducted in children and adults [study duration (6 studies) (<6 months, >12monts) or MAC (4 studies) (5-20 µg, >20 µg).

15 There were no serious adverse observed in any of the 36 studies and five new trials added in the Meadows et al. meta-analysis reported a total of 20 SAEs in a total of 1565 study participants, of which only one, abdominal pain in a placebo-treated patient, was considered likely to be treatment related.

[16] Only 86 events

[17] In the new RCT added in the Meadows et al. meta-analysis the numbers of adverse events were generally not reported. The most commonly reported local reactions were itching, swelling and burning in the oral cavity. Four trials (n = 890), one in children (n= 307) and three in adults (n=583) reported oral pruritus (39% in active group vs. 5% placebo); two trials (n = 782) reported throat irritation (33% active vs. 4% of control), and mild erythema (11% active vs. 1% control); and three trials (n = 863) reported oral paraesthesia (10% in SLIT vs. 2% in placebo) and mouth oedema (9% in SLIT vs. 1% in placebo).

[18] Lower confidence limit was 3.16.

[19] Range: 4 to 24 months.

[20] Only 117 events.

[21] Lower confidence limit was 4.66 21

[22] Only 50 events.

Appendix 4.2: Example of evidence to decision framework

Question 3: Should sublingual specific immunotherapy be used for treatment of allergic rhinitis in adults without concomitant asthma?

Problem: Adults with Allergic Rhinitis
Option: sublingual specific immunotherapy
Comparison: No treatment
Setting: Outpatient
Perspective: Health Care system

Background: Background: Allergic rhinitis (AR) is defined clinically by nasal hypersensitivity symptoms induced by an immunologically mediated (most often IgE-dependent) inflammation after the exposure of the nasal mucous membranes to an offending allergen. Symptoms of rhinitis include rhinorrhea, nasal obstruction or blockage, nasal itching, sneezing, and postnasal drip that are reversible spontaneously or under treatment. Allergic conjunctivitis often accompanies allergic rhinitis.
Allergic rhinitis has been traditionally subdivided into seasonal, perennial, and occupational rhinitis. Perennial allergic rhinitis is most frequently, although not necessarily, caused by indoor allergens such as house dust mites, moulds, cockroaches, and animal dander. Seasonal allergic rhinitis is most often caused by outdoor allergens such as pollens or moulds. As in a 2010 edition of ARIA guideline in this document we retained the terms “seasonal” and “perennial” to enable the interpretation of published studies, and we also include the terms used to classify AR according to the duration of symptoms as “intermittent” rhinitis (symptoms are present less than 4 days a week or for less than 4 weeks) or “persistent” (symptoms are present at least 4 days a week and for at least 4 weeks).
These guidelines do not address the issues related to diagnosis of allergic rhinitis and it is assumed that the correct diagnosis had been established before commencing treatment.

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
PROBLEM	Is the problem a priority?	No ☐ Probably No ☐ Uncertain ☐ Probably Yes ☐ Yes ☑ Varies ☐	1. Overall risk of AR in adults Saudi Arabia is 90 per 1000 (79% SAR) Overall in the Middle East: • Runny nose, nasal and throat itching, postnasal drip, and nasal congestion or stuffed up nose were the most common and bothersome symptoms of AR. • 58% of participants with AR reported that the condition had an impact on their daily private and professional life. • 72% reported that limitations on their work/school activities • 35% reported that interfered with and caused them to miss work or • Sleep disturbances were shown in this survey to be extremely troubling in 15% of AR patients. (Abdulrahman H, 2012. Survey conducted in Middle East including KSA) 2. A high percentage of patients with AR surveyed missed work or had their	The guideline panel estimates a prevalence of 20% to 40% of AR in KSA. They consider that due to the lack of an appropiate data base with this data, the self- reporting studies could underestimate the prevalence (for not recognize the symptoms or not having a medical diagnosis) or overestimate (for considering any kind of rhinitis not only the allergic one).

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
			work performance affected by allergies: work productivity decreasing by 23% in AIA, 24% in AIAP, 33% in AILA and 30% in Middle East when allergy symptoms were at their worst. Nasal allergies also interfered with many patients' sleep, and were associated with feelings of depression, anxiety, irritability and tiredness. (Blaiss 2012, America, Asia pacific, Latin America, and Middle East)	

Seasonal / Intermittent Allergic Rhinitis

BENEFITS & HARMS OF THE OPTIONS

What is the overall certainty of this evidence?

No included studies	Very low	Low	Moderate	High
☐	☐	☐	☒	☐

Is there important uncertainty about how much people value the main outcomes?

Important uncertainty or variability	Possibly important uncertainty or variability	Probably no important uncertainty or variability	No important uncertainty or variability	No known undesirable outcomes
☐	☒	☐	☐	☐

Are the desirable anticipated effects large?

No	Probably No	Uncertain	Probably Yes	Yes	Varies
☐	☐	☒	☐	☐	☐

RESEARCH EVIDENCE

Outcome	Relative importance	Certainty of the evidence (SAR)
Nasal symptoms	Critical	Moderate
Ocular symptoms	Important	Low
Medication score	Important	Moderate
Symptom-medication score	Important	Moderate
Quality of life	Critical	Moderate
Serious adverse effects	Important	High
Withdrawal due to adverse effect	Critical	High
Oral pruritus or burning	Critical	High
Oral oedema	Critical	High
Gastrointestinal adverse effects	Critical	Moderate

Summary of the evidence for patients' values and preferences:

ADDITIONAL CONSIDERATIONS

- There is a concern that some patients in KSA would not accept SLIT with some allergens of animal origin.

