European Reference Networks | ERN GUIDELINES

European Reference Network: Clinical Practice Guidelines And Clinical Decision Support Tools

September 28th 2020

Methodological Handbooks & Toolkit for Clinical Practice Guidelines and Clinical Decision Support Tools for Rare or Low-Prevalence and Complex Diseases

Handbook #4: Methodology for the Development of Clinical Practice Guidelines for Rare or Low-Prevalence and Complex Diseases

Prepared by:
Aragon Health Sciences Institute (IACS)

LEGAL NOTICE

EUROPEAN COMMISSION

Directorate-General for Health and Food Safety

Directorate B — Health systems, medical products and innovation

Unit B.3 —Digital Health, European Reference Networks

Contact: Enrique Terol

E-mail: enrique.terol@ec.europa.eu

European Commission B-1049 Brussels

Authors:

Patricia Gavín Benavent, MD, PhD.
Aragon Health Sciences Institute (IACS, Spain)

Celia Muñoz Fernández, BA in Economics.
Aragon Health Sciences Institute (IACS, Spain)

Authors: Internal Reviewers (in alphabetical order)

María Pilar Blas Diez, Information Specialist
Aragon Health Sciences Institute (IACS, Spain)

María Bono Vega, BSc in Biochemistry.
Aragon Health Sciences Institute (IACS, Spain)

María Soledad Isern de Val, BSc in Biochemistry, PhD.
Aragon Health Sciences Institute (IACS, Spain)

Lucía Prieto Remón, BA in Business and Marketing.
Aragon Health Sciences Institute (IACS, Spain)

Contact Information:

FPS-AETSA
ERNGuideLinesCoordination.fps@juntadeandalucia.es
Fundación Pública Andaluza Progreso y Salud (FPS)
Avd. Américo Vespucio 15, Edificio S-2
C.P. 41092 Seville, Spain
+34 955 006636

This handbook includes a detailed explanation of the process for developing Clinical Practice Guidelines for rare diseases, including:

- ✓ Selecting the CPG topic
- ✓ Determining the CPG scope
- ✓ Preparing the work plan
- ✓ Forming the guideline development group
- ✓ Developing the clinical questions
- ✓ Systematic search for evidence
- ✓ Determining the CPG scope
- ✓ Preparing the work plan
- ✓ Forming the guideline development group (GDG)
- ✓ Developing the clinical questions
- ✓ Selecting relevant evidence
- ✓ Appraising identified research evidence
- ✓ Evidence synthesis and analysis
- ✓ Creating recommendations
- ✓ Final stakeholder consultation
- ✓ Publishing
- ✓ Guideline implementation strategies
- ✓ Updating recommendations.

Purpose:

To provide guidance for the development of Clinical Practice Guidelines for rare diseases.

TABLE OF CONTENTS

ABBREVIATIONS

CDSTs	Clinical Decision Support Tools
CPGs	Clinical Practice Guidelines
CPMS	Clinical Patient Management System
EC	European Commission
EE	Economic evaluation
ERN	European Reference Network
EtD	Evidence to decision
EU	European Union
GDG	Guideline Development Group
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HTA	Health technology assessment
IACS	Aragon Health Sciences Institute
PICO	Patient/Population-Intervention-Comparison-Outcome
QALY	Quality-adjusted life years
ROBINS	Risk Of Bias In Non-randomized Studies of Interventions
SR-EE	Systematic review of economic evaluations

01.

BACKGROUND

There are a number of challenges surrounding the development of CPGs and CDSTs for rare diseases. One of the most relevant barriers is the lack of high-quality evidence, which cutting-edge methodological frameworks like GRADE [1] rely on.

Therefore, there is a need for specific methodological approaches that can provide reliable and useful CPGs and CDSTs for rare diseases. The project also aims to provide a common methodology to harmonise the development of CDSTs and CPGs.

It is worth noting that within the scope of this document, “rare diseases” is the term used to refer to rare diseases, as well as low prevalence complex diseases.

1.1 | Context for the development of Clinical Practice Guidelines in rare diseases

Rare diseases are a global health priority. Though each disease is rare, when taken together the thousands of known rare diseases cause significant morbidity and mortality, impact quality of life, and confer a social and economic burden on families and communities. These conditions are, by their nature, encountered very infrequently by individual clinicians, who may feel unprepared to address their diagnosis and treatment.

Clinical practice guidelines gather existing knowledge and make it available and readily accessible to healthcare professionals, improving effectiveness and quality of care delivered to patients.

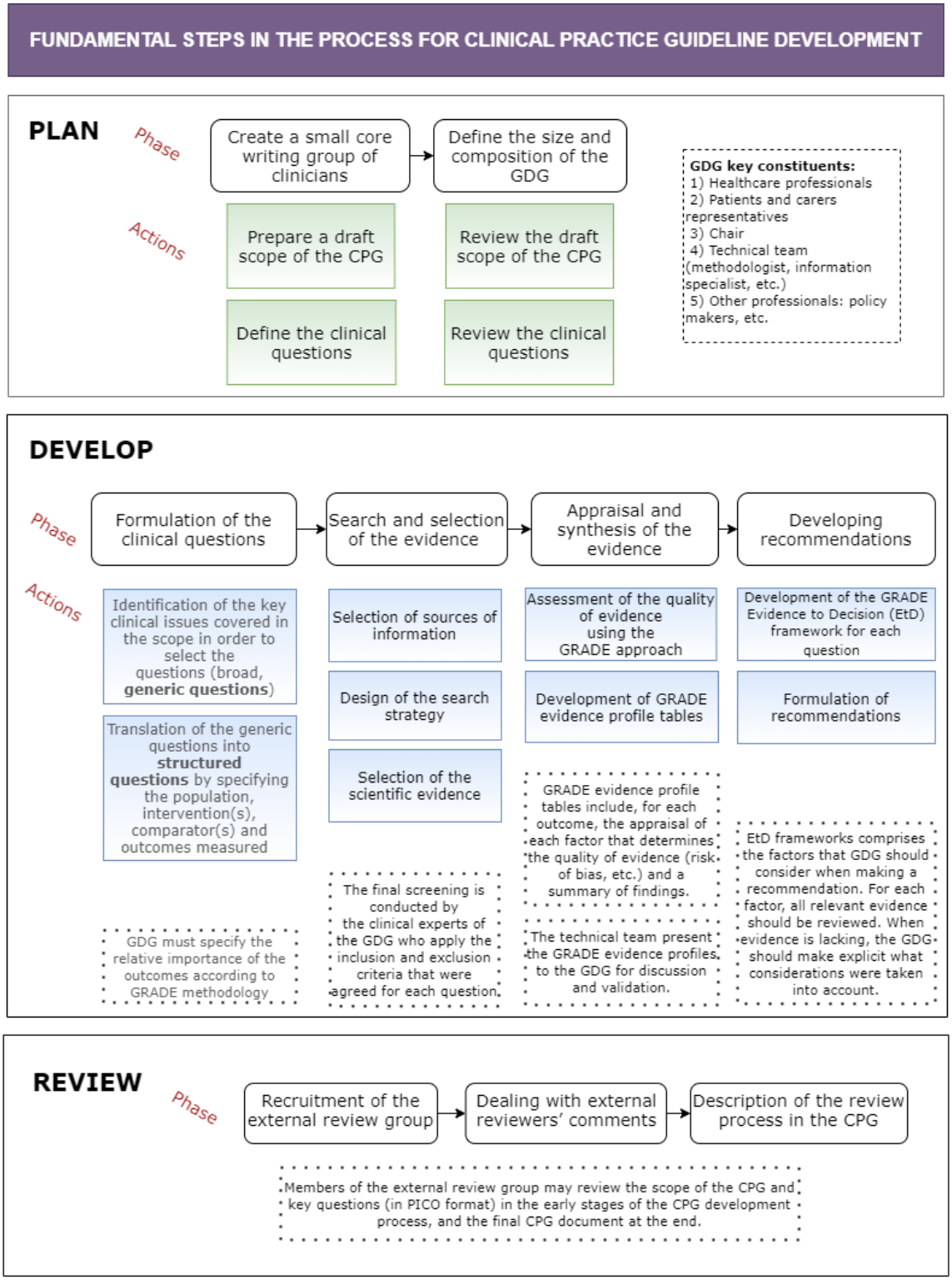
This document seeks to support the development of CPGs for rare diseases. It covers all steps of guideline development and has been designed to meet the reporting standards for trustworthy guidelines[3-5]. Multiple handbooks by guideline developers were reviewed for writing this hanbook[6-11], that uptakes the GRADE system (Grading of Recommendations Assessment, Development and Evaluation) to summarize evidence, grade its quality, and interpret it to make clinical recommendations. It also presents the different resources that discuss the methodological process in greater detail.

1.2 | The development process for clinical practice guidelines (CPG): essential steps

The Institute of Medicine defines clinical practice guidelines as "statements that include recommendations, intended to optimize patient care, that are informed by a systematic review of evidence and an assessment of the benefits and harms of alternative care options"[4].

The key stages in the development of CPG are summarised in Figure 1[6, 12, 13].

Figure 1. Essential steps in Clinical Practice Guideline development



02.

SCOPE AND PURPOSE OF THE CLINICAL PRACTICE GUIDELINE

The objective(s) and scope of the clinical practice guideline (CPG) should be clearly stated. This chapter describes in detail this essential step in the guideline development process. The scope defines the aspects of care that will and will not be covered in the CPG, the target population, the intended users and the context of application.

Preparing the scope and purpose is the first step in developing a CPG. The result of this phase is a document that clearly defines the framework for the development of the guideline. A good scope definition ensures that the approach will meet the objectives of the CPG and facilitates the development of the clinical questions and other parts of the guideline.

2.1 | Steps for determining the Clinical Practice Guideline scope

The first step in defining the scope is to create a small core writing group of clinicians with adequate knowledge in the clinical area of the CPG. They prepare a draft scope of the guideline and define the review questions that cover all areas specified in the scope. To address this task, a preliminary search of scientific literature or scoping review on the condition of interest will be necessary, in order to identify the key clinical issues (see section 2.2).

Once a draft scope is defined, the size and composition of the guideline development group (GDG) can be considered. The final scope and clinical questions of the CPG will require input from the GDG.

An external consultation process with experts in the topic of the guideline is recommended to ensure the relevance of all the issues to be addressed by the CPG (see chapter 9). Patients should also be consulted to define the patient-related elements that need to be addressed.

2.2 | Scoping review

Scoping reviews have been described as a process of mapping the existing literature or evidence base relating to a particular topic [14]. This preliminary search of the literature can be used to:

- ✓ Explore the scope of the literature and identify relevant CPG and systematic reviews.

- ✓ Identify the most important aspects of care that the clinical guideline will cover.
- ✓ Define the target population.
- ✓ Identify gaps or overlaps in current guidance that can justify the need for a guideline.

The search should not be exhaustive. It should be based on the need to reasonably inform the content of the CPG scope. The key phases of this literature review method are listed below:

- ✓ Identifying relevant evidence. Decisions will need to be made on the range of sources (e.g. online databases, key organisational websites) and search terms to be included. In addition to looking through peer-reviewed literature, it is recommended to search government websites and publications, organisational reports and other sources of 'grey' literature.

  Study selection. It may be useful to identify a series of inclusion and exclusion criteria to discard irrelevant documents. These criteria should be broad enough to provide a map of the existing literature. A scoping review may prioritise CPG and systematic reviews.

- ✓ Charting the data. A template may be created to chart relevant data. This will enable review authors to identify commonalities, themes and gaps in the literature. Potential data collection categories include:
  - authors,
  - year of publication,
  - publication type (e.g. CPG, systematic review, randomized controlled trial),
  - target population,
  - scope of the guideline or aims of the study,
  - overview of methods,
  - results

- ✓ Summarising and reporting the results. The scoping review provides an overview of existing literature without assessing quality of included studies and therefore data synthesis is minimal. It is recommended to apply meaning to the results by considering the implications of the findings of the scoping review within the broader practice and policy context, for example, by tagging them.

2.3 | Information to be provided

The document of the scope and purpose has to be structured and clear (see annex I). The components include the following[6]:

- ✓ Reasons for why the guideline is needed (justification)
- ✓ Objectives of the guideline
- ✓ Aspects to be covered
  - Target population
  - Aspects of care that the guideline will cover
  - Aspects related to patients
  - Context of application
  - Issues relevant to special needs groups
- ✓ Aspects not covered by the CPG

- ✓ Considerations with regards to health inequities
- ✓ End users of the CPG

Justification

The document should include an explanation of why the guideline is needed, for example due to a large (unexplained) variability in clinical performance, the presence of areas of uncertainty, or important changes in available evidence.

Objectives of the guideline

The general and specific objectives of the guideline should be stated, together with the benefits that the CPG aims to achieve. Specific objectives describe what will be researched during the study, whereas the general objective is a much broader statement on the overall aims of the study.

Examples of general objectives include the following:

- ✓ Establish a set of evidence-based recommendations to improve the health status of the people affected by the condition addressed by the CPG.
- ✓ Promote efficiency in the choice among all the available diagnostic and therapeutic options.

Examples of specific objectives include:

- ✓ Decrease the variability among clinicians in the diagnosis and therapeutic approach of patients with the condition addressed by the CPG.
- ✓ Decrease the frequency and severity of the adverse effects of a particular treatment, caused by an inappropriate prescription related to dosages, age group or comorbidity.

Aspects to be covered

The aspects to be included in the CPG are listed below:

- ✓ Target population: characteristics of the target population and any subgroups should be described clearly (age group, type of disease or condition, disease or condition severity, or comorbidities).
- ✓ Aspects of care that the guideline will cover: the area of health practice, policy or public/environmental health issue that the guideline addresses. For example, diagnostic tests, surgical treatments, medical and psychological therapies, rehabilitation and lifestyle advice. It is important that the scope is as specific as possible with regard to the interventions the guideline is intended to cover.
- ✓ Aspects related to patients: the way in which the perspective of patients and carers is included should be described, and the development of topic-specific information and support for patients and carers should be stated.
- ✓ Context of application: The health care setting to which the recommendations apply is described, including the health system level (e.g. primary care, acute care) and clinical stage (e.g. whether the guideline covers prevention, screening, assessment, treatment, rehabilitation or monitoring).

Aspects not covered by the CPG

Although the aspects covered and not covered by the CPG are complementary, they should be stated clearly so that the scope is well-defined. If the CPG excludes any clinical stage (e.g. prevention), or certain age groups (e.g. teenagers) or clinical conditions (e.g. hypertensive crisis in a CPG on hypertension), this should be reflected.

Dealing with health inequities

Issues relevant to special-needs groups such as culturally and linguistically diverse communities or groups with low socioeconomic status (e.g. particular risks, treatment considerations or sociocultural considerations) are identified and described.

End users of the CPG

The intended end users of the guideline are clearly defined, and any relevant exceptions are identified. For example, all the health professionals involved in managing the condition, social work professionals, patients and carers, and others.

Key issues

The scope and purpose of the CPG should incorporate the contributions of the institution promoting the development of the guideline, a number of experts on the topic addressed, and the guideline development group. It is also important to take into account the patient and carer perspective.
The final document has to be structured and clear, and it should include at least the following issues:

- Reasons why the CPG is needed (justification)
- Objectives of the CPG
- Target population
- Aspects of care that the CPG will cover
- Aspects not covered by the CPG
- Context of application
- End users of the CPG

03.

GUIDELINE DEVELOPMENT GROUP

This chapter provides information on the size, composition and function of the guideline development group (GDG). These include the roles and responsibilities of the different profiles of the GDG members. It also discusses the practical issues of working in a group to develop a guideline.

3.1 | Composition of the Guideline Development Group

The guideline development group (GDG) must be multidisciplinary and represent the expertise and views relevant to the particular needs of the guideline. Although it is likely that one professional group may dominate, comprehensive stakeholder involvement is as important to the development of guidelines for rare diseases as it is for common diseases [15]. The groups should preferably have 7 to 15 members, apart from the chair and the technical team. More than 15 participants may result in ineffective functioning, whereas less than 7 members may undermine representativeness.

The GDG has four key constituents [6, 16]:

- ✓ Healthcare professionals who are involved at any stage of the care received by patients with rare diseases [17].
    - This implies including at least members of the corresponding European Reference Network (ERN) and, depending on the disease, any other professional usually involved in the care of the patient with the rare condition (e.g. a psychologist). Ideally, members of the ERN should be drawn from different parts of Europe, but this will be influenced by the expertise available.
    - The opinion of a general practitioner, or a paediatrician in the case of a paediatric disease, is imperative.
    - For diseases revealed in paediatric age, the group must not only include paediatric care specialists but also adult care professionals in order to organise the transition from paediatric to adult medicine.
    - Scientific societies or professional national councils concerned can be included.
- ✓ International experts in the guideline topic.
- ✓ Patient and carer representatives. Ideally, the GDG should be supported by a patient advisory group of around 8-10 patients with the disease. The chair of the patient advisory group can be a

formal member of the GDG to represent the views and opinions of the patient group.

- ✓ Technical team.
  - The GDG should include at least one methodologist with expertise in methods to review evidence and develop guidelines, and one information specialist with expertise in scientific literature searching.
  - Ideally the GDG will include an expert on health economics.
- ✓ A chair with leadership capabilities and experience in evidence-based guideline. The chair guides discussions without controlling them and effectively leads and guides the GDG through the tasks of developing the CPG. The chair may be a specialist in the guideline topic, but does not need to be a content expert. The chair should be recruited early to assist in the initial project planning stages and to help select other members of the group.
- ✓ Other professionals: policy makers, healthcare managers, etc.

Table 1. Roles and functions of the GDG members (adapted from NICE).