- Also considered that most people initially do not accept SLIT but when the symptoms do not decrease with all other regular options, they accept this medication with its adverse effects.

- It is considered that the lack of adherence with the medication use is not related with its adverse effects but with the long duration of treatment.

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
	Are the undesirable anticipated effects small?	No ☐ / Probably No ☒ / Uncertain ☐ / Probably Yes ☐ / Yes ☐ / Varies ☐	This recommendation places a relatively high value on alleviating the symptoms of rhinitis, and relatively low value on avoiding adverse effects and resource expenditure. Local adverse effects are relatively frequent (~35%). An alternative choice may be equally reasonable, if patients' values or preferences differ from those described here. ***Summary of findings:*** see evidence table and reference list	
	Are the desirable effects large relative to undesirable effects?	No ☐ / Probably No ☒ / Uncertain ☐ / Probably Yes ☐ / Yes ☐ / Varies ☐		

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
RESOURCE USE	Are the resources required small?	No ☒ / Probably No ☐ / Uncertain ☐ / Probably Yes ☐ / Yes ☐ / Varies ☐	1. SLIT was compared with standard therapy, It was (just) more effective or, in some cases, both more effective and cost-effective - SLIT is likely to be cost-effective at thresholds of £20,000; (Meadows A, 2013. SR) - These studies did not, however, report all of the utility data in a disaggregated form and all were funded by a manufacturer of SIT products (Meadows A, 2013. SR)	**- Average annual cost per patient: around 35 K SAR** **- Average cost per treatment (3 years) and patient: around 100K SAR** Average maintenance vial/ allergen/ month =707 SAR. Average 4 allergens/patient: Annual cost= 707 X 4 X 12 = 33, 936 SAR
	Is the incremental cost small relative to the net	No ☒ / Probably No ☐ / Uncertain ☐ / Probably Yes ☐ / Yes ☐ / Varies ☐		

	CRITERIA	JUDGEMENTS	RESEARCH EVIDENCE	ADDITIONAL CONSIDERATIONS
	benefits?			
EQUITY	What would be the impact on health inequities?	*Increased* ☒ *Probably increased* ☐ *Uncertain* ☐ *Probably reduced* ☐ *Reduced* ☐ ***Varies*** ☐		**Comments from the panel members:** 1. If sublingual immunotherapy use were to be recommended, the health inequity will increase so the indications and the applications of SLIT should be determined: The SLIT should be used only when all other regular options do not work 2. Impact: Few patients will be affected
ACCEPTABILITY	Is the option acceptable to key stakeholders?	*No* ☐ *Probably No* ☐ *Uncertain* ☒ *Probably Yes* ☐ *Yes* ☐ ***Varies*** ☐		Uncertain acceptance from patients and likely not for health care system because of cost consideration reasons
FEASIBILITY	Is the option feasible to implement?	*No* ☐ *Probably No* ☒ *Uncertain* ☐ *Probably Yes* ☐ *Yes* ☐ ***Varies*** ☐		Implementation would require expertise and resources (i.e. skin tests, relevant allergen) not readily available in most areas.

Balance of consequences	Undesirable consequences *clearly outweigh* desirable consequences in most settings	Undesirable consequences *probably outweigh* desirable consequences in most settings	The balance between desirable and undesirable consequences *is closely balanced or uncertain*	Desirable consequences *probably outweigh* undesirable consequences in most settings	Desirable consequences *clearly outweigh* undesirable consequences in most settings
	☐	☐	☒	☐	☐

Type of recommendation	We recommend against offering this option	We suggest not offering this option	We suggest offering this option	We recommend offering this option
	☐	☐	☒	☐

Recommendation (text)	The KSA MoH panel suggests sublingual immunotherapy for treatment of adults with seasonal or intermittent allergic rhinitis (conditional recommendation; Moderate-quality evidence).
Justification	The evidence, with an overall moderate certainty, shows that the desirable effects probably are not large relative to undesirable effects. Furthermore, possibly there is an important variability about how much people value its effectiveness because there is a concern that some patients in KSA would not accept SLIT with some allergens of animal origin, however others would accept it as the last option when the symptoms do not decrease with all other regular options. On the other hand the incremental cost is not small relative to the net benefits, and the implementation would require personnel experts and resources (i.e. skin tests, specific allergen) which are not readily available in most areas. Reasons to formulate a conditional rather than a strong recommendation. It is considered that the lack of adherence with the medication use is not related with its adverse effects but with the long duration of treatment. For this reason in the cases when the SLIT would be the treatment of choice clinicians should provide an adequate educational instruction to the patient.
Subgroup considerations	The SLIT should be used only when all other regular options do not work: It is more appropriate for those with moderate to severe AR who does not respond to first line therapy. The SLIT Should not be started during pregnancy, but could be continued if the woman has already started the treatment.
Implementation considerations	SLIT should only be prescribed by allergy specialists who have expertise in diagnosis of AR, proper identification of the allergens, providing immunotherapy and treatment of potentially serious adverse effects.
Monitoring and evaluation	If patients receiving SLIT do not respond within 6-12 m consider discontinuation SLIT