Group member	Key responsibilities
All members	• Agree on the scope, questions and inclusion and exclusion criteria • Contribute constructively to meetings • Declare all relevant interests • Develop recommendations based on the evidence reviews, or on consensus when evidence is poor or lacking • Identify potential implementation issues and propose steps to overcome them • Assess the acceptability and feasibility of the recommendations • Weigh the potential risks and benefits of the recommendation • Make decisions on what information should be included • Consider and deliberate on public consultation submissions
Chair	• Sets up the rules for how the GDG operates • Assists with the planning of the GDG meetings • Establishes a climate of trust and mutual respect among members • Facilitate group processes and promote balanced participation of group members • Support effective patients and carers involvement • Ensure that the group stays focused and task oriented • Summarises the main points and key decisions from the debate
Content experts (Clinical experts, etc.)	• Use their background knowledge and experience of the guideline topic to provide guidance to the technical team in carrying out systematic reviews and economic analyses • Read all relevant documentation and make constructive comments and proposals at (and between) GDG meetings • Advise on how to identify best practice in areas for which limited evidence is available • Apply their knowledge to improving the identification of relevant evidence • Provide context for the evidence including information about how a recommendation might be received by target audiences
Patients and/or carers	• Advise on the guideline scope and clinical questions • Provide comments on the evidence review and ensure that recommendations address patients' and/or carers' issues and concerns. • Consider the extent to which published evidence reflects outcome measures that patients and carers consider important • Highlight areas where patient preferences and patient choice may need to be acknowledged in the guideline • Participate in formal consensus-building procedures where there are gaps in evidence. • Ensure that the guideline is worded appropriately, and in particular the recommendations
Methodological experts	• Identify, critically appraise and synthesise evidence into a format useful for developing recommendations • Assist the group in understanding the evidence and evidence-to-decision process • Inform the GDG about potential economic issues and to perform economic analyses (health economist). • Maintain comprehensive records

3.2 | Running the Guideline Development Group

The organisation that initiated the guideline development process, or was commissioned to do so, is responsible for recruiting members. Health professionals, international experts, patients and carers can be contacted directly or indirectly through the ERN or scientific societies, or through patient organisations, respectively.

The first meeting of the GDG is very important because the operating rules are set up and the roles and functions of each member are defined. The first meeting can also generate the conditions for developing a good group dynamic.

Table 2. Practical issues for planning the first meeting of the GDG [6, 16]:

Notice convening the meeting	• **It should include the date, time, location and agenda of the meeting.**
	• It must specify the main objective of the meeting, the chair of the GDG and the institution promoting the guideline. • If a scope and purpose and a preliminary list of clinical questions are available, they should be sent out in advance, for example, with the notice convening the meeting.
During the meeting	• The first meeting should focus on providing information for GDG members on the following subjects: ○ the process of clinical guideline development ○ methodology for the elaboration of the CPG (GRADE approach) ○ the role of health economics in decision-making ○ how patient and carer members contribute ○ the role of the healthcare professionals and other content experts (researchers, etc.) ○ the role of the technical team
	• The agenda should include time for agreeing the scope and purpose of the guideline and the clinical questions. Ideally, the GDG will have a draft of the scope and purpose and a preliminary list of clinical questions for potential inclusion in the CPG before the meeting.
	• The GDG should consider including additional members to ensure the right mix of expertise relevant to the particular needs of the CPG.
	• Training needs of the GDG should be identified.
Close of the meeting	• Record the agreements set out by the GDG in the minutes of the meeting
	• Agree on the next meeting date

The specific aspects of the CPG development process may also be covered in the first and second GDG meetings. The second meeting can focus on agreeing the clinical questions, based on the scope.

The extent and complexity of the CPG will influence the frequency of meetings during the

development process. They will be conducted via web conferencing tools, and complemented by face-to-face meetings if possible or feasible. Core responsibilities for all meetings include:

- ✓ Setting meeting or conference call dates, which should be done well in advance
- ✓ Planning agenda items
- ✓ Sending out papers
- ✓ Keeping records of all meetings or conference calls
- ✓ Ensuring that all GDG members have a copy of the current guideline handbook

Relevant materials should be distributed before each meeting, with details of what is required from each member during this process. The chair is responsible for ensuring that the agenda is adhered to and that discussions stay on topic.

The GDG should pay particular attention to the needs of patients with rare diseases when scheduling and organising meetings, as they may have on-going health conditions that will impact their ability to engage.

3.3 | Training needs of the Guideline Development Group

Many members of the group may be unfamiliar with the methods used to develop guideline recommendations. Consideration should be given to providing training to these individuals to help them understand the process and improve participation [18]. The training needs of individual members should be assessed before or when the guideline development group meets for the first time.

Important aspects of the process with which members may need to be familiar include:

- ✓ An overview of GRADE in guideline development
- ✓ Formulating and developing clinical questions using frameworks like PICO, identifying and prioritising outcomes that are important to patients
- ✓ The GRADE approach for assessing the certainty of evidence
- ✓ Presenting evidence summary tables
- ✓ Making recommendations using an evidence to decision (EtD) framework and assigning a 'strength of recommendation' using standard terminology

Patient and carer needs for information, support and training must also be addressed in order to enable and ease their contribution to the CPG development process. They need to receive personalised training focusing on methodological aspects and their participation in the different stages. Likewise, it is important to inform healthcare professionals about the relevance of patient and carer participation to ensure that all parties involved work together.

The level of training required by the healthcare professionals largely depends on whether or not there is a technical team in the GDG with experts in methodology and health economics. With the support of a technical team, the training needs of the healthcare professionals may be covered with the aspects mentioned above. In contrast, the members of the GDG should have expertise or be trained in conducting systematic reviews and in applying or using GRADE methodology

3.4 | Making group decisions and reaching consensus

GDG members need to make group decisions throughout the CPG development process. There are many methods for group decision-making but there is no consensus on which method should be used in which scenario.

In most cases, the GDG reaches decisions through a process of informal consensus. In this case, it is important to ensure that each individual view on the GDG is presented and debated in an open and constructive manner at the GDG meetings.

Some GDGs may choose to use more formal procedures for certain decisions. These include, for example, the Delphi method or the Nominal Group Technique. Efforts should be made to avoid visible voting methods as these can make it less likely for members to change their mind [18, 19].

More information on the development of consensus processes can be consulted in Handbook #5: Methodology for the elaboration of Clinical Consensus Statements for rare diseases.

3.5 | Management of conflicts of interest

Potential conflict of interests within the members of the pathway DG should be carefully identified and duly addressed, following the indications established by our partner FPS.

Key issues

The highest-quality clinical practice guidelines involve a development group consisting of a multidisciplinary team of stakeholders, including healthcare professionals, patients and carers, methodologist, and policy makers.

The first meeting of the GDG is the moment to establish an explicit framework that clarifies the objectives of the work, the specific tasks that need to be carried out, the roles and functions of each member, and the timetable.

The training needs of the GDG should be identified and covered to create the best conditions for group members to contribute equally during group discussions, decision-making and when the group is formulating recommendations.

04.

FORMULATION OF THE CLINICAL QUESTIONS

This chapter provides information on how clinical questions are developed, formulated and agreed. It describes the different types of clinical questions that may be used. It also provides information on the process of selection and prioritisation of relevant outcomes according to GRADE methodology.

Translating the scope of the guideline into a list of specific clinical questions is the next step in the development process. They must be clear, focused and closely define the boundaries of the topic. A good clinical question helps to design the search strategy, sets the limits of the systematic review (inclusion and exclusion criteria to identify relevant studies), and serves as a guide for the development of recommendations [20, 21].

4.1 | Defining and selecting clinical questions

The first step in formulating clinical questions is to prepare a list of generic questions. It may be useful to develop an algorithm that summarises the care components covered in the scope, thus allowing the generic questions at every step of the algorithm to be identified [6]. The appropriate selection of questions ensures that the main questions faced by clinicians will be answered.

Each of these generic questions is subsequently turned into one or more specific questions by articulating them in a structured format (described in section 4.2). The definition of specific questions may be informed by a preliminary search of the literature. In some instances, this search may be performed as part of the scoping review for determining the CPG scope (see section 2.2). The GDG members have relevant expertise and will also contribute significantly to refining the generic questions. Furthermore, a process of external review with experts on the guideline topic may be valuable. An example of a generic question turned into a structured specific question is shown in Table 3 (see section 4.2.1).

The exact number of clinical questions for each CPG depends on the topic and the breadth of the scope. It may also vary considerably according to the number of studies included in each question and the complexity of the analyses required to address them. For example, a single clinical question might involve a complex comparison of several treatment options with many individual studies. At

the other extreme, a question might address the effects of a single intervention and have few relevant studies [6, 20]. The number of clinical questions must be manageable for the GDG within the agreed timescale and must therefore be individualised for each CPG.

The process proposed for developing the list of clinical questions is summarised below:

- ✓ The chair of the GDG, with the support of the technical team, prepare a draft list of clinical questions and send the list out to all GDG members before the first meeting.
    - The draft questions may specify in some detail the particular interventions to be compared and the health outcomes of interest identified during a scoping review (see section 2.2).
- ✓ During the first meeting(s) of the GDG, the content experts (clinical experts, policy makers, etc.) and the patients and carers inform the development of the detailed clinical questions and may contribute additional questions. The list of clinical questions must be agreed by all GDG members.
    - Additional searches may be necessary to frame certain clinical questions.
- ✓ The chair coordinates an external review process on the draft list of clinical questions with external experts (clinicians and patients and carers) who can provide their experience and specific expertise.
    - The GDG assesses and responds to the external review comments.
- ✓ Finally, the list of possible questions is approved by the institution that promotes the CPG.

At the end of this process, the clinical questions will not only address all areas covered in the scope, but also will have the proper structure for identifying the relevant scientific evidence.

4.2 | Structuring clinical questions

A specific and answerable question has several essential components, depending on the nature of the guideline and the questions asked, for example, intervention, diagnosis or prognosis.

4.2.1 | Structuring clinical questions

The most common structure used to articulate intervention questions is based on four anatomic parts (population, intervention, comparison, and outcomes), according to the PICO format:

- ✓ Definition of the **population** of interest, specifying the following issues:
    - Health condition or stages of disease.
    - **Characteristics of the population such as age, gender, comorbidities or risk profiles**.
    - Hospital and/or community setting.

  When the rare diseases does not have clear diagnostic criteria, it may be helpful to use a broad definition of the population by incorporating closely related disease entities to potentially increase the amount of data relevant to the PICO question [22].
- ✓ Description of the **intervention** to be evaluated, specifying timing, delivery, setting and resources. For multi-component interventions or community-level interventions, the core components need to be identified.

  When the patterns of practice differ within a given rare diseases or treatments are not used consistently, thus making it difficult to provide a standardised definition of the intervention, the

use of broad definitions may be an adequate approach (e.g. a class of medication instead of a specific medication) [22].

- ✓ Description of the **comparator** or intervention to be compared. Comparisons of interest may include alternative options, no intervention/exposure (placebo) or varied levels of exposure. However, there is often only one treatment option for any given rare diseases, and the use of placebo for comparison is not an option due to the severe course of the untreated disease. Thus, the comparator may be absent [22].

- ✓ Specify all potential clinically relevant and patient important **outcomes** and decide on their relative importance (which will be discussed in section 4.3). Outcomes may include survival (mortality), clinical events (e.g. strokes or myocardial infarction), patient-reported outcomes (e.g. symptoms, quality of life), adverse events, burdens (e.g. demands on caregivers, restrictions on lifestyle) and economic outcomes (e.g. cost and resource use).

  Indirect or surrogate outcome measures, such as laboratory results are potentially misleading and should be avoided or interpreted with caution because they may not predict clinically important outcomes accurately. Surrogate outcomes may provide information on how a treatment might work but not whether it actually does work [21]. Relying on surrogate outcomes can be even more problematic in rare diseases because the pathophysiology and empiric evidence linking them to patient important outcomes are less likely to be well understood [22].

  Composite outcomes combine two or more single outcomes in one outcome to demonstrate overall treatment effects. They should generally be avoided because their individual constituents are often unreasonably combined and inconsistently defined. In rare diseases, where single outcomes are too rare or occur too late and therefore are not sufficiently informative, the use of composite outcomes can be considered, but it has to be justified in an explicit manner[23].

Table 3. Example of clinical question structured in PICO format.

Generic Question **What options exist for the treatment of retinitis pigmentosa?**		
Structured Specific Question **What is the effectiveness and safety of retinal transplantation for the treatment of retinitis pigmentosa?**		
Population	**Intervention/Comparator**	**Outcomes**
Patients with retinitis pigmentosa	Subretinal transplantation of human embryonic stem cells - derived retinal pigment epithelium	Vision-related quality-of-life Visual acuity Transient multifocal electroretinography (mfERG) response Vitreoretinal surgery complications Rejection Adverse proliferation

Adapted from: Grupo de trabajo de la Guía de Práctica Clínica para las Distrofias Hereditarias de Retina. Guía de Práctica Clínica para las Distrofias Hereditarias de Retina. Ministerio de Sanidad, Servicios Sociales e Igualdad. Servicio de Evaluación del Servicio Canario de la Salud; 2017. Guías de Práctica Clínica en el SNS (https://portal.guiasalud.es/wp-content/uploads/2018/12/GPC_565_DHR_SESCS_compl.pdf).

4.2.2 / Clinical questions on diagnosis

Clinical questions on diagnosis may be approached from two different perspectives [6]:

- ✓ Evaluation of the accuracy (e.g. sensitivity and specificity) of a diagnostic test or test strategy.
- ✓ Evaluation of the clinical value of using the test or test strategy in practice with assessment of direct patient important outcomes (e.g. mortality, symptoms, quality of life).

Although the assessment of test accuracy is an important component for establishing the usefulness of a diagnostic test, the clinical value of a test lies in its usefulness in guiding treatment decisions, and ultimately in improving patient outcomes [20].

The purpose of the test or test strategy should be explicit when deciding on the diagnosis question. Potential applications of a test include, for example, establishing prognosis, monitoring illness and treatment response, screening and diagnosis. The GDG should also clearly establish the role of the test or strategy. A new test may substitute an old one (replacement), or may minimize the need for invasive and expensive testing (triage), or may further enhance diagnostic accuracy beyond the existing diagnostic pathway (add-on) [11].

The format of the diagnosis questions follows the same principles as the format for questions on interventions. When comparing test accuracy, the intervention is the test under investigation (index test), the comparison is the best available test (the reference standard), and the outcome is a measure of the presence or absence of the particular disease or disease stage that the index test is intended to identify (e.g. sensitivity or specificity). Clinical questions aimed at establishing the clinical value of a diagnostic test in practice can be structured in the same way as questions on interventions. In this case, the intervention is the index test and the comparison, the reference standard [20].

4.2.3 / Clinical questions on prognosis

Prognostic questions are useful to inform patients about their prognosis, classify patients into risk categories resulting in different treatment decisions, and define subgroups of patients that may respond differently to an intervention [6, 20].

Addressing prognostic questions involves specifying the population, defining multiple prognostic factors, such as attributes of the patient (e.g. age, gender) or features of the condition, and describing the outcomes (mortality or relapse rate and progression).

4.3 | Prioritise outcomes critical to answering the questions

Before starting an evidence review to answer a clinical question, the GDG should apply an initial rating to the importance of outcomes, according to GRADE methodology [24], in order to identify which outcomes of interest are both critical to decision-making and important to patients. This rating should be confirmed or, if absolutely necessary, revised after completing the evidence review [20].

The relative importance of the outcomes will be rated using an ordinal scale of nine units as proposed by the GRADE group. Using this approach, outcomes are classified into those that are critical, those that are important but not critical, and those of limited importance, as illustrated below [24]:

CRITICAL			IMPORTANT			NOT IMPORTANT		
9	8	7	6	5	4	3	2	1

The first two categories of outcomes, especially the first one, will we considered in developing guideline recommendations, whereas those outcomes rated as 'not important' should not be used for this purpose. The outcomes that are important but not critical for decision making should only be taken into consideration when studies using critical outcomes are not available, or to complement critical outcomes when an important aspect in decision making is not covered and need to be informed. Preferably, each clinical question should address a maximum of seven outcomes [6, 24].

The GDG should consider surrogate outcomes only when high-quality evidence regarding important outcomes is lacking. The necessity to substitute with the surrogate may ultimately lead to rating down the quality of the evidence because of indirectness (see section 6.1.2.3) [24].

The outcomes should include not only those that are favourable but also unfavourable, and if relevant, may include health care costs. It is important to remember that important outcomes must not be excluded because it is anticipated that few studies will be found. The most important outcomes for the GDG may not be those most frequently reported in the research literature.

For some fields of research, core outcome sets have been established to inform researchers, often based on usefulness in decision making and importance to patients and healthcare professionals (see the COMET Initiative for a database of known core outcome sets) [22].

This preliminary classification of outcomes before beginning the review of the evidence should be confirmed later. In exceptional circumstances, the results of the evidence review may modify the selection of relevant outcomes or their relative importance [24].

Key issues

- Formulating precise and well-structured clinical questions allows for efficient literature searches, helps in the review of available evidence and assists in making clear recommendations.
- Determining the clinical questions addressed by the CPG include the following steps:
    - Developing an algorithm to identify the key clinical issues covered in the scope in order to select the questions (broad, generic questions).
    - Translating generic questions into specific, structured clinical questions to define the boundaries of the topic, i.e. by specifying the relevant population, intervention/s (e.g. treatment[s] or diagnostic test[s]), comparator(s) and outcomes measured.
    - Review of the draft list of clinical questions by external experts (healthcare professionals and patients and carers).
- GDG must specify the relative importance of the outcomes. According to GRADE methodology, critical outcomes (rating of 7 to 9) are the essential outcomes for decision making. It is recommended that a maximum of seven outcomes be included in formal analysis for each clinical question.

05.

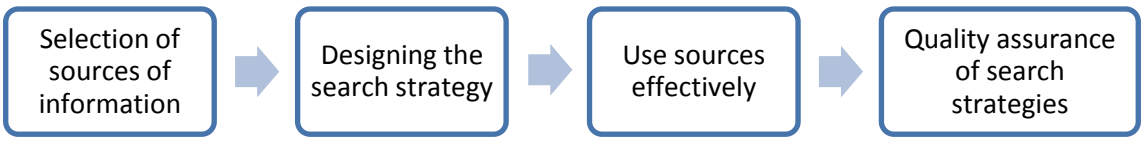
SEARCH AND SELECTION OF SCIENTIFIC EVIDENCE

This chapter focuses on the identification and selection of sources of information, the development of search strategies, database navigation, and how to document the process. Methods for determining the types of studies to be included are also described here.

The systematic identification of evidence is an essential step in CPG development. When appropriately conducted, systematic literature searches should [25]:

- ✓ Identify all or almost all relevant studies and mitigate the risk of omitting significant evidence.
- ✓ Search across multiple bibliographic databases and sources of grey literature.
- ✓ Optimise the balance of sensitivity (the proportion of relevant articles retrieved) and precision (the proportion of irrelevant articles not retrieved).

Performing a systematic literature search involves four major phases:



Literature searches should be accurately recorded for ensuring transparency and reproducibility. They are required to provide enough detail to enable them to be repeated later, tested and updated as necessary.

The GDG should ideally be supported by an information specialist. The role of the information specialist involves [6]:

- ✓ Contributing to the development of clinical questions and their translation into specific searchable questions.

- ✓ Identifying appropriate information sources according to the CPG topic and the type of questions asked.
- ✓ Using or adapting methodological search filters for each question in different databases.
- ✓ Drafting, refining and executing search strategies.
- ✓ Setting up mechanisms to ensure the quality of the searches and the relevance and pertinence of the results.
- ✓ Managing bibliography and acquiring the full text of references.
- ✓ Keeping a log of search results, rationales and strategies.
- ✓ Setting up alerting systems for each clinical question for detecting further evidence relevant for the CPG.

5.1 | Selection of sources of information

In order to ensure adequate coverage of the relevant literature, searches on rare disease conditions should cover at least the core databases and the rare diseases-specific databases listed in Table 4. This proposal comprises major medical databases such as Embase and MEDLINE, used as source of both original studies (clinical trials and observational studies) and systematic reviews, CPG, etc. It is also worthwhile searching Cochrane Library databases, a systematic review-specific resource, and the Health Technology Assessment (HTA) database to access technology assessments. Subject-specific databases should include rare diseases resources such as the Orphanet database, an international data resource dedicated to rare diseases that was co-funded through the European Union's Health Programme and comprises a network of 40 countries [26].

Table 4. Sources of information in CPG development

Core databases
MEDLINE/MEDLINE In-Process
Embase
Cochrane Library:
Cochrane Database of Systematic Reviews (CDSR)
Database of Abstracts of Reviews of Effects (DARE)
Cochrane Central Register of Controlled Trials (CENTRAL)
Health Technology Assessment (HTA) database
Subject-specific databases (non-exhaustive list)
Orphanet
EURORDIS (European Organisation for Rare Diseases)
NORD (National Organisation for Rare Disorders)
RARE-Bestpractices
Gene Reviews
CINAHL (Cumulative Index to Nursing and Allied Health Literature)
PsycINFO

The core databases are predominantly bibliographic databases of peer-reviewed journal articles. They are selected based on a pragmatic strategy for information retrieval proposed by some guideline development organisations (e.g. NICE, GuiaSalud, SIGN).

Other sources that can provide useful information are listed below:

- ✓ Clinical trial registries and repositories to find information about on-going research, for example:
  - European Union Clinical Trials Register
  - ClinicalTrials.gov
  - WHO International Clinical Trials Registry Platform
- ✓ Rare disease patient registries and databases to obtain long-term outcome data in a real-world setting. For information on rare diseases, which is scarce, the following can provide valuable evidence for guideline developers [27]:
  - The Orphanet report on disease registries provides a complete list of the 600 rare disease registers in Europe[28].
  - RD-CONNECT is an integrated platform connecting databases, registries, biobanks and clinical bioinformatics for rare disease research.

  It is worth mentioning that the European Commission advocates the creation of a European Platform for Rare Diseases Registration to cope with the enormous fragmentation of rare disease patient data contained in hundreds of registries across Europe[29].

Information on rare diseases is often scarce and fragmented, and searching for grey literature ensures comprehensive coverage of the topic under consideration. The term 'grey literature' refers to literature protected by intellectual property rights, of sufficient quality to be collected and preserved by library holdings or institutional repositories that it is not controlled by commercial organisations[30]. Grey literature includes materials such as theses and dissertations, working papers, policy statements, technical reports and government documents. A wide variety of methods should be used to search for relevant grey literature and information. A useful approach may be to search grey literature databases (e.g. www.opengrey.eu), websites of relevant organisations and projects (e.g. www.rarebestpractices.eu), and a popular Internet search engine (i.e. scholar.google.com).

In addition to searching bibliographic and grey literature databases, a hand-searching journal is also recommended because not all indexed journal articles can be retrieved from databases. Most hand searches can be performed electronically by scanning journals' electronic tables of contents.

5.2 | Designing the search strategy

Once the clinical question has been framed, key words can be identified for each of its components (e.g. population, intervention, comparator and outcome when using the PICO framework), which will then be translated into subject headings and 'free-text' search terms.

Some databases have lists of controlled vocabulary (e.g. MeSH in MEDLINE and the Cochrane Library, and Emtree in Embase). A term in controlled vocabulary is equivalent to the term itself and all its synonyms. Controlled vocabulary can be used to find all articles on a subject regardless of the word the author has used to describe the topic. Free-text terms are used to complement controlled vocabulary searches. Free-text terms may include, for example, acronyms, synonyms, and brand and generic drug names.

The search strategy consists of a combination of these search terms applying Boolean logical operators such as AND, OR and NOT across the search fields (e.g. title, abstract, keywords).

5.3 | Use sources effectively

Using certain parameters to limit searches can improve precision while barely affecting sensitivity.

Depending on the clinical question, it may be appropriate to limit searches to particular study designs:

- ✓ A clinical question relating to an intervention is usually best answered by a randomised controlled trial (RCT).
- ✓ A clinical question relating to diagnostic test accuracy is usually best answered by a cross-sectional study in which both the index test(s) and the reference standard are performed on the same sample of patients.
- ✓ A clinical question relating to prognosis is best answered using a prospective cohort study.

The use of methodological search filters can help to identify study types. Search filters are pre-tested search strategies that have been designed to retrieve specific types of records and make searching more efficient. The design of search filters should ideally be research-based and presented with data on their performance in finding relevant records. Information on the sensitivity and precision of a search filter is also an important factor for selection of an appropriate filter.

The most comprehensive listing of available search filters can be found on the InterTASC Information Specialists' Sub-Group (ISSG) website, which lists filters by study design, database and interface.

In addition, depending on the clinical question, as well as on practical considerations, CPG developers usually limit the search on the publication language, publication period, search field, etc.

5.4 | Quality assurance of the search strategies

Developing comprehensive search strategies is usually an iterative process in which the information specialist should made efforts to check their quality and accuracy. The following approaches can be used to ensure that the key studies are retrieved [6]:

- ✓ Identify synonyms and related terms to maximise the retrieval of relevant evidence:
    - Search one or two core databases using key terms to identify studies related to the clinical question.
    - Check with GDG members that the search has identified relevant articles. These can be reviewed and additional relevant keywords from within the title, abstract or index may be identified.
- ✓ Run searches with and without certain search terms and assess the differences between the results obtained.
- ✓ Check the bibliographies of the studies included to ensure that all relevant papers have been retrieved by the search strategy used.
- ✓ If relevant papers have not been retrieved by the search strategy, investigate and amend the strategy if appropriate.

Following a web-based survey of experts, the Cochrane Collaboration has published a Peer Review of Electronic Search Strategies (PRESS) evidence-based checklist [31]. This validated checklist is used to evaluate the quality and completeness of an electronic search strategy, and criteria fall into six categories:

- ✓ Translation of the research question.
- ✓ Boolean and proximity operators, which will vary based on the search service.
- ✓ Controlled vocabulary, which is database specific.
- ✓ Text word searching, using free text.
- ✓ Spelling, syntax and line numbers.
- ✓ Limits and filters.

5.5 | Documenting the search process

Thorough documentation of the search process is needed to demonstrate transparency and reproducibility. The following information should be recorded for each search conducted:

- ✓ Details of the question for which the search was conducted.
- ✓ Databases searched (source and provider, e.g. MEDLINE/PubMed).
- ✓ Exact search strategy employed in each database.
- ✓ Any limits applied to the search.
- ✓ Exact date on which the search was conducted.
- ✓ Number of records retrieved from each database.

5.6 | Selection of the scientific evidence

Electronic records of the references retrieved by searches should be stored using a reference management software such as Mendely, EndNote or Zotero. Transferring retrieved citations to a reference manager has the advantages of not only storing and organising the search, but also of providing a relatively straightforward platform for the GDG to review titles, abstracts, and full-text articles.

5.6.1 | Initial screening

The technical team scan titles and abstracts from the retrieved publications in order to exclude publications that are obviously irrelevant to the clinical questions. To increase validity, at least two independent methodologists should be involved in the initial screening.

5.6.2 | Final screening

Final screening is conducted by the members of the GDG who apply the inclusion and exclusion criteria that were agreed for each clinical question. The technical team can prepare a document to support the GDG in this task (see Annex II). Abstracts that do not meet the inclusion criteria are excluded. Any doubts regarding inclusion should be resolved by discussion within the GDG before the results of the study are considered. Studies that fail to meet the inclusion criteria once the full version has been checked are excluded. A list of all excluded studies and those excluded after abstract and full text examination, with the explicit reasons for exclusion concisely stated, should be provided in the CPG.

Key issues

- The search for scientific evidence to develop guidelines implies carrying out systematic and exhaustive searches which require consulting multiple sources of information.
- Search strategies combine key words identified from the components of the PICO, which are translated into subject headings and 'free-text' search terms. Search filters can be used to identify study types.
- Literature searches should be thoroughly documented to ensure transparency and reproducibility.
- The technical team initially screens the references retrieved by titles and abstracts. Final screening is conducted by the GDG applying the inclusion and exclusion criteria that were agreed for each clinical question.

06.

APPRAISAL AND SYNTHESIS OF SCIENTIFIC EVIDENCE

This chapter presents the GRADE approach for assessing the quality of scientific evidence. It describes the process to determine how much confidence can be placed on the effect estimates to support a recommendation. Also, the use of GRADE evidence profiles is suggested for presenting the results of the quality assessment and synthesis of evidence.

Once the evidence to answer the clinical question has been identified, its quality has to be appraised and the results summarised by applying the methodology developed by the GRADE Working Group (Grading of Recommendations Assessment, Development and Evaluation)[11].

The GRADE approach is a structured and transparent method for developing and presenting summaries of evidence, grading its quality, and then transparently interpreting the available evidence to make recommendations. The clinical practice guidelines (CPGs) developed in the European Reference Networks (ERNs) should follow GRADE methodological standards. We include some practical guidance that has been suggested to overcome the challenges that issues specific to rare disease can pose in applying the GRADE approach [22, 27].

The GRADE Working Group provides software - the GRADEpro Guideline Development Tool (GDT) (https://gradepro.org/) - that guides the user through the process of guideline development. It is recommended that guideline development groups (GDGs) use this tool in the ERN context to foster homogeneity between rare disease guidelines produced by different ERN.

More information on the appraisal and synthesis of scientific evidence is provided in a series of articles[32] and an electronic manual [11] published by the GRADE Working Group. Each section of the chapter indicates the articles that describe in more detail the issues addressed in this chapter.

6.1 | Assessing the quality of evidence

The quality of evidence indicates the extent to which we can be confident that an estimate of a treatment effect is adequate to support a particular recommendation[11, 33].

A key issue in the GRADE approach is that the quality of evidence is rated separately for each important outcome across studies (overall quality rating across each outcome). Additionally, an

overall quality rating of the whole body of evidence is assigned when the quality differs across important outcomes (overall quality rating across outcomes). Those outcomes ascertained as being critical for decision making would determine the overall quality of the evidence (see section 3.1.3).

The GRADE approach establishes four categories for rating quality of evidence: high, moderate, low and very low. Table 5 shows what each of the 4 categories represents:

Table 5. GRADE levels of evidence

Quality level	Description
High	It is highly likely that the true effect is similar to the estimated effect
Moderate	It is likely that the true effect is probably close to the estimated effect
Low	The true effect might be markedly different from the estimated effect
Very low	The true effect is probably markedly different from the estimated effect

The GRADE approach for a body of evidence relating to interventions begins by placing studies in one of two categories: randomized controlled trials (RCTs) and observational studies (otherwise known as non-randomized studies). GRADE considers that RCTs begin as high-quality evidence whereas observational studies without important limitations are classed as low quality. In a second stage, GRADE addresses the factors listed in Table 6 in order to either lower or raise the initially allocated level of quality [34-39]. These factors are detailed in section 6.1.2.

Table 6. Factors that may lead to rating down or rating up the quality of evidence

Factors that may lead to rating down the quality of evidence	
Risk of bias **Limitations in study design or execution**	↓1 or 2 quality levels
Inconsistency **Inconsistency in the results of different studies**	↓1 or 2 quality levels
Indirectness **Availability of indirect evidence**	↓1 or 2 quality levels
Imprecision **Imprecision in estimates of effect**	↓1 or 2 quality levels
Publication bias	↓1 or 2 quality levels

Factors that may lead to rating up the quality of evidence*	
Large effect	↑1 or 2 quality levels
Dose-response gradient	↑1 quality level
Plausible residual confounding	↑1 quality level

*For observational studies only

6.1.1 | Additional study design considerations

Case series and case reports are observational studies without controls that should be automatically downgraded to very low quality of evidence.

Expert opinion is not a category of quality of evidence. Expert opinion represents an interpretation of evidence in the context of experts' experiences and knowledge. An expert opinion may be based on the interpretation of studies ranging from uncontrolled case series to randomized controlled clinical trials, thus it is important to describe what type of evidence is being used as the basis for interpretation [11].

Existing systematic reviews are often limited in summarising study limitations across studies. In this case, the assessment should take into consideration the study design and the characteristics of each study included in the review.

6.1.1.1 | Unpublished non-experimental data

Expert-based evidence can be systematically captured from healthcare professionals through structured observation forms to provide clinical observations for questions on therapy and diagnosis. The technical team should help the GDG members identify the evidence underlying their opinions, and judge its quality. This non-experimental and non-comparative data should be collected transparently and systematically, and subjected to the same level of appraisal as other evidence [27].

6.1.2 | Assessing the quality of evidence for questions about interventions

6.1.2.1 | Risk of bias

Limitations in study design or execution (risk of bias) may affect the confidence regarding the estimate of a treatment effect. A risk-of-bias assessment requires the application of the appropriate criteria depending on the study design. Many checklists are available for both clinical trials and observational studies.

- ✓ The Cochrane RoB 2.0 Tool is proposed for assessing the risk of bias of randomised controlled trials. Table 7 summarise the issues addressed by this tool.
- ✓ The ROBINS-I tool (Risk Of Bias In Non-randomized Studies of Interventions) or the Newcastle-Ottawa Scale are proposed for assessing the risk of bias of observational studies [40].
- ✓ The Quality Appraisal Checklist for Case Series of the Institute of Health Economics (IHE) [41] is suggested for the quality appraisal of case series.

To minimise errors and any potential bias in the assessment, two reviewers should independently assess the quality of at least a random selection of studies. Any differences arising from this

assessment should be discussed at a GDG meeting.

Table 7. Key items of the Cochrane RoB 2.0 Tool

Issue	Justification
Random sequence generation (selection bias) **Allocation concealment (selection bias)**	Inadequate generation of a randomised sequence and/or inadequate concealment of allocations prior to assignment, may result in systematic differences between baseline characteristics of the groups that are compared.
Blinding of participants, health care providers (performance bias)	Lack of blinding of health care providers can result in systematic differences in care provided apart from the intervention being evaluated. Lack of blinding of participants may result in systematic differences on how the patients report symptoms.
Incomplete outcome data (attrition bias)	Attrition bias is due to systematic differences between study groups in the number and the way participants are lost from a study. Differences between people who leave a study and those who continue, particularly between study groups, can be the reason for any observed effect and not the intervention itself.
Selective Reporting **(reporting bias)**	Reporting bias arises when only a subset of the original outcomes measured and analysed in a study are fully reported based on the magnitude of the treatment effect or the statistical significance of selected outcomes.

Moving from assessing the risk of bias for each individual study to assessing the risk of bias across a group of studies addressing a particular outcome presents challenges. To deal with this problem, GRADE suggest some principles that can be useful for assessing the risk of bias of an entire body of evidence on a specific outcome[38]. For instance, in deciding whether to rate down for risk of bias, GDG should consider including only the studies with a lower risk of bias rather than taking the average across studies (e.g. when there are some studies with no serious, some with serious and some with very serious limitations).

6.1.2.2 | Inconsistency

Inconsistency refers to the heterogeneity or variability in the estimates of treatment effect across studies for each outcome of interest. GRADE suggests rating down the quality of evidence if large inconsistency (heterogeneity) in study results remains after exploration of a priori hypotheses that might explain heterogeneity, for example, differences in the population (e.g. patients vary in their baseline risk), interventions (e.g. doses, comparison interventions), outcomes (e.g. duration of follow-up) or study design [36].

The following criteria may help decide whether heterogeneity exists [36]:

✓ Point estimates vary widely across studies.

- ✓ Confidence intervals (generally depicted graphically in meta-analysis using horizontal lines) show minimal or no overlap at visual inspection.
- ✓ The statistical test for heterogeneity, which examines in meta-analysis the null hypothesis that all studies are evaluating the same effect, shows a low P-value (usually under 0.10) [21].
- ✓ The I^2, which indicates the percentage of variance in a meta-analysis that is attributable to study heterogeneity, is large (e.g. values of 30% to 60% may represent moderate heterogeneity and values of 50% to 90% may represent substantial heterogeneity) [21].

6.1.2.3 | Indirectness

Direct evidence comes from studies directly addressing the intervention and population of interest which report outcomes important to patients[35].

Evidence can be indirect when:

- ✓ The population tested in the studies differ from the population of interest defined in the PICO question (often referred to as applicability).

  For rare diseases that do not have clear diagnostic criteria, GDG may include extrapolation of data from a population affected by a more common disease that shares certain features with the rare disease [22, 27]. GDG members should judge to what extent the population tested differs from the population of interest and rate down accordingly.

- ✓ The intervention tested differs from the intervention of interest.

  For some interventions, particularly complex interventions, differences in the contextual factors in which the interventions will be offered (e.g. local resources, expertise of the staff) may prevent interventions from being fully implemented, and this requires judgements on indirectness.

- ✓ Differences between the desired outcomes, prioritised by the GDG (see section 4.3), and the outcomes reported by the studies.

  The use of surrogate outcomes (biomarkers) in place of the patient-important outcomes of interest requires rating down the quality of evidence. Considerations on the ability of the surrogate outcome to predict a beneficial effect can be helpful in making a decision about indirectness.

- ✓ Absence of data from head-to-head studies of the options of interest. For example, when data from studies comparing drug A to placebo and drug B to placebo are available, but there is no direct comparisons of the effectiveness of A against B. Evidence is lower quality if comparisons are indirect.

6.1.2.4 | Imprecision

Rating imprecision includes an assessment of both the 95% confidence interval (CI) and the sample size for the body of evidence. In general, the CIs to consider are those around the absolute, rather than the relative effect. GRADE suggests rating down for imprecision if [34, 42]:

- ✓ The CI excludes the clinical decision threshold between recommending and not recommending an intervention; or
- ✓ The sample size is not large enough to reach a sufficient information size. To inform this decision, one can calculate the number of patients required for an adequately powered individual trial (termed the "optimal information size").

6.1.2.5 | Publication bias

There is a tendency for authors to publish studies with significant results. Publication bias occurs when the results of published studies are systematically different from the results of unpublished studies. Although the risk of publication bias may be higher for reviews that are based on small randomized controlled trials, RCTs including large numbers of patients are not immune. As a general rule, GDG should consider rating down for suspicion of publication bias when [39]:

- ✓ The evidence consists of a number of small studies, most of them sponsored by the industry.
- ✓ A systematic review of a novel therapy failed to conduct a comprehensive search (including a search for unpublished studies).

There are several approaches to using available data to provide insight into the likelihood of publication bias that may be useful (e.g. visual inspection of a funnel plot, "trim and fill" test) but all of them have limitations. GRADE recognises the difficulty in assessing the risk of publication bias and suggest rating down a maximum of one level.

6.1.2.6 | Rating up the quality of evidence

There are three factors that might increase the quality of evidence of observational studies[37]:

- ✓ **Large magnitude of effect**: The confidence in the effect estimate may increase when the effect size is large or very large (see table 8).

Table 8. Rating up the quality of evidence (magnitude of effect)

Effect size	Definition	Quality of evidence
Large	RR > 2 o < 0,5 (effect estimate from direct evidence with no plausible confounders)	↑1 quality level
Very large	RR > 5 o <0,2 (and no serious problems with risk of bias or precision [sufficiently narrow CI])	↑2 quality level

RR: risk ratio CI: confidence interval

- ✓ **Dose-response gradient**: The presence of a dose-response gradient supports the judgement of a cause effect relationship, thus increasing confidence in the effect estimates.
- ✓ **Effect of plausible residual confounding**: The term "confounding" refers to a situation when one finds a spurious association or misses a true association between an exposure variable and an outcome variable as a result of a third factor or group of factors referred to as confounding variable(s). Rigorous observational studies apply an adjusted analysis to control for potential confounders. In cases where control of all plausible confounders is unaccounted, and this may result in an underestimate of the apparent treatment effect, the level of evidence can be increased.

6.1.3 | Assessing the overall quality of evidence

GRADE requires an overall rating of confidence in estimates of effect for each important or critical outcome to be made. GDGs will subsequently make an overall rating of confidence in effect estimates across all outcomes based on those outcomes they consider critical to the recommendation [43].

6.1.3.1 | Assessing the overall quality of a single outcome

As mentioned before, the GRADE approach suggests five reasons for rating down the confidence in effect estimates (risk of bias, imprecision, inconsistency, indirectness, and publication bias) and three reasons for rating up the confidence in effect estimates (a large magnitude of effect, a dose-response gradient, and the presence of plausible residual confounding). The levels of evidence quality in GRADE are 'high', 'moderate', 'low', and 'very low'.

These four discrete categories for rating the quality of evidence up or down add information and transparency for guideline users. However, the quality of evidence represents not discrete categories but a continuum from minimal limitations to very serious limitations. That is why GRADE states that contextual decisions are necessary when confidence is near the threshold between categories. In such instances, it is particularly desirable that guideline developers make their judgements explicit to guideline users in the GRADE evidence profiles, when rating the quality of evidence [43].

6.1.3.2 | Assessing the overall quality across outcomes

GDGs must determine the overall quality of the evidence across all the critical outcomes for each recommendation. Because quality of evidence is rated separately for each outcome, the quality frequently differs across outcomes[43].

- ✓ If the quality of the evidence is the same for all critical outcomes, then this is the level of quality that applies to all of the evidence supporting the answer to the key question.
- ✓ If the quality of the evidence differs across critical outcomes, the overall confidence in effect estimates cannot be higher than the lowest level of confidence in the effect estimates for an individual outcome.

Therefore, the lowest quality of the evidence for any single critical outcome determines the overall quality of the evidence.

6.2 | Development of GRADE evidence profiles

The technical team elaborate GRADE evidence profile tables for presenting the results of the quality assessment of the body of evidence supporting a recommendation.

GRADE evidence profile tables include, for each critical and each important outcome, the assessment of each factor that determines the quality of evidence (risk of bias, imprecision, etc.), and a summary of findings.

The technical team will present the GRADE evidence profiles to the GDG for discussion and validation. Whenever necessary, the GDG should explicitly indicate in footnotes their judgements about rating down the quality of evidence in concise and clear text.

GRADEpro GDT software can be used for creating evidence summaries using the GRADE approach (https://gradepro.org).

6.3 | Questions about diagnosis

According to the GRADE approach, the best study design for answering questions about diagnosis is test-treat RCT, in which subjects are randomised to receive the diagnostic test under investigation or the reference standard and that measure outcomes important to patients. When RCTs are available, GRADE recommends applying the approach for questions about intervention [44].

However, when data from RCTs is lacking, studies of diagnostic test accuracy are used as the basis for clinical decisions. Studies of the accuracy of a diagnostic test (or strategy) consider the ability of the test to predict the presence or absence of disease. Thus, the GDG should infer from data on accuracy that using a test improves outcomes that are important to patients [45].

The most valid study design for assessing the accuracy of diagnostic tests is a cross-sectional or cohort study that compares the results of the test under investigation (index test) with an appropriate reference standard in patients with diagnostic uncertainty. The participants undergo both the index test and a reference standard test within a very short time period. These studies start with a high-quality rating, but can be rated down one or two levels for risk of bias, indirectness, imprecision, and publication bias (from high to moderate, low or very low) [46, 47].

GRADE suggests the use of the QUADAS ('Quality Assessment of Diagnostic Accuracy Studies') tool (Annex X) to assess the risk of bias of studies of diagnostic accuracy [48].

6.4 | Qualitative evidence

In the context of qualitative evidence synthesis, the term quality of evidence is used to describe the extent to which one can be confident that the review finding is a reasonable representation of the phenomenon of interest. A review finding is "an analytic output from a qualitative evidence synthesis that, based on data from primary studies, describes a phenomenon or an aspect of a phenomenon"[49].

The GRADE-CERQual approach ('Confidence in the Evidence from Reviews of Qualitative research') provides a framework for assessing the confidence in findings from qualitative evidence synthesis [50]. Each review finding is assessed in terms of four components listed below:

- ✓ Methodological limitations
- ✓ Coherence
- ✓ Adequacy of data
- ✓ Relevance

Initially, the methodological limitations [51] of each study contributing to a finding are assessed using a critical appraisal checklist for qualitative studies (e.g. the Critical Appraisal Skills Programme checklist for qualitative research) [52], along with coherence, which is an assessment of how clear and cogent the fit is between the data from the primary studies and a review finding [53].

Relevance [54] assesses the extent to which the body of data from the primary studies supporting a review finding is applicable to the context specified in the clinical question, and adequacy of data is defined as the degree of richness and the quantity of data supporting each review finding [55].

Quality level	Description
High	It is highly likely that the review finding is a reasonable representation of the phenomenon of interest
Moderate	It is likely that the review finding is a reasonable representation of the phenomenon of interest
Low	It is possible that the review finding is a reasonable representation of the phenomenon of interest
Very low	It is not clear whether the review finding is a reasonable representation of the phenomenon of interest

The CERQual approach sets four categories of level of confidence in a review finding[50]:

This approach assists the GDGs in the use of findings from qualitative evidence syntheses to make judgements about the implementability and acceptability of interventions, which are factors influencing the strength and direction of recommendations (see section 8.2).

When qualitative evidence is lacking, qualitative research methods can be used to generate evidence on patient values and preferences, equity, acceptability, feasibility, and implementability. However, this approach will take additional time and resources as well as the incorporation of a qualitative researcher in the GDG [27].

Key issues

- The quality of evidence is rated separately for each important outcome across the studies. Also, the overall quality across outcomes is determined to inform the recommendations.
- Quality as used in GRADE is more than risk of bias because may also be compromised by other factors. Such factors are subjected to particular specifications according to the type of clinical question and study design.
- The development of clinical questions should be based on a systematic review of the literature and its results should be presented in evidence profiles, which are tables containing the calculated effect estimate for each outcome along with their corresponding quality of evidence.

07.

CONSIDERING RESOURCE USE AND RATING THE QUALITY OF ECONOMIC EVIDENCE

This chapter provides information about the steps to take to incorporate resource use considerations when making recommendations in CPGs. The specific methods for the synthesis of economic evidence are detailed, along with the key elements for developing de novo economic evaluations in the context of rare diseases.

7.1 | Steps for considering resource use in Clinical Practice Guideline (CPG) development

7.1.1 | Guideline Development Group (GDG) considerations

It is important to take into account the economic perspective when reviewing and interpreting economic evidence or when deciding whether to conduct a new economic analysis. Here are some important aspects to consider in relation to the working group [6]:

- ✓ It is recommended that a health economist or a methodologist with training in health economics is included in the working group.
- ✓ It is recommended that working group members have basic training in health economics. Therefore, an initial training in which the main concepts on health economics and key aspects for resource use consideration should be organised.
- ✓ The support or advice from an external health economist may often be necessary.

7.1.2 | Anticipating the impact of resource use in making recommendations

Anticipating the impact of resource use on the recommendations is important for determining the steps to be taken by the working group. Since not all questions lead to recommendations in which

the use of resources is a key aspect, i.e. where the use of resources will not be a defining factor of the recommendation, for every question in the CPG, the working group should assess the following [6]:

- ✓ What influence resource use could have on the future recommendations according to their expertise in the topic
- ✓ Whether the working group has sufficient information to determine said influence and the most appropriate way to obtain this information to contribute to future recommendations.
- ✓ The depth with which it is intended to analyse the economic information to incorporate in future recommendations. For example, it can be addressed through a systematic review if there is sufficient evidence or a *de novo* economic analysis.

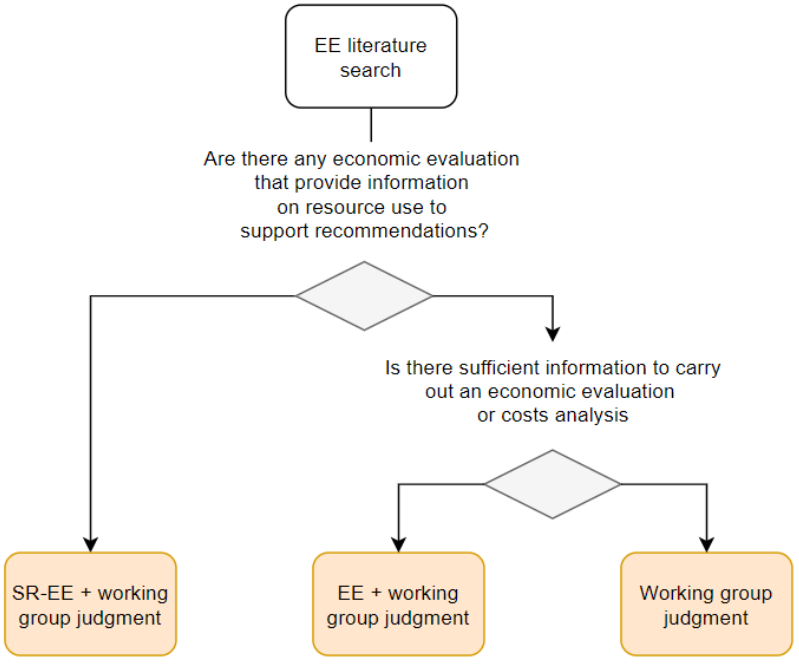
After that, the working group should be able to establish whether resource use is a relevant factor in potential recommendations and how to proceed in incorporating economic information.

It is important to note that, although this analysis must be carried out at the beginning of each question, this is a dynamic process and new factors can influence and modify decisions made previously. For example, new evidence of effectiveness or changes in drug patents.

Once the working group has identified the questions in which resource use is a relevant aspect, a literature search and review of existing economic evidence should be carried out. According to the results of this literature search, the approach may vary (Figure 2):

- ✓ If there is sufficient evidence to inform the recommendations from an economic standpoint, a Systematic Review of Economic Evaluations (SR-EE) will be the preferred option.
- ✓ If there is not enough evidence to inform the recommendations, the decision may be taken to carry out a *de novo* model-based Economic Evaluation (EE). However, this should be a carefully considered decision, which must be made jointly between the working group and the health economist.

Figure 2. Workflow for incorporating resource use considerations into CPG



7.1.3 | Economic Evaluation basics

Two types of EE can be distinguished: full EEs and partial EEs. Table 9 provides a summary.

Full EEs are defined as studies in which two or more alternative interventions are compared, and both costs and effects (consequences and benefits) of at least two alternatives are taken into account. In a partial EE, these requirements (comparison of two alternatives and measurement of both costs and consequences) are not met. Each of these approaches has specific objectives. Although partial EEs are not recommended for analytical purposes, these studies might be considered when there is a lack of knowledge on a specific topic. For example, when an SR-EEs is performed to inform about resource use in CPG recommendations [56].

Table 9. Types of economic evaluations

<table>
<tr><th colspan="4">Are costs and results examined?</th></tr>
<tr><th rowspan="4">Are at least two alternatives examined?</th><th rowspan="2">No</th><th colspan="2">No</th><th>Yes</th></tr>
<tr><td colspan="2">Partial EEs</td><td>Partial EEs</td></tr>
<tr><td>Outcome description</td><td>Cost description</td><td>Cost-outcome description</td></tr>
<tr><td rowspan="2">Yes</td><td colspan="2">Partial EEs</td><td>Full EEs</td></tr>
<tr><td>Efficacy of effectiveness evaluation</td><td>Cost analysis</td><td>Cost-minimisation analysis (CMA)
Cost-effectiveness analysis (CEA)
Cost–utility analysis (CUA)
Cost–benefit analysis (CBA)
Cost–consequence analysis</td></tr>
</table>

Within full EE, five types can be distinguished[8]:

- ✓ Cost-minimisation analysis: a determination of the least costly among alternative interventions that are assumed to produce equivalent outcomes.
- ✓ Cost-effectiveness analysis: a comparison of costs in monetary units with outcomes in quantitative non-monetary units. For example, reduced mortality, years of life gained, conditions measured by biomarkers, etc.
- ✓ Cost–utility analysis: a form of cost-effectiveness analysis that compares costs in monetary units with outcomes in terms of their utility reported by patients, measured in QALYs
- ✓ Cost–benefit analysis: a comparison of costs and benefits, both of which are quantified in monetary terms.
- ✓ Cost–consequence analysis: a form of cost-effectiveness analysis that presents costs and outcomes in discrete categories, without aggregating or weighting them.

Additionally, despite not being considered full EEs, Budget impact analysis can be useful analyses for informing CPG recommendations in the field of RD, since they estimate the expected changes in expenditure in a healthcare system or setting after a new intervention has been implemented (e.g. specific orphan drugs).

7.2 | Using existing evidence to prepare a Systematic Review of Economic Evidence

Multiple resources and recommendations to address the different phases of the SR-EE are presented below[56]:

7.2.1 | Relevant data sources

The main sources for identifying full EEs are general databases, such as PubMed/Medline, Embase and Web of Science. There are also specific databases where it is possible to find EEs:

- ✓ Centre for Reviews and Dissemination in the University of York provides access to the NHS Economic Evaluation Database (NHS EED), which can be used for searches of full EEs up to March 2015[57].
- ✓ Repositories or webpages from Health Technology Assessment (HTA) agencies are also relevant information sources for finding EEs, HTA reports or CPGs that include complete economic evaluations or cost analyses that are accessible for consultation.
- ✓ Other specific sources related to rare diseases may contain EE or resource use information (e.g. Orphanet[26], RARE-BestPractices [58], etc.).

7.2.2 | Development of search strategies

It is not always necessary to develop new search strategies for every new SR-EE. It is recommended to use existing validated search filters. The InterTASC Information Specialists' Sub-Group (ISSG) provides a list of such filters[59].

7.2.3 | Study selection and data extraction

First, the records need to be screened on review title and abstract. Subsequently, the full text records must be screened for compliance with eligibility criteria. An extraction form should be developed for capturing the essential information from the EEs reviewed. Overall, the main items to extract from an EE are:

- ✓ General study characteristics (author, year of publication, objective, intervention and control)
- ✓ Type of EE and perspective
- ✓ Details on EE methods (resource use in both natural and monetary units, costs, effects, outcome valuation methods)
- ✓ Results (incremental cost-effectiveness ratios)

Moreover, for model-based EEs, special attention needs to be paid to:

- ✓ Model structure (e.g. model structure, cycles or iterations, time horizons, etc.)
- ✓ Key assumptions

- ✓ Input data values
- ✓ Uncertainty analyses (sensitivity analysis)

In addition to the extraction, the quality of the studies should be evaluated according to the degree of certainty about each of the resource use estimates that have been identified[60]. The methods used and the authors' assumptions should be verified, and the adequate reporting of results should be assessed. Some examples of checklists for assessing quality and reporting in EEs are provided:

- ✓ British Medical Journal Checklist[61]
- ✓ Consensus on Health Economic Criteria (CHEC)-extended list[62]
- ✓ *Philips* Checklist[63]
- ✓ Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement[64]

7.2.4 | Reporting results

Relevant findings of EEs that have been reviewed should be presented in such a way that makes the reader understand the results and major conclusions. Cost-effectiveness planes or rankings for cost per QALY from different studies, etc. are useful elements for presenting this information.

In order to make comparisons, different currencies reported within the EEs should be converted to a one common currency and the same year should be used as a reference. The Campbell and Cochrane Economics Methods Group (CCEMG) and the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) developed a free web-based tool which automatically adjusts estimates for costs and price year which automatically adjusts estimates for costs and price year [65].

7.2.5 | Discussion and interpretation

More specifically, the following factors can be discussed when using the main findings of the SR-EE to formulate CPG recommendations in the ERN context:

- ✓ Quality of the EEs, analysis of the assumptions made by the authors or identification of possible risks of bias. In accordance with the GRADE methodology, rating the confidence in effect estimates for important outcomes on resource use and its valuation in terms of costs for the specific setting for which recommendations are being made are key steps. The evidence profile tables are proposed as a way to summarise this information [60].
- ✓ Whether the findings of the study show that the new intervention is cost-effective according to the threshold values being used within the CPG development context.
- ✓ Variability and uncertainty of studies should be discussed on whether sensitivity analyses provide robust or variable results depending on contextual parameters or values.
- ✓ Balance between health benefits, side effects, and risks.
- ✓ Whether the EEs results are generalizable or transferable to the ERN context.
- ✓ Whether the incorporation of the EEs into the ERN context poses any implementation problems. For example, if a new intervention may have a large budget impact.

7.3 | Key aspects for conducting a de novo Economic Evaluation

Model-based EEs enable the integration and extrapolation of the results of completed clinical trials, using information from hospital records, databases, expert opinions, medical record reviews, other epidemiological studies, etc. Therefore, they represent a feasible alternative when there is a lack of information or economic evidence in the literature.

When developing a de novo model-based economic evaluation, the following key aspects should be kept in mind. The working group, with the help of the health economist, should make relevant decisions on how to approach each of these elements[66]:

7.3.1 | Perspective

The perspective is the point of view adopted for the evaluation and determines which types of costs and health benefits are to be included in an EE. Typical viewpoints are those of the patient, hospital/clinic, healthcare system or society. Depending on the perspective chosen, the EE may include different resources employed and costs or health outcomes from different stakeholders, so that the results may differ.

- ✓ Rare diseases generally carry a high societal burden, due to the high cost they represent in terms of care, loss of work productivity and quality of life for the patient and their family/caregivers. Therefore, it will be desirable to use the broader perspective (society), whenever possible according to the CPG times and availability of information.

7.3.2 | Costs

The costs associated with rare diseases must also be identified, quantified and valued using monetary units. Cost identification will depend upon the relevant perspective chosen by the working group. In this context, in addition to assessing direct healthcare costs, special emphasis needs to be placed on those costs that are directly borne by families (e.g. out-of pocket medicines, transportation) and society (e.g. productivity losses), both current and future. Cost classification is detailed in Table 10.

Table 10. Cost classification to include in EEs

Type of cost	Examples	Perspectives in which this type of costs is included
Direct healthcare costs	Intervention costs, diagnostic costs, facilities and equipment including hospitalisation and staff	Usually included in the healthcare system or hospital/clinic perspective
Direct non-healthcare costs	Transportation costs, time off work/school for appointments	Usually paid by the patients, often included in the patient and societal perspective
Indirect costs	Lost work/academic productivity by patient or caregiver, lost leisure time	Usually paid by the patients, often included in the patient and societal perspective
Intangible costs	Pain, suffering, grief	Generally not explicitly included, however these costs are usually considered as quality-of-life dimensions in the cost-utility analyses

In order to quantify and assign a monetary value for resources, it is recommendable to use public prices, administrative databases, and official publications, rates applied to benefit contracts or accounting information from centres. Notwithstanding, unit costs may be collected from previously published studies or other sources[6].

- ✓ As previously mentioned, CCEMG-EPPI-Centre Cost Converter tool is a practical resource to convert cost information valued in different currencies or price years[65]. In this case it is helpful for the conversion of the monetary values extracted from literature or administrative databases to be included as model parameters.

7.3.3 | Health outcomes

The EE should reflect to what extent the new intervention modifies the course of the RD or condition analysed, either by increasing life expectancy or improving quality of life by reducing symptoms, improving patient mobility or capabilities or avoiding the side effects of other treatment approaches.

- ✓ In order to capture these improvements Quality-Adjusted Life Years (QALYs) are a fundamental measure for health outcomes, given the high impact that rare diseases have on the quality of life of patients.

There are a few options to obtain utility values for calculating QALYs gained when this information is not available in patient records:

- ✓ The CEA Registry provides a database in which utility weight records from a wide range of EEs have been extracted [67].
- ✓ The InterTASC Information Specialists' Sub-Group (ISSG) provides a list of such filters to identify health state utility values [59].

7.3.4 | Time horizon and discount rate for cost and health outcomes

- ✓ The time horizon applied should be in accordance with the natural course of the disease, which will cover the life expectancy of the patient.
- ✓ It is also recommended to apply both to the costs and health outcomes an annual discount rate of 3% and include other values (0% to 5%) using the sensitivity analysis.

7.3.5 | Modelling

The choice of the most suitable model will depend on the type of problem studied and the availability of data to carry it out. Modelling approaches commonly used are listed below and manuals for their development are provided:

- ✓ Decision trees [68]
- ✓ Markov models [69]
- ✓ Discrete event simulation models [70]

7.3.6 | Presentation of results

Results should be presented in a detailed and transparent way. In this sense, incremental cost-effectiveness (or utility) ratios (ICER) should be calculated for all pairs of alternatives evaluated. ICER links the difference in costs with the difference in health outcomes for two compared options. The following mathematical expression is used to obtain the ICER:

$$ICER = \frac{Cn - Co}{QALYn - QALYo}$$

- ✓ To ensure transparency, costs for the two alternatives evaluated and incremental values must be reported, as well as the outcomes. Finally, the ICER must be presented.
- ✓ It is recommended to present the ICER graphically using a cost-effectiveness plane.

7.3.7 | Sensitivity analysis

EEs in the context of rare diseases are associated with greater uncertainty than those aimed at prevalent diseases. For example, effectiveness parameters may have been collected from clinical trials made up of a low number of patients. Hence, there may be uncertainty regarding multiple parameters and variables, such as long-term effects or complications. Some recommendations are listed below:

- ✓ The EE should include at least univariate methods to count for uncertainty. If possible, multivariate methods are also recommended.
- ✓ Probabilistic sensitivity analyses that handle uncertainty and provide confidence intervals for the ICER are a desirable option when the data included in the model come from patient records, such that it is possible to draw statistical distributions.

When evaluating high-cost interventions, the EE should include an analysis of acceptability curves, according to the willingness to pay thresholds considered in the context of application.

Key issues

Counting for the economic perspective is an important factor in developing CPG. To carry it out, the role of the health economist in the GDG is decisive for:

- ✓ assist the GDG to anticipate the impact of the use of resources in making recommendations,
- ✓ review evidence coming from economic evaluations,
- ✓ develop economic evaluations or de novo cost analysis, if necessary.

In the case of conducting a *de novo* economic evaluation to inform the recommendations, there are particular aspects of rare diseases that must be considered in the following stages:

- ✓ Perspective
- ✓ Costs to include in the analysis
- ✓ Health outcomes
- ✓ Time horizon and discount rate
- ✓ Modeling
- ✓ Results presentation
- ✓ Sensitivity analysis

08.

DEVELOPING RECOMENDATIONS

This chapter provides information about the factors that influence the decisions when moving from evidence to developing recommendations, and how they affect the strength and direction of the recommendations. It also explains how to address special situations when developing recommendations. In addition, some key issues are given to wording and presenting recommendations.

The GRADE Working Group has developed Evidence to Decision (EtD) frameworks to assist GDG in considering all important criteria to inform decisions in the context of clinical recommendations. These frameworks also inform users about the judgements that were made and the evidence supporting those judgements[71].

As mentioned before, the GRADEpro Guideline Development Tool (GDT) (https://gradepro.org/) guides the user through the process of guideline development, which also includes the development and preparation of Evidence to Decision Frameworks.

EtD frameworks are prepared by the technical team for use by GDGs.

8.1 | Moving from evidence to developing recommendations

EtD frameworks are structured in three main sections: question formulation, criteria assessment, and conclusions.

- ✓ The question section includes details of the clinical question in a structured format (see section 4.2), the perspective from which the options to address the question are considered (e.g. health system perspective), relevant subgroups, key background information for understanding the question, and why a recommendation is needed [71].
- ✓ The next section comprises the factors (criteria) that GDGs should consider for making a recommendation. Each criterion must be completed by the judgments made by the GDG, and the research evidence and additional considerations used to inform each judgement. GDGs should explicitly state the perspective that they are taking (individual patient perspective or population perspective), which is especially important for determining which costs (resource use) to consider. These factors influence the direction and strength of recommendations.

The development of these frameworks will require finding and systematically reviewing all relevant evidence on the issues to be addressed (e.g. resources requirements, acceptability or feasibility). When evidence is lacking or resources to conduct systematic reviews are limited, EtD frameworks should explicitly indicate what, if any, evidence was used to inform each judgement and, if no research evidence was available, this fact should be clearly indicated, together with the considerations that were made [71].

The table below presents the criteria assessed in EtD frameworks (Table 11)[71].

Table 11. Criteria included in Evidence to Decision Frameworks

CRITERION	JUDGEMENTS
Is the problem a priority?	The likelihood of being a priority is greater when the consequences of the problem are serious (e.g. high rates of mortality or disability).
How substantial are the desirable anticipated effects?	Substantial desirable effects increase the probability of making a recommendation favourable to the option being considered. GDGs have to make judgements for each outcome for which there is a desirable effect, taking into account the value that patients place on each outcome. If the quality of the evidence is low or very low, or evidence/research is lacking, it is not possible to judge to which extent the desirable effects of the intervention are substantial.
How substantial are the undesirable anticipated effects?	Undesirable effects (adverse effects) decrease the probability of making a recommendation favourable to the intervention being considered. Judgements for this criterion are the same as for desirable effects.
What is the overall certainty of the evidence of effects?	The lower the certainty of the evidence supporting the effects (also referred to as quality of evidence), the less likely it will be to make a recommendation in favour of or against the intervention.
Is there important uncertainty about or variability in how much people value the main outcomes?	If there is significant uncertainty or variability in how much patients value each of the main outcomes and, therefore, it is not possible to know with certainty what decisions well-informed patients would make. In this case, making a strong recommendation will not be justified.
Does the balance between desirable and undesirable effects favour the intervention or the comparison?	The assessment of this criterion requires judgments regarding each of the four preceding criteria. Sometimes one criterion may have a heavier weight than the others. In such cases, the rationale for such inference should be made explicit to guideline users. Also, GDGs may consider the extent to which patients are willing to accept the possibility of adverse effects when they have the probability of obtaining favourable clinical outcomes.
How large are the resource requirements (costs)?	Interventions with larger resource requirements (cost) are less likely to be recommended. When costs are important for decision making, formal economic modelling may be needed. Additional guidance on the consideration of resource use is given in chapter 8.

What is the certainty of the evidence of resource requirements (costs)?	GDGs should identify resource use items that may differ between the options being compared and find economic evidence for such differences. The confidence in effect estimates for each important or critical economic outcome should be appraised, using the same criteria as for health outcomes. GDGs should value resource use in terms of costs for the specific setting for which recommendations are being made. Additional guidance on the consideration of resource use is given in chapter 8.
Does the cost-effectiveness of the intervention favour the intervention or the comparison?	The assessment of this criterion requires judgments regarding each of the six preceding criteria. The intervention being compared is cost-effective when costs are lower and effects are better than the control intervention. However, an intervention that is more expensive but results in higher outcomes in comparison with an existing intervention can also be considered cost-effective. This depends on the threshold values being used. Additional guidance on the consideration of resource use is given in chapter 8.
What would be the impact on health equity?	By explicitly examining the potential impact of the intervention on health equity, GDGs may discover differential effects of the intervention on disadvantaged populations (e.g. health equity in relation to specific characteristics: economic status, employment or occupation, education, place of residence, gender or ethnicity) [72]. GDGs may decide to accompany a general recommendation with subgroup recommendations to promote health equity or even make a separate recommendation for a specific disadvantaged population when evidence of meaningfully different effects for a subgroup is identified [72].
Is the intervention acceptable to key stakeholders?	The less acceptable an intervention is to key stakeholders, the less likely it is that it should be recommended, or if it is recommended, the more likely it is that an implementation strategy will be needed to address concerns regarding acceptability. GDGs should collect information about acceptability based on input from key stakeholders or evidence from the literature.
Is the intervention feasible to implement?	Feasibility determines of how easy it is to carry out the intervention, put it into practice or policy, or stop an existing intervention. The less feasible an intervention, the less likely it should be recommended. Interventions with low feasibility (or high barriers to implementation) may lead to a weak or conditional recommendation.

✓ The conclusions section, based on the judgements made for all of the criteria, include:

- A **summary of the judgements** made for all the criteria.
- The **type of recommendation**, i.e. strength (strong or weak) and direction (in favour of or against the intervention).
- The **recommendation** in concise, clear and actionable text. The wording of recommendations is described in more detail in section 8.3.1.
- The **justification for the recommendation**, explicitly stating the key criteria used in

making the recommendation.

- Any **subgroup considerations** that the GDG took into account when making the decision.
- Key **implementation considerations**, including strategies to address any identified barriers in relation to the acceptability and feasibility of the intervention (see Handbook #12: Implementation and Evaluation of the Uptake of CPGs and CDSTs for Rare Diseases).
- Suggestions for **monitoring and evaluation** if the intervention is implemented, including any important indicators that should be monitored and any needs for further evaluation (see Handbook #10: Methodology for the elaboration of Quality Measures for Rare Diseases).
- **Research priorities** to address any important uncertainties or gaps identified in the research evidence that informed the judgements of the GDG.

The GRADE Working Group has also developed tailored EtD frameworks for making evidence-informed decisions and recommendations on diagnostic and screening tests, coverage, and health system and public health options. Additional information can be consulted in the GRADE Handbook [11] and publications by the GRADE Working Group [45, 73, 74].

8.2 | Strength and direction of recommendations

The GRADE approach classifies the recommendations according to their direction, in favour of or against the use of an intervention and, depending on their strength, into strong and weak. The strength of a recommendation expresses the degree to which the GDG is confident in the balance between the desirable and undesirable consequences of implementing the recommendation [75].

✓ Strong recommendations communicate the message that the GDG is very certain about this balance, i.e. the desirable effects of adherence to the recommendation clearly outweigh the undesirable effects or vice versa. The harm-benefit balance of an intervention is rarely certain, making strong recommendations uncommon. GDGs need to be cautious when considering making strong recommendations on the basis of evidence whose quality is low or very low.

✓ Weak recommendations are made when a GDG is less confident in the balance between the benefits and harms or disadvantages of its implementation. GRADE offers alternative labels for a weak recommendation, including conditional, discretionary, and qualified.

Table 12 provides information on the implications of the recommendations for patients, healthcare professionals, and policy–makers [75].

Table 12. Implications of recommendations to different audiences

AUDIENCE	STRONG RECOMMENDATION	WEAK RECOMMENDATION
Patients	Most individuals in this situation would choose the recommended course of action; only a small proportion would not. Formal decision aids are not likely to be needed to help individuals make decisions consistent with their values and preferences.	Most individuals would choose the recommended course of action, but many would not. Formal decision aids may be useful in the process of shared decision making to ensure that the patient's choice reflects his or her individual values and preferences.
Healthcare professionals	Most individuals should receive the intervention.	Different choices will be appropriate for individual patients, who will require assistance in understanding the implications of the choices they are making.
Policy makers	The recommendation can be adopted as policy in most situations.	Policy-making will require substantial debate and involvement of various stakeholders.

8.3 | Formulating recommendations

Draft recommendations can be prepared either by the technical team before the GDG meet to formulate recommendations, or during the meeting by the GDG itself. To perform this task, the GDG reviews and discusses the GRADE evidence profiles presented by a representative of the technical team. After that, the GDG considers the relevant criteria included in EtD frameworks as discussed above.

Formulating recommendations is an iterative process; the recommendations are likely to be revised several times before the wording is finalised.

8.3.1 | Wording of recommendations

Recommendations should be worded as clear and actionable statements[75]. This means a clear description of the population (or groups) for which the recommendation is intended, the recommended intervention and the alternative options considered. It may also include 'remarks' or 'key considerations' to clarify the 'conditions' needed to balance desirable and undesirable consequences of adopting the recommendation. Wherever possible, language should be consistent across all recommendations in a GPC, which should be written in the active voice.

In order to state the strength of the recommendation, GRADE suggests the use of specific terminology in addition to labels like "strong/weak recommendation". There are different approaches, but we suggest using the term "we recommend" for strong recommendations and the term "we suggest" for weak recommendations. [75].

8.3.2 | Recommendations to use only in research

GDG may face decisions about promising interventions for which evidence of effectiveness is either lacking or insufficient to support their use in clinical practice. In such situation, GDG may recommend the use of an intervention only in the context of research. Only-in-research recommendations will be appropriate when the following conditions are met[75]:

- ✓ There is insufficient evidence supporting an intervention for a GDG to recommend its use.
- ✓ Further research has great potential for providing benefits to patients in a cost-effective manner.
- ✓ The necessary research can realistically be set up or is already planned.

The research recommendations should detail which patient-important outcomes measures to include, and other relevant issues of the research question (patients' characteristics, intervention and comparator[s], study design, etc.).

8.3.3 | Not making any recommendation

When evidence on the effectiveness of an intervention is scarce or desirable and undesirable consequences are closely balanced, GDG may consider not making a recommendation, or issuing a recommendation based on opinion.

In these instances, the GRADE working group encourages GDGs to deal with their uncertainty and not abstain from making any recommendation. It states that it is important to provide advice to healthcare professionals and patients despite the lack of high-quality evidence. Such recommendations will inevitably be weak, and may be accompanied by qualifications [75, 76].

If the GDG finally refrains from providing a recommendation, it should explicitly state that the evidence is insufficient to make recommendations.

Key issues

- The GRADE Working Group has developed Evidence to Decision (EtD) frameworks which comprises the factors that GDG should consider for making a recommendation. For each factor, all relevant evidence should be systematically reviewed. When evidence is lacking, the GDG should make explicit what considerations were taken into account.
- The recommendations are classified according to their direction, in favour or against the use of an intervention and, depending on their strength, into strong and weak. The GRADE approach also contemplates the possibility of making only-in-research recommendations.
- Recommendations should be presented as clear, specific and actionable statements.

09.

EXTERNAL REVIEW

This chapter provides information on the external review process prior to the Clinical Practice Guidelines (CPG) publication. It describes the objectives, method, and composition of the external review group.

Once the CPG has been drafted, a broad-ranging external review should be carried out. An external review is a chance to receive feedback from individuals with expertise and perspectives that may not be represented in the GDG. It presents an opportunity for stakeholders, all those who have a legitimate interest in the guideline, to comment on a full draft, including recommendations, before it is published.

External reviews are mainly intended to assess:

- ✓ The rationale applied in searching for and examining the body of evidence.
- ✓ The quality of the evidence on which the recommendations are based.
- ✓ The rigour of the development process.
- ✓ The usability and acceptability of the recommendations and the overall guideline.

In addition to quality assurance, external reviews can improve CPG uptake by strengthening the legitimacy of recommendations and thus convince users that they are a trustworthy resource.

9.1 | External review group

The external review group is composed of experts interested in the subject of the guideline as well as individuals who will be affected by the recommendations. It may include clinical experts from the ERN, methodological and technical experts, end-users, and individuals affected by the condition addressed in the CPG, among other stakeholders. If important perspectives and stakeholders are missing from the GDG, these should be represented in the external review group.

It is recommended that at least 10–12 reviewers (of whom at least two should be patients and carers) are engaged in the process [6]. However, this number will vary greatly according to the availability of experts in the rare disease targeted by the CPG.

There are different methods for recruitment. External reviewers can be contacted through their organisations. Professionals who are identified as experts in the field may also be asked directly to

participate in the review. The GDG and the institution promoting the guideline can also suggest names.

Potential reviewers should be contacted at least two months in advance of sending them the draft guideline so that they can plan for the work involved. It is important to provide sufficient information to individuals when requesting their involvement in the review process. This material should include specific information on the CPG and the GDG's expectations of the reviewer, including:

- ✓ Basic information about the guideline, including the name of the CPG, the scope and purpose, and the organisations funding and developing it.
- ✓ The scope of the external review, including any specific questions they will be asked to answer and any frameworks they will be asked to refer to.
- ✓ The date they should expect to receive the guideline and the length of time they will have to complete the review (e.g. four weeks).
- ✓ How they will be acknowledged in the published guideline.

External reviewers should be subject to the same declaration of interest policy as members of the guideline development group. Once an individual has confirmed their availability to review the guideline, his or her potential conflict of interests should be carefully identified and duly addressed in accordance with the indications established by our partner FPS.

9.2 | External review methods

Members of the external review group may review the scope of the CPG and key questions (in PICO format) in the early stages of the CPG development process (see chapter 4), and the final CPG document at the end.

Each external reviewer will be sent a draft of the CPG and a template for the review (see annex III), accompanied by the following instructions:

- ✓ Comments received from external reviewers will be supported whenever possible by scientific evidence and accompanied by pertinent references.
- ✓ Given the provisional nature of the document, the external reviewers should make confidential use of same.
- ✓ Comments should be sent before the closing date (which is specified).
- ✓ All external reviewers' comments will be considered by the GDG, which will ultimately decide whether they are incorporated or not.

9.2.1 | Dealing with external reviewers' comments

It is advisable to adopt a systematic process for responding to reviewers' comments. All reviewer comments are collected and recorded. It is not necessary to respond to every single comment individually. However, it is important to let reviewers know how comments were addressed. For example, the GDG might develop a table capturing each commentary from every reviewer, explaining how the guideline was or was not modified accordingly, and describing the rationale for it [77]. The public availability of such information is important to transparency.

The GDG may have to rewrite recommendations and CPG text and reapprove the final document.

An advance copy of the final full guideline should be made available for information purposes to

external reviewers soon before the publication date.

9.3 | Description of the external review process in the CPG

The methodology chapter of the CPG should describe the external review process, including the following information:

- ✓ A description of the multidisciplinary group involved in the external review (clinical experts, methodological experts, patients and/or carers).
- ✓ A description of the methodology used to conduct the external review.
- ✓ A summary of the changes in the CPG after the external review process.

Key issues

The guideline should undergo external review before final publication. If important perspectives and stakeholders are missing from the GDG, these should be represented in the external review group. The GDG should be transparent regarding the handling of comments and changes during this process. After the external review, a second draft may be necessary.

10.

GUIDELINE REPORTING FORMAT

This chapter provides information on the structure, format, content and style of the Clinical Practice Guideline.

There is little information available on the effect that style and presentation have on the adoption and utility of guidelines. Clarity of definitions, language, and format, is likely to be important. Guidelines should, therefore, be written in unambiguous language and should define all terms precisely. Plain language should be used, and unnecessary jargon avoided. Detailed instructions for writing guideline recommendations are given in section 8.3.1.

The most appropriate format for presenting guidelines will vary depending on the target audience (healthcare professionals, policy makers, patients and/or carers), and the intended use of the guideline.

This handbook proposes different formats for the guidelines developed in the context of the European Reference Networks (ERN):

- ✓ Full version of the CPG
- ✓ Short version of the CPG
- ✓ Version for patients and carers

Having a well-developed and defined template for presentation of the final guideline can greatly facilitate the development process. All extra supporting documentation should form a separate resource pack and should include the methodological material of the CPG.

10.1 | Structure of the full CPG

The content of the full version of the guideline should include the following:

- ✓ Title.
- ✓ Authorship and collaboration.
- ✓ A list of all the recommendations of the CPG.
- ✓ Key recommendations.
    - Those considered by the GDG as having the greatest potential impact on patient care.

- ✓ Introduction.
    - Outlining the need for the guideline, including evidence of variation in practice and the potential for the guideline to improve patient care.
- ✓ Scope and purpose.
- ✓ Methodology.
    - Brief information on the methods used in each step of the guideline development process.
- ✓ Clinical chapters, dealing with the review questions and the evidence that led to the recommendations, each with the following content:
    - An introduction to the chapter.
    - The clinical question(s).
    - The recommendation(s).
    - Justification of the recommendation(s) (see section 8.1)
    - The research recommendation(s) (if applicable).
    - The clinical evidence review using GRADE evidence profiles (see sections 6.1 and 6.2). If it is not possible to apply GRADE to the evidence, it may be presented in narrative summaries.
    - A structured summary of all the factors addressed in the GDG discussions to formulate recommendations, which are included in the evidence to decision table(s) (see section 8.1).
- ✓ Algorithm(s) of diagnostic and therapeutic strategies.
- ✓ Dissemination and implementation.
    - Members of the GDG can inform on potential barriers for the use of the CPG in their context and provide counselling for the development of implementation strategies (see Handbook #12: Implementation and Evaluation of the Uptake of CPGs and CDSTs for Rare Diseases).
    - Development of quality measures (see Handbook #10: Methodology for the Development of Quality Measures for Rare Diseases).
- ✓ Future research
    - A list of all the research recommendations.
- ✓ Annexes (declarations of interest, etc.).
- ✓ References.
- ✓ Glossary and abbreviations.

10.2 | Structure of the short CPG

The availability of a short version facilitates the use of the guideline recommendations at the clinical point-of-care. The content of the short version of the guideline should include the following:

- ✓ Clinical questions of the CPG
- ✓ Recommendations of the CPG

✓ Algorithm(s) of diagnostic and therapeutic strategies

10.3 | Structure of patient versions of CPGs

The version for patients and carers describes the condition, and presents options with benefit and risk in easy-to-understand terms. The information is intended to help patients make decisions about treatment alternatives.

Practical guidance on the development of the patients' version of the CPG is provided elsewhere (see Handbook #11: Methodology for the Development of Patient Information Booklets for Rare Diseases).

10.4 | Methodological material of CPGs

The document with methodological material describes the activities and procedures needed for guideline development. The methodological material includes the following components:

✓ Introduction

✓ A list of all the clinical questions of the CPG.

✓ Information on the methods used in each step of the guideline development process.

✓ Template for the development of each clinical question (see Annex IV):

- The review question in PICO format (population, intervention, comparator[s] and outcome).
- Search strategies and results of the bibliographic search (including a flowchart of the selection of the studies and inclusion and exclusion criteria).
- Tables of individual studies.
- Evidence to decision (EtD) table(s), GRADE evidence profile(s) and meta-analysis diagram.
- Full economic report

Key issues

- Clinical Practice Guidelines drafting must be unambiguous, precise, comprehensive and should use plain language, avoiding unnecesary jergon.
- Clinical Practice Guidelines accompanied by patients and carers' versions may empower patients to make more informed healthcare choices.

11.

UPDATING THE CPG

This chapter provides information about key issues related to updating a Clinical Practice Guideline (CPG), including the composition of the CPG updating working group, the assessment of the potential impact of the new evidence on the clinical questions and recommendations of the original CPG, and how to modify the recommendations. It also offers guidance on the reporting of updated CPGs.

Because scientific knowledge is continually developing and improving, the emergence of new studies requires on-going reviews of clinical practice. Updating CPGs is therefore an essential matter to be addressed in order to ensure the validity and quality of CPG recommendations. Based on the results of studies that evaluated the validity of CPGs [78-80], most methodological handbooks for the development of CPGs propose two to three years as a reasonable time frame to update their guidelines [6].

The updating of a CPG is an iterative process that involves an explicit and systematic methodological approach for the identification and assessment of new evidence not included in the original CPG.

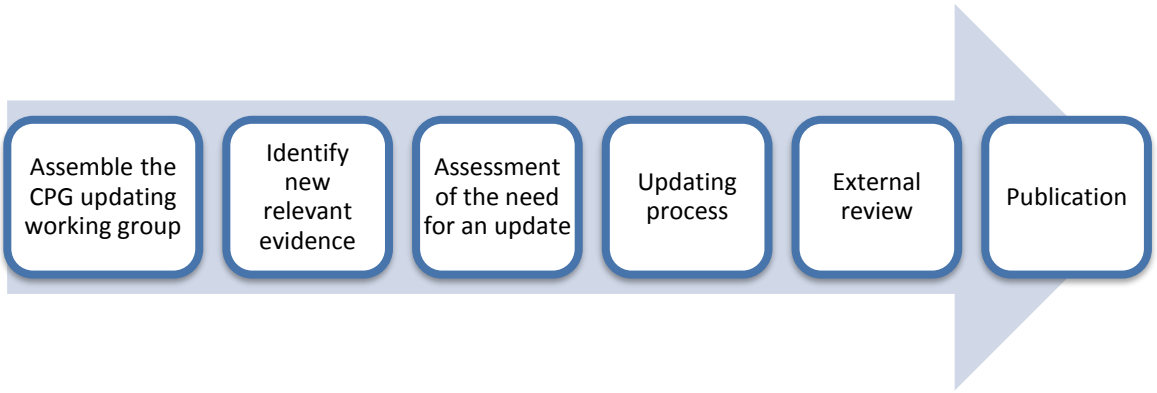
11.1 | Strategies for updating the CPG

This handbook presents two different strategies for keeping guidelines up-to-date[6]:

- ✓ Updating when a pre-specified interval has elapsed. In this case, we suggest assessing the validity of CPGs every three years [78].
- ✓ Continuous surveillance and updating strategy to identify new relevant evidence and evaluate the need to update (periodic scans of new evidence every 6 or 12 months).

The time to update a CPG should be considered according to several factors and it can be different for each clinical question. Some key issues are the topic addressed by the CPG, the volume of bibliographic production and the body of evidence previously published on the topic. The GDG of the original guideline should evaluate, once the CPG has been developed, the most appropriate approach for updating the clinical questions (continuous surveillance or periodic identification and assessment of new evidence).

Figure 3. CPG updating process



11.2 | Composition of the group responsible for updating the CPG

The composition and responsibilities of the CPG updating working group are very similar to those described for the guideline development group (see chapter 3).

The CPG updating working group should have a similar structure to that of the GDG which contributed to the development of the original CPG, unless changes in the scope due to the identification of new clinical areas raise the need to recruit new professional profiles. Members of the original GDG should be invited to participate in the updating process. The technical team plays a key role in the identification and assessment of new evidence that could modify the recommendations.

The strategy for keeping guidelines up-to-date will determine the qualitative and quantitative composition of the working group and the resources necessary for its implementation.

- ✓ Continuous surveillance and updating (every 6 or 12 months) needs a smaller group of individuals to identify and assess the new evidence, and the participation of the entire working group only in specific circumstances (e.g. when new evidence suggests the need for modification of a recommendation). This approach demands the availability of an on-going updating group and resources.
- ✓ Periodic updating of the CPG (every 3 years) requires an updating working group very similar to that of the original CPG, and the resources usually assigned to the development of a CPG project during a specified period of time.

11.3 | Identification of new evidence

The first step is to perform a restrictive literature search, prioritising precision over sensitivity, to identify new evidence that will trigger a recommendation update. The efficiency and feasibility of the restrictive approach described below has been validated previously [81].

- ✓ Run searches in a source selected for the topic of the CPG (e.g. MEDLINE) and prioritise the retrieval of systematic reviews.
- ✓ Develop the restrictive search strategies considering the minimum number of Medical Subject Headings (MeSH) terms and text words required from the original exhaustive search strategies. Limit the search to keywords that only appear in the title and abstract.

- ✓ Validate the strategies by checking that all key references supporting the recommendations in the original CPG are retrieved and refine them if necessary.
- ✓ Apply validated filters to improve precision, e.g. PubMed Clinical Queries (www.ncbi.nlm.nih.gov/pubmed/clinical).

11.3.1 | Search for recently published guidelines

It is recommended to identify good quality guidelines that are up to date by searching the repositories, institutions and databases listed below:

ECRI Guidelines Trust®	https://guidelines.ecri.org/
G-I-N international guideline library	www.g-i-n.net/library/international-guidelines-library
GuíaSalud	www.guiasalud.es
NICE (National Institute for Health and Care Excellence) clinical guidelines	www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-guidelines/nice-clinical-guidelines
Orphanet	www.orpha.net
RARE-Bestpractices	www.rarebestpractices.eu
Scottish Intercollegiate Guidelines Network (SIGN)	www.sign.ac.uk
CMA Infobase: Clinical Practice Guidelines Database (CPGs)	www.cma.ca/En/Pages/clinical-practice-guidelines.aspx
Australia's Clinical Practice Guidelines Portal	www.clinicalguidelines.gov.au
Tripdatabase	www.tripdatabase.com
MEDLINE and EMBASE by using methodological filters	

The most recent guidelines on the topic are a source of new relevant evidence that may be useful to update the recommendations.

11.3.2 | Questionnaire for the CPG updating working group

It is recommended to complement searches by sending a questionnaire to the CPG updating working group to identify new evidence that could have an impact on the CPG. The questionnaire should cover the different areas of the CPG including the scope, new potential aspects not included in the original version, or new relevant evidence assessing the effectiveness and safety of the interventions. It should also include questions about other relevant factors such as changes in the relative importance of the outcomes, changes in the resource use and cost of the interventions, equity, acceptability, or feasibility issues that might have arisen since the publication of the CPG. Information about on-going research studies should also be sought in the questionnaire (see Annex

V) [6, 82].

11.3.3 | Alerts for drugs and medical devices

It is important to plan and implement strategies rigorously to collect alerts issued by regulatory authorities, such as the European Medicines Agency. This is necessary to obtain new information on the adverse effects of treatments, so that recommendations on drugs and other healthcare interventions can be removed from the CPG and risk/benefit ratios modified.

11.4 | Assessment of the need for an update

The aim of this step is the identification of new evidence that could potentially change any component of the clinical questions (patients, intervention, comparison or outcomes of interest), or any factor influencing the strength and direction of the recommendations (e.g. the balance between benefits and risks or resource use and costs).

An initial qualitative assessment of the evidence identified should be performed, with the support of the technical team, by achieving consensus within the updating working group on the potential impact of the new evidence on the recommendations. The updating working group should decide if recommendations are still up to date or if they need to be reviewed [6].

11.4.1 | Screening and classification of references

The references are classified according to their relationship with the elements of the clinical question (PICO) and the adequacy of the study design, as follows [6, 82]:

- ✓ Pertinent references: topic-related references that met the study design criteria.
- ✓ Relevant references: pertinent references that could be used when considering an update to a recommendation, but that would not necessarily trigger a potential update.
- ✓ Potential key references: relevant references that could potentially trigger an update.

Qualitative criteria to classify relevant references into key references:

- ✓ The new study causes a change in the scope of the clinical question (patients, intervention, comparison, or outcomes of interest). For instance, new evidence can result in refining the recommendation for a subgroup or in including new interventions in a recommendation on the use of a diagnostic strategy or treatment.
- ✓ The new study leads to significant changes in the factors that influence the formulation of recommendations, which may modify the strength or direction of the recommendations.

11.4.2 | Classification of clinical questions

As a result of the assessment of the impact of the new evidence, the clinical questions are classified as follows:

- ✓ Clinical questions to be reviewed: with potential key references and with different relevant references or important pharmacological alerts.
- ✓ Valid clinical questions: without potential key references associated.
- ✓ New clinical questions.

11.5 | Updating process

The updating process will follow the methodology of development of CPGs described in this handbook. If restrictive searches have been conducted to evaluate the need for updating, it will be necessary to redesign the search strategies and broaden the sources of information to perform more exhaustive searches, as described in chapter 5.

It is noteworthy that whenever a clinical question has been developed with a methodological approach other than GRADE, it will be necessary to elaborate ex novo the evidence profiles, which means that both the body of evidence of the original question and the new references (relevant and key references) should be assessed, as described in chapter 5. This issue applies particularly in the process of adaptation of CPGs (see Handbook #3: Adaptation and Adoption of CPGs and CDSTs).

The recommendation development procedure will be assisted by the development of evidence to decision frameworks, as described in chapter 8. Finally, the following types of recommendations can be distinguished:

- ✓ New recommendations
- ✓ Reviewed and modified recommendations (with key references)
- ✓ Reviewed and unmodified recommendations (with relevant references)
- ✓ Excluded recommendations

The additional material based on the guideline recommendations such as patient information booklets or quality measures should be updated accordingly.

The final draft with the corresponding modifications will be submitted to a external review process as described in chapter 9.

11.6 | Exceptional updates

CPGs should be reviewed earlier than planned if significant evidence emerges (changes in treatment indications, new diagnostic tests, alerts by healthcare authorities or significant changes regarding safety) that requires one or several recommendations to be updated in a way that will likely change clinical practice substantially. These exceptional updates usually need to be published rapidly.

11.7 | External review

The updated CPG should be externally reviewed prior to publication by a full spectrum of stakeholders as described in chapter 9.

11.8 | Edition of the update

The edition of an update will follow the same style principles as for the development of ex novo recommendations (see chapter 10). However, guideline updating requires some different methodological considerations and unique communication procedures. The Checklist for the Reporting of Updated Guidelines (CheckUp) can support guideline developers in the reporting of updated CPGs and ensure compliance with guideline methodological standards. CheckUp was developed by a partnership formed by the Iberoamerican Cochrane Centre, the AGREE Collaboration, and the G-I-N Updating Guidelines Working Group [83].

The tool consists on 16 items that address the following aspects [83]:

- ✓ The updated version is distinguished from the previous version of the guideline.
- ✓ The sections reviewed in the updating process are described.
- ✓ The recommendations are clearly presented and labelled as new, modified, or no change. Deleted recommendations are clearly noted.
- ✓ The panel participants in the updated version are described.
- ✓ The rationale for updating the guideline is reported.
- ✓ Changes in the scope and purpose between the updated and original version are described and justified.
- ✓ Changes in the original recommendations are reported and justified.
- ✓ The methods used for searching and identifying new evidence in the updating process are described.
- ✓ The methods used for evidence selection in the updating process are described.
- ✓ The methods used to assess the quality of the evidence included in the updating process are described.
- ✓ The methods used for the evidence synthesis in the updating process are described.
- ✓ The methods and plan for implementing the changes of the updated version in practice are described.
- ✓ The methods used for externally reviewing the updated version are described.
- ✓ The plan and methods for updating the new version in the future are reported.
- ✓ The conflicts of interests of the group responsible for the updated version are recorded.
- ✓ The role of the funding body for the updated guideline is identified and described.

A user-friendly version of CheckUp is provided in the authors' publication[83]. The explanations and examples for each item can also be accessed in the authors' publication as a supportive information article (S1 Appendix)[83].

Key issues

- Guidelines can be updated on a continuous surveillance basis or at pre-established regular intervals (e. g. every three years).
- The composition of the CPG updating working group is similar to that of the original guideline.
- For each clinical question, the main steps of the process are: 1) identification of new evidence; 2) assessment of the evidence with qualitative criteria to determine the need for an update; 3) review of the recommendations.
- The Checklist for the Reporting of Updated Guidelines (CheckUp) can support guideline developers in the reporting of updated CPGs.

12.

BIBLIOGRAPHY

1. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1.Introduction-GRADE evidence profiles and summary of findings tables. Journal of clinical epidemiology. 2011;64(4):383-94.

2. Vandevelde, Nathalie M. 'Strategies to Improve the Quality of Reference Networks for Rare Diseases'. 1 Jan. 2019 : 1 – 22.

3. Schünemann HJ, Wiercioch W, Etxeandia I, Falavigna M, Santesso N, Mustafa R, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. CMAJ. 2014;186(3):E123-42.

4. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines. Clinical practice guidelines we can trust [Internet]. Washington (DC): National Academies Press (US); 2011 [cited 15/06/2020]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK209539/.

5. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42.

6. Grupo de trabajo para la actualización del Manual de Elaboración de GPC. Elaboración de Guías de Práctica Clínica en el Sistema Nacional de Salud. Actualización del Manual Metodológico. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad; Zaragoza: Instituto Aragonés de Ciencias de la Salud (IACS); 2016.

7. National Institute for Health and Care Excellence (NICE). Developing NICE guidelines: the manual [Internet]. Manchester: NICE, 2014 [cited 4-06-2020]. Available from: https://www.nice.org.uk/process/pmg20/chapter/introduction-and-overview.

8. National Institute for Health and Care Excellence (NICE). Incorporating economic evaluation. In: Developing NICE guidelines: the manual [Internet]. Manchester: NICE, 2014 [cited 4-06-2020]. Available from: https://www.nice.org.uk/process/pmg20/chapter/incorporating-economic-evaluation.

9. WHO handbook for guideline development. 2nd ed. Geneva: World Health Organization; 2014. [4 December 2015]. http://www.who.int/kms/handbook_2nd_ed.pdf.

10. Scottish Intercollegiate Guidelines Network (SIGN). A guideline developer's handbook. Edinburgh: SIGN; 2019. (SIGN publication no. 50). [November 2019]. Available from URL: http://www.sign.ac.uk.

11. Schünemann H, Brozek J, Guyatt G, Oxman A (eds.). GRADE Handbook: Introduction to GRADE Handbook [Internet]. The GRADE Working Group; 2013 [cited 15/06/2020]. Available from: https://gdt.gradepro.org/app/handbook/handbook.html.

12. Qaseem A, Forland F, Macbeth F, Ollenschläger G, Phillips S, van der Wees P. Guidelines International Network: toward international standards for clinical practice guidelines. Annals of Internal Medicine. 2012;156(7):525-31.

13. Ansari S, Rashidian A. Guidelines for guidelines: are they up to the task? A comparative assessment of clinical practice guideline development handbooks. PloS One. 2012;7(11):e49864.

14. Armstrong R, Hall BJ, Doyle J, Waters E. Cochrane Update. 'Scoping the scope' of a Cochrane review. Journal of Public Health. 2011;33(1):147-50.

15. Hilton Boon M, Harbour J, Ritchie K, Thompson L. and RARE-Bestpractices Consortium. Report of an international workshop to explore the utility of the AGREE II instrument for appraisal of rare disease guidelines. Rare Diseases and Orphan Drugs. 2015; 2(1):11 - 5.

16. National Institute for Health and Care Excellence (NICE). Decision-making committees. In: Developing NICE guidelines: the manual [Internet]. Manchester: NICE; 2014 [cited 4-06-2020]. Available from: https://www.nice.org.uk/process/pmg20/chapter/decision-making-committees.

17. Haute Autorité de Santé (HAS). Méthode d'élaboration d'un protocole national de diagnostic et de soins pour les maladies rares: Guide Methodologique. 2012.

18. Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 5. Group processes. Health research policy and systems. 2006;4:17.

19. Black N, Murphy M, Lamping D, McKee M, Sanderson C, Askham J, et al. Consensus development methods: a review of best practice in creating clinical guidelines. Journal of health services research & policy. 1999;4(4):236-48.

20. National Institute for Health and Care Excellence (NICE). Developing review questions and planning the evidence review. In: Developing NICE guidelines: the manual [Internet]. Manchester: NICE; 2014 [cited 4-06-2020]. Available from: https://www.nice.org.uk/process/pmg20/chapter/developing-review-questions-and-planning-the-evidence-review

21. Higgins J, Thomas J (ed.). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 [Internet]. Cochrane; 2019 [cited 15/06/2020]. Available from: www.training.cochrane.org/handbook.

22. Pai M, Iorio A , Meerpohl J, Taruscio D, Laricchiuta P, Mincarone P, Morciano C, Leo CG, Sabina S, Akl E, Treweek S, Djulbegovic B, Schunemann H; Consortium for the RARE-Bestpractices. Developing methodology for the creation of clinical practice guidelines for rare diseases: A report from RARE-Bestpractices. Rare Dis. 2015;3(1).

23. Assessment EUnetHTA. Endpoints used for Relative Effectiveness Assessment: Composite endpoints. 2015. p. 22.

24. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. Journal of Clinical Epidemiology. 2011;64(4):395-400.

25. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to

searching: an efficient and complete method to develop literature searches. JMLA. 2018;106(4):531-41.

26. Pavan S, Rommel K, Mateo Marquina ME, Höhn S, Lanneau V, Rath A. Clinical Practice Guidelines for Rare Diseases: The Orphanet Database. PloS One. 2017;12(1):e0170365.

27. Pai M, Yeung CHT, Akl EA, Darzi A, Hillis C, Legault K, et al. Strategies for eliciting and synthesizing evidence for guidelines in rare diseases. BMC medical research methodology. 2019;19(1):67.

28. Rare disease registries in Europe [Internet]. Orphanet; 2019 [cited 15/06/2020]. Orphanet Report Series. Available from: https://www.orpha.net/orphacom/cahiers/docs/GB/Registries.pdf.

29. European Commission [Internet]. [cited 15/06/2020]. Rare diseases. Available from: https://ec.europa.eu/health/non_communicable_diseases/rare_diseases_es.

30. Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, et al. Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. Systematic Reviews. 2016;5(1):164.

31. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. Journal of Clinical Epidemiology. 2016;75:40-6.

32. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. Journal of Clinical Epidemiology. 2011;64(4):380-2.

33. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of Clinical Epidemiology. 2011;64(4):401-6.

34. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence-imprecision. Journal of Clinical Epidemiology. 2011;64(12):1283-93.

35. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence-indirectness. Journal of Clinical Epidemiology. 2011;64(12):1303-10.

36. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence-inconsistency. Journal of Clinical Epidemiology. 2011;64(12):1294-302.

37. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. Journal of Clinical Epidemiology. 2011;64(12):1311-6.

38. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias). Journal of Clinical Epidemiology. 2011;64(4):407-15.

39. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence-publication bias. Journal of Clinical Epidemiology. 2011;64(12):1277-82.

40. Schünemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized

studies should be used to rate the certainty of a body of evidence. Journal of Clinical Epidemiology. 2019;111:105-14.

41. Institute of Health Economics (IHE). Quality Appraisal of Case Series Studies Checklist. Edmonton: Institute of Health Economics; 2014.

42. Zhang Y, Coello PA, Guyatt GH, Yepes-Nuñez JJ, Akl EA, Hazlewood G, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences-inconsistency, imprecision, and other domains. Journal of Clinical Epidemiology. 2019;111:83-93.

43. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, 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. 2013;66(2):151-7.

44. Brozek JL, Akl EA, Jaeschke R, Lang DM, Bossuyt P, Glasziou P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines: Part 2 of 3. The GRADE approach to grading quality of evidence about diagnostic tests and strategies. Allergy. 2009;64(8):1109-16.

45. Schünemann HJ, Mustafa RA, Brozek J, Santesso N, Bossuyt PM, Steingart KR, et al. GRADE guidelines: 22. The GRADE approach for tests and strategies-from test accuracy to patient-important outcomes and recommendations. Journal of Clinical Epidemiology. 2019;111:69-82.

46. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. Journal of Clinical Epidemiology. 2020;122:129-41.

47. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. Journal of Clinical Epidemiology. 2020;122:142-52.

48. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine. 2011;155(8):529-36.

49. Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gülmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). PLoS Medicine. 2015;12(10):e1001895.

50. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implementation science. 2018;13(Suppl 1):2.

51. Munthe-Kaas H, Bohren MA, Glenton C, Lewin S, Noyes J, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 3: how to assess methodological limitations. Implementation science. 2018;13(Suppl 1):9.

52. Critical Appraisal Skills Programme (CASP). CASP Qualitative Checklist [Internet]. 2018 [cited 22/06/2020]. Available at: https://casp-uk.net/wp-content/uploads/2018/03/CASP-Qualitative-Checklist-2018_fillable_form.pdf.

53. Colvin CJ, Garside R, Wainwright M, Munthe-Kaas H, Glenton C, Bohren MA, et al. Applying

GRADE-CERQual to qualitative evidence synthesis findings-paper 4: how to assess coherence. Implementation Science. 2018;13(Suppl 1):13.

54. Noyes J, Booth A, Lewin S, Carlsen B, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 6: how to assess relevance of the data. Implementation Science. 2018;13(Suppl 1):4.

55. Glenton C, Carlsen B, Lewin S, Munthe-Kaas H, Colvin CJ, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings-paper 5: how to assess adequacy of data. Implementation Science. 2018;13(Suppl 1):14.

56. van Mastrigt GA, Hiligsmann M, Arts JJ, Broos PH, Kleijnen J, Evers SM, et al. How to prepare a systematic review of economic evaluations for informing evidence-based healthcare decisions: a five-step approach (part 1/3). Expert Rev Pharmacoecon Outcomes Res. 2016;16(6):689-704.

57. Centre for Reviews and Dissemination. University of York [Internet]. [cited 22/06/2020] Available from: https://www.crd.york.ac.uk/CRDWeb/

58. RARE-Bestpractices: A platform for sharing best practices for the management of rare diseases [Internet]. [cited 22/06/2020] Available from: http://www.rarebestpractices.eu/.

59. ISSG Search Filter Resource [Internet]. The InterTASC Information Specialists' Sub-Group Search Filter Resource; [cited 15/06/2020]. Filters to Identify Economic Evaluations. Available from: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-find-i.

60. Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, et al. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. J Clin Epidemiol. 2013;66(2):140-50.

61. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ (Clinical research ed). 1996;313(7052):275-83.

62. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. Int J Technol Assess Health Care. 2005;21(2):240-5.

63. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355-71.

64. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Value Health. 2013;16(2):e1-5.

65. CCEMG-EPPI-Centre Cost Converter [Internet]. [cited 15/06/2020]. Available from: http://eppi.ioe.ac.uk/costconversion/default.aspx].

66. Zozaya N V, R, Hidalgo A. et al. Guía metodológica de evaluación económica aplicada a medicamentos huérfanos: recomendaciones del grupo de expertos RADEEV. Madrid: Instituto Max Weber; 2015.

67. Center for the Evaluation of Value and Risk in Health. Tufts Medical Center [Internet]. [cited 15/06/2020]. Cost-Effectiveness Analysis (CEA) Registry. Available from: https://cevr.tuftsmedicalcenter.org/databases/cea-registry.

68. Detsky AS, Naglie G, Krahn MD, Naimark D, Redelmeier DA. Primer on medical decision analysis: Part 1--Getting started. Med Decis Making. 1997;17(2):123-5.

69. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics. 1998;13(4):397-409.

70. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J, et al. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--4. Value Health. 2012;15(6):821-7.

71. Alonso-Coello P, Schünemann HJ, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. BMJ. 2016;353:i2016.

72. Pottie K, Welch V, Morton R, Akl EA, Eslava-Schmalbach JH, Katikireddi V, et al. GRADE equity guidelines 4: considering health equity in GRADE guideline development: evidence to decision process. Journal of Clinical Epidemiology. 2017;90:84-91.

73. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. Journal of Clinical Epidemiology. 2013;66(7):726-35.

74. Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. Journal of Clinical Epidemiology. 2016;76:89-98.

75. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. Journal of Clinical Epidemiology. 2013;66(7):719-25.

76. Scharpf J. The Challenge of Guideline Development When Evidence Is Sparse. Otolaryngology--head and neck surgery. 2017;157(3):383-4.

77. 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: official journal of American Academy of Otolaryngology-Head and Neck Surgery. 2013;148(1 Suppl):S1-55.

78. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA. 2001;286(12):1461-7.

79. The validity of recommendations from clinical guidelines: a survival analysis. CMAJ. 2017;189(1):E33.

80. Alderson LJ, Alderson P, Tan T. Median life span of a cohort of National Institute for Health and Care Excellence clinical guidelines was about 60 months. Journal of Clinical Epidemiology. 2014;67(1):52-5.

81. Martínez García L, Sanabria AJ, Araya I, Lawson J, Solà I, Vernooij RW, et al. Efficiency of pragmatic search strategies to update clinical guidelines recommendations. BMC medical research methodology. 2015;15:57.

82. Martínez García L, Pardo-Hernández H, Sanabria AJ, Alonso-Coello P. Continuous surveillance of a pregnancy clinical guideline: an early experience. Systematic reviews. 2017;6(1):143.

83. Vernooij RW, Alonso-Coello P, Brouwers M, Martínez García L. Reporting Items for Updated

Clinical Guidelines: Checklist for the Reporting of Updated Guidelines (CheckUp). PLoS Med. 2017;14(1):e1002207.

13.

ANNEXES

ANNEX 13.1 | Template for the definition of the scope and purpose of the CPG

SECTION		CONTENT
Justification		*Explain why the guideline is needed*
Objectives		*Define general and specific objectives and the benefits that the guideline aims to achieve.*
Aspects to be covered	Target population	*Describe the characteristics of the target population and any subgroups (age group, type of disease or condition, disease severity, or comorbidities)*
	Aspects of care covered by the CPG	*Area of health practice, policy or public/environmental health issue that the guideline addresses.*
	Aspects related to patients	*Explain the way in which the perspective of patients and carers will be included. The development of topic-specific information and support for patients and carers should be stated.*
	Context of application	*Describe the health care setting to which the CPG applies, including the health system level (e.g. primary care, acute care) and clinical stage (e.g. prevention, screening, assessment, treatment, etc.).*
Aspects not covered by the CPG		*Define the aspects not covered by the CPG, for example, the exclusion of any clinical stage (e.g. prevention), or certain age groups (e.g. teenagers).*
Dealing with health inequities		*Explain the way in which potential health inequities will be identified.*
End users of the CPG		*Specify the intended end users of the CPG.*

ANNEX 13.2 | Template for presenting the results of the initial screening of the evidence for a clinical question

Title of the Clinical Practice Guideline – Bibliographic search. Clinical Question no. X

Wording of the clinical question.

Example:

In patients with suspected hereditary retinal dystrophies, which test are the most accurate confirm or refute the diagnosis?

Content

General comments:

The section "Results" includes the 'title and abstract' of the studies selected after the first screening. Below of each 'title and abstract' you will find a space to indicate if the study is included or excluded. In case of exclusion, the reasons must be stated. If you want to consult the full-text article before deciding about its inclusion or exclusion, please mark the corresponding space.

Search strategies used in each database is presented in Annex 1.

Databases searched:

Database	Results (number of studies retrieved)
Medline	
Embase	
-	
-	
Total number (without duplicates)	

Publications that passed initial screening (total number)	

Date of search:

General limits:

Limits by publication period:

Limits by publication language:

Other:

Limits by study design:

Results:

1. Bravo-Gil N, Méndez-Vidal C, Romero-Pérez L, et al. Improving the management of Inherited Retinal Dystrophies by targeted sequencing of a population-specific gene panel. Sci Rep 2016; 6: 23910.

Abstract

Next-generation sequencing (NGS) has overcome important limitations to the molecular diagnosis of Inherited Retinal Dystrophies (IRD) such as the high clinical and genetic heterogeneity and the overlapping phenotypes. The purpose of this study was the identification of the genetic defect in 32 Spanish families with different forms of IRD. With that aim, we implemented a custom NGS panel comprising 64 IRD-associated genes in our population, and three disease-associated intronic regions. A total of 37 pathogenic mutations (14 novels) were found in 73% of IRD patients ranging from 50% for autosomal dominant cases, 75% for syndromic cases, 83% for autosomal recessive cases, and 100% for X-linked cases. Additionally, unexpected phenotype-genotype correlations were found in 6 probands, which led to the refinement of their clinical diagnoses. Furthermore, intra- and interfamilial phenotypic variability was observed in two cases. Moreover, two cases unsuccessfully analysed by exome sequencing were resolved by applying this panel. Our results demonstrate that this hypothesis-free approach based on frequently mutated, population-specific loci is highly cost-efficient for the routine diagnosis of this heterogeneous condition and allows the unbiased analysis of a miscellaneous cohort. The molecular information found here has aid clinical diagnosis and has improved genetic counselling and patient management.

- **Inclusion:** Yes __ No__

- **Reasons for exclusion:** Yes __ No__

- **Request for full-text article:** Yes __ No__

Annex 1. Search strategies:

Medline (Pubmed)

Example:

(Pigmentary retinopathy or Rod-cone dystroph* or Rod cone dystroph* or Retinal Dystroph* or Retinitis pigmentosa or pigmentary retinosis or retinosis pigmentary or retinosis pigmentosa or North Carolina Macular Distrophy or Stargardt-Fundus flavimaculatus or stargardt's fundus flavimaculatus or Stargardt's disease or Stargardt Macular Degeneration or Stargardt disease or fundus flavimaculatus or Sorsby dystrophy or Gyrate Atrophy or Atrophia Gyrata or Enhanced S-cone Syndrome or Goldman-Favre or Wagner-Stickler or vitreoretinal dystroph* or X-linked Juvenile retinoschisis or Occult Macular Dystrophy or Macular dystroph* or Choroideremia or Congenital Stationary Night Blindness or central areolar choroidal dystrophy or Bestrophinopathy or Bestrophinopathies or Best Vitelliform Macular Dystrophy or Vitelliform Macular Dystrophy or Familial Exudative vitreoretinopathy or adult-onset foveomacular Dystrophy or Butterfly-shaped pattern dystrophy or Pattern dystrophies in Retinal Pigment Epithelium or Autosomal dominant Stargardtlike macular dystrophy or Stargardt Macular Degeneration or Stargardt disease).ab,ti.

3. 1 or 2

4. *Early Diagnosis/

5. *Symptom Assessment/

6. (Early detection of disease or primary

diagnosis or Main diagnosis or predictor or

clinical feature or Symptom or Sign).tw.

7. 4 or 5 or 6

8. 3 and 7

9. limit 8 to "all child (0 to 18 years)"

10. limit 9 to humans

11. limit 10 to (English or Spanish)

12. remove duplicates from 11

ANNEX 13.3 | Template for external reviewers Barriers

Given the different profiles of external reviewers (clinical experts, methodologists and patients and carers or patient representatives), two different models of the external review form are proposed: one aimed at clinical experts and methodologists (model 1); and another aimed at patients or patient representatives (model 2).

Model 1: External review form for healthcare professionals

Clinical Practice Guideline on... (Please, enter the title)

EXTERNAL REVIEWER'S NAME: .....

ASPECTS TO TAKE INTO ACCOUNT IN THE EXTERNAL REVIEW OF A CPG

A. Applicability and relevance

Explore whether the topic addressed by the CPG provides useful and relevant information for the healthcare context, and whether the social relevance of the topic is considered.

B. Contents and structure of the CPG

Assess the recommendations formulated in the CPG. Observe whether they are related and based on scientific evidence and evaluate if they are clear. Also assess whether the CPG is structured following an adequate and logical order, and whether the figures, algorithms, tables and annexes are useful. Also consider information to carers (and whether this information is adequate, useful and sufficient).

C. Formal aspects

Assess whether the language used in the information provided to healthcare professionals and patients and carers is correct and clear.

External review (please, check one of the following boxes):

1. Are the objectives of the CPG specifically described?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

2. Is the definition of the health condition clearly described?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

3. Are the clinical management criteria clearly described in every situation?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

4. Are the recommendations specific and unambiguous?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

5. Is there an explicit relationship between the recommendations and the scientific evidence that supports them?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

6. Please, rate (1-10) the applicability of this CPG in your healthcare context:

Additional comments

Please, make comments, observations and suggestions that you deem appropriate to improve this CPG in the following sections. If you do not agree with any of the recommendations, please provide the evidence that could modify them together with the corresponding bibliographic references (enter the first author's last name, study title, journal, volume, pages and publication date).

OVERALL COMMENTS ON THE CPG:

COMMENTS BY SECTIONS:

1. Introduction

2. Scope and objectives

3. Methodology

4. Chapter 4

5. Chapter 5

Model 2: External review form for patients and carers or patient representatives

Clinical Practice Guideline on... (Please, enter the title)

EXTERNAL REVIEWER'S NAME: .....

ASPECTS TO TAKE INTO ACCOUNT IN THE EXTERNAL REVIEW OF A CPG

Please, check one of the following boxes. Provide comments if considered necessary.

CONTENTS			
● ● In your opinion, do you think that any relevant content for the CPG is missing?	Yes ☐	No ☐	DK/NA ☐
Comments:			
● Is there any information or recommendation that you agreed with (e.g. due to possible benefits or disadvantages mentioned in the CPG)?	Yes ☐	No ☐	DK/NA ☐

Comments:			
• Is the information provided appropriate for those affected by the condition covered in the CPG?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Does the CPG consider collaboration between healthcare professionals and patients?	Yes ☐	No ☐	DK/NA ☐
Comments:			
LANGUAGE AND FORMAT			
• Is the CPG clear and easy to read?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Does the CPG use understandable and appropriate language?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Are the recommendations specific and unambiguous?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Do you think that the wording is respectful to patients and carers?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Do you think that the document format is adequate (e.g. regarding length, presentation of the recommendations, algorithms, figures, etc.)?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Overall, do you consider that the edition and format of this CPG facilitate its use?	Yes ☐	No ☐	DK/NA ☐
Comments:			
OTHER ISSUES			

• Do you consider that this CPG could be useful?	Yes ☐	No ☐	DK/NA ☐
Comments:			
• Do you think that there is room for improvement in any aspect?	Yes ☐	No ☐	DK/NA ☐
Comments:			
OBSERVATIONS AND SUGGESTIONS			
Please, detail all those aspects that should be improved or corrected in the CPG reviewed or in the external review process itself. Also make suggestions that you deem appropriate and that have not been considered in the previous sections.			
Comments:			

ANNEX 13.4 | Template for the methodological material of a clinical question

CONTENT:

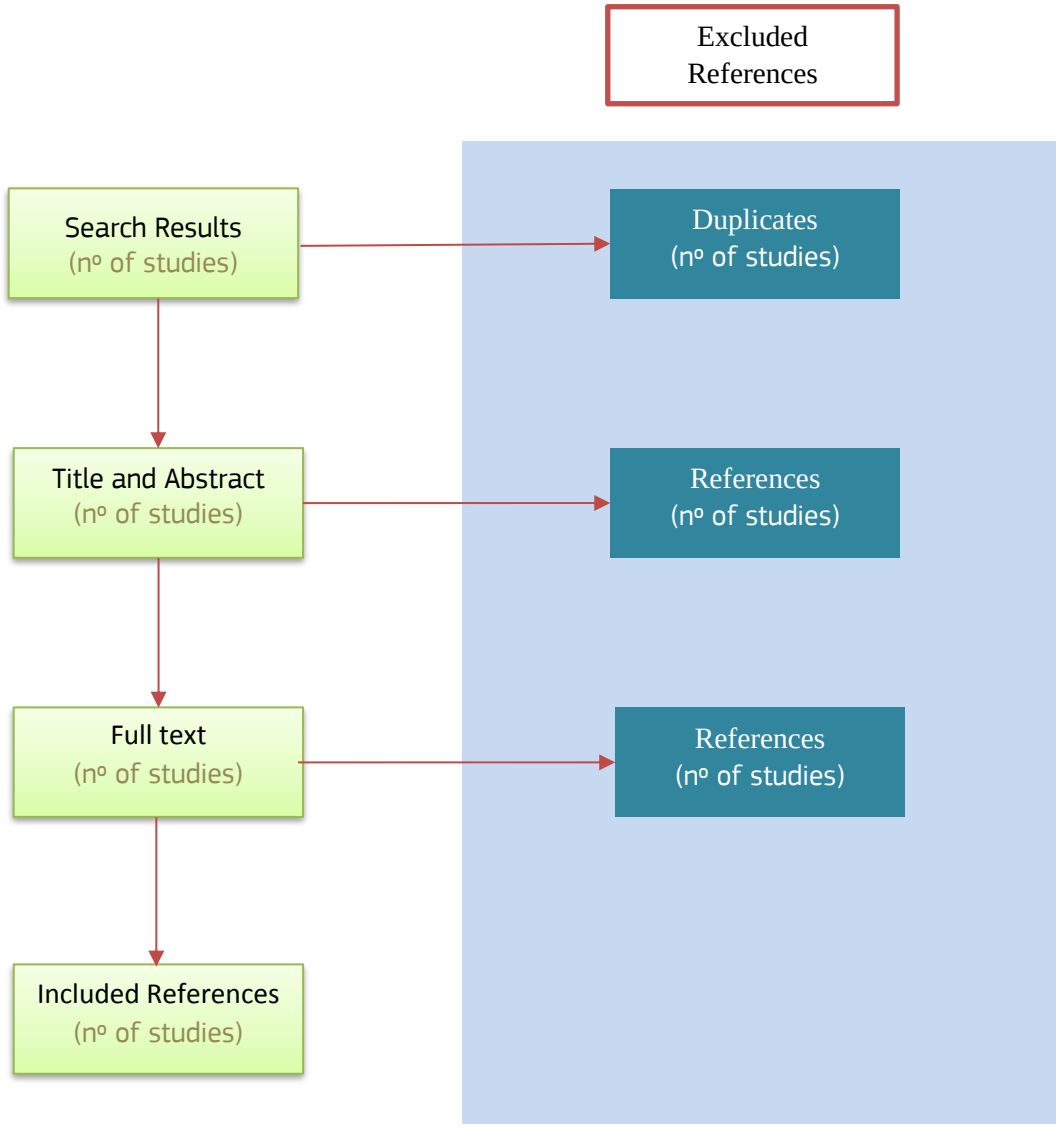
1. STRUCTURED CLINICAL QUESTION (PICO FORMAT)

	For example: in adult patients is the use of preservative-free drugs recommended against the use of preservative drugs for the treatment of open-angle glaucoma?
Patient, Population or Problem	
Intervention or Exposure	
Comparison	
Outcomes	

2. SEARCH STRATEGY

Database: MEDLINE (Pubmed)
Search Date: 27/09/2013
Database: The Cochrane Library (includes NHS EED, DARE, CENTRAL)
Search Date: 27/09/2013

3. STUDY SELECTION FLOWCHART AND INCLUSION/EXCLUSION CRITERIA



Out of the total number of bibliographic references retrieved for the development of the entire CPG, it is necessary to explain the number of studies selected and design types and subtypes (e.g. systematic review, meta-analysis, observational study, etc.) for the each clinical question, according to the inclusion and exclusion criteria:

Inclusion Criteria	
Exclusion Criteria	

4. TABLES OF INDIVIDUAL STUDIES

Example of an individual table corresponding to a clinical trial

Randomised Clinical Trial	
ID: Lewis, 2017 PMID: 17224758	
Characteristics	**Bias risk**
Patients:	Generation of the randomisation sequence:
Intervention:	Concealment of the randomisation sequence:
Comparison:	Blinding:
Outcomes:	Follow-up losses:
Follow-up time	Others:

5. EVIDENCE TO DECISION (EtD) TABLES, GRADE EVIDENCE PROFILE(S) AND META-ANALYSIS DIAGRAM

For each clinical question, evidence profiles should be developed using GRADEPro-GDT software (http://www.guidelinedevelopment.org/) or iEtD (http://ietd.epistemonikos.org/#/).
The relevant evidence profile(s) and meta-analysis diagrams that have been developed are included .

ANNEX 13.5 | Questionnaire for gathering evidence from experts and members of the Guideline Development Group (GDG)

Scope and purpose of the Clinical Practice Guideline (CPG)

1. **Are you aware of new aspects that should be addressed in the CPG but are not?**
   - ☐ Yes
   - ☐ No
   - ☐ Does not know/ Does not answer

   If that is the case, please specify:

Efficacy of the interventions

2. **Are you aware of any relevant new studies published regarding the efficacy of the interventions included in the guide?**
   - ☐ Yes
   - ☐ No
   - ☐ Does not know/ Does not answer

If you know of any new relevant studies published since the development of the CPG (month/year), please specify, as exactly as possible, the bibliographic reference of each study, ideally in the following format:
First author. Title of the article. Title of the journal. Year; volume (number): first page-last page.

Study 1:
Study 2:
Study 3:
Study 4:
Study 5:

Safety of the interventions

3. **Are you aware of any relevant aspect related to the safety of the interventions included in the CPG?**
   - ☐ Yes
   - ☐ No
   - ☐ Does not know/ Does not answer

   If that is the case, please specify:

Relative importance of the outcomes

The relative importance of the outcomes refers to the value that patients and the general public place on each of the outcomes related to an intervention, i.e. the outcomes which patients or the general public consider to be more or less important..

4. **Do you know if there has been any change in the way patients or the general public value the outcomes of interest of the interventions included in the CPG?**

- ☐ Yes
- ☐ No
- ☐ Does not know/ Does not answer

If that is the case, please specify:

5. Are you aware of any relevant new studies published on how patients or the general public assess the outcomes of interest of the interventions included in the guide?

- ☐ Yes
- ☐ No
- ☐ Does not know/ Does not answer

If that is the case, please specify:

Differences in the use of resources and costs of interventions

The evaluation of the differences in the use of resources and costs of the interventions refers to the impact that these differences can have in the development of recommendations.

6. Do you know if there have been any changes in the use of resources and costs of the interventions included in the guide (such as the termination of a patent for a medicine)?

- ☐ Yes
- ☐ No
- ☐ Does not know/ Does not answer

If that is the case, please specify:

7. Are you aware of new published economic evaluations of the interventions included in the guide (such as cost-effectiveness analyses)?

- ☐ Yes
- ☐ No
- ☐ Does not know/ Does not answer

If that is the case, please specify:

Equity

Sometimes the recommended interventions can give rise to inequities in certain subgroups of the population. These inequities are a consequence of the impact of the intervention on a population with individuals that have characteristics associated with a disadvantage (e.g. according to economic status, type of job or occupation, education, place of residence, gender or race, among others).

8. Do you know if there has been any change in equity due to the implementation of the interventions included in the guideline?

- ☐ Yes

- ☐ No
- ☐ Does not know/ Does not answer

If that is the case, please specify:

Acceptability

An intervention can be considered more or less acceptable to the different parties involved (such as health service administrators, health providers or users), due to their opinions and / or experiences regarding the intervention and the relative importance they attribute to its consequences.

9. **Do you know if there has been any change in the acceptability of the interventions included in the guide by stakeholders?**
   - ☐ Yes
   - ☐ No
   - ☐ Does not know/ Does not answer

   If that is the case, please specify:

Feasibility

An intervention may be more or less feasible to implement (that is, capable of being implemented or applied). The feasibility of implementing an intervention may be related to the nature of the intervention itself, to the health providers involved in its application, or users and other stakeholders, to elements of the health system, or to social and political factors.

10. **Do you know if there have been any changes in the feasibility of implementing the interventions included in the guideline?**
    - ☐ Yes
    - ☐ No
    - ☐ Does not know/ Does not answer

    If that is the case, please specify:

On-going research projects

11. **Do you know of any research project currently in progress whose results have not been published yet, that is relevant and is related to any of the previous aspects included in this questionnaire (efficacy, safety, patient values, resource use and costs, equity, acceptability or feasibility)?**
    - ☐ Yes
    - ☐ No
    - ☐ Does not know/ Does not answer

    If that is the case, please specify:

European Reference Networks

European Reference Networks | ERN GUIDELINES

European Commission